

International Society for Knowledge Organization  
ISKO España – Portugal



XII Congreso ISKO España  
II Congreso ISKO España-Portugal

**Organización del conocimiento:  
sistemas de información abiertos**

ACTAS



UNIVERSIDAD DE  
MURCIA

Facultad de Comunicación  
y Documentación  
19 y 20 de noviembre de 2015

**CENTUM**  
CIEN AÑOS DE LA UNIVERSIDAD DE MURCIA  
1915 | 2015



International Society for Knowledge Organization  
ISKO España – Portugal

# II Congreso ISKO España-Portugal XII Congreso ISKO España

Organización del conocimiento:  
sistemas de información abiertos

ACTAS

Murcia  
Universidad de Murcia  
2015

**Ficha técnica:**

Título: *Organización del conocimiento: sistemas de información abiertos*

Autor: II Congreso ISKO España y Portugal / XII Congreso ISKO España

Editores: José Vicente Rodríguez Muñoz, Isidoro Gil Leiva, Pedro Manuel Díaz Ortuño, Francisco Javier Martínez Méndez.

Edición: Facultad de Comunicación y Documentación. Universidad de Murcia.

ISBN: 978-84-608-3558-5

UNIVERSIDAD DE  
MURCIA



# ISKO – Ibérico 2015 Murcia

---

## XII Congreso ISKO España / II Congreso ISKO España-Portugal Universidad de Murcia, 19 y 20 de noviembre de 2015.

**Lema:** Organización del conocimiento: Sistemas de información abiertos

---

*Si aceptamos al conocimiento como el sujeto que trabaja para observar e interpretar la realidad de los objetos, esto es, los datos, podemos interpretar a la información como el elemento catalizador que permite alcanzar un nuevo orden, dicho de otro modo, como un proceso que transforma la incertidumbre en realidad, el caos en orden. En este contexto, para la sostenibilidad de cualquier sistema, entendido este en su expresión más amplia, se vislumbra como necesario que su estado sea el de un sistema abierto, de forma que permita el libre intercambio y por ende, una evolución sostenida.*

### Hitos

---

- Modelos de organización del conocimiento. Modelaje de sistemas y ciclo de vida de la información.
- Entornos y sistemas para la organización del conocimiento. Diseño, normalización e interoperabilidad de sistemas de información abiertos.
- Instrumentos para el acceso y transparencia de la información: gestión de documentos, plataformas digitales y archivos.
- Usuarios y usos de la información. Patrones vinculados a la transparencia y el gobierno abierto.
- Gobernanza y Auditoría de información en la organización del conocimiento.
- Representación del conocimiento: procesos y herramientas en sistemas abiertos y multilingües.
- Epistemología en la organización del conocimiento.
- Igualdad, género y diversidad en la organización del conocimiento.

### Coordinadores Generales

---

- José Vicente Rodríguez Muñoz. *Universidad de Murcia.*
- Rosa San Segundo Manuel. *Universidad Carlos III de Madrid.*

### Comité Organizador

---

- Isidoro Gil Leiva. *Universidad de Murcia.*
- Francisco Javier Martínez Méndez. *Universidad de Murcia.*
- Pedro Manuel Díaz Ortuño. *Universidad de Murcia.*
- José Antonio Gómez Hernández. *Universidad de Murcia.*
- José Antonio Frías Montoya. *Universidad de Salamanca.*
- Rosana López Carreño. *Universidad de Murcia.*
- Juan Antonio Pastor Sánchez. *Universidad de Murcia.*
- Tomás Saorín Pérez. *Universidad de Murcia.*

---

**Comité Científico**

---

- Emilia Currás Puente. *Miembro de Honor ISKO. España.*
- Adolfo Alonso Arroyo. *Universidad de Valencia.*
- María Luisa Alvite Díaz. *Universidad de León.*
- Virginia Bentes Pinto. *Universidade Federal do Ceará, Brasil.*
- María da Graça de Melo Simões. *Universidade de Coimbra, Portugal.*
- Carlos Cândido de Almeida. *Universidade Estadual Paulista, Brasil.*
- Dulce Amélia de Brito Neves. *Universidade Federal da Paraíba, Brasil.*
- Vera Lucia Doyle Louzada de Mattos Dodebei. *Universidade Federal do Estado do Rio de Janeiro, Brasil.*
- Ana Extremeño Placer. *Universidad de Alcalá.*
- José Pablo Gallo León. *Universidad de Murcia.*
- Jesús Gascón García. *Universitat de Barcelona.*
- Blanca Gil Urdiciáin. *Universidad Complutense de Madrid.*
- Ricardo César Gonçalves Santana. *Universidade Estadual Paulista, Brasil.*
- José Augusto Chaves Guimarães. *Universidade Estadual Paulista, Brasil.*
- Cristina Herrero Pascual. *Universidad de Murcia.*
- Manuela Moro Cabero. *Universidad de Salamanca.*
- José Antonio Moreira González. *Universidad Carlos III.*
- Luis Miguel Moreno Fernández. *Universidad de Murcia.*
- Antonio Paños Álvarez. *Universidad de Murcia.*
- Fernanda Peset Mancebo. *Universidad Politécnica de Valencia.*
- Antonio Pulgarín Guerrero. *Universidad de Extremadura.*
- Fernanda Antunes Ribeiro. *Universidade de Porto, Portugal.*
- Blanca Rodríguez Bravo. *Universidad de León.*
- Mariângela Spotti Lopes Fujita. *Universidade Estadual Paulista, Brasil.*
- Jualianne Texeira e Silva. *Universidade Federal da Paraíba, Brasil.*
- Jesús Tramullas Saz. *Universidad de Zaragoza.*
- Crispulo Travieso Rodríguez. *Universidad de Salamanca.*

---

# Presentación

---

En primer lugar, felicitarnos todos por haber hecho posible una nueva edición de este Congreso y en particular a los todos los que han participado en la organización del mismo, sin cuyo esfuerzo no se hubiera hecho realidad. Todo ello, con una especial mención a los profesores, alumnos y personal de administración y servicios, por su intenso trabajo.

Y, debemos felicitarnos todos pues dado el contenido temático del congreso significa que nos va a permitir conocer la labor de numerosísimas personas que trabajan en el desarrollo de algo tan importante para el ser humano como son las Ciencias de la Información (bibliotecología, archivología, documentación y todo lo que comprende la gestión de la información y el conocimiento).

Siempre es importante que las personas se reúnan para compartir viejas, actuales y nuevas ideas, pues de seguro que de ese vínculo siempre se sale mejorado como científicos y como personas. Por ello, reivindico los tradicionales métodos epistolares en forma de comunicaciones, así como las exposiciones orales presenciales, nada sustituibles por la tecnología en su versión online. Nada que ver tanto en el modo del mensaje como en el modo de comunicación.

Lo que no quiero dejar de señalar es que además servirá para la exposición de ideas bajo diferentes puntos de vista, para el despliegue de nuevas ideas, para que todos, en definitiva, podamos compartir conocimiento.

Vamos a desarrollar el Congreso bajo el lema: “*Organización del conocimiento, sistemas de información abiertos*”. Puesto que ISKO es organización del conocimiento resulta ocioso señalar que en el actual devenir estamos en un continuo desarrollo de estructuras y artefactos que no sólo sean capaces de tener las capacidades de la legibilidad, sino que también tengan las de la inteligibilidad a modo de mimesis con el pensamiento humano, en la idea del manejo de grandes volúmenes de información y conocimiento que puedan ser usables.

La idea es que la meta de la Ciencia de la Información sea percibir y controlar el conocimiento en sus distintas representaciones. Hoy afrontamos el ciclo de vida de la información desde diversas perspectivas, tanto representacionales como tecnológicas. Y en este asunto, las Tecnologías de la Información y de la Comunicación han sido y están siendo causa y efecto de este fenómeno, como consecuencia de su naturaleza disruptiva.

Tampoco es menos cierto que en el desarrollo e incorporación de las Tecnologías de la Información y de la Comunicación existe un riesgo inherente sobre ellas, lo que podríamos llamar el lado oscuro de la globalización, y que no deja de ser un punto débil que amenaza el horizonte, esto es, el de crear mayores desigualdades entre los seres humanos, desigualdades que es denominada, no sin cierto eufemismo “brecha digital”.

Pero, a mi entender, no creo que resulte banal la segunda parte de ese lema, los sistemas abiertos. Quiero indicar con la idea de abiertos, no sólo la plétora de conceptos, procedimientos, técnicas, etc. que ellos implican y que en ellos están imbricados; entiéndase: software abierto, aplicaciones abiertas, metadatos, datos abiertos y/o enlazados; así como todo aquello que más o menos de forma colateral está relacionado, como, interoperabilidad, visibilidad, usabilidad, accesibilidad, transparencia, etc.

Lo que interesa es resaltar el concepto intrínseco que sostiene la idea de sistema abierto, frente a sistema asilado y cerrado. Frente a estos dos últimos, cuya cuyo final es predecible desde su nacimiento, esto es la muerte, los sistemas abiertos son la incertidumbre, pero la incertidumbre creativa, el caos determinista que ha permitido y permite entre otros fenómenos la biodiversidad.

Es en este extremo donde se manifiesta la importancia de la organización del conocimiento y los sistemas abiertos, pues sólo desde este enfoque, sólo desde esta ideología seremos capaces, ya no sólo de avanzar, sino de seguir creando. ¿Qué es sino la creatividad? La posibilidad de enriquecernos intelectualmente a través de entornos de incertidumbre.

El ser humano ha creado el concepto abstracto de dato como medio para establecer equilibrios con los objetos susceptibles de ser descritos, equilibrios que en los sistemas abiertos no son permanentes, lo que nos obliga a un continuo ajuste para adaptar los escenarios que queremos representar. Unas veces

los adaptamos, en otras no es posible, esto obliga a construir otros nuevos donde volver a un equilibrio, lo que sucede es que estos nuevos son tan diferentes a los iniciales que representan una nueva visión de la realidad. Y este es el gran papel del concepto de abierto.

Es por ello, que entendemos que la sustanciación de la organización del conocimiento y los sistemas abiertos no es vana, pues de ese modo no sólo incorporamos transparencia, elemento central en las sociedades modernas y exigencia incuestionable en la actualidad. Sino que, como tales sistemas abiertos, permitamos que su esencia, la evolución y la diversidad, se haga patente en la superación de las amenazas y debilidades actuales y generemos nuevas fortalezas y oportunidades para que el conocimiento humano sirva a sus fines. No apartemos del horizonte esta idea, integrémosla en todo aquello que nos incumba y no sólo me refiero en lo científico, sino en todas nuestras actividades.

Ahora procede que sean los congresistas los que tomen la palabra con las ideas y propuestas que han traído y que quedan expresadas en el presente documento. Espero y deseo que les sea de lo más gratificante y fructífero posible, que de seguro lo será.

Murcia, 19 de Noviembre de 2015

José Vicente Rodríguez Muñoz

## TABLA DE CONTENIDO

|   |     |
|---|-----|
| Presentación.....   | 7   |
| <b>PONENCIA INAUGURAL</b> .....   | 11  |
| Perfis de Aplicação de Metadados: o quê, porquê, quando e como.....   | 12  |
| <b>Hito 1: Modelos de organización del conocimiento. Modelaje de sistemas y ciclo de vida de la información</b> .....   | 15  |
| O modelo científico no domínio de organização do conhecimento: um estudo a partir dos trabalhos apresentados nos congressos internacionais da ISKO .....          | 16  |
| Catologação sob a perspectiva dos artigos publicados em revistas científicas nacionais avaliados nos estratos B1, B2 e B3 no sistema Qualis a partir de 1998..... | 24  |
| O profissional da informação na elaboração de políticas públicas de informação e tecnologia .....   | 37  |
| A Informação Contábil na construção do conhecimento organizacional: uma modelagem a partir da DRE .....   | 41  |
| Aplicação de reengenharia de tesauro: modelagem do THESAGRO .....   | 50  |
| The Future of Knowledge Organization and Information Organization .....   | 61  |
| HotWheels: catalogação e recuperação da informação em miniaturas de carros colecionáveis .....  | 67  |
| <b>Hito 2: Entornos y sistemas para la organización del conocimiento. Diseño, normalización e interoperabilidad de sistemas de información abiertos.</b> .....    | 82  |
| Bibliotecas eclesiásticas de Castilla y León. Visibilidad y acceso web.....   | 83  |
| Modelagem de dados para representação descritiva no domínio da História da Arte: uma proposta para a construção de um banco de dados .....                        | 96  |
| OntoSmart: proposta de um modelo de recuperação de informação baseado em ontologia .....  | 105 |
| Estudo da comunicabilidade da interface e do comportamento de busca da informação no Portal de Periódicos da CAPES .....  | 116 |
| Elaboração de um vocabulário taxonómico sobre História Medieval portuguesa: problemas e desafios .....  | 130 |
| Apropriação do Conhecimento na Web de Dados - Gestão de Recursos Computacionais em Sistemas Abertos ...   | 138 |
| Organización del conocimiento en entornos wiki: una experiencia de organización de información sobre lecturas académicas .....                                    | 152 |
| Sistema de informação aberto como recurso de organização do conhecimento e de mediação da aprendizagem..  | 160 |
| <b>Hito 3: Instrumentos para el acceso y transparencia de la información: gestión de documentos, plataformas digitales y archivos</b> .....                       | 172 |
| A justiça na <i>web</i> : estudo sobre a gestão da informação em processos judiciais eletrónicos no Brasil.....   | 173 |
| Abordagem sobre vocabulários controlados para arquivos: conceitos, aplicações e metodologias.....   | 187 |
| Clasificación y indización de la información en Registros y Archivos de Gestión de las universidades .....  | 195 |
| La búsqueda de información en herramientas de descubrimiento y OPAC's: fortalezas y debilidades.....  | 206 |
| Gestão do Conhecimento nas Organizações de Ensino: recorrência de registros na <i>Web of Science</i> .....  | 220 |
| <b>Hito 4: Usuarios y usos de la información. Patrones vinculados a la transparencia y el gobierno abierto</b> ....   | 233 |
| Competências infocomunicacionais para acesso e uso da informação.....   | 234 |
| <b>Hito 5: Gobernanza y Auditoría de información en la organización del conocimiento.</b> .....   | 243 |
| Uso de taxonomia sobre privacidade para identificação de atividades encontradas em termos de uso de redes sociais .....   | 244 |
| <b>Hito 6: Representación del conocimiento: procesos y herramientas en sistemas abiertos y multilingües</b> ...   | 257 |
| Aplicabilidad del mapa conceptual en la representación del conocimiento registrado en las historias clínicas del paciente.....                                    | 258 |
| Armonización de ISO 25964 y SKOS en software de gestión de tesauros. Propuesta de análisis .....  | 266 |
| Organização do conhecimento e tratamento da informação: analisando a subjetividade no processo de indexação e implicações na recuperação da informação.....       | 279 |
| Metadatos en noticias: un análisis internacional para la representación de contenidos en periódicos.....  | 290 |
| Perspectivas metodológicas para a representação da informação do documento de arquivo .....   | 304 |
| Construção de macroestrutura de categorias de linguagem documentária em economia: proposta de inovação metodológica.....  | 319 |

|  |            |
|--|------------|
| Política de indexação para bibliotecas universitárias: um estudo com docentes da área de Organização e Representação do Conhecimento.....  | 332        |
| Compatibilidade entre linguagens documentais para construção, atualização e adequação de vocabulário de bibliotecas universitárias.....  | 345        |
| Organização e Representação do Conhecimento sobre Seca no Brasil: o caso do Núcleo Temático da Seca e do Semiárido da UFRN.....  | 357        |
| Organização da Informação em Contexto Dinâmico e Sensível: Tesouro de Inteligência Policial.....   | 362        |
| Terminología del huipil triqui.....  | 373        |
| Representação de Contexto em Acervos Digitais.....   | 387        |
| Aplicação da Teoria da Análise Facetada e Mapas Conceituais para navegação facetada em uma ontologia de domínio: resultados preliminares.....  | 398        |
| Abordagens cognitivas da ciência da informação no Brasil: mapeamento conceitual.....   | 411        |
| Organização da Informação no ambiente Web: produzindo conhecimento a partir de grandes volumes de dados.....   | 419        |
| A representação da imagem do bibliotecário nas histórias em quadrinhos (HQs): um estudo descritivo.....  | 430        |
| Estudo da Produção e Divulgação Científica do Instituto Federal de Educação, Ciência e Tecnologia de São Paulo: Uso do software livre Script Lattes.....   | 438        |
| Cognição, neuromarketing e neuroeconomia: entendendo as decisões dos atores baseados em informações.....   | 446        |
| Organização do conhecimento arquivístico: um estudo terminológico comparativo (português, espanhol, francês, inglês) sobre classificação e descrição na <i>Multilingual Archival Terminology – ICA</i> ..... | 461        |
| Políticas de indización en bibliotecas escolares de Brasil y Portugal: análisis comparativo.....   | 470        |
| Categorización de revistas científicas españolas en las bases de datos ISOC.....   | 485        |
| <b>Hito 7: Epistemología en la organización del conocimiento.....</b>  | <b>497</b> |
| Aportaciones Semióticas de la Documentación en España a la Organización Del Conocimiento: un análisis preliminar.....  | 498        |
| Referentes Teóricos em Organização do Conhecimento: uma Análise de Domínio na ISKO Espanha (2005-2013).....  | 508        |
| Referentes e interlocuções teóricas em análise documental no contexto brasileiro de Ciência da Informação.....   | 518        |
| Exaustividade e especificidade na indexação: uma análise de conteúdo como perspectiva de investigação sobre o tema.....  | 527        |
| La construcción de la Biblioteconomía como ciencia y su relación con la clasificación.....   | 533        |
| A dimensão conceitual da organização do conhecimento no universo científico da ISKO: uma análise de domínio a partir dos congressos de ISKO-Brasil e ISKO- Espanha.....                                      | 544        |
| Organização da informação e do conhecimento no contexto da Ciência da informação: da análise terminológica à reflexão epistemológica.....  | 556        |
| Referentes teóricos basilares na ISKO-Ibérico e suas contribuições para a Organização do Conhecimento (OC).....  | 570        |
| Organização e representação da informação em arquivos: uma análise a partir da função classificação.....   | 579        |
| O “nome próprio” no contexto da Epistemologia da Organização do Conhecimento: um debate filosófico-teórico.....  | 584        |
| <b>Hito 8: Igualdad, género y diversidad en la organización del conocimiento.....</b>  | <b>597</b> |
| Sesgo y universalidad: un enfoque histórico-conceptual.....  | 598        |
| Interações entre Semiótica da Cultura e Organização do Conhecimento: conceitos integradores.....   | 603        |
| Información y crianza con apego en España.....   | 618        |
| Formação discursiva, sistemas de classificação e o discurso estrangulado de Charlie Hebdo: impactos, desdobramentos e implicações na representação da informação da ciência da religião no século XXI.....   | 629        |
| <b>Presentación comercial.....</b>   | <b>636</b> |
| TesaurVAL: software para la creación y gestión de tesauros.....  | 637        |
| <b>Índice de autores.....</b>  | <b>638</b> |

# PONENCIA INAUGURAL

---

# Perfis de Aplicação de Metadados: o quê, porquê, quando e como

*Metadata Application Profiles: what, why, when and how*

---

**Prof. Dra. Ana Alice Baptista**

Universidade do Minho, Portugal. analice@dsi.uminho.pt

## Resumo extendido

Atualmente existem aplicações de metadados por todo o mundo. Algumas das mais conhecidas são as dos Repositórios Digitais (e.g., DSpace e Eprints) e as dos sistemas de gestão de periódicos e conferências científicos (e.g. OCS e OJS). No entanto, apesar de serem sintaticamente interoperáveis, muitas destas aplicações descumpram a interoperabilidade semântica. Por exemplo, há diversos casos em que a semântica dos elementos de metadados (ou propriedades) não é respeitada, relacionando-os com valores que não se enquadram no previamente estabelecido ou nas boas práticas recomendadas. No caso-exemplo da Figura 1, Figura 2 e Figura 5 (Repositórios da Universidade do Minho, do IBICT, e da Universidade nacional da Austrália, respetivamente) associam-se ao elemento `relation` e `identifier` valores que não são identificadores, conforme a boa prática recomendada (c.f. Dublin Core Metadata Initiative Usage Board, 2012; Dublin Core Metadata Initiative, 2008).

```
<dc:relation>
  Documentos de Trabalho = Working Papers / NEGE ; 4/2003
</dc:relation>

<dc:relation>
  Working Papers / NEGE ; 4/2003
</dc:relation>
```

Figura 1 - Valores associados ao elemento `relation` no Repositório da Universidade do Minho.

```
<dc:type>
  article
</dc:type>

<dc:identifier>
  Pesquisa Brasileira em Ciência da Informação e Biblioteconomia, v.1,n.1, 2006.
</dc:identifier>

<dc:identifier>
  1981-0695
</dc:identifier>
```

Figura 2 – Elemento `type` e valores associados ao elemento `identifier` no repositório do IBICT.

Noutros casos, por exemplo, valores muito específicos são associados a elementos cuja semântica é demasiado lata. No caso-exemplo da Figura 3 (Universidade de Cambridge), existem quatro elementos `date` aos quais estão associados dois valores distintos, sem outra informação que permita distinguir a semântica destas datas, i.e., a que tipo de data se refere cada um daqueles valores (data de depósito? Data de publicação?...)

```
<dc:date>
  2004-05-10T12:48:42Z
</dc:date>

<dc:date>
  2004-05-10T12:48:42Z
</dc:date>

<dc:date>
  2002-12-19 15:45:00
</dc:date>

<dc:date>
  2004-05-10T12:48:42Z
</dc:date>
```

Figura 3 - Repetição de elementos `date` sem refinamento semântico no repositório da Universidade de Cambridge.

Na maioria dos casos, no entanto, os próprios valores não se conseguem interpretar e, por isso, não são interoperáveis, relacionáveis nem processáveis em conjunto. Por exemplo, na Figura 4, como deverão ser interpretados os valores E5122 ou u9115357 associados ao elemento `contributor` (Repositório da Universidade Nacional da Austrália)? Ou qual a diferença entre “article” e “Journal article” (Repositório do IBICT e da Universidade Nacional da Austrália) - Figura 2 e Figura 5? O que significa o valor relacionado com o elemento `publisher` (repositório do POSTECH, República da Coreia)? Como devem estes valores ser processados?

```

<dc:contributor>
  E5122
</dc:contributor>

<dc:contributor>
  u9115357
</dc:contributor>

<dc:contributor>
  E2888
</dc:contributor>

<dc:contributor>
  E30351
</dc:contributor>

```

Figura 4 - Valores associados ao elemento *contributor* no repositório da Universidade Nacional da Austrália.

```

<dc:type>
  Journal article
</dc:type>

<dc:type>
  Published Version
</dc:type>

<dc:identifier>
  Optics Express 15.9 (2007): 5730-5741
</dc:identifier>

<dc:identifier>
  1094-4087
</dc:identifier>

<dc:identifier>
  http://hdl.handle.net/10440/322
</dc:identifier>

```

Figura 5 - Alguns elementos e valores a eles associados no repositório da Universidade Nacional da Austrália.

```

<dc:publisher>
  포항공과대학교
</dc:publisher>

```

Figura 6 – Valor associado ao elemento *publisher* no repositório do POSTECH (República da Coreia).

Algumas destas questões resolvem-se através da utilização e combinação de elementos de metadados oriundos de diferentes esquemas. Outras resolvem-se através da obediência estrita a normas. Outras, ainda, através da identificação clara dos esquemas de codificação utilizados e, em muitos casos, da substituição de valores literais por valores não literais, em particular IRI (Internationalized Resource Identifier).

O paradigma de base de muitas das aplicações de metadados é o LOD (Linked Open Data – Dados Abertos Ligados), no qual a interoperabilidade é vital. De modo a maximizar a interoperabilidade destas aplicações sem perder acuidade semântica, é necessário obedecer a normas produzidas por organismos próprios, a normas de facto, ou ter em conta o anteriormente produzido no âmbito de outras aplicações no mesmo domínio ou em domínios relacionados. Assim, as aplicações resultantes poderão vir a utilizar elementos de metadados oriundos de vários esquemas de metadados adaptados e otimizados para os seus casos específicos. É neste contexto que surge o conceito de Perfil de Aplicação de Metadados (PAM).

Segundo Heery & Patel (2000) um PAM é um esquema que consiste em “elementos de dados procedentes de um ou mais namespaces, combinados pelos implementadores e otimizados para uma aplicação local em particular”. O Enquadramento de Singapura (Nilsson, Baker, & Johnston, 2008) fornece uma visão mais global, considerando que um PAM descreve “o conjunto de linhas orientadoras, regras de descrição e restrições usadas para a criação de um conjunto específico de registos de metadados”. Neste âmbito, um PAM é definido como um pacote de documentação constituído pelos seguintes componentes: Requisitos Funcionais, Modelo de Domínio, Perfil de Conjuntos de Descrições, Guias de Utilização e Guias de Codificação, sendo os três primeiros obrigatórios e os dois últimos opcionais.

O Enquadramento de Singapura define os componentes de um PAM e estabelece alguma ordem no seu desenvolvimento ao identificar quais se baseiam em quais, mas não define nenhum método nem técnicas para o desenvolvimento de cada um desses componentes. Algumas informações mais detalhadas sobre o processo são veiculadas em duas outras recomendações da Dublin Core Metadata Initiative (DCMI): “Guidelines for Dublin Core Application Profiles” (Coyle & Baker, 2009) e “Description Set Profiles: A constraint language for Dublin Core Application Profiles” (Nilsson, 2008). Outras recomendações relacionadas podem ser encontradas no site da DCMI. No entanto, nenhuma destas recomendações é suficientemente detalhada quanto aos métodos e técnicas a utilizar.

Até agora a comunidade de metadados tem desenvolvido PAM ora utilizando o bom senso, ora adaptando partes de métodos e técnicas oriundos de outros contextos, em particular os do desenvolvimento de software. O esforço de desenvolvimento de métodos e técnicas especí-

ficos para o desenvolvimento de PAM começa agora a ver a luz do dia.

As duas palestras keynote apresentadas na conferência DC-2015 (International Conference on Dublin Core and Metadata Applications), realizada recentemente no Brasil, centraram-se no processo de desenvolvimento de PAM. Adicionalmente, esta conferência contou com a realização de um workshop, várias apresentações e posters sobre PAM. Não tendo sido o tema base pré-definido para esta edição da conferência, os PAM foram, contudo, o grande tema transversal, o que parece ser um indicador da adesão da comunidade de metadados ao tema.

As perspetivas sobre o processo de desenvolvimento de PAM apresentados no âmbito das duas palestras keynote são radicalmente diferentes. Uma dirige-se a PAM num domínio algo vasto, que abrange comunidades diversas a vários níveis (linguísticos, culturais, sociais, geográficos...) e em que as preocupações com a interoperabilidade são fundamentais. No âmbito desta palestra foi apresentado genericamente o método Me4MAP (Method for the Development of Metadata Application Profiles) (Malta & Baptista, 2013; Malta, 2014), métodos baseados no Rational Unified Process (RUP), da engenharia de software. A outra perspectiva debruça-se sobre aplicações de metadados específicas no âmbito de uma comunidade com fronteiras bem definidas e em que a interoperabilidade, sendo desejável, não é fundamental. O palestrante apresentou o RIOXX, um perfil de aplicação cujo desenvolvimento se inspirou nos métodos ágeis da engenharia de software (Walk, 2015).

Ambas as abordagens, embora com perspetivas diferentes, são interessantes e assumem-se como primeiros esforços da comunidade no sentido de sistematizar o conhecimento sobre o processo de desenvolvimento de PAM adquirido até à data. É importante que a comunidade de metadados participe neste esforço, testando devidamente estes métodos e dando retorno atempado aos investigadores envolvidos no seu desenvolvimento. É, também, importante que a comunidade de metadados, especialmente a Ibero-Americana, se envolva ativamente no esforços que estão a ser feitos a nível internacional, especialmente os relacionados com a DCMI, em cujo sítio Web se poderá encontrar informação sobre comunidades, grupos de trabalho, presença nas redes sociais, entre outros.

Esta palestra incidirá sobre os PAM: o que são, porque são necessários, quando se devem desenvolver e como se podem desenvolver. Rela-

tivamente a este último ponto, apresentar-se-ão e contextualizar-se-ão as tendências de processo de desenvolvimento que atualmente emergem na comunidade de metadados.

## Referencias

- Coyle, K., & Baker, T. (2009). *Guidelines for Dublin Core Application Profiles*. Retrieved from <http://dublincore.org/documents/2009/05/18/profile-guidelines/>
- Dublin Core Metadata Initiative. (2008). *Dublin Core Metadata Element Set, Version 1.1*. Retrieved from <http://dublincore.org/documents/dces/>
- Dublin Core Metadata Initiative Usage Board. (2012). *DCMI Metadata Terms. Dublin Core Metadata Initiative*. Retrieved from <http://dublincore.org/documents/dcmi-terms/>
- Heery, R., & Patel, M. (2000). Application profiles: mixing and matching metadata schemas. *Ariadne*, (25). Retrieved from <http://www.ariadne.ac.uk/issue25/app-profiles/>
- Malta, M. C. (2014). *Contributo metodológico para o desenvolvimento de perfis de aplicação no contexto da Web Semântica*. University of Minho. Retrieved from <http://hdl.handle.net/1822/30262>
- Malta, M. C., & Baptista, A. A. (2013). A Method for the Development of Dublin Core Application Profiles (Me4DCAP V0.2): detailed description. In Muriel Foulonneau & K. Eckert (Eds.), *International Conference on Dublin Core and Metadata Applications*. Lisbon, Portugal: Dublin Core Metadata Initiative.
- Nilsson, M. (2008). *Description Set Profiles: A constraint language for Dublin Core Application Profiles*. Retrieved from <http://dublincore.org/documents/2008/03/31/dc-dsp/>
- Nilsson, M., Baker, T., & Johnston, P. (2008). *The Singapore Framework for Dublin Core Application Profiles*. Retrieved from <http://dublincore.org/documents/singapore-framework/>
- Walk, Paul. (2015, September). Exploiting the value of Dublin Core through pragmatic development. Slides of the Keynote Speech presented at the *International Conference on Dublin Core and Metadata Applications, São Paulo, Brazil*. Retrieved from <http://www.paulwalk.net/exploiting-the-value-of-dublin-core-through-pragmatic-development/>

**Hito 1: Modelos de organización del conocimiento. Modelaje de sistemas y ciclo de vida de la información**

**Hito 2: Entornos y sistemas para la organización del conocimiento. Diseño, normalización e interoperabilidad de sistemas de información abiertos.**

**Hito 3: Instrumentos para el acceso y transparencia de la información: gestión de documentos, plataformas digitales y archivos**

**Hito 4: Usuarios y usos de la información. Patrones vinculados a la transparencia y el gobierno abierto**

**Hito 5: Gobernanza y Auditoría de información en la organización del conocimiento.**

**Hito 6: Representación del conocimiento: procesos y herramientas en sistemas abiertos y multilingües**

**Hito 7: Epistemología en la organización del conocimiento**

**Hito 8: Igualdad, género y diversidad en la organización del conocimiento**

---

# O modelo científico no domínio de organização do conhecimento: um estudo a partir dos trabalhos apresentados nos congressos internacionais da ISKO

*Scientific model in knowledge organization domain: a study from the papers presented at ISKO's international congresses*

---

**Dra. Aline Elis Arboit (1), Prof. Dra. Ely Francina Tannuri de Oliveira (2), Prof. Dra. Leilah Santiago Bufrem (3), Dr. Rene Faustino Gabriel Junior (4)**

- (1) Analista de Controle do Tribunal de Contas do Estado do Paraná, Praça Nossa Senhora da Salete, s/n, CEP 80.530-910 - Curitiba – Brasil, aline.arboit@tce.pr.gov.br
- (2) Pesquisadora 2 do CNPq e Livre-docente do Programa de Pós-Graduação em Ciência da Informação da UNESP, Av. Hygino Muzzi Filho, 737 ? Cx.P. 181, CEP 17525-900, Marília, Brasil, etannuri@gmail.com
- (3) Pesquisadora 1 do CNPq e Docente do curso de Pós Graduação em Ciência da Informação da Universidade Estadual Paulista (UNESP) - Av. Hygino Muzzi Filho, 737 ? Cx.P. 181, CEP 17525-900, Marília, Brasil, e da Universidade Federal de Pernambuco, santiagoobufrem@gmail.com
- (4) Docente e Pesquisador da Pontifícia Universidade Católica do Paraná (PUCPR)-Estado do Paraná- Rua Imaculada Conceição, 1155, CEP: 80215-901 – Curitiba - Brasil, renefgj@gmail.com

## Resumo

As pesquisas em Organização do Conhecimento (OC) e demais áreas do saber são determinadas tanto pelo contexto social e histórico quanto pelos marcos teóricos e metodológicos orientadores de conceitos e pressupostos científicos. A partir dessa concepção, a presente comunicação apresenta uma discussão teórica sobre o conceito e a utilização da palavra “modelo”, realizada com base em análise de um corpus composto por 51 trabalhos apresentados nas Conferências Internacionais da *International Society for Knowledge Organization* (ISKO) que ocorrem bianualmente entre 1990 e 2014. O corpus foi delimitado a partir da seleção de trabalhos categorizados por enunciados de subtemas relacionados ao termo “modelo” e suas variações. Tais enunciados foram formulados pelos pesquisadores que integraram os grupos responsáveis pela organização de cada edição do evento em questão. Destaca-se a contribuição de Badiou (2007), em que o autor propõe uma noção de modelo considerando a proposta do positivismo lógico e do empirismo, introduzindo uma nova relação ciência e ideologia. Como etapas do procedimento metodológico foram realizadas as seguintes ações: análise dos resumos e das palavras-chave e, na sequência, cotejamento desse conteúdo com os conceitos de modelo, conjunto então relacionado com as opções metodológicas dos estudos, a fim de compreender as opções teóricas dos autores, permitindo reconhecer as características dos modelos propostos. Foram categorizados, assim, os

trabalhos constituintes do corpus em “modelos conceituais” e “modelos computacionais” como tentativa de representar a realidade impressa nos eventos da ISKO.

**Palavras chave:** Conceito de Modelo, Modelos em OC, Produção científica em OC, ISKO.

## Abstract

Research in Knowledge Organization (KO) and other areas of knowledge are determined by both social and historical contexts and theoretical and methodological frameworks of guiding concepts and scientific assumptions. In this scenario, this study presents a theoretical discussion on the concept and use of the word “model”, carried out on the analysis of a corpus composed by 51 papers presented at the International Conferences of *International Society for Knowledge Organization* (ISKO) held biannually between 1990 and 2014. The corpus was delimited from the selection of papers categorized by subtheme statements related to the term “model” and its variations. Such statements were formulated by the researchers integrating the groups responsible for organizing each edition of the event. Badiou (2007) is highlighted as the author proposes a notion of model considering the proposal of logical positivism and empiricism, introducing a new relationship between science and ideology. As methodological steps, the following were performed: analysis of abstracts and keywords, and subsequently, investigation of the content on the concepts of model. This set was then related with the

methodological options of studies in order to understand the authors' theoretical options, which allowed to recognize the characteristics of the proposed models. Thus, we characterized the papers constituting the corpus in "conceptual models" and "computational

models" as an attempt to represent reality in ISKO's events.

**Keywords:** Concept of Model, Models in KO, Scientific production in KO, ISKO.

## 1. Contextualizando a pesquisa

Indagações sobre os modos que norteiam a pesquisa em Organização do Conhecimento (OC) suscitam a necessidade de construir uma análise não reducionista, mas crítica e multidimensional para compreender os aspectos componentes das pesquisas. Isso porque a atividade científica requer planejamento e atividades coerentes aos fins que a orientam e, como consequência, as produções resultantes dessas atividades descrevem modalidades científicas pertinentes às propostas e aos domínios de conhecimentos específicos, portanto, diferenciando-se de outras áreas nas suas expressões concretas.

Considerando que a pesquisa em OC, assim como ocorre em todas as demais áreas científicas, resulta de um contexto dinâmico de produção e reelaboração de conhecimentos, enfatiza-se aqui que este contexto se transforma continuamente em razão da sua aplicação concreta e da comunicação científica. Ao reorganizar-se constantemente na tentativa de sistematizar conceitos e, desse modo, interferir na realidade, a OC contribui para a dinâmica dos demais domínios na medida em que estes dependem de informações e conhecimentos, elementos cruciais para incessante construção e revisão do conhecimento científico. Tal demanda, acentuada pela crescente ramificação e inter-relação dos saberes, requer processos mais complexos para o estudo de seus objetos, pois as atividades de pesquisa são determinadas tanto pelo contexto social e histórico quanto pelos marcos teóricos e metodológicos orientadores de conceitos e pressupostos científicos.

O histórico de trabalhos referentes a pesquisas concretas que vêm sendo realizadas no campo específico da OC é compreendido como resultado de buscas intencionais por respostas a questionamentos sobre a realidade. Uma das constantes discussões presentes no cenário da literatura é a característica interdisciplinar da área, que torna nebulosas as fronteiras entre domínios e não raro discutível sua sistematização, especialmente a partir dos trabalhos apresentados nas Conferências Internacionais da *International Society for Knowledge Organization* (ISKO) (Zherebchevsky, 2010; Friedman, 2010; Arboit, Grácio, Oliveira e Bufrem, 2012, Arboit, 2014). Extrapolando a própria ISKO,

atualmente, o estudo de seu domínio tem sido foco de várias pesquisas, tais como: Guimarães, Santos, Cândido e Pinho (2014), Smiraglia (2014) e Arboit (2012), entre outros, que se propõem a definir e conceituar seu escopo, seu objeto de estudo, seus fundamentos teóricos e, principalmente, delinear metodologias que melhor possam compreender o conhecimento construído. Entende-se a ISKO e os conteúdos produzidos e registrados durante as realizações do seu principal evento como representativos da área em função do papel de centralidade que a entidade adquire.

Apresentando concepções teóricas diversas por parte dos autores, caracterizações e tipologias, em suas modalidades de aplicação, destaca-se aqui uma ampliação do referencial semântico relativo ao conceito de "modelo" e de suas possibilidades de aplicação, razão da curiosidade acadêmica que preside este estudo. Em relação aos trabalhos metodológicos, encontrados em corpus específico da área de OC (Freitas, Nascimento e Bufrem, 2014), percebeu-se um aprofundamento teórico mais intenso em relação ao objeto estudado, como propostas de criação de modelos, metodologias e estruturas para a organização do conhecimento em contextos diversos, o que representa um indício de amadurecimento da pesquisa, graças ao estudo das relações específicas entre variáveis que envolvem os estudos. Em alguns casos, os estudos metodológicos remetem à necessidade de conhecimento sobre tecnologias da informação e comunicação, quando não são exclusivamente modelos conceituais. Mas os modelos também são encontrados em estudos avaliativos, cujo foco pode ser um objeto, serviço, fluxo, ou unidade de informação apresentando critérios ou parâmetros a partir de princípios de avaliação.

Em tese sobre as modalidades de pesquisa e opções concretas de dissertações da área de CI, verificou-se que, quanto às técnicas e instrumentos mais utilizados por autores de estudos de caráter exploratório-metodológico, destacam-se os dados colhidos anteriormente ou registros existentes e, em segundo plano, a entrevista, a utilização de modelo e o questionário (Bufrem, 1996).

As pesquisas citadas, seguidas de leituras exploratórias realizadas, originaram a curiosidade dos autores no sentido de mais especificamente

reconhecerem noções acerca dos conceitos de “modelos” empregados no domínio, percebendo-se que, tanto como objeto, foco de estudo ou como instrumento, os modelos se fazem presentes com mais intensidade nos trabalhos de pesquisadores de OC. O conseqüente questionamento sobre os tipos de modelos presentes, seja como objeto, seja como instrumento de pesquisa, nos trabalhos apresentados nas Conferências Internacionais da ISKO e registrados nos seus *proceedings* encontrou corpus suficiente para o estudo e, portanto, capaz de responder às expectativas dos objetivos da pesquisa.

Com foco na palavra “modelo” e suas utilizações na OC, entende-se relevante a discussão de alguns aspectos da natureza e do papel desempenhado pelos modelos no domínio, além de discutir suas relações com as teorias em meio a diferentes enfoques do conhecimento científico. Esta comunicação apresenta uma discussão teórica sobre o assunto, proporcionando reflexões a respeito do conceito e da utilização de “modelos”, seguida de uma análise de um corpus representativo do domínio específico da OC, composto por trabalhos apresentados nas Conferências Internacionais da ISKO realizadas bianualmente entre 1990 e 2014.

Tentando superar o tom superficial que poderia ser dado a uma definição dicionarizada, pela qual o termo ‘modelo’ pode ser definido como uma representação de um sistema por outro, usualmente mais familiar, cujo funcionamento se supõe ser análogo ao do primeiro (Dicionário Oxford de Filosofia Blackburn, 1997) e ainda, respeitando a complexidade que envolve a aceitação de uma definição terminológica, além da evolução dos significados que o termo adquire historicamente, a contribuição de Badiou (2007) foi a mais significativa para a dimensão teórica desta pesquisa.

Em sua obra editada originalmente em 1969, *Le concept de modèle*, Badiou propõe uma noção de modelo considerando, por um lado, a proposta do positivismo lógico e, por outro, a do empirismo, introduzindo uma nova relação ciência e ideologia. Ao argumentar que a filosofia é a prática de uma relação impossível entre a ciência e a ideologia, Badiou investe contra o empirismo e o idealismo, pois argumenta que ambos são apoiados pela suposta relação transparente entre objeto de estudo e sua representação.

Assim, o autor utiliza o termo modelo como conceito científico do ramo da lógica matemática. A questão epistemológica sobre o uso de modelos surge, no entender do autor, quando se trata de entender a diferença entre o modelo e o real

(empírico). A ciência é construída a partir do enfrentamento com o objeto real, que se pretende investigar e, por outro lado, com o objeto artificial destinado a reproduzir e a imitar o objeto real: a ciência é o confronto entre estes dois objetos.

Ele propõe uma tríplice distinção entre **noções** (*notions*), ou unidades do discurso ideológico, **conceitos** (*concepts*), relativos ao discurso científico e **categorias** (*catégories*) ao discurso filosófico (Badiou, 2007, p. 51).

Ao atribuir à filosofia o papel de recuperação ideológica da ciência, ele se refere a uma intervenção epistemológica (por meio das matemáticas) para expor as ideologias implícitas nas diversas apropriações filosóficas dos conceitos científicos, defendendo sua substituição por categorias materialistas. Procura, com essa posição, superar a distinção entre uma imagem da ciência definida como representação formal do objeto dado e outra, cuja predominância provém da anterioridade dos dispositivos formais, ou código matemático pelo qual o objeto é representado. O autor argumenta que empirismo e formalismo são constituintes de uma conjugação ideológica.

## 2. Objetivos e Metodologia

Esta comunicação objetiva, de forma geral, apresentar uma discussão teórica sobre o tema, por meio de reflexões a respeito do conceito e da utilização de “modelos”, seguida de uma análise de um corpus representado por 51 trabalhos apresentados e publicados nos *proceedings* das Conferências Internacionais da ISKO, entre os anos de 1990 a 2014.

A delimitação deste corpus foi realizada a partir da seleção de trabalhos categorizados por enunciados de subtemas – elaborados pela Sociedade durante a organização de cada evento - que se referissem ao termo “modelo” e suas variações.

Tal opção foi adotada em razão da concepção acolhida pelos autores de que a ISKO, sendo considerada a entidade mais representativa do domínio no âmbito internacional, congrega grupos de pesquisadores que tanto produzem conhecimento como dirigem a produção de conhecimento nesse ambiente, conforme constatações feitas em tese sobre a entidade (Arboit, 2014). Os enunciados temáticos e subtemáticos representariam, portanto, quais os rumos trilhados pelo domínio, decisão esta tomada em função dos interesses de pesquisa do grupo responsável pela direção da entidade, variáveis conforme o contexto espaço-temporal.

Assim, dentro de um universo de 24 anos de conferências foram identificados os subtemas com as respectivas quantidades de trabalhos por eles agrupados na tabela abaixo.

| Ano  | Enunciados subtemáticos  | Qtd. |
|------|--|------|
| 1998 | “Abordagens cognitivas para OC (A): entidades conceituais e modelagem de ontologias” | 3    |
| 1998 | “Modelagem conceitual”   | 4    |
| 1998 | “Modelos computacionais”   | 3    |
| 1998 | “Modelagem de perfil do usuário”   | 3    |
| 2002 | “Modelos e métodos para OC. Ferramentas e sistemas”                                  | 5    |
| 2002 | “Modelos e métodos para OC e relações conceituais”                                   | 6    |
| 2002 | “Modelos e métodos para OC e recuperação”  | 5    |
| 2002 | “Modelos e métodos para representação do conhecimento”                               | 6    |
| 2002 | “Modelos e métodos para integração do conhecimento em sistemas de informação”        | 6    |
| 2008 | “Modelos e métodos em OC”  | 10   |

Tabela I. *Corpus de pesquisa*

Desdobram-se como objetivos específicos:

- analisar o conteúdo das pesquisas descritas nos trabalhos;
- descrever sistematicamente o modo pelo qual são empregados os conceitos de “modelos” e suas variações;
- analisar e tecer inferências sobre os enfoques teórico-metodológicos e sobre os conhecimentos que vêm sendo construídos sobre os diferentes modelos.

### 3. Descrição e análise dos resultados

O universo temático que compõe o *corpus* de análise passou por um conjunto de movimentos de aproximação, análise e interpretação e, então, foi correlacionado com a teoria de Badiou, de modo a haver um alinhamento entre os temas do *corpus* e os princípios teóricos do autor de referência.

Para tanto, foram analisados os resumos e as palavras-chave e cotejados com os conceitos de modelo, conjunto então relacionado com as opções metodológicas dos estudos.

Com esse procedimento, pretendeu-se compreender as opções teóricas e metodológicas dos autores, por meio de um diálogo com o texto, que permite reconhecer as características e

natureza dos modelos, utilizados ou objetos de estudo.

Com efeito, foi necessário estabelecer uma nova categorização relacionada à tipologia de modelos. Verificou-se que os trabalhos selecionados giram em torno de dois enfoques principais: modelos conceituais e modelos computacionais. Isso fica evidenciado nos textos empregados pelos enunciados expostos na Tabela I.

Esses enfoques, por sua vez, foram identificados nos conceitos específicos de cada categoria tendo em vista a linguagem empregada, os objetivos, os objetos de estudo e as opções metodológicas de cada um dos 51 estudos selecionados, conforme dados dispostos na Tabela II no Anexo 1.

Inicialmente, dos 51 trabalhos, foram considerados 34 como pertencentes à categoria “Modelos conceituais” e dezessete à categoria “Modelos computacionais”. Esse resultado demonstra a crescente preocupação em incorporar tecnologias da informação na área, haja vista a quantidade de trabalhos voltados especificamente para a modelagem computacional. Também foi observada a presença significativa de conceitos oriundos da área de tecnologia da informação em trabalhos categorizados com “Modelos conceituais”. A necessidade de absorver a tecnologia da informação na OC aparece como aqui justificada, uma vez que, se percebe os usos dessa tecnologia para promover o melhor compartilhamento e usabilidade de sistemas de conceitos. Foi identificada uma pequena parcela dos 34 trabalhos se destinava a teorizar sobre os modelos conceituais.

Depois de categorizados, estes conceitos foram reunidos conforme a semelhança e similaridade. Como existem conceitos convergentes categorizados tanto como modelo conceitual como modelo computacional, foi estabelecida a relação entre eles por meio de cores, admitindo, desse modo, a existência de ambiguidade em sistemas de OC, acatando o argumento de Garcia Gutiérrez (2007).

Na tipologia “Modelos conceituais” foram identificados oito grupos distintos, enquanto na tipologia “Modelos computacionais” foram inseridos nove.

O primeiro grupo ligado à categoria nomeada “Modelos conceituais”, reúne conceitos relacionados à Ontologia e Tesaurus e encontra-se representado pela cor laranja. Na sequência, em azul claro estão organizados os conceitos ligados pela cor ao sexto grupo da categoria “Modelos computacionais”, por ambos se tratarem de questões terminológicas. O terceiro, de

cor vermelha, agrupa conceitos que fazem referências aos tipos de relações e estruturas conceituais ou entre entidades.

Quanto aos conceitos concernentes a tipos de modelos conceituais já construídos, esses foram utilizados para compor o grupo em amarelo. Já o próximo conjunto, de cor roxa, juntou termos que se referiam a sistemas de organização e representação de informações audiovisuais, digitais e eletrônicas.

O próximo conjunto (preto) une conceitos relacionados ao enfoque semântico. Na mesma cor, porém categorizado em “Modelos computacionais”, encontram-se também termos relativos a esse enfoque, como disposto no segundo grupo dessa categoria.

Em seguida, tem-se o conjunto de cor rosa que agrupa termos concernentes a conceitos ranganathianos como facetas, criando outro vínculo com a segunda categoria onde, no último grupo, há termos da mesma natureza, todavia relacionados a questões de buscas em ambientes eletrônicos e representações em linguagens de programação.

O último grupo, categorizado como modelo conceitual, destacado pela cor verde, é composto por conceitos relativos ao contexto de uso e produção do conhecimento, bem como ao que se refere às questões locais do uso da informação e do desenvolvimento de uma coleção. Este grupo se relaciona com o quinto da mesma cor que foi considerado pertencente à categoria “Modelos computacionais”, que, por sua vez, também aborda questões sobre o modo de produção dos documentos e a relação entre eles.

São seis os grupos inseridos na categoria “Modelos computacionais” onde não foram encontradas relações com os grupos já mencionados. O primeiro dessa categoria em azul escuro se refere a conceitos voltados para o processamento da linguagem natural. Já o terceiro, em cinza claro, congrega conceitos relacionados para construção de redes. O próximo, destacado pela cor marrom, é composto por termos que fazem alusão à manutenção automática de sistemas em função do dinamismo das representações. O sétimo grupo, em cinza escuro, aborda conceito relativos às ferramentas e técnicas de inteligência artificial. O conjunto seguinte, em branco, traz conceitos atinentes a linguagens específicas de representação e transferência de informação.

#### 4. Modelos em OC

Retomando o exposto por Badiou (2007), é possível afirmar que o modelo proposto no Anexo 1

representa um sistema que agrega noções, conceitos e categoria, reunindo desse modo unidades dos discursos ideológico, científico e filosófico.

As categorias estabelecidas com base nos enunciados dispostos na Tabela I “Modelos conceituais” e “Modelos computacionais” são encaradas como unidades do discurso filosófico, ao passo que os conceitos nelas inseridos são tratados como fruto de um discurso científico e as noções adquiridas nesse entorno, provém do discurso ideológico.

São por meio das noções contextualmente adquiridas que os conjuntos e as relações conceituais foram construídos. Tal construção busca estabelecer um sentido entre os termos, os conceitos e as categorias na tentativa de representar para compreender o domínio estudado. Durante esse processo, foram atribuídos valores ao discurso científico e filosófico, razão pela qual não se deve dissociá-los do discurso ideológico, como postula Badiou.

Para se atribuir valor a qualquer unidade do discurso, parte-se de uma postura avaliativa e apreciativa. Porque considerar um conceito como parte de um determinado grupo ou de uma determinada categoria e não de outro ou de ambos? Essa decisão é tomada sempre com base em uma apreciação determinada pela posição de um sujeito ou grupos de sujeitos inseridos em uma coletividade contextualmente localizada. Tal posição, por sua vez, é gradativamente construída e valorada porque é concomitantemente fruto de processos científicos, filosóficos e ideológicos.

A partir das configurações observadas, deduz-se que: a) os modelos resultam de um contexto teórico e uma situação histórica empírica ou conjunto de características e relações que permitem o seu surgimento; b) uma decisão cognitiva posiciona esse modelo como parte de um conjunto de conhecimentos adquiridos e em relação com o que ainda está por se adquirir; c) graças a sua visibilidade e inteligibilidade, é possível conceber o modelo como conjunto de elementos existente em relação a outros modelos pois, se ele advém da historicidade, a primeira inclinação é pensá-lo em relação algo que a ele se compara; d) o modelo é resultante de um procedimento genérico, um cálculo, ou um processo condicionado, o que resulta na imbricação entre o ver e o pensar.

Opções teórico-metodológicas são realizadas também a partir de processos avaliativos. Porque adotar uma linha teórica ou determinados métodos e não outros? Esses julgamentos são igualmente resultados de uma construção ideo-

lógica fundamentada no sentido axiológico que já carregam e que são atribuídos a determinados autores ou métodos.

O ato de categorizar os trabalhos constituintes do corpus selecionados em “Modelos conceituais” e “Modelos computacionais” como tentativa de representar a realidade impressa nos eventos da ISKO demonstra uma tendência à incorporação acrítica das tecnologias da informação por parte dos atores sociais que povoam esse domínio, não somente dos autores dos trabalhos, mas dos dirigentes dos eventos, uma vez que foram eles os elaboradores dos enunciados subtemáticos dos eventos e os responsáveis por aceitar os trabalhos submetidos. A ênfase não deveria residir nos modelos conceituais? A TI, nesse contexto, não deveria estar a serviço dos modelos conceituais? Quais são os objetivos da OC, fabricar sistemas de informação ou sistemas de OC? Esse posicionamento pró-tecnologia da informação é considerado resultado da aderência de parte do domínio a um discurso mais ideológico do que científico.

Vários autores têm sinalizado sobre o fato da tecnologia da informação ser pensada de modo hegemônico no domínio de OC. Hjørland (2003) alerta para a carência de mapeamento teórico na OC, fato que, segundo ele, a torna uma disciplina subjugada às mudanças tecnológicas. Barité (2011), em pesquisa que cobriu a produção científica controlada pelo *Classification Scheme for Knowledge Organization Literature* (CSKOL) entre os anos de 1994 e 2009, obteve resultados que reforçam a indicação de Hjørland, uma vez que mostram um crescimento regular das pesquisas em torno dos temas relativos aos sistemas de recuperação da informação on-line, à aplicação da classificação e indexação e à influência do ambiente na aplicação da OC, enquanto temas voltados para questões teóricas diminuíram.

Diante desse panorama, García Gutiérrez (2011) sugere que a tecnologia da informação e a tecnologia digital devam ser incorporadas em qualquer domínio, a partir da análise das posições específicas vinculadas aos contextos aos quais pretende atender. Ou seja, defende que essa tecnologia seja reinventada e reapropriada a partir do pluralismo lógico e usos culturais decorrente de necessidades de organização do conhecimento e da memória registrada, e não o contrário. Isso demanda de acolhimento de concepções etnológicas, participativas, diacrônicas, autonarrativas, críticas e autocríticas.

Em função dos debates constantemente ocorridos na coletividade, assim como os fatos e as decisões decorrentes deles, os discursos cientí-

ficos, filosóficos e ideológicos se alteram. Por isso, nenhum modelo pode ser permanente, uma vez que é sempre uma representação provisória espacial e temporalmente localizada. Como defende Badiou, nenhuma relação entre o objeto de estudo e sua representação é transparente, pois qualquer representação é atravessada por concepções de mundo distintas, valores morais e ideologias, variáveis de acordo com o contexto em que se encontram o sujeito e/ou grupo de sujeitos que estabelecem ou estabeleceram essa relação. Isso fica registrado no discurso e, com efeito, nos documentos, como é o caso dos anais da ISKO, objeto do presente estudo.

Uma análise entre a relação dos conceitos representativos dos trabalhos que compõe o corpus dessa pesquisa e as demais produções dos seus autores, nos moldes do estudo já realizado por Arboit (2014), reforçaria a não neutralidade da relação entre os objetos de estudo e sua representação. Esse confronto entre conceitos sistematicamente defendidos por um determinado número de autores trata-se de uma proposta para estudos futuros, uma vez que fortaleceria o argumento de que as relações axiomáticas entre os objetos de estudos elegidos pelos seus autores e a representação construída por eles enquanto participantes de um dínamo social.

## Referências

- Arboit, A. E. (2014) *O processo de institucionalização socio-cognitiva do domínio de Organização do Conhecimento a partir dos trabalhos científicos dos congressos da ISKO*. 286 f. Tese (Doutorado) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília.
- Arboit, A. E. (2012). *O processo de (re)construção da teoria do conceito no domínio de Organização do Conhecimento: Uma visão dialógica*. *Scire*, v. 18, n. 2, p. 129-134.
- Arboit, A. E.; Grácio, M. C. C.; Oliveira, E. T.; Bufrem, L. S. (2012). Relationship between authors and main subject categories in the Knowledge Organization domain: a bibliometric approach. In: Neelameghan, A.; Raghavan, K. S. (eds.). (2012). *Categories, contexts and relations in Knowledge Organization: proceedings of the twelfth international ISKO conference*. Würzburg: Ergon Verlag, 44-50.
- Badiou, A. (2007). *Le concept de modèle*. Paris: Fayard Ouvertures.
- Barité, M. (2011). *La garantía literaria como herramienta de revisión de sistemas de organización del conocimiento: modelo y aplicación*. 2011. 360 f. Tese (Doutorado em Información Científica) - Facultad de Comunicación y Documentación. Universidad de Granada.
- Blackburn, S. (1997) *Dicionário Oxford de Filosofia*. Tradução de Desiderio Murcho, et al. Rio de Janeiro: Jorge Zahar.
- Bufrem, L. S. (1996) *Linhas e tendências metodológicas na produção acadêmica discente do mestrado em ciência da informação do IBICT/UFPRJ*. Curitiba, 1996. (Tese pa-

- 
- ra Concurso de Professor Titular) – Universidade Federal do Paraná, Escola de Biblioteconomia.
- Friedman, A. (2006) Concept mapping a measurable sign. In: Budin, G.; Swertz, C.; Mitgutsch, K. (eds.). (2006) *Proceedings of the Ninth International ISKO Conference* (4-7 July 2006, Vienna, Austria). Würzburg: Ergon Verlag, 131-139.
- Freitas, J. L.; Nascimento, B. S.; Bufrem, L. S. (2014) A Organização do Conhecimento na dinâmica da pesquisa em artigos da literatura científica da Brapci. *Transinformação*, v. 26, p. 295-303.
- García Gutiérrez, A. (2007) *Desclasificados: pluralismo lógico y violencia de la clasificación*. Barcelona: Anthropos.
- García Gutiérrez, A. (2011). *Epistemología de la documentación*. Barcelona: Stonberg, 2011.
- Guimarães, J. A. C.; Santos, A. Y. A., Cândido, G. G. A., Pinho, F.A. (2014). The subject dimension of knowledge organization research: a domain analysis of the Brazilian, Spanish and North American ISKO chapters' conference. *Scire*, v.20, n. 2, p. 19-25.
- Hjørland, B. (2003). Fundamentals of knowledge organization. *Knowledge Organization*, v.30, n.2.
- Smiraglia, R.P. (2014). II Congresso brasileiro em representação e organização do conhecimento: Knowledge organization in Rio 2013 - An editorial *Knowledge Organization*, v.41, n.2, p. 105-112.
- Zherebchevsky, S. (2010) Formalism in knowledge organization: thematic analysis of ISKO 10 proceedings. In: Gnoli, C.; Mazzocchi, F. (eds.) (2010) *Paradigms and conceptual systems in knowledge organization*. Würzburg: Ergon Verlag, 98-105.

## Anexo 1

| Modelos conceituais  |   | Modelos computacionais  |
|--|---|---|
| Ontologia<br>Ontologia por diferenciação<br>Ontologia para recursos legais<br>Tesouros vs. Ontologia<br>Modelagem ontológica   | Tesouros simbólicos<br>Tesouros e cabeçalho de assuntos em rede<br>Tesouro pós-coordenado<br>Tesouros gráficos<br>Reavaliação de tesouros   | Processamento de linguagem natural<br>Técnicas de processamento de linguagem natural<br>Análise da linguagem natural em vocabulários controlados  |
|  |   | Rede de proximidade semântica<br>Formalismo semântico<br>Semântica intertextual   |
| Terminologia construída<br>Intercâmbio de dados terminológicos e lexicográfico<br>Vocabulários de especializados de domínios específicos   | Conotação de termos<br>Vocabulário controlado<br>Abordagem hermenêutica na OC   | Análise de componente principal<br>Redes artificiais neurais<br>Escalonamento multidimensional<br>Mapas baseado em dados de cocitação   |
| Estrutura conceitual<br>Relações interconceituais<br>Relacionamento de entidades<br>Atributos e relacionamentos<br>Relações de equivalência<br>Modelos fragilmente estruturados<br>Causalidade e coerência conceitual<br>Reestruturação da CDU | Modelos fortemente estruturados<br>Relações hierárquicas, associativas e de equivalência<br>Pontos de acesso e relacionamento entre entidades<br>Demarcação e discriminação entre conceitos e entidades<br>Visibilidade e invisibilidade de relações<br>Relações na CDD<br>Subdivisão e conjunção | Representações dinâmicas<br>Modelo computacional para manutenção de classificações<br>Manutenção automática de tesouros<br>Alocação automática de ligações referenciais   |
|  |   | Nova estrutura para registros bibliográficos<br>Relacionabilidade de documentos   |
| Modelos científicos<br>Modelos de classificação<br>Modelos conceituais de dados<br>Metadados<br>Metadados simbólicos e metafóricos<br>Classificacionista como autor<br>Modelos de representação múltipla                                       | FRBR<br>Análise de esquemas organizacionais acadêmicos<br>Integração do conhecimento interdisciplinar<br>Taxonomia de Bloom<br>Classificação de objetos de aprendizagem   | Papel da terminologia   |
|  |   | Ferramentas e técnicas de inteligência artificial   |
|  |   | Linguagem de representação de conhecimento narrativo<br>Linguagem de transferência de informação  |
|  |   | Abordagem analítico-sintética<br>Representações XML de facetas fundamentais<br>Vocabulário controlado em XML<br>Sistemas facetados para buscas por termos múltiplos e resultados flexíveis<br>Catálogo online baseado em decisões do usuário<br>Futuro dos catálogos<br>Modelos mentais e comportamento de busca<br>Comportamento do usuário<br>Preferências dos usuários<br>Indexação e busca<br>Design de ferramentas de busca<br>Representação e recuperação da informação |
| Imagem como recurso de indexação<br>Informação digital audiovisual<br>Informação analógica   | Organização de recursos digitais<br>Portal informativo na web<br>Biblioteca para arranjar recursos eletrônicos na web   |   |
| Interoperabilidade semântica das classificações bibliográficas<br>Web semântica<br>Representação semântica   | Interoperabilidade do e-government<br>Rede interconectada de informação   |   |
| Análise facetada<br>Esquema baseado em facetas<br>Estrutura facetada de conhecimento   | Terminologia facetada<br>Classificação facetada   |   |
| Metodologia para tomada de decisões sobre desenvolvimento de coleções<br>Uso de informações locais<br>Contextos de produção e uso  | Contexto cultural de uso ( <i>cultural marker</i> )<br>Localização de site cultural   |   |

Tabela II. Mapa conceitual dos temas dos trabalhos categorizados como tipos de modelos

---

# Catologação sob a perspectiva dos artigos publicados em revistas científicas nacionais avaliados nos estratos B1, B2 e B3 no sistema Qualis a partir de 1998

*Cataloging from the perspective of articles published in national journals evaluated in strata B1, B2 and B3 in the Qualis system from 1998*

---

Helena Campos Assumpção (1), Zaira Regina Zafalon (2)

Universidade Federal de São Carlos, Rod. Washington Luís, km 235, São Carlos/SP, Monjolinho, CEP 13565-905. (1) Email: helena.assumpa@yahoo.com.br. (2) Email: zzafalon@gmail.com.

## Resumo

Centra-se na análise de artigos sobre catalogação em periódicos científicos brasileiros, a partir de 1998, marco na área diante da publicação da obra *Functional Requirements for Bibliographical Records* (FRBR). O recorte dos periódicos científicos foi feito considerando-se os estratos B1, B2 e B3 no Qualis, definidos segundo critérios que orientam o desempenho e o julgamento dos programas de pós-graduação no Brasil. Definiu-se como objetivo analisar os artigos, com estudos de catalogação, nos periódicos científicos brasileiros, com estratos Qualis B1, B2 e B3, no período de 1998 a 2013. Com abordagem quali-quantitativa, a pesquisa configura-se como exploratória e descritiva e recorre-se à pesquisa documental e bibliográfica, para o estudo do arcabouço teórico, ao formulário, para a coleta de dados, e à análise de conteúdo, para definição da pertinência temática. Foram estudados aspectos teóricos da catalogação, identificados os periódicos científicos e os artigos, segundo critérios de estratificação, de adequação à área e de pertinência temática. Como resultados observou-se que, apesar do aumento em 600% na publicação de artigos sobre catalogação nos periódicos científicos brasileiros, estes não apresentam correlação com a publicação dos FRBR, e centram-se em Ontologia, Organização da Informação e Metadados.

**Palabras clave:** Catalogação. Organização e representação da informação. FRBR. Ontologia. Metadados.

## 1. Introdução

Esta pesquisa está centrada na catalogação, em especial no período de 1998 a 2013, e considera em seu bojo o fato de que, na representação de informações, a padronização é essencial, tendo em vista a ampliação promovida, inclusive, pela adoção de recursos tecnológicos, e deve-se levar em conta as necessidades dos

## Abstract

It focuses on the analysis of articles on cataloging in Brazilian scientific journals, from 1998 landmark in the area before the publication of Functional Requirements work for Bibliographical Records (FRBR). The clipping of scientific journals was done considering the layers B1, B2 and B3 in the Qualis, defined according to criteria that guide the performance and the judgment of graduate programs in Brazil. It was decided to analyze the articles, with cataloging studies in Brazilian scientific journals, with strata Qualis B1, B2 and B3, from 1998 to 2013. With Qualiquantitative approach, research is characterized as exploratory and descriptive and uses is directed to the documentary and bibliographical research, to study theoretical framework, to form, to collect data, and content analysis, to define the thematic relevance. Theoretical aspects of cataloging were studied, identified scientific journals and articles, according stratification criteria of suitability to the area and thematic relevance. As a result it was observed that, despite the increase by 600% in publishing articles on cataloging the Brazilian scientific journals, they do not sit appecorrelation with the publication of FRBR, and focus on Ontology, Organization of Information and Metadata.

**Keywords:** Cataloging. Organization and representation of information. FRBR. Ontology. Metadata.

usuários. Diz respeito, portanto, ao estudo da representação de recursos informacionais, do desenvolvimento tecnológico e de aspectos teóricos e metodológicos da catalogação.

A disponibilização de recursos informacionais está permeada de estruturas computacionais e de métodos de definição de metadados. Diante dessa constatação observa-se que a produção,

a organização, a distribuição, o acesso, o armazenamento, a preservação, o uso e o reuso dos recursos informacionais foram reconfigurados. Está posto, portanto, novo contexto sociocultural, educacional e econômico. Aliado a esses fatores há o crescimento exponencial de acervos documentais nas mais variadas comunidades que fazem uso corrente ou potencial de padrões de descrição. Incluem-se tanto as comunidades voltadas ao patrimônio cultural (arquivos, bibliotecas e museus) quanto as da indústria da informação (organizações que compõem tanto o público quanto a web comercial), cada qual com demanda informacional própria. No âmbito técnico-funcional-administrativo dessas instituições, o emprego de tecnologias de informação e comunicação tornou-se constante, o que favorece tanto o surgimento quanto a inovação de várias atividades cooperativas.

Nesse sentido, analisar o cenário brasileiro das pesquisas desenvolvidas na área de Catalogação torna-se necessário, uma vez que, para Ranganathan e Gopinath (1967), recorrer a mecanismos computacionais nas atividades do dia-a-dia das comunidades, como substitutos das atividades manuais e mecânicas, exige viabilidade econômica e, em última instância, indica o estágio de desenvolvimento de um país.

Manter o foco da representação nas necessidades informacionais do público a que se propõe atender caracteriza o princípio do objetivo bibliográfico da catalogação, uma vez que, por meio de registros bibliográficos disponíveis em catálogos, é possível promover a recuperação da informação, tanto por humanos quanto por máquinas, posto que veiculam mensagens codificadas sobre os recursos informacionais de modo a atender e a se nutrir das informações demandadas pelas comunidades.

Dada a relação entre itens documentais e demandas informacionais, o esquema de comunicação proposto por Shannon (1948; 1969) pode ser (re)configurado, conforme Figura 1.



Figura 1. Esquema do processo comunicativo entre recursos informacionais e comunidades por meio de metadados

A catalogação, pautada, portanto, nos objetivos bibliográficos, foi discutida por Panizzi, em 1841, Cutter, em 1876, Lubetzky, em 1960, o que culminou na definição de *Statement of principles* (Princípios de Paris), com o propósito de

servir como bases para a normalização internacional na catalogação. Entretanto, Svenonius (2000), discute o fato de que tais princípios limitavam a função do catálogo uma vez que restringiam-se à organização dos documentos por autor e título.

Diante da necessidade de se reconfigurar os princípios de catalogação indicados em 1961 a IFLA indica, em 1997, sua revisão. Entretanto, Svenonius (2000) apresenta críticas e destaca que os quatro princípios indicados no documento (*encontrar, identificar, selecionar e obter*) são insuficientes para a identificação das necessidades dos usuários e propõe a inserção do princípio da *navegação*, princípio definido com destaque para o meio pelo qual se navega em um universo bibliográfico no processo de recuperar o documento desejado. A autora argumenta essa inserção partindo do fato de que alguns usuários sabem exatamente o que querem ao realizar a busca por informação, enquanto outros não sabem bem ou não são capazes de reconhecer o objeto de sua pesquisa, ou, ainda, não são capazes de reconhecer imediatamente quando encontram. Ela também está amparada quanto à inserção do princípio da *navegação* devido ao fato de que os códigos das regras bibliográficas utilizadas para organizar os documentos assumem essa existência e estabelecem relacionamentos bibliográficos, especificando a relação entre obras e também as relações envolvendo os atributos da obra, como autores e assuntos. Desse modo, a autora afirma que o objetivo da navegação fornece a garantia ao usuário e também ao código.

A IFLA reconheceu a importância da contribuição de Svenonius e publicou, em 2009, a Declaração dos Princípios Internacionais de Catalogação, a qual estabelece como princípios:

**4.1 Encontrar** recursos bibliográficos numa coleção como resultado de uma pesquisa, utilizando atributos e relações entre recursos:

**4.1.1** Para **encontrar** um determinado recurso;

**4.1.2** Para **encontrar** conjuntos de recursos representando

todos os recursos que pertencem à mesma obra;

todos os recursos que representam a mesma expressão;

todos os recursos que exemplificam a mesma manifestação;

todos os recursos associados a determinada pessoa, família ou coletividade (entidade);

todos os recursos sobre um determinado assunto;

todos os recursos definidos por outros critérios (língua, lugar de publicação, data de publicação,

tipo de conteúdo, tipo de suporte, etc.), normalmente como uma delimitação secundária de um resultado de pesquisa;

**4.2 Identificar** um recurso bibliográfico ou agente (ou seja, confirmar que a entidade descrita corresponde à entidade procurada ou distinguir entre duas ou mais entidades com características similares);

**4.3 Selecionar** um recurso bibliográfico que seja apropriado às necessidades do usuário, (ou seja, escolher um recurso que esteja de acordo com as necessidades do usuário, no que diz respeito ao conteúdo, suporte, etc. ou rejeitar um recurso que seja inadequado às necessidades do usuário);

**4.4 Adquirir** ou **obter** acesso a um item descrito (ou seja, fornecer informação que permitirá ao usuário adquirir um item por meio de compra, empréstimo, etc. ou acessar eletronicamente um item por meio de uma ligação em linha a uma fonte remota); ou acessar, adquirir ou obter dados bibliográficos ou de autoridade;

**4.5 Navegar** num catálogo ou para além dele através da organização lógica dos dados bibliográficos e de autoridade e da apresentação de formas claras de se navegar, incluindo a apresentação de relações entre obras, expressões, manifestações, itens, pessoas, famílias, entidades coletivas, conceitos, objetos, eventos e lugares.

A reconfiguração da declaração exerce, portanto, forte impacto nos catálogos, uma vez que destaca a recuperação da informação pela comunidade. Estes, por sua vez, devem pautar-se na funcionalidade exercida pelos registros bibliográficos.

Estudos que buscavam compreender a funcionalidade dos registros no universo bibliográfico foram desenvolvidos por um grupo de trabalho nomeado pela IFLA, o qual tinha o objetivo de propor uma arquitetura estruturada e claramente definida na qual pudessem ser criadas relações entre os dados presentes nos registros bibliográficos e as necessidades dos usuários, e recomendar um nível básico da funcionalidade dos registros. Tais estudos culminaram, em 1997, na aprovação de relatório intitulado *Functional Requirements for Bibliographic Records* (identificado internacionalmente como FRBR) pelo *Standing Committee on Cataloguing* da IFLA.

O modelo FRBR é um modelo caracterizado como entidade-relacionamento e mantém o usuário como ponto de partida para a definição do modelo. Em um modelo entidade-relacionamento devem ser reconhecidos as entidades, os atributos e as relações entre as entidades.

Desse modo, no modelo FRBR, destacam-se três grupos de entidades:

- a) Grupo 1, no qual identificam-se os produtos do trabalho intelectual ou artístico: obra, expressão, manifestação e item;
- b) Grupo 2, com identificação dos responsáveis pelo conteúdo intelectual: pessoas e instituições;
- c) Grupo 3, com destaque para os assuntos: conceito, objeto, acontecimento, lugar e todas as entidades dos grupos 1 e 2.

Segundo Oliver (2010), os atributos, definidos para cada entidade identificada, podem ser intrínsecos (descobertos ao analisar a entidade) e extrínsecos (fora da entidade, definidos por um identificador, por exemplo). Fundamental na definição do modelo é a identificação e o mapeamento das relações entre as entidades. As definições, os exemplos e a especificidade dos atributos e dos relacionamentos, não são objeto de discussão nesta pesquisa, e devem ser consultados em IFLA (2009), publicação aprovada pelo grupo de estudo em 1997 e que conta com alterações e correções datadas de 2009.

Assim, examina-se que, a partir de 1998, com a publicação de *Functional Requirements for Bibliographical Records* retoma-se a relevância do usuário, definem-se outros modelos conceituais para a catalogação e tecnologias usadas forma reconfiguradas. Essas mudanças refletiram nas atividades dos bibliotecários.

Desse modo, devido às grandes mudanças ocorridas no cenário internacional de estudos teóricos e práticos da catalogação, interessou-se estudar: qual a repercussão que os estudos de catalogação tiveram nos artigos publicados em revistas científicas nacionais com estrato B1, B2 e B3 no sistema Qualis, a partir de 1998?

Optou-se por pesquisar no sistema Qualis para avaliar a produção intelectual presente nas publicações avaliadas pela Coordenação de Aperfeiçoamento de Nível Superior (CAPES). O Qualis envolve uma série de procedimentos utilizados para estratificação da qualidade da produção intelectual dos programas de pós-graduação e, como resultado, disponibiliza uma lista com a indicação estratificada das publicações utilizadas para divulgação científica dos programas. Os estratos indicativos de qualidade em publicações periódicas científicas são A1, A2, B1, B2, B3, B4, B5 e C.

O recorte feito na pesquisa, para o estudo para análise das publicações periódicas científicas classificadas nos estratos B1, B2 e B3 foi feito a partir da análise da tabela 1.

| Estrato | Pontuação | Qtde. Periódicos | Proporção |
|---------|-----------|------------------|-----------|
| A1      | 100       | 46               | 2,93 %    |
| A2      | 85        | 61               | 3,99 %    |
| B1      | 70        | 167              | 10,90 %   |
| B2      | 55        | 99               | 6,46 %    |
| B3      | 40        | 174              | 11,35 %   |
| B4      | 25        | 259              | 16,90 %   |
| B5      | 10        | 425              | 27,74 %   |
| C       | 0         | 301              | 19,64 %   |
| Total   |           | 1.532            | 100 %     |

Tabela I – Quantidade de periódicos científicos identificados nos estratos do Qualis

Considerou-se a opção pelos estratos B1, B2 e B3 visto que estes são responsáveis por mais de um quarto das publicações (28,71%) na área de avaliação. Outros estratos foram desconsiderados dado que: os estratos A1 (nota 100) e A2 (nota 85) representam a menor proporção das publicações científicas, que as notas obtidas pelos estratos B4 (nota 25) e B5 (nota 10) indicam notas bem menores que a nota média, e, ainda, que o estrato C indica publicações que não são consideradas científicas.

Diante da questão de pesquisa proposta determinou-se, como objetivo principal, analisar os artigos sobre catalogação publicados em periódicos científicos nacionais com estratos B1, B2 e B3 no Qualis, no período de 1998 a 2013.

Para atingir este objetivo estabeleceram-se os seguintes objetivos específicos:

- estudar a configuração da catalogação;
- identificar os periódicos nacionais classificados nos estratos B1, B2 e B3 do Qualis na área de Ciências Sociais Aplicadas I;
- identificar os artigos publicados no período de 1998 a 2013 sobre catalogação.

Tendo em vista o cenário internacional, o surgimento de novas tecnologias, as publicações internacionais de catalogação, justificou-se o desenvolvimento desta pesquisa, uma vez que pode enriquecer as contribuições profissionais e acadêmicas que os resultados possam identificar. Destaca-se, como contribuição profissional, o fato de que a divulgação das publicações faz parte das necessidades de atualização profissional e o acompanhamento e discussão dos estudos.

Como contribuição acadêmica vislumbra-se a compilação do estado da arte, neste universo de pesquisa, da catalogação na ciência nacional.

Tendo em vista a apresentação dos aspectos teóricos da pesquisa apresenta-se o próximo tópico.

## 2. Aproximações teóricas

Com o intuito de se estabelecer o cunho teórico no qual a pesquisa foi desenvolvida, o tópico dedica-se ao cenário da catalogação e das publicações científicas.

A catalogação é um processo utilizado para representar informações de modo a garantir a recuperação por comunidades, uma vez que cientistas da informação entendem como relevante a necessidade de se transmitir as informações contidas nos documentos aos usuários. Para que seja possível manter um padrão na representação é requerido o uso de regras e códigos de catalogação.

Smiraglia (2003) afirma que o catálogo surgiu com o propósito de armazenar e facilitar o trabalho do bibliotecário mas, com o passar do tempo e com o aumento da produção documental, notou-se a necessidade de uma nova organização nas bibliotecas. Devido à esses fatores, Thomas Hyde criou em 1674 o primeiro catálogo de autores para a biblioteca de Bodleian, em Oxford, o que facilitou a busca para os usuários.

Denton (2007), por sua vez, aponta que, no século XIX começam ocorrer mudanças importantes no cenário da catalogação. Em 1841, o italiano Panizzi defende a ideia de que o catálogo deve reunir e diferenciar os itens semelhantes e que o catálogo deve ter como função, além de indicar se há o livro na biblioteca, se há outras edições e obras do autor. Sendo assim, estabelece regras para a elaboração do catálogo do Museu Britânico que se tornaram conhecidas como *As 91 regras de Panizzi*.

Barbosa (1978) destaca que, em 1852, o americano Jewett reconheceu a catalogação coletiva e elaborou um código de catalogação, com regras baseadas no código de Panizzi, para as bibliotecas americanas. Foi o primeiro americano a utilizar o autor como ponto de entrada. Ainda no século XIX, em 1868, Charles Cutter passa a observar, na Biblioteca de Athenæum, as necessidades dos usuários. Após alguns anos de trabalho, já em 1876, ele desenvolve um código, *Rules for a dictionary catalog*, que indica a autoria das obras literárias.

O Joint Committee for Development of RDA (2009) destaca que, em 1893, publicaram-se *Cataloguing Rules*, criada pela *Library Association* (LA) localizada na Inglaterra. Essa obra adotava a autoria como ponto de acesso principal.

Em 1899, foram publicadas as *Prussian Instructions*, regras adotadas na Alemanha e em muitos países da Europa e traduzidas para o inglês em 1938, conforme afirma Barbosa (1978), e trad. Essas regras abordavam a autoria coletiva e as obras anônimas, reconhecidas como necessárias a um código internacional somente em 1936.

Segundo Denton (2007), a American Library Association (ALA) começou a criar, em 1900, um código de catalogação padrão, o que contou com a ajuda de Cutter, que veio a falecer em 1902. Sendo assim, John Dewey, membro fundador da ALA, propôs parceria aos membros da Library Association para a publicação de *Catalog Rules, Author and Title Entry*, que ocorreu em 1908.

Figueiredo (1992) indica a relevante contribuição do indiano Shiyali Ramamrita Ranganathan para a catalogação, em decorrência da publicação das cinco leis da biblioteconomia em 1931. Estas leis são importantes por conta do destaque para o atendimento às necessidades dos usuários. Em suma, a primeira lei, *Os livros são para uso*, onde está escrito livro, pode ser aplicada obra. A segunda, que profere *A cada leitor o seu livro*, indica a necessidade de se conhecer o público a que se destinam os registros, dado que a catalogação deve ter significado aos usuários. Na lei seguinte *A cada livro seu leitor*, indica o destaque que o acervo tem para o leitor e de que é preciso divulgá-lo. A quarta lei, *Poupe o tempo do leitor*, apresenta a economia de tempo como imperiosa para o sucesso do atendimento às necessidades do público, uma vez que a catalogação, por caracterizar-se como processo de representação do documento, de indicação da localização e pela definição de pontos pelos quais o mesmo pode ser acessado, abrevia o tempo de busca e obtenção da informação desejada. A última lei *A biblioteca é um organismo em crescimento*, constata a atualização da biblioteca, tanto no que se refere aos seus processos, quanto ao seu acervo e, principalmente, as mudanças nas demandas informacionais.

As *Catalog Rules, Author and Title* foram revistas e, em 1941, lançou-se a segunda edição, com 324 regras e dividida em duas partes. Entretanto, a edição foi criticada dada a complexidade e ao alto grau de detalhamento. De acordo com Joint Steering Committee For Development of RDA (2009), as regras publicadas pela ALA não abordavam a temática da catalogação descritiva. Assim, a ALA adota *Rules for Descriptive Cataloging*, publicada em 1949 por Lubetzky, bibliotecário na Library of Congress. Estas regras foram muito criticadas por serem muito

detalhadas. Sendo assim, no mesmo ano, foi publicada a revisão *Cataloging Rules for Author and Title Entries*, que continha apenas regras de entrada e de título. Em 1953, Lubetzky publicou *Cataloging Rules and Principles*.

Em 1961, com o intuito de se estabelecer princípios para a criação de catálogos com pontos de acesso de autores e títulos, a IFLA (International Federal of Library Associations and Institutions) promove um encontro em Paris, que passou a ser identificado como Conferência de Paris. Esse evento é considerado como um dos primeiros esforços para uma normalização internacional na catalogação e um marco para a definição de que os códigos de catalogação desenvolvidos após esta data deveriam seguir os princípios estabelecidos na conferência realizada pela IFLA. (IFLA, 2009).

Corrêa (2008) destaca que, em 1971, com a publicação da International Standard for Bibliographical Description (ISBD), pela IFLA, indicou-se um conjunto de regras para a criação de descrições bibliográficas, uma vez que indica os elementos necessários para a descrição de diferentes tipos de documentos. Por ser considerada de âmbito internacional foi adotada por vários países, contribuindo, desse modo, para que discrepância na representação de informações fossem evitadas.

Conforme Oliver (2011), as AACR foram revistas em 1967 e, em 1978, publicou-se a segunda edição, identificada como AACR2. O autor destaca que, desde então, foram as mais utilizadas no mundo, inclusive em países de língua não-inglesa. Inicialmente, as AACR foram desenvolvidas como um código de catalogação de livros, periódicos impressos e outros documentos em papel. Sucessivamente foram anexadas regras sobre outras mídias, mas sua estrutura continua remetendo a dos catálogos tradicionais em papel.

Segundo Moreno (2009), com o passar do tempo e a criação de novas tecnologias, os Princípios de Paris não eram mais suficientes para representar as mudanças ocorridas nas bibliotecas e nos catálogos. Dessa forma, foram realizados encontros para atualizar os princípios de catalogação com o objetivo de satisfazer as novas necessidades dos usuários. Esses encontros, chamados de IFLA Meetings of Experts on an International Cataloguing Code (IME-ICC), tinham, como o próprio nome sugere, o objetivo de promover o desenvolvimento de um código de catalogação internacional para descrição bibliográfica e acesso aos recursos. Recorreram, então, à comparação de semelhanças e diferenças entre os códigos adotados em vá-

rios países para, futuramente, tivesse condições de propor um código de catalogação internacional.

De acordo com a IFLA (2014), o primeiro encontro foi realizado em 2003 em Frankfurt, Alemanha. Além dos países Europeus, os autores das regras Anglo-Americanas também participaram. Conforme Tillet (2007), em 2004 realizou-se o segundo encontro, em Buenos Aires, Argentina, com a participação de integrantes de 14 países da América Latina. Nesta edição o objetivo era de rever e atualizar a declaração de princípios realizada na Europa e comparar o código utilizado na Europa com o código usado na América Latina (TILLET, 2007). Segundo a IFLA (2005), em Cairo, no Egito, foi realizado o terceiro encontro, em 2005, que contou com a presença dos países árabes do Oriente Médio. Eles examinaram os códigos utilizados por eles para confrontar com os Princípios de Paris. De acordo com Tillet (2007), o quarto encontro foi realizado para estudos pelos países asiáticos, em Seul, na Coreia do Sul, no ano de 2006. Esse encontro contou com a participação de representantes de 12 países asiáticos. Segundo a IFLA (2007), o quinto e último encontro foi realizado na África do Sul, em Pretória no ano de 2007. Através desse encontro, pode-se verificar as semelhanças e diferenças dos códigos utilizados na África Subsaariana com os Princípios de Paris.

A partir destes encontros foram retomados os esforços para o estabelecimento de um código internacional de catalogação e o RDA foi apresentado como tal. A RDA foi desenvolvida por uma comissão de 4 países, entre eles, Austrália, Canadá, Grã-Bretanha e Estados Unidos, nomeada como Joint Steering Committee for Development of RDA. O RDA, com o propósito de ser utilizado num contexto internacional, tem o objetivo de, além de substituir a AACR2, focar a atenção no usuário e atender às novas exigências de representação e organização da informação (SILVA, 2012). Discutir o atingimento ou não do objetivo a que se propõe não é o foco nesta pesquisa.

De acordo com Silva (2012), a RDA, nome dado ao código de catalogação intitulado Resource Description and Access, foi desenvolvido com base nos Functional Requirements for Bibliographic Records (FRBR) e tem uma abordagem diferenciada para a descrição de recursos tecnológicos. A RDA apresenta um propósito mais abrangente uma vez que foi desenvolvida para atender às novas exigências de representação e organização da informação nos meios tecnológicos.

Assim sendo, as mudanças nos catálogos começaram a ocorrer em 1841. De acordo com Panizzi, os objetivos dos catálogos eram reunir e diferenciar os itens similares. Já em 1876, Cutter elabora objetivos fundamentados nas necessidades dos usuários. E foi a partir de então que os catálogos começaram a apresentar melhorias e mudanças (ZAFALON, 2011).

Com o surgimento de novas tecnologias, notou-se a necessidade de alterar os objetivos dos catálogos, isto é, focar novamente no usuário para que seja possível encontrar, identificar e selecionar a informação desejada. Para atingir esses objetivos recorreram-se aos modelos Requisitos Funcionais para Registros Bibliográficos (FRBR) e Requisitos Funcionais para Dados de Autoridade (FRAD), que tem como função permitir que o usuário navegue pelo catálogo diferenciando os registros e facilitar a busca (OLIVER, 2011; SILVA, 2012).

Dado o contexto necessário ao núcleo da pesquisa recorre-se às discussões sobre publicações periódicas científicas, visto que sua relevância está atrelada à difusão do conhecimento produzido nas universidades e ao objetivo de promover e registrar o progresso da ciência. Pelas publicações científicas também é possível calcular o desempenho dos docentes e discentes da universidade e o desenvolvimento de uma área. "A comunicação científica é o processo que acontece dentro de uma comunidade científica e/ou acadêmica para a transferência e socialização do conhecimento científico" (COSTA; AMARO; ASSIS, 2012, p. 37).

De acordo com Ziman (1979), até metade do século XVII, a troca de informações entre pesquisadores era feita através de cartas, livros e panfletos.

De fato, do ponto de vista da pesquisa, era razoável que as ideias inicialmente circulassem por meio de cartas manuscritas entre um pequeno círculo de amigos que poderiam analisá-las e, quando conveniente, testá-las e depois enviar a resposta (MEADOWS, 1999, p. 5).

Com o passar dos anos, a importância de se registrar as informações foi aumentando. Após a restauração da monarquia em Londres, em 1660, alguns grupos se reuniram para discutir questões filosóficas. Em 1662 os pesquisadores começaram a formalizar essas reuniões, dando surgimento à Royal Society, que objetivava coletar e analisar as informações importantes. A necessidade de registro de informações fez com que, em 1665, Denis de Sallo criasse um periódico com o intuito de publicar notícias sobre os acontecimentos da Europa (MEADOWS, 1999).

O fato é que a publicação de trabalhos científicos não é feita irrestritamente. Um artigo publicado numa revista conceituada não representa apenas a opinião do autor; leva também o selo da autenticidade científica através do imprimatur dado pelo editor e os examinadores que ele possa ter consultado (ZIMAN, 1979, p. 124).

Com o crescimento das publicações científicas, aumentou-se também o trabalho das editoras comerciais, uma vez que o método mais utilizado para expandir a comunicação científica é por meio dos periódicos (COSTA; AMARO; ASSIS, 2012).

Devido ao avanço das publicações científicas, surgiu o movimento *Open Access*, que defende o acesso livre às publicações, com destaque para a *Budapest Open Access Initiative* que lançou, em 2002, uma campanha mundial para acesso aberto a todas as novas pesquisas.

“Acesso aberto” à literatura científica revisada por pares significa a disponibilidade livre na Internet, permitindo a qualquer usuário ler, fazer download, copiar, distribuir, imprimir, pesquisar ou referenciar o texto integral desses artigos, recolhê-los para indexação, introduzi-los como dados em software, ou usá-los para outro qualquer fim legal, sem barreiras financeiras, legais ou técnicas que não sejam inseparáveis ao próprio acesso a uma conexão à Internet. As únicas restrições de reprodução ou distribuição e o único papel para o direito autorial neste domínio é dar aos autores o controle sobre a integridade do seu trabalho e o direito de ser devidamente reconhecido e citado (BOAI, 2012).

O movimento de acesso aberto é uma proposta de enfrentamento aos direitos de reprodução, que reportam ao século XVIII, visto que, em 1710, Ana, a Rainha da Inglaterra, criou o “Copyright Act” para incentivar a produção cultural da época. Este ato previa a remuneração de autores que eram explorados pelos editores detentores dos meios técnicos de reprodução e de privilégios de publicação das obras (BARACAT; RIGOLIN, 2012). Entretanto, essas regras não foram aplicadas às revistas científicas.

Uma forma de proteger os direitos autorais e, ao mesmo tempo, garantir o acesso e uso da obra, é a utilização da licença *Creative Commons*. A *Creative Commons* é uma organização sem fins lucrativos que permite o compartilhamento e o uso da criatividade e do conhecimento por meio de instrumentos jurídicos gratuitos (CREATIVE COMMONS, 2013, tradução nossa).

Este movimento está longe do fim, mas a BOAI se compromete a realizar progressos nos próximos anos e já determinaram a nova meta para os próximos dez anos “ser o método normal e padrão para distribuir os novos resultados de pesquisa com revisão por pares, em todos os

domínios científicos e em todos os países.” (BOAI, 2012).

### 3. Procedimentos metodológicos

Recorre-se, para o atingimento dos objetivos, à pesquisa com enfoque quantitativo e qualitativo uma vez que se complementam tendo em vista a busca e a compreensão das causas do problema. É qualitativa visto que trabalha com observações, descrições, comparações e interpretações. Gerhardt e Silveira (2009, p. 32) explicam que “Os pesquisadores que utilizam os métodos qualitativos buscam explicar o porquê das coisas, exprimindo o que convém ser feito [...]. O caráter quantitativo da pesquisa está presente dado que também é um estudo numérico que se utiliza técnicas estatísticas para descrever as características de uma situação. Faz uso, portanto, de instrumentos, como por exemplo formulários, para coletar dados. Para Assis (2001, p. 20), a pesquisa quantitativa se utiliza de números para deduzir informações.

Com abordagem descritiva busca-se descrever na pesquisa as características de um grupo ou de um acontecimento pois, segundo Assis (2001, p. 18), visa observar, registrar, analisar, classificar e interpretar os dados sem interferência, sem manipulação do pesquisador. Caracteriza-se, também, como exploratória uma vez que busca habituar-se com o que se deseja estudar. Definiu-se o tipo de pesquisa exploratória por ser uma pesquisa no campo social. Segundo Gonsalves (2011), a pesquisa exploratória apresenta dados que são a base do estudo e propõe uma visão panorâmica devido à elucidação e o desenvolvimento de ideias.

Como técnica de pesquisa apela para a pesquisa bibliográfica e documental; bibliográfica para identificar estudos sobre o tema pesquisado e, nesta pesquisa, para a seleção do material teórico. A vantagem por escolher a pesquisa bibliográfica se encontra na vasta gama de opções de fontes bibliográficas. Gil (2002, p. 44) realça que a pesquisa bibliográfica:

[...] é desenvolvida com base em material já elaborado, constituído principalmente de livros e artigos científicos. Embora em quase todos os estudos seja exigido algum tipo de trabalho dessa natureza, há pesquisas desenvolvidas exclusivamente a partir de fontes bibliográficas. Boa parte dos estudos exploratórios pode ser definida como pesquisas bibliográficas.

A pesquisa documental, também adotada nesta pesquisa, investiga a veracidade das informações nos documentos, dado que recorre a fontes mais diversificadas. Segundo Gil (2002, p. 45), a “pesquisa documental vale-se de materi-

ais que não recebem ainda um tratamento analítico, ou que ainda podem ser reelaborados de acordo com os objetos da pesquisa.”

O formulário, instrumento utilizado para a coleta de dados, foi aplicado nesta pesquisa com o intuito de se obter informações sobre os periódicos, objeto de pesquisa para a realização do estudo. O formulário é um instrumento com campos pré-definidos, nos quais são preenchidos dados e informações levantados na pesquisa, o que permite a formalização das comunicações e o registro destes dados (CURY apud KAUARK; MANHÃES; MEDEIROS, 2010). No formulário foram estabelecidos os seguintes campos para a identificação dos periódicos científicos: área de avaliação, estrato, nacional (sim ou não), título do periódico, ISSN e escopo (informado no editorial do periódico). Para a coleta de dados sobre os artigos os seguintes campos fizeram parte do formulário: título do periódico, ISSN, volume, número e ano em que o artigo foi publicado, título do artigo, link de acesso, autor, palavras-chave, resumo e categoria de análise.

A análise dos dados coletados foi feita por meio de análise de conteúdo que, de acordo com Gil (2002, p. 89), “possibilita a descrição do conteúdo manifesto e latente das comunicações.”

Para Bardin (1977):

A análise de conteúdo desenvolve-se em três fases. A primeira é a pré-análise, onde se procede à escolha dos documentos, à formulação de hipóteses e à preparação do material para análise. A segunda é a exploração do material, que envolve a escolha das unidades, a enumeração e a classificação. A terceira etapa, por fim, é constituída pelo tratamento, inferência e interpretação dos dados.

A coleta concentrou-se nos anos de 1998 e 2013, com estrato Qualis B1, B2 e B3 no sistema Qualis, na área de Ciências Sociais Aplicadas I, que abrangem estudos de Biblioteconomia e Ciência da Informação foi possível reunir dados de diversos artigos publicados em diferentes periódicos, identificar e quantificar os artigos nacionais publicados sobre catalogação.

Para coletar os dados dos periódicos acessou-se o sistema WebQualis (disponível em: <http://qualis.capes.gov.br/webqualis/principal.seam>) e, em Consulta, Por Classificação / Área de Avaliação, selecionou-se Ciências Sociais Aplicadas I, e indicou-se o estrato B1. O resultado foi exportado para PDF, uma vez que seria necessário tratamento dos dados. O mesmo processo foi feito com os estratos B2 e B3. Os dados coletados no sistema foram registrados em um formulário específico.

Para obter as informações dos artigos para análise foi preciso procurar o site de cada periódico para se ter acesso ao país de publicação, ao escopo, aos artigos, aos anos, às palavras-chave e ao resumo. Feito isso foi possível analisar os artigos e estabelecer as categorias de pertencimento temático à área de catalogação.

#### 4. Análise dos resultados

Os resultados indicam a análise quantitativa e, posteriormente, a análise qualitativa dos dados. A tabela 2 indica a quantidade de periódicos científicos identificados nos estratos B1, B2 e B3 do Qualis.

| Estrato | Números de Periódicos |
|---------|-----------------------|
| B1      | 167                   |
| B2      | 99                    |
| B3      | 174                   |

Tabela II – Quantidade de periódicos científicos identificados nos estratos B1, B2 e B3 do Qualis

Observa-se que as publicações concentram-se nos estratos B1 e B3. O gráfico 1 apresenta a distribuição das revistas nos estratos e reflete, percentualmente, os dados apresentados na tabela II.

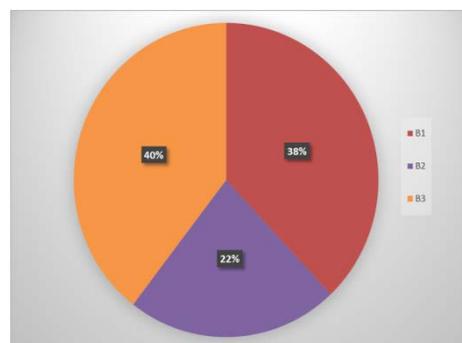


Gráfico 1. Distribuição de periódicos científicos segundo estratos B1, B2 e B3 do Qualis

Tendo sido identificados os títulos dos periódicos científicos da área de Ciências Sociais Aplicadas, procedeu-se outra análise, na qual identificou-se a nacionalidade das publicações e dividiu os periódicos entre os estratos B1, B2 e B3, em nacionais e internacionais. Esta análise poder ser vista na tabela e nos gráficos a seguir:

| Estrato | Nacional     | Internacional | Total         |
|---------|--------------|---------------|---------------|
| B1      | 112<br>(67%) | 55<br>(33%)   | 167<br>(100%) |

|    |              |             |               |
|----|--------------|-------------|---------------|
| B2 | 52<br>(53%)  | 47<br>(47%) | 99<br>(100%)  |
| B3 | 115<br>(66%) | 59<br>(34%) | 174<br>(100%) |

Tabela III – Distribuição da publicação dos periódicos científicos segundo estratos B1, B2 e B3 do Qualis

De acordo com a tabela a cima, nota-se que a maioria dos periódicos divididos nos estratos B1, B2 e B3, são nacionais.

O gráfico a baixo representa a porcentagem dos periódicos nacionais divididos entre os estratos B1, B2 e B3.

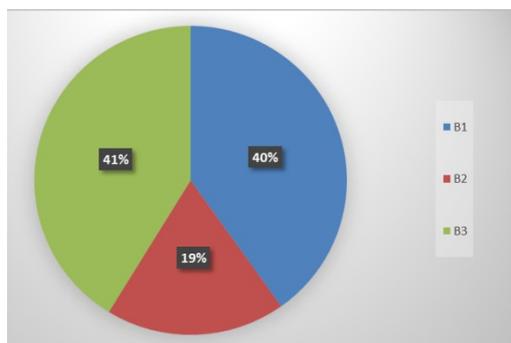


Gráfico 2. Distribuição de periódicos científicos nacionais segundo estratos B1, B2 e B3 do Qualis

A partir do gráfico, observa-se que os periódicos se concentram mais nos estratos B1 e B3.

Mostra-se neste gráfico a porcentagem dos periódicos internacionais divididos entre os estratos B1, B2 e B3.

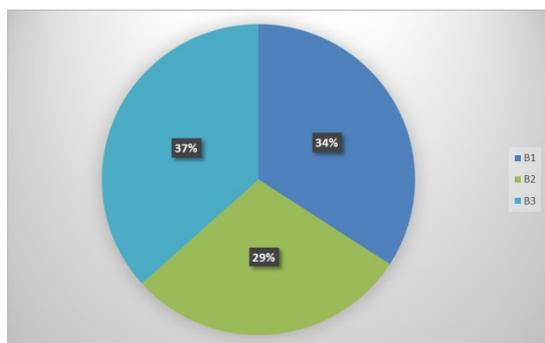


Gráfico 3. Distribuição de periódicos científicos internacionais segundo estratos B1, B2 e B3 do Qualis

Nesse caso, os periódicos internacionais são mais distribuídos entre os três estratos.

Tendo sido identificados os periódicos nacionais, utilizou-se os periódicos nacionais e os subdividiu em escopos gerais e de Biblioteconomia e Ciência da Informação. Confira na tabela a baixo:

| Estrato B1 | Números de Periódicos |
|------------|-----------------------|
| Geral      | 94                    |
| Ci. Inf.   | 18                    |

Tabela IV – Quantidade de periódicos científicos nacionais do Estrato B1 com escopo em Biblioteconomia e Ciência da Informação

No gráfico a baixo pode ser observado a grande diferença de porcentagem entre os periódicos nacionais da área de Biblioteconomia e Ciência da Informação e os periódicos gerais.

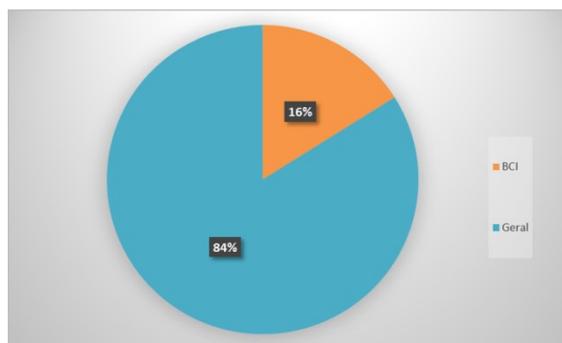


Gráfico 4. Distribuição de periódicos científicos nacionais com estratos B1 com escopo em Biblioteconomia e Ciência da Informação

A seguir, exhibe-se uma tabela com o número de periódicos nacionais, com estratos B2 e escopo em Biblioteconomia e Ciência da Informação.

| Estrato B2 | Números de Periódicos |
|------------|-----------------------|
| Geral      | 48                    |
| Ci. Inf.   | 4                     |

Tabela V – Quantidade de periódicos científicos nacionais do Estrato B2 com escopo em Biblioteconomia e Ciência da Informação

O gráfico 5 simboliza a porcentagem de periódicos nacionais com estrato B2, que são da área de Biblioteconomia e Ciência da Informação.

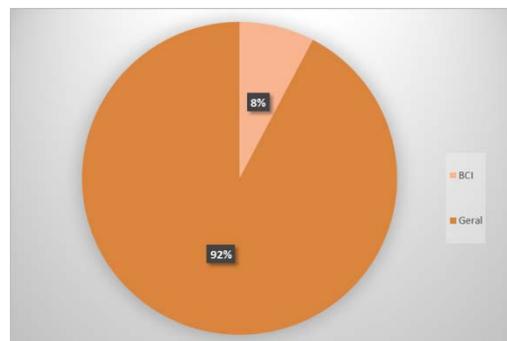


Gráfico 5. Distribuição de periódicos científicos nacionais com estratos B2 com escopo em Biblioteconomia e Ciência da Informação

A tabela a seguir tem como objetivo representar os periódicos nacionais, com estrato B3, que são da área de Biblioteconomia e Ciência da Informação.

| Estrato B3 | Números de Periódicos |
|------------|-----------------------|
| Geral      | 110                   |
| Ci. Inf.   | 5                     |

Tabela VI – Quantidade de periódicos científicos nacionais do Estrato B3 com escopo em Biblioteconomia e Ciência da Informação

Para finalizar, o gráfico a seguir mostra o percentual dos periódicos nacionais, com estrato B3 na área de Biblioteconomia e Ciência da Informação.

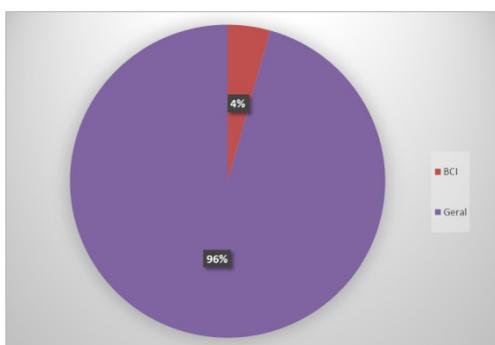


Gráfico 6. Distribuição de periódicos científicos nacionais com estratos B3 com escopo em Biblioteconomia e Ciência da Informação

#### 4.1. Análise qualitativa dos dados

A análise qualitativa da pesquisa foi desenvolvida a partir da análise de conteúdo, proposta por Bardin (1977).

Segundo Bardin (1977) a análise de conteúdo se organiza em três momentos: [1] pré-análise para escolha e preparo do material para análise; [2] exploração e classificação do material; [3] tratamento dos resultados pela interpretação dos dados.

A pré-análise é responsável por escolher os documentos a serem analisados e organizar o material. Nesta pesquisa a pré-análise foi feita ao acessar o sistema WebQualis, consultar os dados por classificação e selecionar a área Ciências Sociais Aplicadas I e os estratos B1, B2 e B3.



Figura 1. Acesso ao sistema WebQualis

Após feito isso, os dados foram sistematizados em um formulário contendo as informações: título do periódico, área de avaliação, escopo, estrato e nacionalidade, para posteriormente explorá-los.

A exploração do material é a etapa mais longa e trabalhosa e tem-se três sub etapas: a) escolha do tema, b) seleção das regras a serem utilizadas e c) escolha de categorias. Nesta etapa ocorre a realização das decisões tomadas na pré-análise, isto é, é aqui que os dados começam a ser tratados e classificados conforme o interesse da pesquisa.

Neste trabalho a exploração do material foi feita a partir dos dados agrupados no formulário. Inicialmente buscou o site dos periódicos nas bases MIAR(1) e ROAD(2). Os periódicos que não foram encontrados nessas bases foram procurados nas bases do SCIELO, IBICT e Latindex. Para os periódicos que não foram encontrados nessas bases, as buscas foram feitas no Google.

Após identificar os sites dos periódicos, fez-se uma varredura e descartou-se os periódicos internacionais, restando apenas os periódicos brasileiros para serem analisados. Nos sites dos periódicos nacionais, examinou-se a partir da descrição do periódico qual seria o escopo. Após esse diagnóstico, selecionou-se apenas os periódicos com escopo de Biblioteconomia e Ciência da Informação para, posteriormente, analisá-los. Com o material selecionado, deu-se início à exploração do mesmo. Buscou-se todos os artigos científicos publicados desde 1998 até 2013 nos periódicos nacionais com escopo de catalogação. Em seguida coletou-se os dados relevantes, tais como: título do artigo, ano de publicação, volume, número, link de acesso, autor, palavras-chave e resumo, em outro formulário. Terminada a coleta de dados, analisou-se as palavras-chave para identificação temática dos artigos, depois analisou os resumos para esclarecer possíveis dúvidas e por último agru-

pou-se em categorias os artigos relacionados à catalogação.

Finalizando a análise de conteúdo, foi feito o tratamento dos resultados, fase responsável por compreender e interpretar os dados. Após a classificação dos dados em categorias foi feita uma interpretação dos resultados de acordo com o tema proposto e criou-se uma nuvem de tags usando o wordle.

A análise da nuvem de tags foi feita para identificar as palavras mais utilizadas como palavras-chave, que pode ser vista a seguir:



Figura 2. Nuvem de tags criada a partir das ocorrências identificadas nos artigos

O tamanho das palavras relaciona-se à quantidade de vezes que ela foi citada. Sendo assim, a palavra mais citada aparecerá maior, que é o caso da expressão “Recuperação da Informação”, citada 47 vezes. Em seguida, temos as 10 palavras-chave mais citadas em ordem decrescente e a frequência que ela foi citada: Ontologia (43), Organização da Informação (37), Metadados (31), Organização do conhecimento (28), Indexação (24), Representação do conhecimento (21), Representação da informação (19), Classificação (16), Tesauro (16) e Linguagem Documentária (15). Todas essas palavras estão ligadas à catalogação.

#### 4.2. Análise qualitativa dos dados

Ao analisar os dados, observou-se que o número de artigos publicados em periódicos científicos nacionais, da área de Biblioteconomia e Ciência da Informação, com temática sobre catalogação, com estratos B1, B2 e B3 no sistema da Qualis, aumentou consideravelmente.

O acesso aos artigos utilizados nesta pesquisa pode ser consultado no apêndice C.

O crescimento entre 1998 e 2013 foi de 600% e pode ser visto na tabela a seguir:

| Ano  | Qtde. artigos |
|------|---------------|
| 1998 | 7             |
| 1999 | 5             |
| 2000 | 4             |

|      |    |
|------|----|
| 2001 | 13 |
| 2002 | 11 |
| 2003 | 7  |
| 2004 | 21 |
| 2005 | 16 |
| 2006 | 21 |
| 2007 | 32 |
| 2008 | 30 |
| 2009 | 31 |
| 2010 | 37 |
| 2011 | 35 |
| 2012 | 37 |
| 2013 | 42 |

Tabela VII – Distribuição de artigos de periódicos científicos nacionais estrato B1, B2 e B3 e com escopo em Biblioteconomia e Ciência da Informação entre os anos de 1998 e 2013

O gráfico a seguir indica a distribuição de artigos ao longo dos anos.

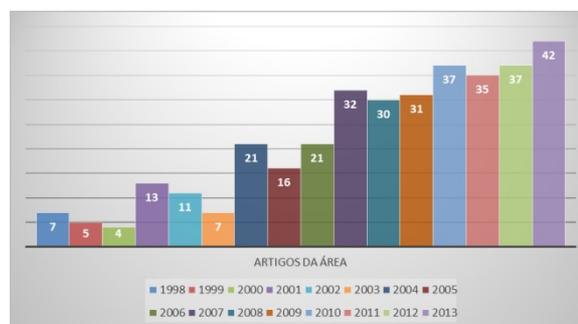


Gráfico 7. Distribuição de artigos de periódicos científicos nacionais estrato B1, B2 e B3 e com escopo em Biblioteconomia e Ciência da Informação entre os anos de 1998 e 2013

Observa-se no gráfico que o crescimento do número de artigos da área desde 1998 até o ano de 2013 não foi constante, teve ano que esse número diminuiu em relação ao ano anterior. Foi um avanço muito grande e significativo na área.

#### 5. Considerações finais

A relação entre as necessidades dos usuários e a criação de novas tecnologias se mistura de tal modo que não há como definir o sentido da influência. O que faz com que novas ferramentas e tecnologias se alterem.

A Biblioteconomia tem passado por intensas modificações, repercutindo nos estudos do cenário nacional da catalogação. Constatou-se que o forte avanço na área de catalogação a partir de 1998, foi devido às mudanças nos catálogos.

Os resultados desta pesquisa mostram que houve um crescimento considerável nos estudos da área de catalogação após 1998. No início desta pesquisa, acreditava-se que essa evolução na área de catalogação era em virtude das publicações dos FRBR, porém, não foi confirmada, uma vez que foi pouco mencionado nos artigos. Constatou-se que este avanço tenha se dado pelas novas ferramentas e tecnologias, ao fato da mudança ocorrida nos catálogos e na reconfiguração dos objetivos bibliográficos, com destaque à preocupação com os usuários. Na atualidade os catálogos são voltados aos usuários, sendo assim, as mudanças devem acompanhar suas necessidades, que envolve buscas nos catálogos on-line, por isso o sistema deve ser mais flexível, de fácil entendimento e de rápida navegabilidade.

Os procedimentos metodológicos foram válidos para a realização dessa pesquisa. Por meio deles, tornou-se exequível realizar o objetivo proposto neste trabalho. Com os objetivos específicos foi possível estudar a evolução do desenvolvimento da catalogação ao longo dos anos a partir da publicação dos FRBR em 1998 por meio de coleta de dados; identificar os periódicos nacionais da área de Ciências Sociais Aplicadas I que abrangem estudos de Biblioteconomia e Ciência da Informação com Qualis, por meio do escopo editorial e da análise dos dados coletados; e identificar os artigos publicados a partir de 1998 sobre catalogação a partir da análise das palavras-chave e do resumo de cada artigo.

Após a análise dos resultados, conclui-se que, houve um aumento de 600% no número de artigos publicados nas revistas científicas nacionais com estratos B1, B2 e B3 sobre catalogação no sistema Qualis, desde 1998. Verificou-se que, a repercussão dada à publicação dos FRBR nas publicações não se afirmou, isto é, a expectativa não se confirmou e deduz-se que o avanço na catalogação está voltado às necessidades dos usuários, às mudanças ocorridas nos catálogos e ao surgimento de novas ferramentas e tecnologias.

## Notas

- (1) Base de dados que por meio do Índice Composto de Difusão Secundária (ICDS) mostra em quais bases de dados estão os periódicos científicos e exibe alguns dados do periódico. Disponível em: <http://miar.ub.edu/>.
- (2) Base de dados que conta com o apoio da UNESCO e ao pesquisar nela você encontra todos os dados do periódico, inclusive o site do periódico. Disponível em: <http://road.issn.org/>

## Referências

- BAPTISTA, Dulce. Catalogação: uma revolução em curso. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 9.. 2008. São Paulo. Anais... São Paulo, 2008. p. 1-9. Disponível em: <http://repositorios.questoesemrede.uff.br/repositorios/bitstream/handle/123456789/789/7.pdf?sequence=1>. Acesso em 19 abr. 2014.
- BARACAT, Alyssa Cecilia; RIGOLIN, Camila Carneiro Dias. Proteção do Conhecimento e Movimento Open Access: Discussões no Âmbito da Organização Mundial da Propriedade Intelectual | Knowledge protection and open access movement: discussions in the scope of the World Intellectual Property Organization. Liinc em Revista, v. 8, n. 2, 18 dez. 2012a.
- BARBOSA, Alice Príncipe. Novos rumos da catalogação. Rio de Janeiro: BNG/Brasilart, 1978.
- BARDIN, Laurence. Análise de conteúdo. [L'analyse de contenu]. Luis Antero Reto (Trad.). São Paulo: Edições 70, c1977. 225 p.
- BIBLIOTECA NACIONAL DE PORTUGAL. ISBD: Descrição Bibliográfica Internacional Normalizada. [s.d.]. Disponível em: [http://www.bnportugal.pt/index.php?option=com\\_content&view=article&id=491:isbd&catid=78:normalizacao&Itemid=541&lang=pt](http://www.bnportugal.pt/index.php?option=com_content&view=article&id=491:isbd&catid=78:normalizacao&Itemid=541&lang=pt). Acesso em: 11 mai. 2014.
- BOAI. Budapest Open Access Initiative | Portuguese (Brazilian) Translation. Disponível em: <http://www.budapestopenaccessinitiative.org/boai-10-translations/portuguese-brazilian-translation>. Acesso em: 08 nov. 2014.
- COORDENAÇÃO DE APERFEIÇOAMENTO DE NÍVEL SUPERIOR. Qualis. 2009. Disponível em: <http://www.capes.gov.br/servicos/sala-de-imprensa/36-noticias/2550-capes-aprova-a-nova-classificacao-do-qualis>. Acesso em: 19 abr. 2014.
- CORRÊA, Rosa Maria Rodrigues. Catalogação descritiva no século XXI: um estudo sobre RDA. 2008. 65 f. Dissertação (mestrado em Ciência da Informação) – Universidade Estadual Paulista, Marília. 2008. Disponível em: [http://www.marilia.unesp.br/Home/Pos-Graduacao/CienciadaInformacao/Dissertacoes/correa\\_rmr\\_me\\_mar.pdf](http://www.marilia.unesp.br/Home/Pos-Graduacao/CienciadaInformacao/Dissertacoes/correa_rmr_me_mar.pdf). Acesso em: 19 abr. 2014.
- COSTA, M.; AMARO, B.; ASSIS, T. Serviços de difusão de políticas para o acesso aberto: Sherpa/Juliet, Melibea e ROARMAP. Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação, v. 17, n. esp. 2, p. 36–54, 2012.
- CREATIVE COMMONS. Get Creative Commons updates, 2013. Disponível em: <http://creativecommons.org/about>. Acesso: 09 nov. 2014.
- DENTON, William. FRBR and the history of cataloging. In: TAYLOR, Arlene. Understanding FRBR: What it is and how it will affect our retrieval tools. Westport: Libraries Unlimited, 2007. Disponível em: <http://yorkspace.library.yorku.ca/xmlui/bitstream/handle/10315/1250/denton-frbr-and-the-history-of-cataloging.pdf?sequence=1>. Acesso: 11 mai. 2014.
- FIGUEIREDO, Nice Menezes de. A modernidade das cinco leis de Ranganathan. Ciência da Informação. Brasília, v. 21, n. 3, p. 186-191, set/dez. 1992. Disponível em: <http://revista.ibict.br/ciinf/index.php/ciinf/article/viewFile/1277/911>. Acesso em 28 mai. 2014.
- GERHARDT, Tatiana Engel; SILVEIRA, Denise Tolfo (Org.). Métodos de pesquisa. Porto Alegre: Editora da UFRGS, 2009. (Série Educação a Distância).
- GIL, Antônio Carlos. Como elaborar projetos de pesquisa. 4. ed. São Paulo: Atlas, 2002.

- GIL, Antônio Carlos. Métodos e técnicas de pesquisa social. 6. ed. São Paulo: Atlas, 2008.
- GONSALVES, *Elisa Pereira. Conversas sobre a iniciação à pesquisa científica. 5. ed. Campinas: Alínea, 2011.*
- HUNTER, E. J.; BAKEWELL, K. G. B. Cataloguing. 2nd. rev. exp. ed. London: Clive Bingley, 1983.
- INTERNATIONAL FEDERATION OF LIBRARY ASSOCIATION. IFLA meeting of experts on an international cataloguing code. 2014. Disponível em: <<http://www.ifla.org/node/576>>. Acesso: 29 mai. 2014.
- INTERNATIONAL FEDERATION OF LIBRARY ASSOCIATION. (2009). Functional Requirements for Bibliographic Records (FRBR): final report. [s.l.]: IFLA.
- INTERNATIONAL FEDERATION OF LIBRARY ASSOCIATION. Fifth IFLA meeting of experts on an international cataloguing code. 2007. Disponível em: <<http://www.imeicc5.com/>>. Acesso: 29 mai. 2014.
- INTERNATIONAL FEDERATION OF LIBRARY ASSOCIATION. Fourth IFLA meeting of experts on an international cataloguing code. 2006. Disponível em: <<http://www.nl.go.kr/icc/icc/main.php>>. Acesso: 29 mai. 2014.
- INTERNATIONAL FEDERATION OF LIBRARY ASSOCIATION. Second IFLA meeting of experts on an international cataloguing code. 2004. Disponível em: <<http://archive.ifla.org/imeicc2/>>. Acesso: 29 mai. 2014.
- INTERNATIONAL FEDERATION OF LIBRARY ASSOCIATION. Third IFLA meeting of experts on an international cataloguing code. 2005. Disponível em: <<http://archive.ifla.org/imeicc3/>>. Acesso: 29 mai. 2014.
- INTERNATIONAL FEDERATION OF LIBRARY ASSOCIATIONS. Declaração de princípios internacionais de catalogação. 2009. Disponível em: <[http://www.ifla.org/files/assets/cataloguing/icp/icp\\_2009-pt.pdf](http://www.ifla.org/files/assets/cataloguing/icp/icp_2009-pt.pdf)>. Acesso em: 19 abr. 2014.
- JOINT STEERING COMMITTEE FOR DEVELOPMENT OF RDA. RDA Resource description and access. 2009. Disponível em: <<http://www.rda-jsc.org/rda.html>>. Acesso em: 11 mai. 2014.
- KAUARK, Fabiana da Silva; MANHÃES, Fernanda Castro; MEDEIROS, Carlos Henrique. Metodologia da pesquisa: um guia prático. Itabuna, 2010.
- LE BOEUF, Patrick. O admirável mundo novo do FRBR. 2007. Disponível em: <[http://www.imeicc5.com/download/portuguese/Presentations2c\\_BraveNewFRBRWorld\(PR\)\\_Port.pdf](http://www.imeicc5.com/download/portuguese/Presentations2c_BraveNewFRBRWorld(PR)_Port.pdf)>. Acesso em: 19 abr. 2014.
- MEADOWS, Arthur Jack. A comunicação científica. Tradução Antonio Agenor Briquet De Lemos. Brasília: Briquet De Lemos, 1999.
- MORENO, Fernanda Passini. Modelo conceitual FRBR: discussões recentes e um olhar sobre as tarefas do usuário. Ciência da Informação, Florianópolis, v. 14, n. 27, p. 47-68, 2009. Disponível em: <<http://www.brapci.ufpr.br/documento.php?dd0=0000006843&dd1=ccf71>>. Acesso em 19 mai. 2014.
- OLIVER, C. (2010). Introducing RDA: a guide to the basics. Chicago: ALA.
- RANGANATHAN, S. R. As cinco leis da Biblioteconomia. Tradução Tarcisio Zandonade. Brasília: Briquet De Lemos, 2009.
- SHEOLA, Noah. Charles Ammi Cutter. 2010. Disponível em: <<https://www.bostonathenaeum.org/library/book-recommendations/athenaeum-authors/charles-ammi-cutter>>. Acesso em: 18 mai. 2014
- SILVA, Eliana Barboza de Oliveira et al. Conceituação e aplicação do novo padrão para descrição bibliográfica Resource Description and Access. CRB-8 Digital, São Paulo, v. 1, n. 5, p. 113-123, jan. 2012. Disponível em: <<http://revista.crb8.org.br/index.php/crb8digital/article/viewFile/74/76>>. Acesso em: 19 abr. 2014.
- SMIRAGLIA, Richard. The history of the work in the modern catalog. Cataloging & Classification Quarterly, v. 35, n. 3/4, p.553-567, ago. 2003.
- TILLET, Bárbara. Visão Geral do Esboço da Declaração dos Princípios de Catalogação. 2007. Disponível em: <[http://www.imeicc5.com/download/portuguese/Presentations4\\_Cataloguing%20principles-IME%20ICC%20overview\\_BT\\_Port.pdf](http://www.imeicc5.com/download/portuguese/Presentations4_Cataloguing%20principles-IME%20ICC%20overview_BT_Port.pdf)>. Acesso em 29 mai. 2014.
- ZAFALON, Zaira. Objetivos bibliográficos. São Carlos, 2011. Resenha de: SVENONIUS, Elaine. Bibliographic objectives. In: \_\_\_\_\_. The intellectual foundation of information organization. Cambridge: MIT Press, 2000. p. 15-30.
- ZIMAN, J. Conhecimento público. São Paulo: Itatiaia, 1979.

---

# O profissional da informação na elaboração de políticas públicas de informação e tecnologia

---

*Professional information in policy making public information and technology*

---

**Ângela Maria Grossi de Carvalho (1), Thabyta Giraldeili Marsulo (2)**

(1) Universidade Estadual Paulista Júlio de Mesquita Filho – Faculdade de Filosofia e Ciências, Campus de marília, Av. Hygino Muzzi Filho, 737, Marília –SP, Brasil, [angela@carvalho.jor.br](mailto:angela@carvalho.jor.br)

(2) [thabytagm@hotmail.com](mailto:thabytagm@hotmail.com)

## Resumo

O presente estudo tem como tema o desempenho dos profissionais da informações sobre o ciclo de vida e o desenvolvimento de políticas públicas e tecnologia da informação. Assumimos que Tecnologia da Informação (TIC), e não são apenas aplicando instrumentos no contexto das ciências da informação de investigação. Afeta significativamente o comportamento das pessoas e contribuir para as profundas mudanças na maneira de ser e de pensar da humanidade na sociedade da informação e, assim, transformar o processo político de construção de um espaço para a colaboração e participação social.

**Palavras-chave:** Políticas públicas de informação e tecnologia. O papel do profissional da informação. Avaliação de políticas públicas. Tecnologias de informação e comunicação.

## 1. Introdução

A preocupação com a estruturação, representação, organização, disseminação e uso de conteúdos informacionais acompanha o desenvolvimento cultural da humanidade. Contudo, é na contemporaneidade, com a ruptura das barreiras de tempo e espaço propiciadas pelas Tecnologias de Informação e Comunicação (TIC) que essa preocupação adquire importância ainda maior.

Essas tecnologias, cujas interfaces interferem nos modos de produzir e acessar informação propiciando o surgimento de um novo sistema hibridizado, que modificam, o fluxo informacional aumentando-o exponencialmente, tornam prementes ações que favorecem o efetivo acesso à informação e ao conhecimento, bens indispensáveis para o desenvolvimento global da humanidade.

Segundo Davenport e Prusak (1998) O conhecimento é a informação agregada, pois existe diante do contexto vivenciado por alguém, que

## Abstract

The present study has as its theme the professional performance information on the life cycle and development of public policy and information technology. We assume that Information and Communication Technologies (ICTs) are not only applying tools in the context of research of Information Science. They greatly affect the behaviour of individuals and contribute to the profound changes in the way of being and thinking of humanity in the information society and thus transform the policy building process in a collaborative space and social participation.

**Keywords:** Informed public policies and technology; Role of information professional. Evaluation of public policies. Information and communication technologies.

deu a aquela informação sua interpretação e significado de modo a tornar-se mais valioso.

A sociedade do conhecimento abarca todas as ferramentas e todas as oportunidades que as tecnologias vem proporcionando as pessoas nos últimos anos, colocando a disposição uma grande quantidade de informações e múltiplas formas de acesso a elas.

Esta sociedade se baseia no uso compartilhado de recursos, na construção coletiva de conhecimento, na interação livre de restrições de espaço e tempo e, na valorização do direito à informação, às tecnologias de informação e comunicação e à educação, como um bem comum, assim a população passa a ter mais poder diante da reivindicação de seus direitos, uma vez que, a informação e o conhecimento, matérias-primas indispensáveis para a construção da nova sociedade.

Dziekaniak e Rover (2011, P. 7) apontam que a “evolução das tecnologias da informação e comunicação e a sua invasão nos países periféri-

cos, sem um compromisso sério dos Governos em repensar a nova economia, ou proposição e aplicação de novas formas de manutenção e gestão da sociedade como um todo, levando em consideração o grande impacto que as tecnologias da informação e comunicação trazem (*principalmente na geração de empregos e desenvolvimento social*), faz com que os inúmeros aspectos negativos ganhem terreno e coloquem a Sociedade da Informação em crise. Uma crise sem precedentes na história, dado que os aspectos que a desencadearam tendem a crescer exponencialmente e se desenvolvem. “

Assim as políticas públicas deviam traduzir em seu processo de elaboração e implantação e, sobretudo, em seus resultados as formas de exercício do poder político desta população, envolvendo a distribuição e redistribuição de poder (e do conhecimento), o papel do conflito social nos processos de decisão e a distribuição de custos e benefícios sociais. (Teixeira, 2002)

Entretanto para que isso ocorra Lévy (2010) aponta que na sociedade do conhecimento as pessoas devem ser preparadas e educadas para a interação com as informações: *“Não se deve fabricar pessoas que consomem informação previamente empacotadas por terceiros”*. Para o autor, a população tem de ser formada e habilitada para a interação crítica, a fim de compreender qual informação possui fonte fidedigna e ser capaz de encontrar a informação que procuram enquanto produzem informação para ser consumida, através da inserção do material contida nas fontes em seu contexto social específico, resultando em um movimento de troca, colaboração e complementação de conhecimentos, que alimenta o sistema.

Entretanto como o poder é uma relação social que envolve vários atores com projetos e interesses diferenciados e até contraditórios, há necessidade de mediações sociais e institucionais, para que se possa obter um mínimo de consenso e, assim, as políticas públicas possam ser legitimadas e obter eficácia.

No âmbito da mediação o profissional da informação apresenta-se qualificado para atuar, sendo assim o presente projeto busca inserir o profissional da informação e suas habilidades mediadoras no processo de formulação e avaliação de políticas públicas, aproximando governantes e governados em uma discussão produtiva de melhoria da sociedade em que estão locados.

## 2. Metodología

O estudo apresenta características de pesquisa fundamental e aplicada. No primeiro caso nos interessamos pelas investigações do campo teórico e metodológico da Ciência da Informação que contribui para a validação das teorias propostas neste projeto. A compreensão dos sujeitos sociais no processo de atuação do profissional da informação no âmbito das políticas públicas, através da conexão do percurso histórico da ciência da informação e seus profissionais e o cenário político nacional, reforçando a importância do envolvimento do profissionais da ciência da informação na elaboração destas políticas. Dessa forma, se faz necessária a investigação das políticas de Informação e Tecnologias no âmbito da Ciência da Informação, objetivando o desenvolvimento da sociedade da informação e contribuindo com melhorias de acesso à informação e construção do conhecimento.

## 3. Objetivos

Analisar a inserção do profissional da informação no âmbito do ciclo das políticas públicas como agente central na compreensão das demandas informacionais, e apontar os benefícios dessa atuação na elaboração, execução e avaliação das políticas públicas, para a socialização da informação para os cidadãos envolvidos no processo.

## 4. Referencial teórico

As políticas públicas vêm sendo definida historicamente como um campo do conhecimento que busca, ao mesmo tempo, cobrar ações de governos democráticos e constantemente avaliar essas ações propondo mudanças e adequações sempre que necessárias.

Considera-se que este ramo do conhecimento contou com quatro grandes fundadores: H. Laswell, H. Simon, C. Lindblom e D. Easton, que apresentaram conceitos norteadores dos estudos de ação do poder público; regras e procedimentos para as relações entre poder público e sociedade, mediações entre atores da sociedade e do Estado, nesse caso, políticas explicitadas, sistematizadas ou formuladas em documentos (leis, programas, linhas de financiamentos) que orientam ações que normalmente envolvem aplicações de recursos públicos).

Os objetivos das políticas têm uma referência valorativa e exprimem as opções e visões de mundo daqueles que controlam o poder, mesmo que, para sua legitimação, necessitem contemplar certos interesses de segmentos sociais

dominados, dependendo assim da sua capacidade de organização e negociação.

Assim as políticas públicas se desenvolvem através de um ciclo deliberativo, formado por vários estágios e constituindo um processo dinâmico e de aprendizado.

## 5. Discussão

A evolução dos recursos tecnológicos, nas últimas décadas, trouxe facilidades no acesso a uma grande quantidade de informações, o que era no passado exclusividade de pequenos grupos, foi se tornando cada vez mais acessível, atingindo assim indivíduos de diversos âmbitos da sociedade.

A partir do final dos anos 90 começou a ser moldado um novo modelo de financiamento e gestão de ciência e tecnologia no Brasil, considerando que as atividades de produção de indicadores quantitativos em ciência, tecnologia e inovação vêm se fortalecendo no país na última década, com o reconhecimento da necessidade, por parte dos governos federal e estaduais e da comunidade científica nacional, de dispor de instrumentos para definição de diretrizes, alocação de investimentos e recursos, formulação de programas e avaliação de atividades relacionadas ao desenvolvimento científico e tecnológico no país.

Segundo o IBICT (2011) a nova política de ciência e tecnologia deve implementar tarefas aparentemente contraditórias: estimulando a liberdade, iniciativa e criatividade dos pesquisadores e, ao mesmo tempo, estabelecer um forte vínculo entre o que eles fazem e as necessidades da economia, do sistema educacional e da sociedade como um todo.

As políticas destes últimos vinte e cinco anos podem ser entendidas como desdobramentos das mudanças ocorridas na sociedade e economia brasileiras nas décadas anteriores..

Na área de ciência e tecnologia, o maior desafio no Brasil é a elaboração e a implementação de uma política de longo prazo que permita ao desenvolvimento científico e tecnológico alcançar a população e que efetivamente tenha um impacto determinante na melhoria das condições de vida da sociedade.

Sendo assim associamos o profissional da informação ao contexto das políticas públicas quando ressaltamos as formas de atuação do Estado com relação aos elementos estruturais da Sociedade, uma vez que suas políticas podem traçar o horizonte e definir os modos de interação dos indivíduos, grupos, organizações

e instituições públicas e privadas, tanto no interior do Estado quanto fora de seus limites institucionais. (MIRANDA, 2000,p.4).

Para a maioria dos analistas, as mudanças no conteúdo e na metodologia das políticas públicas só se efetuarão com mudanças nas elites políticas, na composição do poder político. É certo que mudanças mais substantivas só podem ocorrer quando efetivamente se muda a composição do poder, mas pode-se obter conquistas sociais através da mobilização social, da ação coletiva, sobretudo quando esta passa a ter um conteúdo de proposição, de debate público de alternativas e não de mera crítica. Para isso, é necessário que as proposições sejam legitimadas por um amplo consenso e que tenham uma abrangência maior que os interesses corporativos ou setoriais.

Observa-se que o desenho de muitas políticas públicas são verdadeiras caixas pretas, no sentido de abstração. Para se entender como certos bens e serviços são obtidos e como os atores estão envolvidos no processo é necessário que opere uma transparência para que a população entenda como os produtos e serviços anunciados correspondem as transformações do meio.

Nesse aspecto a atuação de um profissional capacitado para mediar informações e construir uma cultura para compartilhar conhecimento, criar ambientes para transferência de conhecimento tácito (Sveiby, 2000) de acordo com a necessidade de um público específico se torna essencial.

Segundo REMEIKIS (1996), o profissional mediador atua como um “canal” para distribuir conhecimento pelas organizações e tem habilidades específicas de sua área e atendem às diferentes interações das várias áreas desta instituição.

Os profissionais da informação usam diariamente essas habilidades nas suas atividades, como esclarecem Tarapanoff, Suaiden e Oliveira (2002) ao afirmar que este profissional tem a habilidade de lidar com a informação e conhecimento, agregando valor aos mesmos, a fim de trabalhar com pessoas incentivando-as a participar da sociedade e exercer a cidadania.

Na sociedade civil também há uma diversidade de interesses e de visões que precisa ser debatida, confrontada, negociada, buscando-se um consenso mínimo. Essa formulação hoje se torna complexa devido à fragmentação das organizações, apesar de algumas iniciativas de articulação em alguns setores.

Esses elementos precisam ser traduzidos contudo em parâmetros objetivos, para que possam nortear a elaboração, implementação e avaliação das políticas propostas, neste ponto o profissional da informação precisam, focar sua formação na transdisciplinariedade, permitindo que suas habilidades sejam potencializadas para atender uma determinada demanda social.

## 6. Considerações

A Ciência da Informação (CI) desenvolveu-se historicamente porque os problemas informacionais modificaram completamente sua relevância para a sociedade. Transmitir o conhecimento para aqueles que dele necessitam é uma responsabilidade social, e essa responsabilidade social parece ser o verdadeiro fundamento da Ciência da Informação.

A era digital vem impactando drasticamente o planejamento das atividades de gestão pública, exigindo que se efetue novas configurações para acomodar o uso da tecnologia, tanto para a armazenagem quanto para prover acesso às informações digitais disponíveis em redes de computadores e na web. Nesse processo de mudanças, a presença cada vez mais ativa da sociedade civil nas questões de interesse geral, torna fundamental a necessidade do debate público, da transparência, da sua elaboração em espaços públicos e não nos gabinetes governamentais.

Os bibliotecários nos países do Terceiro Mundo precisam ampliar seu senso crítico sobre a política de informação e tecnologia, principalmente no que se refere à questão da transferência da informação. (AMARAL, 1995) Além disso, penso que os profissionais da área devem combinar sua experiência na linha de frente da prática com a perspectiva teórica dos estudiosos para a formulação conjunta de uma economia política mais inclusiva.

As mudanças sociais associadas as tecnologias mostram que as relações profissionais estão sujeitas a mudanças viscerais e todas elas afetam a questão comportamental do indivíduo. Percebe-se que já não há limites entre a esfera pública e o espaço privado.

Uma atuação consciente se faz valer pelo princípio de fornecer informação correta e fazer uso da disseminação da informação, o requisito de o profissional ter cultura geral sólida para saber dar a cada usuário a informação desejada, atender a todos com imparcialidade, respeitando a pessoa humana.

O imperativo tecnológico que vem expondo o desenvolvimento e aplicação de uma crescente

gama de produtos e serviços de informação vindos de dentro e de fora da Ciência da Informação, com crescente competição tornando como o ponto crucial do manejo da informação, a responsabilidade básica dos especialistas da informação constantemente inovar produtos e serviços para prover o conhecimento de direito a todo cidadão. Nesse caminho, as condições de desenvolvimento da própria sociedade indicarão suas possibilidades de assimilação das novas tecnologias e do seu nível de entendimento do valor dessa absorção para o progresso científico e tecnológico tornando-o um profissional multifacetado capaz de transformar realidade sociais através dessa mediação política.

## Referências

- AMARAL, S. A. (1995) Serviços bibliotecários e desenvolvimento social: um desafio profissional. *Ciência da Informação*, Brasília, v.24, n.2
- DAVENPORT, T. H., PRUSAK, L.(1998) *Conhecimento empresarial*. Rio de Janeiro: Campus,
- DZIEKANIAK, G.; ROVER, A. (2011) *Sociedade do Conhecimento: características, demandas e requisitos*. Data-GramaZero – Revista de Informação, v. 12, n.5, [http://www.dgz.org.br/out11/F\\_I\\_art.htm](http://www.dgz.org.br/out11/F_I_art.htm)
- LEVY, P. (2010) *Cibercultura*. Rio de Janeiro: p. 34
- MIRANDA, A.(2000) *Sociedade da informação: globalização, identidade cultural e conteúdos*. *Ci. Inf.*, Brasília, v. 29, n. 2, Aug
- REMEIKIS, L A. (1996, winter.) *Knowledge management – roles for information professionals*. *Business & Finance Division Bulletin*, Chicago, n. 101, p. 41-43,
- SVEIBY, K.E. (2000) *Métodos para avaliar ativos intangíveis*. Disponível em: [http://www.unifae.br/publicacoes/pdf/revista\\_da\\_fae/fae\\_v7\\_n2/rev\\_fae\\_v7\\_n2\\_01.pdf](http://www.unifae.br/publicacoes/pdf/revista_da_fae/fae_v7_n2/rev_fae_v7_n2_01.pdf)
- TEIXEIRA, E. C. (2014) *O Papel das Políticas Públicas no Desenvolvimento Local e na Transformação da Realidade*. Disponível em <http://www.aatr.org.br/papelpp.pdf>

---

# A Informação Contábil na construção do conhecimento organizacional: uma modelagem a partir da DRE

*The accounting information to build organization knowledge: modelling from the Income Statement*

---

**Maria Amélia Duarte Oliveira Ferrarezi (1), Ângela Cristina Basílio de Freitas (2), Marinês Santana Justo Smith (3), Bárbara Fadel (4).**

(1) Mestranda Uni-FACEF; Docente SENAC. (16) 3724-1862, maria.adoliveira.sp.senac.br

(2) Mestre Uni-FACEF, Docente SENAC. (16) 3713-4618, angelafreitas@facef.br

(3) Doutora em Ciências da Informação, Docente titular Uni-FACEF. (16) 3713-4661, mari-nessmith@netsite.com.br

(4) Doutora em História Social, Docente e Coordenadora do PPG-DR Uni-FACEF. (16) 3713-4619, bafadel@facef.br

## Resumo

O ambiente organizacional requer conteúdos informacionais para o alcance de metas que podem divergir em função dos valores culturais. Para o alcance das metas as organizações necessitam gerenciar variáveis como: receita, custo, produtividade, tributação, despesa, que estão associadas à Demonstração de Resultado do Exercício (DRE). O objetivo desta pesquisa é contribuir com a construção de uma modelagem a partir da DRE capaz de orientar as decisões no tocante a planejamentos e análises com ênfase na lucratividade. Cenário que conduz responder, como a DRE pode contribuir na construção de uma modelagem que orienta de forma sistêmica planejamentos e decisões no tocante a lucratividade e que, ainda, possa promover o conhecimento organizacional. Busca-se apresentar a DRE na perspectiva informacional; destacar o uso da informação na construção do conhecimento e elaborar a proposta de uma modelagem decisória a partir da DRE. No que se refere ao método, trata-se de discussão teórica, com abordagem qualitativa dos dados, cujos objetivos são de caráter descritivo, com estratégia bibliográfica e documental. Como resultado, a partir da abordagem sistêmica foi possível construir a MODRE (Modelagem decisória a partir da DRE no contexto da informação contábil), como um sistema de informação que subsidia as decisões estratégicas, táticas e operacionais através dos indicadores informacionais.

**Palavras-chave:** Demonstração de Resultado do Exercício. Modelagem sistêmica. Conhecimento organizacional. Informação contábil. Tomada de decisões.

## 1. Introdução

No contexto organizacional, considerando que os sistemas e subsistemas são apoiados por informações e conhecimentos, é possível observar que o desenvolvimento organizacional

## Abstract

The organizational environment requires effective information contents for achieving goals that may differ depending on cultural values. To achieve this goal, organizations need to manage variables such as revenue, cost, productivity, taxation and expenditure, which are associated with the Statement of Income for the Year (DRE). Given this context, the objective of this research is to contribute to building a modeling from the DRE which is able to guide decisions regarding planning and analysis with an emphasis on profitability. Thus, we are led to answer how the DRE can contribute in building a modeling that systemically guides the planning and the decisions related to profitability and that can also promote organizational knowledge. Therefore, we seek to present the DRE in the informational perspective; highlight the use of information in the construction of organizational knowledge and finally prepare a proposal for a decisional modeling from the DRE in the context of accounting information. In relation to the method, it is theoretical discussion, with qualitative approach, the objectives of which are descriptive, with bibliographical and documentary strategy. As a result, from the systemic approach it was possible to build MODRE (Operative Modeling from the DRE in the context of accounting information) as an information system that supports the strategic, tactical and operational decisions through the informational indicators.

**Keywords:** The Income Statement. Systemic modeling. Organizational knowledge. Accounting information; Decision-making.

depende do acesso e uso da informação dentro do seu ciclo de vida. O uso das informações em ambientes organizacionais pode contribuir para o desenvolvimento da organização por meio da aprendizagem e auxiliar no processo decisório. O processo decisório transita pelos níveis estratégicos, táticos e operacionais, os quais reque-

rem conteúdos informacionais específicos. No que tange a problemática desta pesquisa, a necessidade informacional que se busca, passa pelos três níveis do processo decisório.

Diante desse contexto este trabalho tem como objetivo geral contribuir com a construção de uma modelagem a partir da DRE capaz de orientar as decisões no tocante a planejamentos e análise com ênfase na lucratividade. A fim de alcançar este objetivo apresenta-se a DRE na sua perspectiva informacional e seus produtos informacionais, destaca o uso da informação na construção do conhecimento organizacional e elabora a modelagem decisória a partir da DRE no contexto da informação contábil.

Como procedimentos metodológicos utiliza-se a pesquisa bibliográfica e documental de abordagem qualitativa e de caráter descritivo, com o intuito de responder como a DRE contribui na construção de uma modelagem que orienta planejamentos e decisões no tocante a lucratividade, promovendo o conhecimento organizacional.

Para tanto este estudo pretende contribuir de forma interdisciplinar por meio da gestão econômico-financeira para a tomada de decisões no tocante à DRE.

## 2. Demonstração do Resultado do Exercício e seus aspectos informacionais

A Demonstração do Resultado do Exercício (DRE) é uma demonstração contábil que evidencia a apuração do resultado de uma organização, com o objetivo de identificar o lucro ou prejuízo num determinado período, respeitando as características qualitativas fundamentais e de melhoria da informação contábil-financeira, evidenciando os aumentos ou reduções ocasionados no Patrimônio Líquido da empresa. Essa demonstração contábil encontra-se devidamente regulamentada no Brasil pelo Comitê de Pronunciamento Contábil - CPC 26 R1 e a Norma Brasileira de Contabilidade Técnica Geral – NBC TG 26, aprovada pela Resolução 1.185/09, as quais atendem os padrões do *International Financial Reportins Standards* (IFRS), especificamente a *Internacional Accounting Standards - IAS 1*. Todavia o processo de adequação ao padrão internacional encontra-se em fase de adaptação, principalmente no que tange às pequenas empresas.

Nesse cenário necessário de regulamentação no Brasil a DRE é assim definida por Ludícibus:

A Demonstração do Resultado do Exercício é um resumo ordenado das receitas e despesas da empresa em determinado período de (12 meses). É

apresentada de forma dedutiva (vertical), ou seja, das receitas subtraem-se as despesas, e, em seguida indica-se o resultado (lucro ou prejuízo) (IUDÍCIBUS, 2012, p. 37-39).

Ludícibus e Marion (2002) afirmam que a DRE é imprescindível para a tomada de decisão, visto que a mesma destaca todas as receitas e despesas, os impostos, os diversos tipos de lucro e outras informações relevantes.

Para os pesquisadores Ludícibus, et al., (2010, p. 475), a Demonstração do Resultado do Exercício (DRE), tem como objetivo “demonstrar o resultado líquido do período”, a partir das operações de compras, vendas, gastos com a manutenção do patrimônio e gastos com esforços para geração das receitas, realizado pelas empresas durante o exercício social. Operações planejadas e controladas pela gestão, todavia mutuamente influenciadas pelo ambiente econômico, mas também pela cultura organizacional, conforme estudo de Smith (2013).

Mediante essa regulamentação apresenta-se abaixo a estrutura da Demonstração de Resultado do Exercício (DRE), para que a mesma sirva como base informacional na construção dos indicadores aos usuários internos e externos:

|   |
|---|
| Receita Operacional Bruta   |
| (-) Deduções e abatimentos da Receita bruta   |
| (-) Devolução de vendas/ Impostos e Contribuições incidentes sobre venda                  |
| (-) Descontos/abatimentos   |
| (=) Receita Operacional Líquida   |
| (-) Custos produtos vendidos  |
| (-) Custo das Mercadorias Vendidas  |
| (-) Custo dos Serviços prestados  |
| (=) Lucro Operacional Bruto (Margem ou Resultado)   |
| (-) Despesas Operacionais   |
| (-) Despesas com vendas   |
| (-) Despesas Administrativas  |
| (-) Despesas Tributárias  |
| (-) Despesas Operacionais Financeiras   |
| (-) Despesas Financeiras  |
| (+) Receitas Financeiras  |
| (+/-) Variações Cambiais  |
| (+/-) Outras receitas ou (despesas)   |
| (=) Lucro Líquido Antes do Imposto de Renda e da Contribuição Social                      |
| (-) Provisão para Imposto de Renda e Contribuição Social sobre o Lucro                    |
| (=) Lucro Líquido pós provisão e Antes das Participações                                  |
| Administradores e Partes beneficiárias e Fundo de Assistência e Previdência de Empregados |
| (=) Lucro Líquido do Exercício  |

### Quadro 1 - Demonstração do Resultado do Exercício

Fonte: elaborado pelas autoras conforme CPC 26 R1, NBC TG 26 – Resolução 1.185/09.

Na perspectiva da DRE como base informacional para processos decisórios e planejamentos organizacionais, Smith (2013) destaca a contabilidade classificada em três ramos, a Societária, Gerencial e Fiscal, sob esta perspectiva a DRE surge como um produto informacional da contabilidade Societária, com a finalidade de “Informar o resultado econômico gerado pela dinâmica operacional em determinado período”, capaz de subsidiar o alcance dos propósitos informacionais citado por Smith (2013), em cada um dos ramos da contabilidade. Questão que se concretiza em conformidade com a finalidade da informação da DRE, mas principalmente como a informação pode ser aplicada e utilizada.

A DRE é relevante para análise, avaliação de quocientes e ou indicadores financeiros para a tomada de decisões nas organizações. Ludícibus (2012), afirma sobre a importância do uso de quocientes de análise para identificar o que ocorreu no passado, como também projetar o que poderá ocorrer no futuro.

Dessa forma, entende-se os indicadores como os recursos-informacionais a partir da DRE, ou seja, os mesmos representam o produto informacional da DRE.

Dentre os relevantes indicadores, ora classificados como produtos informacionais da DRE, foram selecionados, como ponto de referência, os indicadores usuais no cálculo do ranking das empresas maiores e melhores, pesquisados anualmente pela Revista Exame, S/A (2014), compilada segundo a análise técnica da equipe Fundação Instituto de Pesquisas Contábeis, Atuariais e Financeiras (FIPECAFI), sendo eles: Vendas Líquidas (R\$); Crescimento das Vendas (%); EBITDA (R\$); Salários e Encargos (R\$); Tributos (R\$); Exportação - % das Vendas (%); Rentabilidade das Vendas (%); Margem das Vendas (%). Adicionalmente devido à relevância, foram incluídos o Ponto de Equilíbrio (R\$) e Alavancagem Operacional.

Os produtos informacionais da DRE, de acordo com a Revista Exame S/A (2014) e Assaf Neto e Lima (2014, p.289) são os seguintes:

As vendas líquidas (VL) em valor refere-se ao valor das vendas brutas (VB) da empresa, deduzidas das devoluções (DEV) e dos abatimentos (AB), bem como dos impostos (IMP) sobre vendas.

$$VL = VB (-) DEV (-) AB (-) IMP$$

Crescimento das Vendas em porcentagem (CV) equivale a evolução da receita líquida de vendas (VL) em reais, descontada a inflação média (I) em (%) apontada pela variação do IPCA – IBGE.

$$CV = VL \text{ ano } 2 - VL \text{ ano } 1 - I (\%)$$

Já o EBITDA em valor equivale ao lucro antes de descontar os juros, os impostos sobre o lucro, a depreciação e a amortização (LAJIDA), correspondendo em essência ao caixa gerado pela operação da empresa.

$$EBITDA = LAJIDA$$

Os Salários e Encargos (SE) é o valor referente à remuneração (R) dos empregados mais os encargos sociais (ES).

$$SE = R + ES$$

Tributos é valor dos impostos (IMP) incidentes sobre as vendas (IPI, ICMS, ISS, PIS e COFINS), os tributos (T) incidentes sobre o lucro (Imposto de Renda Pessoa Jurídica e Contribuição Social), a contribuição social sobre a folha de pagamentos (CSF) e outros sobre atividades ou propriedades específicas (OPE).

$$T = IMP \text{ S/Vendas} + T \text{ s/Lucro} + CSF + OPE$$

A Exportação da Venda (EV) em porcentagem refere-se as parcelas das vendas líquidas realizadas para o exterior (VLE), obtida das demonstrações contábeis publicadas.

$$EV = VLE$$

Rentabilidade das Vendas (RV) em porcentagem é a parcela das vendas líquidas, depois de deduzir os custos e as despesas fixas e variáveis.

$$RV = LL / VL$$

A Margem das Vendas (MV) em porcentagem é diferença entre Receita (R) e soma de Custo (CV) e Despesas Variáveis (DV).

$$MV = R - (CV + DV)$$

O Ponto de Equilíbrio Contábil (PEC) é quando o lucro é igual a zero. Formado a partir das relações dos custos e despesas fixos (CF) e variáveis unitário (CVU) com o preço de vendas (P).

$$PEC = \frac{CF}{(P - CVU)}$$

Ponto de Equilíbrio Econômico (PEE) é aquele no qual se projeta um lucro mínimo (LM) para compensar o investimento realizado.

$$PEE = \frac{CF + LM}{(P - CVU)}$$

Ponto de Equilíbrio Financeiro (PEF): custo e despesas efetivamente gastos, desconsiderando a depreciação (D) contida nos custos e despesas fixos adicionando-se a amortização (A).

$$PEF = \frac{CF - D + A}{(P - CVU)}$$

Ao destacar os indicadores como produto informacional da DRE e elementos integrantes da modelagem que orienta decisões no tocante a planejamentos e análises com ênfase na lucratividade, percebe-se o aspecto integrador aos diversos aspectos do uso da informação contábil, condição que torna contundente pontuar a DRE na perspectiva sistêmica e informacional na construção do conhecimento organizacional.

### 3. O uso da informação como conhecimento organizacional na perspectiva sistêmica e informacional da DRE

O conceito contábil, ao longo dos anos passou por várias fases históricas em seu desenvolvimento, “desde o estado de simples método de escrituração baseado nas partidas dobradas até sua maturação como ciência social aplicada de forte fundo econômico” (Ludícibus, Martins e Carvalho, 2005, p.1). Nesta jornada de desenvolvimento o pesquisador Sá (2006, p.46) destacou um conceito que muito se correlaciona com a teoria de sistema, na perspectiva de amplitude de visão do todo, e não só na visão pontual do patrimônio, quando escreve: “Contabilidade é a ciência que estuda os fenômenos patrimoniais, preocupando-se com realidades, evidências e comportamentos dos mesmos, em relação à eficácia funcional das células sociais”. Assim, se a contabilidade preocupa-se com o comportamento dos fenômenos patrimoniais inseridos nas células sociais, e estas se inserem no todo social, é incisivo inferir que a contabilidade tem uma abordagem sistêmica ao correlacionar este conceito a visão de Capra (1996):

O pensamento sistêmico, ao contrário do pensamento analítico, caracteriza-se pela busca do entendimento da totalidade integrada por meio da conexão das relações e do contexto. Analisar significa isolar alguma coisa a fim de tentar entendê-la; o pensamento sistêmico significa buscar a sua compreensão no contexto de um todo mais amplo. (CAPRA, 1996 página 1).

O pensamento sistêmico é um estudo de Ludwing Von Bertalanffy, que desenvolveu a teoria geral dos sistemas, na qual o destaque é que ao se tratar de sistemas refere-se no todo, e não unicamente à cada parte que integra um todo, conforme Rosini & Palmisano (2012). Bio (2008) pontua que os sistemas são compostos sucessivamente por subsistemas que se relacionam entre si, compondo um sistema maior, com intermináveis processos de adaptação as mudanças nos ambientes, com a finalidade de promover resultados desejados.

Partindo destes pressupostos as organizações podem ser entendidas como um sistema, um

sistema aberto, que conduz recursos materiais, humanos tecnológicos e principalmente recursos informacionais, que resultam em bens e serviços para a sociedade.

Nesta concepção Bio (2008), classifica os sistemas em: Sistema Físico-Operacional, Sistema de Informação, Sistema Social e Sistema de Gestão. Abaixo segue a conceituação do autor para cada um dos sistemas:

Sistema Físico-Operacional (ou sistema logístico e de produção): processos interdependentes de natureza física, que vão desde a obtenção de materiais, peças e componentes nos fornecedores até a entrega do produto final ao cliente (BIO, 2008, p. 21-22).

O Sistema Físico-Operacional abrange todo o processo produtivo da organização, e o Sistema de Informação têm como objetivo auxiliar todos os processos na geração de relatórios e armazenagem dos dados obtidos, assim como no processo de tomada de decisão.

Sistema de Informação: é uma rede de subsistemas, em que cada qual se decompõe em procedimentos que coletam dados, os processam, e produzem e distribuem as informações resultantes. Assim o sistema de informação, “aciona e controla” o andamento dos processos físicos (BIO, 2008, p. 21-22).

Os sistemas citados acima se inter-relacionam de forma a contribuir com o Sistema Social que fará parte da cultura organizacional, que por sua vez influenciará no Sistema de Gestão para a obtenção das metas e objetivos da empresa.

Sistema Social (ou sociopolítico), no qual os indivíduos interagem em grupos, grupos interagem entre si, afetam e são afetados pela cultura da organização e pelos processos políticos. Sistema de Gestão: conjunto, interdependente, dos processos decisórios – gerenciais (planejar, organizar, controlar, etc.) que visa levar a empresa aos resultados desejados (BIO, 2008, p. 21-22).

No tocante ao sistema de informação destaca-se a visão do pesquisador Choo (2003, p. 65), ao retratar que “a busca da informação são fundamentais em muitos sistemas sociais”. Já Moody e Walsh (1999) apud Beal (2004, p 23) discute que a replicação da informação não agrega valor, “apenas custos adicionais”. Interligando estas visões com as informações geradas pela DRE, é possível inferir que esta não é replicada em nenhuma outra demonstração contábil, garantindo neste cenário seu propósito e relevância enquanto informação do resultado líquido do exercício.

A informação tornou-se um dos requisitos essenciais para a gestão empresarial, e necessita de planejamento para ser utilizada como ferra-

menta no processo organizacional. Esta informação necessita da realização do ciclo de vida de cada informação, pois esta pode se alterar com base nas modificações constantes no processo organizacional.

*Ponjuán Dante* (2007, p. 31), retrata a relevância no contexto organizacional do ciclo de vida da informação:

Domínio do ciclo de vida de cada informação, sendo relevante entender quando precisa ser gerada, seu prazo de retenção para cada nível da organização, quando perde seu valor e pode ser descartada ou se deve ser conservada por possuir valor permanente (PONJUÁN DANTE, 2007, p. 31).

Para reafirmar ainda mais a abordagem sistêmica da contabilidade, mediante a DRE, pontua-se segundo *Werkema* (1996, p.1) que as organizações demandam ferramentas para solução de problemas inerentes aos processos organizacionais, o que leva ao encontro do Ciclo PDCA, “um método gerencial de tomada de decisões para garantir o alcance das metas necessárias à sobrevivência de uma organização”. Sobrevida essa que está diretamente relacionada ao uso da informação no momento propício (ciclo de vida).

Mediante a visão de *Campos* (2004), o ciclo PDCA (*Plan, Do, Check, Action*) é composto por atividades específicas para cada etapa do ciclo, assim como *Werkema* (1996) também o conceituou por etapas. Correlacionando este referencial ao ciclo e vida da informação (*Lyra*, 2008) com a DRE, projetou-se o Quadro 2, conforme consta no Anexo 1.

**Anexo 1 (Quadro 2 - Produtos Informacionais da DRE)**

Pensando em projetar cenários para atingir metas e ou objetivos com foco na lucratividade, surgem movimentos de mudanças e reassignificação. Mediante estes movimentos, e a relevância do afloramento destas questões, as organizações deverão aprender e reaprender novas formas de traçar planos e objetivos para manter-se em um mercado de economia globalizada que muda com frequência.

Como parte destas mudanças, as organizações requerem profissionais que desenvolvam as competências necessárias aos seus fazeres funcionais, mas que também sejam competentes em informação, ou seja, que saibam lidar com as tecnologias da informação e com o ciclo informacional para gerar e compartilhar conhecimento, promover aprendizado para alcance do objetivo da organização. (SMITH, 2013, p.111).

*Senge* (1999) já pontuava a importante participação das pessoas nas organizações que

aprendem, inferindo que as organizações que aprendem são aquelas que:

[...] nas quais as pessoas expandem continuamente sua capacidade de criar os resultados que realmente desejam, onde se estimulam padrões de pensamento novos e abrangentes, a aspiração coletiva ganha liberdade e onde as pessoas aprendem continuamente a aprender juntas.

Assim a criação da nova informação, resultante da competência informacional, ou seja, competência em perceber a necessidade informacional e sua extensão, saber buscar e avaliar criticamente as fontes, bem como usar a informação entendendo o aspecto econômico, legal e social, e a partir dela gerar nova informação “(ACRL, 2000, tradução nossa) em conjunto com ALA publicou um documento titulado “*Literacy Competency Standards for Higher Education*”, é o que sensivelmente contribui com a construção do conhecimento organizacional.

Para o desenvolvimento da competência informacional como alavanca ao aprendizado organizacional apoia-se nas cinco disciplinas de *Senge* (1999), que retrata a organização que aprende, sendo estas: Domínio pessoal (possibilita expandir a visão pessoal); Modelos mentais (relacionado as modificações administrativas); Visão compartilhada (compartilhamento mútuo entre os colaboradores); Aprendizagem em equipe (utiliza-se do diálogo); Pensamento sistêmico (contempla a visão do todo, vinculando teoria, prática e as demais disciplinas).

#### **4. Modelagem decisória a partir da DRE no contexto da informação contábil**

A partir da discussão conceitual, foi possível a construção de uma modelagem a partir da DRE capaz de orientar as decisões no tocante a planejamento e análise com ênfase na lucratividade, conforme proposta deste estudo.

Como metodologia, pontuou-se a DRE na perspectiva informacional e sistêmica, como produto informacional, tomou-se com base os indicadores para avaliação das Melhores e Maiores divulgados na Revista Exame S/A (2014). Destacou-se o uso da informação na construção do conhecimento organizacional a partir da estratégia bibliográfica, que nos conduziu a inferir de forma descritiva, a modelagem sistêmica da DRE.

*Bio* (2008, p. 29) esclarece que “um sistema de informação é um conjunto de partes interdependentes no seu todo por ser parte de um todo maior”. A interligação entre os diversos sistemas resulta no processamento dos recursos de entradas e geração de lucro, sendo seu uso de-

pendente da competência informacional, que fortalece e influencia o conhecimento organizacional.

A abordagem da pesquisa descritiva e bibliográfica levou em consideração as entradas relacionadas às Receitas, Impostos, Custos, Despesas (INPUTS) que processadas pela contabilidade através do sistema DRE produzirão informações fidedignas sobre a lucratividade gerada (OUTPUTS), suscitando decisões estratégicas, táticas e operacionais, realimentando o próprio sistema, sendo influenciado e influenciando o ambiente organizacional oportunizando o conhecimento organizacional.

**Anexo 2 (Figura 1: Modelagem decisória a partir da DRE no contexto da informação contábil (MODRE))**

A Figura 1, Anexo 2, representa a MODRE, proposta que possibilita uma análise sob a perspectiva sistêmica ao apontar que a partir das entradas dos diversos recursos, tais como receitas, impostos, custos e despesas; é possível acionar o sistema físico-operacional para processamento dos dados e alimentação dos diversos sistemas. Com destaque para a DRE como sistema de informação e do sistema de gestão para imprimir as prioridades e estabelecer metas, resultar a lucratividade e o produto informacional, ora destacados como indicadores, representando as saídas que irão subsidiar as decisões.

A DRE, informação contábil cerne desta modelagem, pôde através do seu viés informacional estabelecer a inter-relação entre os sistemas, influenciando os mutuamente, definindo as características do sistema social, que por sua essência cultural e de conhecimento dialoga com o sistema físico-operacional, sistema de informação e de gestão, tornando os congruentes aos objetivos organizacionais.

A congruência da modelagem proposta deve subsidiar as decisões estratégicas, táticas e operacionais. O Quadro 3, constante no Anexo 3, retrata as decisões inerentes aos níveis organizacionais a partir do produtos informacional gerado pela perspectiva sistêmica da DRE.

**Anexo 3 (Quadro 3 Decisões inerentes aos níveis organizacionais a partir da modelagem sistêmica da DRE.)**

Assim, ao observar esta totalidade, integrada por meio das relações entre receitas, custos e despesas, é possível chegar ao lucro e seus efeitos, no resultado de uma organização através de um sistema informacional que apoia decisões e ações estratégicas, táticas e operacionais suscitando a modelagem com ênfase na

lucratividade, promovendo a eficiência econômica.

As atividades do ciclo PDCA, em ação apresentam propostas para melhoria contínua dos resultados da organização, suscitando aprimoramento e aprendizados, redirecionando novas decisões, através da retroalimentação inerentes aos sistemas, promovendo a gestão do conhecimento organizacional.

## 5. Considerações finais

A DRE demonstra de forma sistêmica as informações contábeis e o resultado obtido pela organização. Por ser uma base de informações ela orienta o planejamento e as decisões e, a partir dessa discussão foi possível construir a modelagem capaz de orientar os gestores para o uso da informação na construção do conhecimento organizacional e participação nas decisões estratégicas, táticas e operacionais no contexto da informação contábil.

A modelagem permite o processamento de dados e a alimentação dos diversos sistemas o que possibilita imprimir as prioridades e metas a serem alcançadas pela organização. A modelagem retrata os níveis organizacionais a partir do produto informacional gerado pela perspectiva sistêmica da DRE.

A presente pesquisa justifica-se na medida em que a articulação dos temas pertinentes às questões da pesquisa proporciona um construto teórico para os estudos da Ciência da Informação no âmbito da informação contábil. E, a partir da articulação interdisciplinar a contribuição deve se estender à gestão da informação econômica financeira na construção do conhecimento organizacional por meio da divulgação da relevância de interações e conhecimentos para tomada de decisões a partir da DRE.

Este estudo não tem a intenção de esgotar o tema, e sim contribuir com os estudos interdisciplinares da Ciência da Informação e da Ciência Contábil, sugerindo a continuidade, para geração de novas informações e novos conhecimentos em outros aspectos da gestão da informação contábil.

## Referências

- ASSAF NETO, Alexandre; LIMA, Fabiano GUSTI (2014). *Curso de administração financeira*. São Paulo: Atlas.
- ASSOCIATION OF COLLEGE AND RESEARCH LIBRARIES – ACRL. (2000) *Information literacy competency for higher education*. Chicago: ALA. Disponível em: <<http://www.ala.org/ala/mgrps/divs/acrl/standards/informationliteracycompetencystandards.cfm>>. Acesso em: 05 abr. 2015.

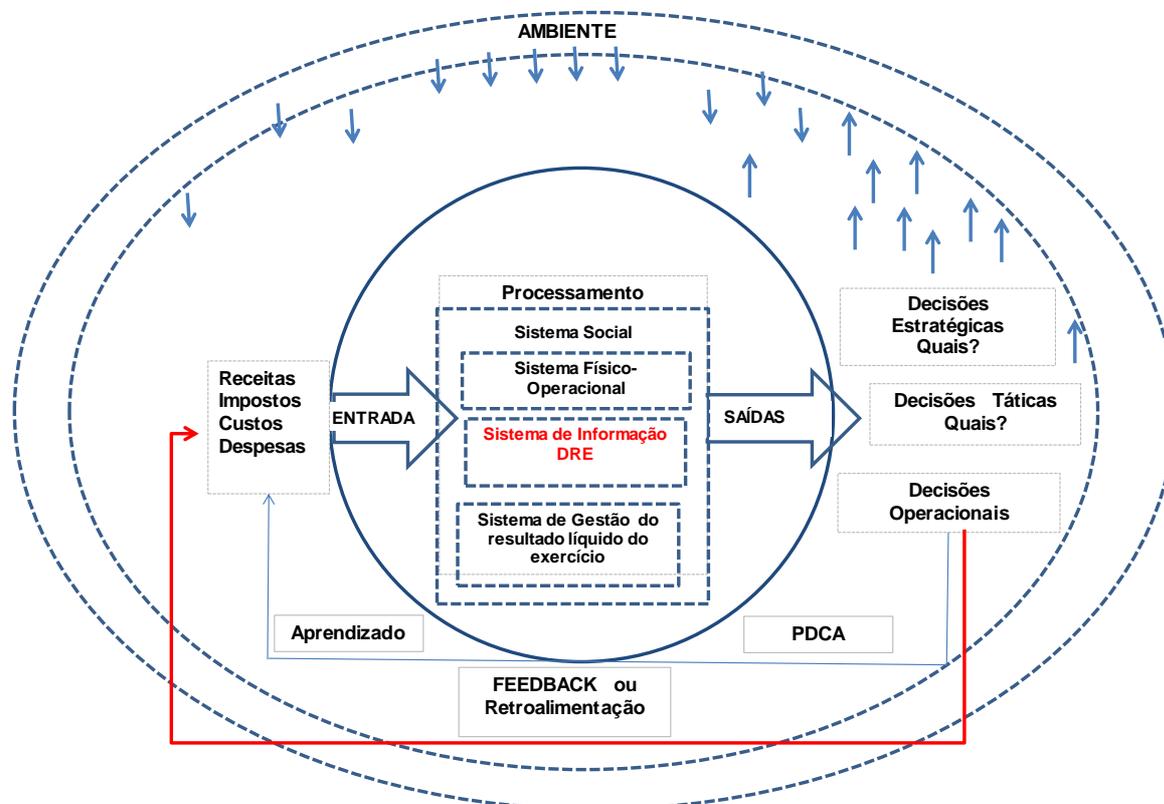
- BEAL, A. (2004). *Gestão estratégica da informação: como transformar a informação e a tecnologia da informação em fatores de crescimento e de alto desempenho nas organizações*. São Paulo: Atlas.
- BIO, S. R. (2008). *Sistemas de informação: um enfoque gerencial*. São Paulo: Atlas.
- CAMPOS, Vicente Falconi (2004). *Qualidade Total: padronização de empresas*. Nova Lima: INDG Tecnologia e Serviços Ltda.
- CAPRA, F. (1996). *A teia da vida: uma nova compreensão científica dos sistemas vivos*. (Newton Roberval Eichenberg, Trad.) São Paulo: Cultrix.
- CHOO, Chun Wei (2003). *A Organização do Conhecimento: como as organizações usam a informação para criar significado, construir conhecimento e tomar decisões*. São Paulo: SENAC.
- COMITÊ DE PRONUNCIAMENTOS CONTÁBEIS (2011). *Pronunciamento conceitual básico (R1): estrutura conceitual para elaboração e divulgação de relatório contábil-financeiro*. CPC\_00\_R1. Disponível em: <[http://static.cpc.mediatgroup.com.br/Documentos/312\\_CPC\\_26\\_R1\\_rev%2006.pdf](http://static.cpc.mediatgroup.com.br/Documentos/312_CPC_26_R1_rev%2006.pdf)>. Acesso em: 11 mar. 2015.
- IUDÍCIBUS, S. (2012). *Análise de Balanços*. S. Paulo: Atlas, 2012.
- IUDÍCIBUS, S. de; et al. (2010). *Manual de contabilidade societária: aplicável a todas as sociedades, de acordo com as normas internacionais e do CPC*. São Paulo: Atlas.
- IUDÍCIBUS, Sérgio de; MARTINS, Eliseu; CARVALHO, L. Nelson (2005). *Contabilidade: aspectos relevantes da epopéia de sua evolução*, São Paulo, v. 16, n. 38, p. 7-19, May/Aug. Disponível em: [http://www.scielo.br/scielo.php?pid=S1519-70772005000200002&script=sci\\_arttext](http://www.scielo.br/scielo.php?pid=S1519-70772005000200002&script=sci_arttext). Acesso em 15 mai. 2015.
- IUDÍCIBUS, Sérgio de; MARION, José Carlos (2002). *Introdução à teoria da contabilidade para o nível de graduação*. 3. ed. São Paulo: Atlas.
- LYRA, Maurício Rocha (2008). *Segurança e Auditoria em Sistema de Informação*. Rio de Janeiro: Ciência Moderna.
- PONJUÁN DANTE, G (2007). *Gestión de información: dimensiones e implementación para el éxito organizacional*. Gijón: Trea.
- ROSINI, A. M & PALMISANO, A (2012). *Administração de sistemas de informação e a gestão do conhecimento*. São Paulo: Cengage Learning.
- SÁ, A. L. de (2006). *Teoria da contabilidade*. 4. ed. São Paulo: Atlas.
- SENGE, Peter M. (1999). *A quinta disciplina: Arte e prática da organização que aprende*. São Paulo.
- SMITH, M. S. J. (2013). *A cultura informacional como alicerce de gestão contábil nas pequenas e médias empresas*. 255 f. (Tese apresentada ao Programa de Pós-Graduação em Ciência da Informação). Faculdade de Filosofia e Ciências – UNESP de Marília - SP.
- WERKEMA, Maria Cristina Catarino (1996). *Avaliação da qualidade de medidas*. Belo Horizonte: Fundação Christiano Ottoni, Escola de Engenharia da UFMG.
- FUNDAÇÃO INSTITUTO DE PESQUISAS CONTÁBEIS, ATUARIAIS E FINANCEIRAS (2014) <<http://www.fipecafi.org/downloads/projeto-exame-melhores-maiores/mmm-criterios-indicadores.pdf>> Acessado em 04/06/2015.
- REVISTA EXAME S.A. <<http://exame.abril.com.br/negocios/melhores-e-aiiores/empresas/ficha/toyota/> 2013> Acessado em 04/06/2015.

## Anexo 1

| PDCA       | Atividades Específicas   | Conceito   | Ciclo de Vida   | Exemplo (DRE)  |
|------------|--|--|---|--|
| P (Plan)   | Ação de preparo para atingir o objetivo.                           | Identificação do problema;<br>Reconhecimento das características do problema;<br>Descoberta das causas principais;<br>Plano de Ação. | <b>Identificação das necessidades de dos requisitos:</b> conhecer suas necessidades de informação, bem como suas características.                 | Ações necessárias para planejar a eficiência econômica com base nas Receitas, Custos Despesas, e Resultados.               |
| D (Do)     | Execução das ações necessárias para que o objetivo seja alcançado. | Atuação de acordo com o Plano de Ação.   | <b>Obtenção</b> (recepção da informação); <b>Tratamento</b> (formatação e classificação); e <b>Armazenamento</b> (integridade e disponibilidade). | Registros Contábeis; DRE conforme estrutura e Comparativo pelo histórico.  |
| C (Check)  | Ações de verificação se o objetivo está sendo ou foi alcançado.    | Confirmação da efetividade da ação.  | <b>Distribuição</b> (informação aos usuários para a tomada de decisão).   | Análise comparativa por meio de quocientes entre planejado (orçado) e realizado.   |
| A (Action) | Ações de correção da execução (D) ou correção do padrão (P).       | Revisão das atividades e planejamento para trabalho futuro.  | <b>Uso</b> (gerar valor para a organização); ou <b>Descarte</b> (perde a utilidade para a organização).   | Replanejamento em função dos cenários prospectivos para projeção de novas metas e ou objetivos, com foco na lucratividade. |

**Quadro 2:** PDCA e o ciclo de vida informacional na DRE. **Fonte:** Elaborado pelas autoras com base em Campos (2004), Werkema (1996) e Lyra (2008).

## Anexo 2



**Figura 1** Modelagem decisória a partir da DRE no contexto da informação contábil (MODRE). **Fonte:** Elaborado pelas autoras (2014) adaptado de BIO (2008, pg. 23 e 183)

## Anexo 3

| Níveis organizacionais | Decisões a partir da modelagem sistêmica da DRE   |
|------------------------|---|
| Estratégico            | Vendas Líquidas (R\$);<br>Crescimento das Vendas (%);<br>EBITDA (R\$);<br>Salários e Encargos (R\$);<br>Tributos (R\$);<br>Exportação - % das Vendas (%);<br>Rentabilidade das Vendas (%);<br>Margem das Vendas (%)<br>Ponto de Equilíbrio Contábil (R\$)<br>Ponto de Equilíbrio Econômico (R\$)<br>Ponto de Equilíbrio Financeiro (R\$) <div data-bbox="1090 387 1350 707" style="border: 1px solid black; padding: 5px; margin-left: 200px;"> <p><b>DECISÃO:</b><br/>O que e quanto produzir?<br/>(Planejamento dos objetivos)</p> </div> |
| Tático                 | Custos e despesas fixas;<br>Custos e despesas variáveis;<br>Carga tributária; <div data-bbox="890 853 1398 943" style="border: 1px solid black; padding: 5px; margin-left: 100px;"> <p><b>AÇÕES:</b> Definição e execução das Metas.<br/>(Controle)</p> </div>  |
| Operacional            | Registros;<br>Monitoramentos;<br>Auditorias;<br>Guarda de arquivos e documentos <div data-bbox="1066 983 1340 1144" style="border: 1px solid black; padding: 5px; margin-left: 200px;"> <p><b>TAREFAS:</b><br/>Garantir o planejamento.<br/>(Execução)</p> </div>   |

**Quadro 3** Decisões inerentes aos níveis organizacionais a partir da modelagem sistêmica da DRE. **Fonte:** Elaborado pelas autoras.

---

# Aplicação de reengenharia de tesouro: modelagem do THESAGRO

*Reengineering thesaurus application: modeling THESAGRO*

---

**Benildes C. M. S. Maculan (1), Gercina A. B. O. Lima (2),  
Ivo Pierozzi Jr. (3), Leandro H. M. Oliveira (4)**

(1) Escola de Ciência da Informação, Universidade Federal de Minas Gerais (ECI/UFMG), Av. Presidente Antônio Carlos, número 6627, Gabinete 4030, Pampulha, Belo Horizonte, MG, Brasil, CEP 31270-901, benildes@gmail.com (2) limagercina@gmail.com

(3) Embrapa Informática Agropecuária, Av. André Tosello, 209, Cidade Universitária, Campinas, SP, CEP 13083-886, ivo.pierozzi@embrapa.br. (4) leandro.oliveira@embrapa.br

## Resumo

Este estudo investigou e aplicou um modelo de reengenharia de tesouros tradicionais para tornar o sistema de relações entre os conceitos em uma estrutura semântica rica. O modelo selecionado para a aplicação foi desenvolvido por Soergel *et al.* (2004) e Lauser *et al.* (2006). Esse modelo é composto por três etapas e envolve a melhoria e explicitação formal das relações semânticas em um tesouro. A reengenharia foi aplicada no tesouro brasileiro THESAGRO, do domínio da Agricultura, no recorte temático da Intensificação Agropecuária. A metodologia incluiu o uso das ferramentas: (a) Sistema e-Termos, para o gerenciamento da terminologia; e (b) Extrator de Termos, *software* que permite a comparação automática da terminologia de diferentes tesouros. Os resultados demonstraram a viabilidade da utilização do modelo aplicado na reengenharia de tesouros tradicionais, uma vez que permitiu modelar e obter uma estrutura mais semanticamente enriquecida. Concluiu-se que o refinamento das relações ajudou a organização do conhecimento da subárea temática modelada, o que pode facilitar a sua exploração pelo usuário final, assim como ser um importante elemento para a promoção da interoperabilidade entre diferentes tesouros.

**Palavras-chave:** Reengenharia tesouros. Sistema de organização do conhecimento. Tesouros. Modelagem conceitual. Modelo de conversão de tesouros.

## 1. Introdução

No campo da Biblioteconomia e Ciência da Informação (BCI), os distintos tipos de instrumentos de representação, tais como taxonomias, listas de cabeçalho de assunto, tesouros, redes semânticas e ontologias, têm sido agrupados sob a designação de Sistemas de Organização do Conhecimento (SOC). Esse termo foi cunhado em 1998, durante a primeira Conferência da ACM *Digital Libraries*, em Pittsburgh (Pennsylvania), quando o *Networked Knowledge Organization Systems Working Group* propôs o uso do

## Abstract

Research that aimed to study and apply a model of reengineering thesauri making it a conceptual structure semantically enriched. The method included the selection of the reengineering model developed by Soergel *et al.* (2004) and Lauser *et al.* (2006). This model consists of three stages and consists of the improvement and formal explanation of the semantic relationships in the conceptual structure of a traditional thesaurus. The reengineering was applied to the Brazilian thesaurus Thesagro from the Agriculture domain more specifically it was applied to the thematic frame of the agriculture intensification. In methodological course some tools were used: (a) e-Terms system: for the terminology management; and (b) Terms Extractor, a software that allows automatic comparison of the terminology of different thesauri. The results showed the viability of the use of the analyzed model in the reengineering of tradition thesauri because it allowed the imprinting of more semantics to the structure of the modeled thesaurus specifying the kind of relationship existing between concepts and terms. It was concluded that the possibility of refining the relations between concepts helped in the organization of the modeled domain and it can facilitate the exploration by the final users.

**Keywords:** Reengineering of thesaurus. Knowledge Organization System. Thesaurus. Conceptual Modeling. Thesaurus model conversion.

termo “*Knowledge Organization System*” (KOS). Para Vickery (2007, *on-line*) os SOCs “são vistos como esquemas que visam organizar, gerenciar e recuperar informações”, para aplicações em ambientes digitais. Carlan (2010) afirma que os SOCs representam uma “denominação nova para as linguagens documentárias que agregam elementos incorporados nas inovações tecnológicas da era digital” (CARLAN, 2010, p. 29-30).

Hodge (2000) acrescenta que os SOCs são utilizados para “organizar conteúdos para apoiar

a recuperação de itens relevantes, disponibilizados na base de dados de uma biblioteca digital” (HODGE, 2000, p. 9). Sendo assim, os SOCs são instrumentos, já tradicionais na área da Biblioteconomia, que podem ser utilizados para a representação e na recuperação de informações junto a aplicações tecnológicas em ambiente digital.

Segundo Soergel (1999), um SOC tem a função de ser um dicionário mono, bi ou multilíngue, para uso humano ou como base de conhecimento em uma aplicação em ambiente digital, para ser compreendido pela máquina. O autor afirma que os SOCs têm como objetivos:

- (1) mapear domínios individuais, sendo um mapa semântico capaz de indicar os relacionamentos entre conceitos no domínio mapeado e servindo como uma ferramenta de referência;
- (2) dar apoio a professores e alunos ao criar estruturas conceituais na elaboração de materiais didáticos, aprimorando a comunicação do conhecimento de um dado domínio e, assim, o seu aprendizado;
- (3) apoiar a implantação de projetos de pesquisa ou de atividades profissionais ao criar uma base de conhecimento de auxílio à criação de um contexto conceitual de estudo;
- (4) proporcionar classificações para diferentes finalidades, tais como classificação de doenças e de competências para atribuição de tarefas;
- (5) oferecer uma base de conhecimento para a construção de mecanismos de buscas de apoio à recuperação de informação;
- (6) auxiliar o desenvolvimento de *software* ao fornecer uma base conceitual para a definição de elementos de dados e de hierarquias de objetos.

Os diversos tipos de SOCs possuem distintos níveis de controle terminológico de um determinado domínio (campo do saber ou assunto, atividade corrente ou tarefa). Eles também são construídos com diferentes abordagens de modelagem, sobretudo no que diz respeito à indicação de relacionamentos entre os conceitos de sua estrutura.

Segundo Hodge (2000), os diversos tipos de SOCs podem ser sistematizados da seguinte maneira: (1) grupo de instrumentos compostos por listas de termos: arquivo de autoridade, glossários, gazetteers, e dicionários; (2) grupo de instrumentos compostos por classificações e categorizações: lista de cabeçalhos de assunto, sistemas de classificação bibliográfica, taxonomias e sistemas de classificação bibliográfica

facetados; (3) grupo de instrumentos compostos por listas de termos e relacionamentos: tesouros, redes semânticas e ontologias.

Entre os diferentes tipos de SOCs, os tesouros são linguagens de indexação, construídos a partir de um conjunto de regras pré-estabelecidas e constituídos por descritores preferidos e não-preferidos, que representam conceitos que podem ser combinados no momento de seu uso (pós-coordenação) e usualmente são restritos a uma única especialidade.

Em geral, os tesouros são apresentados na forma alfabética e na forma sistemática, que oferece elementos de significação, permitindo ao usuário a apreensão do conhecimento de um domínio por meio das relações estabelecidas entre conceitos. Tradicionalmente, eles têm dois planos de trabalho: o plano das ideias e o plano verbal (TRISTÃO; FACHIN; ALARCON, 2004). Alguns tesouros também podem oferecer o plano notacional (tal como o existente nos sistemas de classificações), o que possibilita a localização de recursos informacionais.

Os projetos de construção de tesouros possuem três etapas básicas: (1) inicial: composição de uma equipe de trabalho; planejamento; resolução sobre os objetivos; definição do público-alvo; levantamento de terminologia; (2) desenvolvimento: concepção da estrutura conceitual; compilação e seleção do conjunto de termos; definições dos conceitos; determinação dos descritores (preferidos e não-preferidos); agrupamentos em classes básicas e facetadas; elaboração dos de mapa conceitual; atribuição de relações entre conceitos; (3) edição: construção da estrutura conceitual; seleção de *software* para edição do tesouro; elaboração de notas de escopo; determinação da forma de apresentação.

Assim, ao construir um tesouro, cria-se um sistema de conceitos que é composto por um conjunto de conceitos relacionados semântica e genericamente entre si, permitindo diferentes tipos de organização (por exemplo: alfabética, relacional, estruturada por campos semânticos, entre outros). Nesse sentido, Currás (1995) afirma que em ambiente organizacional os tesouros desempenham as funções de representação de assuntos e de apoio às consultas de busca dos usuários, auxiliando o processo de recuperação da informação.

Tradicionalmente, os tesouros possuem uma estrutura semântica constituída por uma rede de três distintos tipos de relações: (1) de equivalência, com controle de termos em sinonímia e controle de variações linguísticas; (2) hierárquicas, com agrupamentos constituídos por

conceitos ordenados em níveis diferentes de generalidade e de especificidade; e (3) associativas, com a atribuição de ligações não-hierárquicas entre conceitos.

Essa estrutura conceitual dos tesouros vem evoluindo (MOTTA, 1987; CAMPOS, 1995; CAMPOS; GOMES, 2003) e já há evidências de diferentes desdobramentos para cada um desses tipos de relações. Como, por exemplo, no caso da relação de equivalência, ela pode variar desde uma equivalência ortográfica, total ou parcial, até ao uso de abreviaturas, nome fantasia ou equivalência em outro idioma. Essa situação se repete, também, para os relacionamentos hierárquicos e associativos.

Estruturalmente, um tesouro é também composto por: (1) uma terminologia (descritores preferidos e não-preferidos); (2) uma estrutura gramatical (forma de apresentação e de composição dos descritores); (3) uma rede paradigmática (*a priori*); (4) uma rede sintagmática (*a posteriori*). O conjunto desses elementos tem como consequência a não existência de qualquer descritor em um tesouro sem que esteja diretamente relacionado à significação de outro descritor na sua estrutura (Svenonius, 2000). Entretanto, não se pode deixar de considerar que a construção de tesouros deve ser sempre dependente do propósito de sua elaboração, de seu uso e do domínio modelado.

A estrutura conceitual do tesouro é modelada já visando a minimizar a ambiguidade (imprecisão do significado) e a polissemia (pluralidade de significados) da linguagem natural. Apesar de a estrutura conceitual do tesouro possuir essa semântica bastante rica, a falta de especificação dos distintos tipos de relacionamentos existentes entre os conceitos e termos ainda pode ser considerada uma limitação para o uso do tesouro em ambiente digital e para que ele seja classificado como um tipo de SOC.

Nesse contexto, o objetivo deste estudo foi aplicar um modelo de reengenharia de tesouro tradicional para torná-lo um instrumento mais formalizado, de tal forma que os relacionamentos semânticos entre os conceitos do sistema estejam identificados e explicitados para o usuário.

## 2. Fundamentos teóricos

As bases teóricas utilizadas neste estudo deram subsídios para o seu desenvolvimento e para a aplicação do modelo de reengenharia escolhido. Foram feitas descrições e reflexões sobre os fundamentos empregados na construção de tesouros, a partir da literatura da BCI, Terminologia, Pragmática e Semântica.

Da BCI foram utilizados os princípios teóricos para a organização e sistematização dos conceitos de um domínio, sobretudo com as bases da Teoria da Classificação Facetada, de Ranganathan (1967) e a Teoria do Conceito, de Dahlberg (1978). Na primeira teoria, Ranganathan apresenta o conceito como uma unidade do pensamento,

[...] um corpo de ideias organizado ou sistematizado, cujas extensão e intensão devem ser coerentes com o domínio de interesse e ajustadas à competência intelectual e ao campo especializado de qualquer indivíduo (RANGANATHAN, 1967, p. 82).

Para Ranganathan, o conceito é concebido no plano das ideias, por meio de distintos procedimentos, a saber:

(1) processo de definição do assunto; (2) seleção das características que constituem o assunto; (3) seleção de um modelo para o mapeamento da informação sobre os conceitos; (4) agrupamento e divisão destes conceitos conforme suas características comuns e diferentes; (5) organização e o arranjo de grupos e subgrupos (LIMA, 2007, p. 32).

O conjunto de procedimentos se refere ao processo de análise conceitual e, ao final, há a reprodução mental do objeto e a sua representação na forma de um termo (RANGANATHAN, 1967). Kobashi e Francelin (2011) alegam que a

[...] lógica subjacente à sua teoria indica que o conceito se estabelece em uma sequência de etapas, recortes, enfim, 'fatias' que determinam o movimento contínuo e infinito de sua Espiral do Universo do Conhecimento (KOBASHI; FRANCELIN, 2011, p. 10).

Para as autoras, as relações entre termos e conceitos são estabelecidas nessa espiral, cuja estrutura conceitual criada representa percepções individuais sobre o mundo real. Sobre Ranganathan, Campos e Gomes (2003) afirmam que

Ranganathan elabora uma série de princípios que visam a permitir que os conceitos de um domínio de saber possam ser estruturados de forma sistêmica, isto é, os conceitos se organizam em renques e cadeias, essas estruturadas em classes abrangentes, que são as facetas, e estas últimas dentro de uma dada categoria fundamental. A reunião de todas as categorias forma um sistema de conceitos de uma dada área de assunto e cada conceito no interior da categoria é também a manifestação dessa categoria (CAMPOS; GOMES, 2003, p. 158).

Com esses princípios, Ranganathan (1957) aponta que na espiral há diversos tipos de relacionamentos entre conceitos e termos, o que torna possível também representar realidades complexas (sentidos multidirecional e multidirecional).

mensional). Com isso, o conhecimento de um domínio é representado como um organismo vivo (metáfora da árvore Baniana), que é dinâmico, pois está em constante desenvolvimento e suas unidades de conhecimento se inter-relacionam entre si.

Sobre a Teoria do Conceito, Dahlberg (1992) afirma que ela compõe o campo da Organização do Conhecimento, juntamente com os campos filosóficos da Lógica, Teoria da Ciência, Epistemologia, Ontologia, Fenomenologia, Aleitologia e Metafísica.

Dahlberg (1978a) desenvolveu a sua teoria tendo por base os princípios analíticos de Aristóteles e os princípios analítico-sintéticos da teoria de Ranganathan. Os princípios analíticos auxiliam a faturação do objeto representado em suas partes constituintes (elementos individuais), determinando uma hierarquia de proposições verdadeiras sobre o objeto, dando origem a uma classe mais genérica. Já os princípios analítico-sintéticos permitem a integração desses elementos sistêmicos (proposições) que culminam na identificação do termo que representa o conceito. Assim, na Teoria do Conceito, o

conceito é uma unidade do conhecimento, compreendendo afirmações verdadeiras sobre um dado item de referência, representado numa forma verbal [sendo que:] afirmação verdadeira é a componente de um conceito que expressa um atributo do seu item de referência; item de referência é o componente de um conceito para o qual sua afirmação verdadeira e sua forma verbal estão diretamente relacionadas, sendo assim seu 'referente'; forma verbal (termo/nome) de um conceito é o componente que resume convenientemente ou sintetiza e representa um conceito com o propósito de designar um conceito em comunicação (DAHLBERG, 1978b, p. 147, grifos da autora citada).

Dessa forma, nota-se que para Dahlberg (1978) o conceito é formado pela tríade: (1) referente (objeto a ser conceitualizado); (2) características (todos os enunciados verdadeiros a respeito do referente); (3) forma verbal (termo), conforme Figura 1.

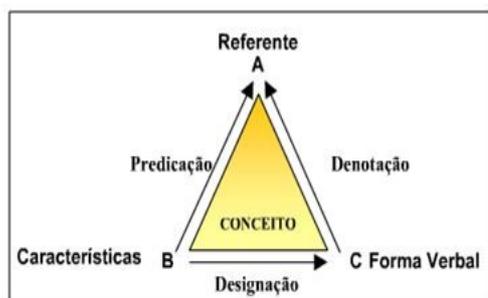


Figura 1. *Triângulo do Conceito* (Dahlberg, 1978b, p. 149).

Assim, o objeto (referente) é fatorado por predicações (características) que são as afirmações verdadeiras sobre esse objeto, cujo conjunto será designado por uma forma verbal (ou outro símbolo). A teoria desenvolvida por Dahlberg possui forte natureza analítica e lógico-positivista, pois utiliza uma abordagem na qual se determina a univocidade do significado de um termo.

Essa natureza lógico-positivista da Teoria do Conceito tem sua origem no campo da Terminologia, advinda dos princípios normativos desenvolvidos por Wüster na Teoria Geral da Terminologia (TGT). Em seus estudos, Eugen Wüster, fundador da Escola Terminológica de Viena, criou métodos para a compilação, padronização e organização da terminologia da área da eletrotécnica, que tinha como objetivo permitir a comunicação entre os profissionais dessa área (KRIEGER; FINATTO, 2004). A TGT pode ser sintetizada pelas características:

- a) a prioridade do conceito em detrimento do termo;
- b) a precisão do conceito, o que retoma, de certo modo, a eliminação da ambiguidade e a busca da univocidade;
- c) a consequente abordagem onomasiológica, já que toda a atividade terminológica parte do conceito;
- d) a proeminência do nível lexical em detrimento dos demais níveis de descrição linguística (morfológico, sintático, textual, discursivo);
- e) e finalmente; e) a prescrição (ALMEIDA, 2006, p. 86).

Nessa perspectiva, esse conjunto de características torna o produto terminológico bastante rígido, o que pode ser elucidado porque a TGT

[...] tem como foco principal o componente conceitual, em detrimento do significado [...] os conceitos, nesta perspectiva, são estáveis, paradigmáticos e universais, como idealiza o lógico-positivismo (KAMIKAWACHI, 2010, p. 21).

Cabré (1999) aponta que esse caráter prescritivo e normativo pode ser válido em alguns contextos de controle terminológico. Porém, a partir de 1990 houve a expansão das reflexões sobre a área da Terminologia e críticas às bases teóricas da TGT. Com isso, novas abordagens surgiram e, entre elas, a Teoria Comunicativa da Terminologia (TCT), desenvolvida por Cabré.

As bases da TCT estão fundamentadas em três teorias: (1) do conhecimento, de natureza cognitiva, para o entendimento sobre os conceitos, suas designações e as inter-relações estabelecidas entre eles; (2) da comunicação, que estabelecem situações comunicativas, assim como as suas características, perspectiva, propósito e limitações; e (3) da linguagem, que permite compreender as unidades léxicas, na linguagem comum e de especialidade, e os seus contextos de uso (Cabré, 1999).

Com a combinação dos princípios dessas três teorias é possível perceber as unidades terminológicas a partir de seus aspectos linguísticos e comunicativos, que têm comportamento semelhante às palavras do léxico de uma língua geral. Para Cabré (1999), a linguagem de especialidade é regida pelas mesmas regras e é caracterizada pelos mesmos fenômenos de sinonímia e variação linguística, presentes na linguagem geral.

Portanto, o que difere um termo de uma palavra é, principalmente, porque os termos que compõem uma linguagem especializada possuem características específicas que os tornam termos advindos de uma temática única, utilizados por um grupo específico de interlocutores e em um contexto de comunicação. Assim, “um termo é uma unidade linguística que tem uma função comunicativa e pragmática” (Maculan, 2015, p. 74).

As teorias da Pragmática proporcionaram as bases para a exploração de relações mais contextualizadas ao ambiente de uso. Segundo os princípios da Pragmática, não existe uma verdade absoluta, que possa ser generalizada a toda situação de uso. As verdades são um construto social de sentido e estão sujeitas às mudanças no tempo e no espaço. Assim, a verdade é estabelecida para

[...] dar conta especificamente da consideração da linguagem como ação, como produzindo efeitos e consequências em contextos determinados (MARCONDES, 2000, p. 41).

Nesse sentido, cada verdade é validada a partir dos efeitos práticos que é capaz de produzir para atender a uma comunidade de usuários ou domínio. Essa verdade é representada por meio de uma linguagem que, segundo Wittgenstein, é compartilhada por uma comunidade discursiva, mas não própria de um único indivíduo. Dessa forma,

[...] a importância do uso ganha uma dimensão mais complexa em Wittgenstein porque não se refere apenas à inserção de palavras em frases, mas a uma situação de ação com finalidade prática, como um exercício de influência de uns sobre os outros em um ambiente complexo. A esse ambiente, o autor denominou “jogo de linguagem” ou Sprachspiel, uma atividade regulada e partilhada (SOUZA; HINTZE, 2010, p. 115).

Nota-se, assim, que o “jogo” é composto por regras que devem ser seguidas entre os sujeitos, a partir de um consenso no uso dessa linguagem, visando à comunicação.

Da Semântica foram aplicados os princípios dos campos semânticos, uma vez que os conceitos,

termos e relações têm suas significações dependentes do valor que lhes são impressos.

Para criar a taxonomia dos relacionamentos foi preciso entender o valor semântico que os verbos abarcam no português brasileiro. Os verbos carregam um valor semântico, e compreender esse valor é importante, conforme é salientado por Soergel *et al.* (2004). Para os autores, os relacionamentos em um tesauro devem ser representados por expressões verbais. Para isso, buscamos respaldo em duas teorias: na Teoria da Valência (TV), criada por Francisco S. Borba (1996), e em parte da Teoria do Léxico Gerativo, desenvolvida por Pustejovsky (1995), especificamente do elemento que esclarece sobre a definição dos papéis *Qualia* para os verbos.

Borba (1996, p. xxi) conceitualiza a TV como um “conjunto de relações estabelecidas entre o verbo e seus argumentos ou constituintes indispensáveis”. Essa teoria foi desenvolvida a partir da combinação de duas outras teorias: a Gramática de Valências (Tesnière, 1966; Chafe, 1970; Vilela, 1992) e a Gramática de Casos (Fillmore, 1968, 1969, 1977; Anderson, 1971; Jackendoff, 1972; Cook, 1979, 1989).

Para Borba, a valência de um verbo representa o número de argumentos que ele necessita preencher para completar seu sentido e essa valência pode ser classificada em três níveis:

(1) Valência quantitativa (lógica ou lógico-semântica): número de argumentos que um predicado pode ter: avalente (zero), monovalente (um), divalente (dois), trivalente (três) e tetravalente (quatro);

(2) Valência qualitativa (sintática ou morfossintática): características dos actantes (relações gramaticais e/ou funcionais), das propriedades morfológicas, das funções sintáticas, das propriedades sintáticas e das classes que preencham os argumentos;

(3) Valência semântica: traços semânticos das categorias (+humano, +animado, +contável), das funções ou dos papéis temáticos (agente, causativo, beneficiário) e das restrições relacionais de coocorrência ou exclusão.

Para cada situação específica, em um dado domínio, é possível determinar qual a classe dos verbos (ação, processo, atividade e estado). Assim, a ausência de um agente permite que se distinga um verbo de processo de um verbo de ação, ou um verbo de atividade de um verbo de estado (Schwarze, 2001, p. 97).

Os verbos podem exprimir diversas operações, que se referem às funções ou propósitos de

objetos reais, imaginários ou abstratos. Dada essa multiplicidade de dimensões de significados dos verbos, a estrutura dos papéis *Qualia*, que é parte da Teoria do Léxico Gerativo, pode ser utilizada como passo inicial na identificação do significado semântico de um verbo. A determinação dos papéis *Qualia* para as expressões verbais, que representam as relações semânticas na estrutura de um tesouro, permite identificar os fenômenos de hiperonímia/hiponímia (relação de gênero-espécie) e de holonímia/meronímia (relação todo-parte).

A estrutura *Qualia* é composta pelos papéis Formal, Constitutivo, Télico e Agentivo, conforme descrito a seguir:

(a) Formal: generalização de uma operação descrita através de outra operação representada; distingue um objeto em um domínio mais amplo ou geral;

(b) Constitutivo: constituição de uma operação descrita, expondo-a por meio de outras operações que são necessárias para efetivá-la; indica uma relação entre um objeto e suas partes constituintes;

(c) Agentivo: especificação da entrada de uma operação, na forma de argumentos, representados por objetos reais, imaginários ou abstratos; indica elementos ou fatores que estão envolvidos na origem do objeto ou as causas para o objeto acontecer, existir ou ocorrer;

(d) Télico: especificação da saída de uma operação, na forma de objetos reais, imaginários ou abstratos; expressa o propósito e a função do objeto.

Tanto o sentido verbal (entre o verbo e o sujeito) quanto o sentido nominal (entre o substantivo e o adjetivo) dos papéis *Qualia* estabelecem as relações existentes entre os termos em um determinado contexto comunicativo. Assim, os papéis *Qualia* que interessam ser representados no refinamento das relações semânticas em tesouros são os sentidos verbais das relações que expressam a ligação que existe entre dois conceitos em um dado domínio.

### 3. Metodologia

O modelo de reengenharia de tesouros tradicionais estudado foi desenvolvido por Soergel *et al.* (2004) e, depois, detalhado por Lauser *et al.* (2006), cuja estrutura conceitual pode ser observada na Figura 2.

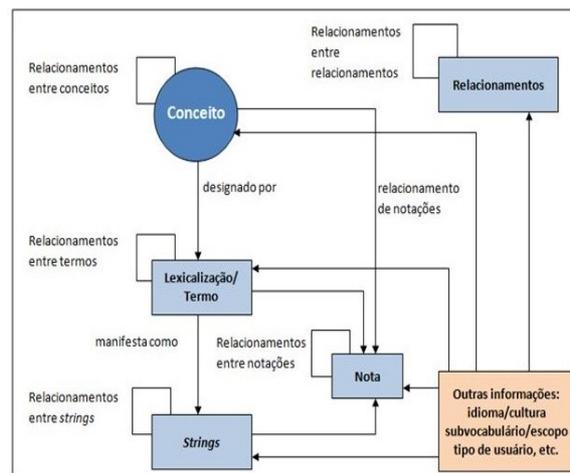


Figura 2. Modelo para reengenharia de tesouro

Observando a Figura 2 é possível perceber que os autores propuseram um modelo de reengenharia de tesouro cuja principal característica é a individualização da modelagem em cinco distintos níveis de entidades: conceito, termo ou lexicalização, *string* ou variações linguísticas, notas de escopo e relacionamentos. Com isso, os relacionamentos podem ocorrer entre entidades de mesmo tipo e entre entidades de tipos diferentes. Ademais, uma importante particularidade do modelo é a distinção entre os termos que designam os conceitos e as manifestações desses termos, representadas pelos strings, que se referem às variações linguísticas (singular/plural; variação regional).

|                        |  |
|------------------------|--|
| conceito para conceito | é_uma (hierarquia); praga_de   |
| termo para termo       | é_sinônimo_de; é_tradução_de   |
| conceitos para termos  | tem_lexicalização (liga os conceitos a seu representante lexical)                                  |
| termo para string      | tem_acrônimo; tem_variação_ortográfica; tem_abreviatura (liga os termos com suas formas variantes) |

Figura 3. Estrutura relacional do Thesagro

Dessa maneira, os relacionamentos são a espinha dorsal do tesouro.

Antes da aplicação do modelo de reengenharia de tesouro, houve um planejamento inicial no qual foi selecionada uma equipe de trabalho, o domínio modelado, da Agropecuária, e selecionou o tesouro *Thesaurus Agrícola Nacional* (Thesagro), único vocabulário controlado em português brasileiro, para a aplicação do modelo de reengenharia de tesouro.

No percurso metodológico foram utilizadas duas ferramentas: o sistema Termos Eletrônicos (e-Termos) e o Extrator de Termos e Estruturas Conceituais Agrícolas Multilíngue (ETECAM).

O sistema e-Termos é um ambiente computacional colaborativo *web*, de utilização gratuita e restrita aos usuários cadastrados. A ferramenta é composta por seis etapas, possuindo um conjunto de procedimentos automatizados e semi-automatizados, que têm como objetivo dar suporte à criação e gestão de produtos terminológicos para distintos fins (ensino, glossários, vocabulários controlados). O sistema e-Termos tem por base os fundamentos teóricos da Teoria Comunicativa da Terminologia (TCT), desenvolvida por Cabré (1999), que foi adotada no desenvolvimento deste estudo.

A ferramenta ETECAM foi criada para uso privativo da Embrapa, sendo utilizada para executar uma comparação automática sobre a existência de termos coincidentes entre terminologias de distintos tesouros. Essa ferramenta realiza as tarefas: (a) entrada de uma lista de termos, (b) a ferramenta verifica a existência de termos coincidentes à lista, (c) extrai os termos e seu *cluster* semântico e (d) os equivalentes em outros idiomas.

#### 4. Aplicação das etapas e resultados

O modelo escolhido para a reengenharia é composto por três etapas básicas: (1) definição da estrutura do tesouro, (2) coleta de terminologia e (3) edição do tesouro.

##### 4.1. Primeira etapa

A definição da estrutura do tesouro envolveu o mapeamento das características estruturais do tesouro Thesagro e a determinação como recorte temático a subárea da Intensificação Agropecuária.

O conceito de Intensificação Agropecuária adotado neste estudo segue a teoria de Boserup (1965), ou seja, é a relação entre o crescimento ou não da população de uma região e a determinação do “aumento da produção total agrícola em uma mesma unidade de área ou, ainda, a manutenção de certa produção com uma menor quantidade de insumos” (OLIVEIRA, 2011, p. 4).

O Thesagro possui cerca de 9.400 descritores, todos identificados por um ID (identificador) numérico.

A sua estrutura conceitual é composta pelos três relacionamentos básicos de qualquer tesouro tradicional, ou seja, relações de equivalência, hierárquicas e associativas.

|                           |  |
|---------------------------|--|
| Relações de equivalência: | USE e USED FOR (UF)                    |
| Relações hierárquicas:    | BROADER TERM (BT) e NARROWER TERM (NT) |
| Relações associativas:    | RELATED TERM (RT)                      |

Figura 4. Estrutura relacional do Thesagro

Nota-se que os símbolos dos diferentes tipos de relacionamentos são representados na língua inglesa.

Segundo Maculan (2015):

Foram identificadas cerca de 12.000 representações de relações associativas (RT). A versão impressa do THESAGRO apresenta uma nota explicativa que esclarece e justifica o uso de um alto número de ligações associativas (RT): para os responsáveis pela manutenção do THESAGRO, essa foi uma decisão intencional (Maculan, 2015, p. 200).

O objetivo desse tipo de abordagem é manter associações entre conceitos que não são intuitivos de serem percebidos, principalmente pelos usuário não especialista.

A modelagem da subárea da Intensificação Agropecuária teve como ponto de partida uma taxonomia já existente, elaborada pelos especialistas da Embrapa, que foi composta por 639 conceitos da temática.

Essa taxonomia foi estruturada em nove classes básicas: (1) agricultura extensiva; (2) agricultura intensiva; (3) material e métodos; (4) ambiente; (5) agronomia; (6) território e paisagem; (7) socioeconomia; (8) espaço e tempo; (9) instituições.

Uma vez que a taxonomia inicial era muito extensa, houve a necessidade de aplicar a modelagem de reengenharia do Thesagro usando uma amostra, conforme Figura 5.

|   |  |
|---|--|
| 1. INTENSIFICAÇÃO AGROPECUÁRIA<br>2. AGRICULTURA EXTENSIVA<br>3. AGRICULTURA INTENSIVA  | MATERIAL E MÉTODOS<br>4. cultura<br>5. sensoriamento remoto<br>6. sistema de informação geográfica   |
| AMBIENTE<br>7. meio ambiente<br>8. ambiente físico<br>9. solo   | AGRONOMIA<br>10. manejo da cultura<br>11. manejo do solo<br>12. pousio<br>13. período de pousio<br>14. pesticida<br>15. ciclo da cultura<br>16. cultura anual<br>17. pecuária<br>18. adubo verde<br>19. produto agropecuário<br>20. biomassa |
| TERRITÓRIO E PAISAGEM<br>21. posse da terra<br>22. escassez de terra<br>23. cobertura da terra<br>24. mudança de cobertura da terra |  |
| SOCIOECONOMIA<br>25. densidade demográfica<br>26. crescimento populacional<br>27. pressão populacional                              | ESPAÇO E TEMPO<br>28. mudança agrícola<br>29. análise de séries temporais  |
| INSTITUIÇÕES<br>30. Embrapa   |  |

Figura 5. Amostra de estudo e análise

Assim, a validação do modelo foi realizada pela modelagem e análise de uma amostragem intencional, composta por 30 conceitos representativos do conjunto de classes básicas e da área temática escolhida.

#### 4.2. Segunda etapa

A coleta de terminologia teve como insumos terminológicos a taxonomia da Intensificação Agropecuária e outros três tesauros: o Thesagro, o Agrovoc e o *National Agricultural Library* (NAL).

Nessa atividade, ocorreu a comparação entre os conceitos da amostra, oriundos da terminologia da taxonomia, e a terminologia existente em cada um dos três tesauros selecionados. Para essa comparação, a listagem original dos conceitos da amostra foi subdividida em duas listas:

Lista Um: composta pelos 30 conceitos da amostra, em português brasileiro, adicionando-se as expressões desses termos no singular e plural;

Lista Dois: composta pelos 30 conceitos da amostra, traduzidos para o inglês, adicionando-se as expressões desses termos no singular e plural, assim como na sua forma inversa (adjetivo + substantivo), por essa inversão ser comum na língua inglesa.

O procedimento da comparação terminológica foi realizado com a ferramenta ETECAM, que permitiu recuperar os termos coincidentes com as duas listas, comparando os termos de cada uma delas com a terminologia existente nos três tesauros, separadamente. Depois dessa comparação, foi realizada também uma comparação intelectual, possibilitando adicionar os conceitos e seus *clusters*, que não haviam sido resgatados com a comparação automática.

#### 4.3. Terceira etapa

A edição do tesouro Thesagro incluiu a atividade intelectual da modelagem do recorte temático da Intensificação Agropecuária e a inserção da estrutura conceitual no sistema e-Termos, para a gestão da terminologia. Assim, essa etapa foi constituída pelos seguintes procedimentos: compilação da base definicional, elaboração de glossário, confecção das fichas terminológicas, determinação de notas de escopo e construção do sistema de conceitos.

A compilação da base definicional foi realizada de forma constante e dinâmica, e recolheu e armazenou, no sistema e-Termos, contextos explicativos e/ou definitórios sobre o domínio modelado. Esse recurso informacional auxiliou a elaboração das definições terminológicas dos conceitos da amostra, dando origem ao glossário.

Para cada um dos conceitos da amostra foi confeccionada uma ficha terminológica, composta por 38 campos semânticos, para preenchimento. Dentre esses campos se destacam: definições dos conceitos (do especialista, modelador e final), informações enciclopédicas e de glosa, notas de escopo, termos em relação de equivalência, de variação linguística, assim como os conceitos em relação hierárquica e associativa. Destaca-se que foram elaboradas notas de escopo para alguns dos conceitos da amostra, a partir da avaliação feita pela equipe de trabalho quanto à necessidade dessa nota explicativa.

A construção do sistema de conceitos foi realizada a partir dos conteúdos das definições e dos registros das fichas terminológicas. Essa atividade envolveu a aplicação de 44 diferentes relações, que criou uma rede semântica para os conceitos da amostra e seus *clusters* semânticos, gerando um desdobramento que totalizou cerca de 600 relacionamentos.

Foram representados os relacionamentos: de gênero e suas espécies, do todo e suas partes, de equivalências, de *strings* (variações) e associativas.

Para exemplificar e comparar a atual estrutura do Thesagro e a modelagem realizada utilizando o modelo de reengenharia de tesouro, a seguir apresenta-se a atual modelagem do descritor GATO no Thesagro e, depois a estrutura refinada:

|  |  |
|--|--|
| GATO<br>BT MAMÍFERO DOMÉSTICO<br>NT GATO ANGORÁ<br>NT GATO DO MATO<br>RT FELIS CATTUS DOMESTICUS<br>RT FELIS DOMESTICA | FELIS CATTUS DOMESTICUS<br>RT GATO<br><br>FELIS DOMESTICA<br>RT GATO |
|--|--|

Figura 6. Estrutura atual do Thesagro

Na estrutura atual do Thesagro, percebe-se que os relacionamentos são ainda bastante genéricos. Além disso, os descritores FELIS CATTUS DOMESTICUS e FELIS DOMESTICA são nomes científicos para GATO, mas essa informação fica perdida na estrutura atual do Thesagro, pois os descritores estão ligados por uma relação associativa e não por uma relação de equivalência.

Aplicando o modelo de reengenharia de tesaurus, a reformulação da modelagem do descritor GATO ficou com a seguinte configuração:

|  |  |
|--|--|
| GATO<br>temTermoGenérico MAMÍFERO DOMÉSTICO<br>temTermoEspecifico GATO ANGORÁ<br>temTermoEspecifico GATO DO MATO<br>temNomeCientífico FELIS CATTUS DOMESTICUS<br>temNomeCientífico FELIS DOMESTICA | FELIS CATTUS DOMESTICUS<br>temNomePopular GATO<br><br>FELIS DOMESTICA<br>temNomePopular GATO |
|--|--|

Figura 7. Estrutura reformulada do Thesagro

Nota-se que houve o refinamento das relações estabelecidas entre os conceitos e termos, com a explicitação do tipo de ligação há entre eles, facilitando o entendimento da estrutura semântica. Com a reformulação da estrutura do Thesagro ficou claro ao usuário quais são os nomes científicos utilizados para representar o conceito GATO, que é o nome popular desse animal.

Outro exemplo é a estrutura original do Thesagro para o descritor PESTICIDA e a estrutura remodelada, utilizando o modelo de reengenharia de tesaurus:

| (1) ESTRUTURA ATUAL   | (2) REENGENHARIA DA ESTRUTURA  |
|---|--|
| PESTICIDA<br>NT ACARICIDA<br>NT CARRAPATICIDA<br>NT FUNGICIDA<br>NT GERMICIDA<br>NT INSETICIDA<br>NT MOLUSCICIDA<br>NT NEMATICIDA<br>NT PERSISTÊNCIA DE PESTICIDA<br>NTRATICIDA<br>NT REPELENTE | PESTICIDA<br>temoGenéricoGênero (TGG) SUBSTÂNCIA QUÍMICA<br>temoEspecificoGênero (TEG) ACARICIDA<br>temoEspecificoGênero (TEG) CARRAPATICIDA<br>temoEspecificoGênero (TEG) FUNGICIDA<br>temoEspecificoGênero (TEG) GERMICIDA<br>temoEspecificoGênero (TEG) INSETICIDA<br>temoEspecificoGênero (TEG) MOLUSCICIDA<br>temoEspecificoGênero (TEG) NEMATICIDA<br>temoEspecificoGênero (TEG) RATICIDA<br>temoEspecificoGênero (TEG) REPELENTE<br>temPropriedade (TR) PERSISTÊNCIA DE PESTICIDA |

Figura 8. Estrutura conceitual: descritor PESTICIDA

É possível notar que na estrutura original (1) do THESAGRO o descritor PESTICIDA não está ligado a um conceito superordenado (hiperônimo), que indique o seu pertencimento a uma classe mais geral. Também observamos as relações hierárquicas entre PESTICIDA e os outros 10 descritores, mas são ligações genéricas que não identificam se são relações do tipo gênero-espécie, todo-parte ou de instância.

Ao aplicar a reengenharia (2) nessa estrutura do Thesagro, atribuímos uma relação hierárquica, do tipo gênero-espécie, entre PESTICIDA (TEG) e SUBSTÂNCIA QUÍMICA (TGG), que indica a classe mais geral de PESTICIDA. Com exceção do descritor PERSISTÊNCIA DE PESTICIDA, foram refinadas as relações hierárquicas entre PESTICIDA e os outros descritores, identificadas como relações do tipo gênero-espécie (os conceitos subordinados carregam as mesmas características do conceito PESTICIDA).

Com o descritor PERSISTÊNCIA DE PESTICIDA foi percebido um problema de abstração conceitual na estrutura original do Thesagro, pois ele não é um tipo de PESTICIDA, mas uma propriedade que indica o período durante o qual a toxicidade do pesticida permanece inalterada (longevidade do produto), afetando o ambiente no qual foi aplicado. Assim, criamos uma relação associativa, do tipo <tem\_propriedade>, entre PESTICIDA e PERSISTÊNCIA DE PESTICIDA, facilitando a compreensão do domínio.

Na representação e refinamento dos relacionamentos na estrutura conceitual dos conceitos da amostra, houve predominância dos relacionamentos hierárquicos, totalizando 286 ocorrências, sendo 225 relações de gênero-espécie (com 52 termos gerais e 173 termos específicos) e 61 relações todo-partes (com 22 termos gerais e 39 termos específicos). Esse resultado demonstrou, na nova estrutura semântica construída para o Thesagro, a manutenção da sua natureza específica de origem.

Quanto às relações associativas, foram representados 232 relacionamentos, evidenciando a complexidade da subárea da Intensificação Agropecuária.

Os resultados demonstraram que a expressão explícita das relações entre pares de entidades (conceitos, termos, strings e notas de escopo) refinou a semântica da estrutura do tesouro, dando subsídios para facilitar a interoperabilidade entre diferentes tesouros ou sistemas.

## 5. Considerações finais

Este estudo teve como ambientação a Embrapa Informática Agropecuária (Embrapa), Unidade Campinas/SP, sendo o primeiro resultado do convênio firmado entre essa instituição, a Universidade Federal de Minas Gerais (UFMG) e o Grupo de Pesquisa Protótipo Mapa Hipertextual (MHTX).

O desenvolvimento deste estudo partiu do pressuposto de que os tesouros tradicionais já possuem uma representação semântica bastante consistente, mas ainda insuficiente para classificá-lo como um SOC, que abarca instrumentos que podem ser entendidos pela máquina.

Dessa maneira, neste estudo foi priorizada a explicitação dos relacionamentos entre conceitos e termos, que torna a estrutura conceitual do tesouro semanticamente mais rica.

Ao se desenvolver uma nova estrutura semântica para o Thesagro, ficou evidente a necessidade da adoção de normas e padrões internacionais na criação de tesouros, que oferece parâmetros que permitem maior formalidade nas representações dos relacionamentos, auxiliando a interoperabilidade com outros vocabulários.

Nesse sentido, o uso do modelo de dados *Simple Knowledge Organization Systems* (SKOS) possibilita a explicitação de relações que podem ser interpretadas (legíveis) por máquinas, sendo um elemento importante que pode auxiliar na interoperabilidade entre diferentes vocabulários e sistemas.

Por fim, a aplicação do modelo de reengenharia selecionado auxiliou a organização e a ampliação da visão acerca do conhecimento da área modelada, o que poderá facilitar a sua compreensão e exploração pelo usuário, uma vez que tal refinamento os tornou mais claros e específicos.

## Referências

Anderson, J. M. (1971). *The grammar of case: towards a localist theory*. London: CUP.

Borba, F. S. (1996). *Uma gramática de valências para o português*. São Paulo: Ática.

Boserup, E. (1965). *The conditions of agricultural growth: the economics of agrarian change under population pressure*. Chicago: Aldine.

Cabré, M. T. (1999). *La terminología: representación y comunicación*. Barcelona: Institut Universitari de Lingüística Aplicada, Universitat Pompeu Fabra.

Campos, M. L. A. (1995). Linguagens documentárias: núcleo básico de conhecimento para seu estudo. *R. Esc. Biblioteconomia UFMG*, Belo Horizonte, 24:1, (jan./jun.) 52-62.

Campos, M. L. A.; Gomes, H. E. (2003). Organização de domínios de conhecimento e os princípios rangathanianos. *Perspectivas em Ciência da Informação*, Belo Horizonte, 8:2 (jul./dez.).

Chafe, W. L. (1970). *Meaning and the structure of language*. Chicago: University of Chicago.

Cook, W. A. S. J. (1979). *Case grammar: development of the matrix model (1970-1978)*. Washington, D.C.: Georgetown University.

Cook, W. A. S. J. (1989). *Case grammar theory model*. Washington, D.C.: Georgetown University.

Dahlberg, I. (1978a). Teoria do conceito. Tradução Astério Tavares Campos. *Ciência da Informação*, Rio de Janeiro, 7:2, 101-107.

Dahlberg, I. (1978b). A referent-oriented, analytical concept theory of Interconcept. *International Classification*, 5:3, 122-151.

Dahlberg, I. (1992). Knowledge organization and terminology; philosophical and linguistic bases. *International Classification*, 19:2, 65-71.

Fillmore, C. J. (1968). The case for case. In: BACH, E.; HARMS, R.T. (Org.) *Universals in linguistic theory*. New York: Rinehard and Winston. p. 1-88.

Fillmore, C. J. (1969). Types of lexical information. In: KIEFER, F. (Ed.). *Studies in syntax and semantics*. Dordrecht-Holland: D. Reidel.

Fillmore, C. J. (1977). *A semântica na linguística moderna: o léxico*. Tradução de Lúcia M. Lobato. Rio de Janeiro: Francisco Alves.

Hodge, G. (2000). *Systems of knowledge organization for digital libraries: beyond traditional authorities files*. Washington, DC: Council on Library and Information Resources. Recuperado 24-06-2013, de <http://www.clir.org/pubs/re-ports/pub91/contents.htm>.

Jackendoff, R. (1972). *Semantic interpretation in generative grammar*. Massachusetts: MIT Press, Cambridge.

Kobashi, N. Y.; Francelin, M. M. (2011). Conceitos, categorias e organização do conhecimento. *Informação e Informação*, Londrina, 16:3, (jan./jun.), 1-24.

Lauser, B. et al. (2006). *From Agrovoc to the Agricultural Ontology Service: Concept Server an OWL model for creating ontologies in the agricultural domain*. En International Conference on Dublin Core and Metadata Applications, 2006, Colima, Mexico. México: DCMI.

Lima, G. A. B. O. (2007). *A análise facetada na modelagem conceitual para organização hipertextual de documentos acadêmicos: sua aplicação no protótipo MHTX (mapa hipertextual)*. *Informação e Sociedade: Estudos*, João Pessoa, 17:1, (jan./abr.), 31-41.

Maculan, B. C. M. S. (2015). *Estudo e aplicação de metodologia para reengenharia de tesouro: remodelagem do THESAGRO*. 345f. Tese (Doutorado) – Universidade Federal de Minas Gerais, Escola de Ciência da Informação, Belo Horizonte, Brasil.

Marcondes, D. (2000). *Filosofia, linguagem e comunicação*. São Paulo: Cortez.

- Motta, D. F. da. (1987). *Método relacional como nova abordagem para a construção de tesouros*. 1987. Dissertação (Mestrado em Ciência da Informação) – Instituto Brasileiro de Informação em Ciência e Tecnologia, Rio de Janeiro.
- Oliveira, J. R. (2011). Sustentabilidade e intensificação produtiva da agricultura familiar: um estudo comparativo entre duas comunidades em Itapejara D'Oeste, Sudoeste do Paraná. *Synergismus Scyentifica*, Universidade Tecnológica Federal do Paraná, Pato Branco, 6:1.
- Pustejovsky, J. (1995). *The generative lexicon*. Cambridge, MA: MIT Press.
- Ranganathan, S. R. (1967). *Prolegomena to library classification*. Bombay: Asia Publishing House.
- Soergel, D. (1999). The rise of ontologies or the reinvention of classification. *Journal of the American Society of Information Science*, 50:12, 1119-1120.
- Soergel, D. et al. (2004). Reengineering thesauri for new applications: the AGROVOC example. *Journal of Digital Information*, 4:4.
- Vickery, B. C. (2007). *A note on knowledge organization*. Site Lifeboat for Knowledge Organization. Recuperado 11-06-2013, de [http://www.iva.dk/bh/lifeboat\\_ko/concepts/Vickery-a\\_note\\_on\\_knowledge\\_organisation.htm](http://www.iva.dk/bh/lifeboat_ko/concepts/Vickery-a_note_on_knowledge_organisation.htm).
- Wüster, E. (1998). *Introducción a la teoría general de la terminología y a la lexicografía terminológica*. Barcelona: IULA.
- Schwarze, C.(2001). La sémantique do verbe. In: \_\_\_\_\_. *Introduction à la sémantique lexicale*. Tübingen: Nar. p. 89-113.
- Svenonius, E. (2000). *The intellectual foundations of information organization*. Cambridge: The MIT Press.
- Tesnière, L. (1966). *Éléments de syntaxe structurale*. 2. ed. Paris: Klincksieck.
- Tristão, A. M. D.; Fachin, G. R. B.; Alarcon, O. E. (2004). Sistema de classificação facetada e tesouros: instrumentos para organização do conhecimento. *Ciência da Informação*, Brasília, 33:2, (ago), 161-171.
- Vilela, M.(1992). *Gramática de valências: teoria e aplicação*. Coimbra: Almedina.

---

# The Future of Knowledge Organization and Information Organization

---

H. Peter Ohly

International Society for Knowledge Organization ISKO, peter.ohly@gmx.de

## Abstract

On the 13<sup>th</sup> conference of the International Society for Knowledge Organization (ISKO) 2014 in Krakow a panel was held on 'The Future of Knowledge Organization and ISKO'. Here a synopsis with additional information and more dense presentation is given by a panelist. The main items were: What is knowledge organization (KO)? What will be the most challenging

for ISKO and KO in the future? What is your ideal picture of ISKO and KO of the future? The results were inter alia that to ameliorate its standing not only the fundamentals have to be worked out but also its diverse application fields and openness to new objects and methods.

**Keywords:** ISKO. Knowledge Organization. KOS. Library Science.

## 1. Introduction

The 13<sup>th</sup> international ISKO conference titled 'Knowledge Organization in the 21st Century: Between Historical Patterns and Future Prospects' was held in Krakow in May 2014. With respect to the 25th Anniversary of ISKO and Knowledge Organization, Rebecca Green, Assistant Editor for the Dewey Decimal Classification at Online Computer Library Center (OCLC), moderated a panel on 'The Future of Knowledge Organization and ISKO'. She understood this panel as a platform for discussing knowledge organization in the past, present, and future within ISKO. The statements of the panelists and the audience were worked out and published in the bi-monthly Journal of the International Society for Knowledge Organization (KO) by Green (2014). Here a synopsis with additional information and a more dense presentation of the published part will be given by one of the panelists.

Participants of the panel were: Joseph Tennis, Univ. Washington, Information School, Assoc. Professor (new ISKO President 2014-2018), Vera Dodebei, Chair of the Brazilian ISKO, Univ. Fed. Rio de Janeiro, Graduate Program in Lib. Sci., PhD in Comm. and Culture, Rosa San Segundo, Chair of the Spanish ISKO, Univ. Carlos III Madrid, Director o. DeLib. and Inform. Sci., Associate Professor, Wiesław Babik, Chair of the Polish ISKO, Jagiellonian Univ. Krakow, Inst. of Inform. and Lib. Sci, Assoc. Professor, Peter Ohly, ISKO President 2010-2014, Social Scientist (formerly at GESIS – Leibniz-Institute for the Social Sciences in Germany), Amos David, Chair of the French ISKO, Lorraine Univ. Nancy, Res. Lab. ComSci. and Appl., Prof. Inform. Sci.

and Comm., and Claudio Gnoli, Webmaster of ISKO, Univ. Pavia, Mathematics Department Library, Natural Scientist.

The discussion was amended by the audience, namely: Ingetraut Dahlberg, Founder of German Classification Society and of ISKO, PhD in Linguistics, Grant Campbell, Univ. Western Ontario, Faculty of Inform. and Media Studies, Assistant Professor, Dagobert Soergel, University of Buffalo, Department of Library and Information Studies, Prof. Emeritus, Inform. Studies, Maria Lopez-Huertas, ISKO President 2006-2010, Univ. Granada, School of Library and Information Science, Professor, Jill McTavish, Librarian at London Health Sciences Centre, Ontario, and Laura Ridenour, University of Wisconsin–Milwaukee, MA.

The *International Society for Knowledge Organization (ISKO)* denotes itself "the premier international scholarly society devoted to the theory and practice of knowledge organization" (1). ISKO charter lists as ISKO's aims: "...to promote research, development and application of all methods for the organization of knowledge in general or of particular fields by integrating especially the conceptual approaches of classification research and artificial intelligence. The Society stresses philosophicological, psychological and semantic approaches for a conceptual order of objects..." (2). With reference to these aims Rebecca Green posed more or less these three questions to the panelists (3):

1. What is knowledge organization (KO)?
2. What changes do you foresee in the future that will prove to be the most challenging for ISKO?

3. What is your ideal picture of what the IS-KO of the future could be? How do we get there?

## 2. What is knowledge organization (KO)?

Green introduces this question by a critical statement: "In Dewey [Decimal Classification] the rule of application instructs us to class a work on, say, a thesaurus of architecture – that is, the making of a thesaurus applied to architecture – with other works on architecture. But developing a thesaurus on architecture doesn't make the developer an architect." Grant Campbell takes up this statement directly by stressing out, that disciplines and domains have their own terms, practices, traditions, canonical texts but KO is outside of special domains and has more an iterant role by communicating information between different groups. This reminds me to the fiction of a 'Troubadour of Knowledge' as de Beer (2010) had stated it for the knowledge worker in a new knowledge age with reference to 'Le Tiers-Instruit' by Michel Serres (1997). In so far this might be already a position for the third question.

Dahlberg explains the origin of the naming of 'Knowledge Organization', which is not at least a part of the name of the society which was founded in 1989 in Germany. They took up the wording 'Organization of Knowledge [in libraries]' of Henry Bliss (1933) but changed it into the shorter 'Knowledge Organization', what is in German an allowable collocation of words. It should describe order activities in classification. In other languages this might be misunderstood, as organization refers to institutions and business aspects. The earlier journal *International Classification* of the German Society for Classification, which was founded mainly by Dahlberg, was accordingly renamed to *Knowledge Organization* but the coding system for the bibliography remained exactly the same.

In so far Ohly regards ISKO and KO within its historical and structural boundaries. It emerged from library science cataloging. But the GfKI (German Society for Classification) was founded in contrast to the DGD (German Society for Documentation), with the GfKI as more theoretical and methodological oriented and less stress on documentation praxis. In contrast the ISKO was founded by the non-statistician part of GfKI as a society with less orientation to business informatics. Hereby problems are arising as it lost its connections to computer-oriented fields, including knowledge management. Hence ISKO and KO have to claim a focus that is not already occupied by other established scientific neighbor

communities, e.g., artificial intelligence, neuroscience.

Dahlberg sees KO as a subdiscipline of the *Science of Science*, what brings it in my view on a critical meta level to other sciences but underlines in my understanding as well its more descriptive research orientation (like Scientometrics) and less its fundamental approach. Hjørland instead sees KO as a meta science like the science of science as it has a unique focus, but also with a dependency from subject knowledge (2013).(4) Dahlberg underlines that the application fields of KO are not only in the Information Sciences but also in all subject fields (domains) needing taxonomies (classification systems of objects). She mentions as examples: statistics, commodities, utilities, weapons, patents, museology, however with the above stated limitation of classification aims. In KO the scientific objects are '(all kinds of) knowledge' and the scientific methods here are the 'organization' principles, that create order of the given kinds of knowledge and its activities.

Babik states the many different definitions of 'knowledge organization' and recommends etymological definitions of its constituent parts 'knowledge' and 'organization'. Only then it becomes more clear what is meant with 'knowledge organization'. In my opinion this includes time and culture dependent definitions. Like Dahlberg he sees the subjects of KO as compositions of knowledge. Here information is seen as a raw material for knowledge, but it becomes its meaning from the viewpoint of its organization. As a science of [various aspects of] knowledge KO is for him indispensable to science, education and research, as well as to information science.

Ohly complains that ISKO as a society does not attract a well-defined established profession, like 'Knowledge Organizer' or 'Semantic Worker'. But it has its main application and acceptance in library science. When the focus of KO should be more general one should speak of 'arranging of knowledge' instead of ordering, classification or organization, what has some connotation with rigidity and stability. How far extraction, connection, reasoning, or interpretation of knowledge should be included in the focus of ISKO has carefully to be considered. Are these only some aspects of knowledge, which are treated well in other disciplines or are these inevitable implications of classification? For him the definition of KO is missing: the economic dimension (high quality information implies this), the scientific background of the applied field (which changes the principles of classification) (5), and the sociological aspects (which are relevant for the

acceptance of ordering systems, for the social dynamics of use and misuse, and for the development of social software models).

Soergel as well asserts that most ISKO members come out of a library and bibliographic systems tradition, what “presents somewhat a barrier to bring the KO expertise to the much wider arena where it is applicable and where it would be beneficial”. But he rates documentary information only as a part of the information landscape. As other applications of high importance, Soergel names: electronic health records (EHR), scientific data, research networking systems, business information systems, linked data as a format. “To enable transfer of ISKO expertise into these wider application areas and the associated communities requires a re-orientation. ISKO members need to work in other areas ...”. López-Huertas (2014) and Rodríguez-Bárceñas/López-Huertas (2013) refer e.g., to the importance of KO in decision making processes. Even more impetus on management relations of KO is given by the report of the ISKO-Maghreb chapter (Sidhom, 2014a) with respect to the themes of its annual conferences: “The challenge of integration of business applications in dematerialized flows is fundamental to avoid breaks in the information processing. The other challenge is to manage in a unified way the whole relationship with the customer or user, regardless of the exchange modes he used, and regardless of the requested resources: data, information, knowledge, know-how and skill”(6). Even more the program of the ISKO Maghreb conference 2014 states the important role of context conditions for KO (7), what could be named as ‘Order of Knowledge’ or ‘Knowledge Order’ (c.f. Spinner, 1094)

Gnoli suggests to analyze ISKO resources to identify the scope of KO: the journal Knowledge Organization (8), the online KO bibliography (9), and the forthcoming online dictionary/glossary of KO. In his definition e.g. ‘KO events’ should deal with the subject content of documents in a broad sense of document, not just in libraries but independently from the technical means and carriers it addresses. He refers to Buckland (2014) who sees a trend to “ubiquitous recording, pervasive representations, simultaneous interaction regardless of geography, and powerful analysis and visualization of the records resulting from that ubiquitous recording”. He votes for ‘conceptual interoperability’ (conceptual mapping, SKOS, OWL, etc.) as field of KO, in contrast to technical interoperability. Therefore developments of the semantic web should be included in KO. KO is often named with other terms in the field of ontology, taxonomy, terminology, topic

maps, information architecture, etc. Here KO is often not identified as a field in itself, as basic logical components of knowledge (classes, hierarchies, terms, etc.) are often taken for granted. What is lacking is a common, consistent terminology in this field. He prefers the view of ‘dimensions of KO’ (ontological, epistemological, pragmatic, etc.).

As summary of the question ‘What is KO?’ we can state that the main focus is order activity or classification. As such it has a meta view on the subjects it deals with and has some universal aspects like Science of Science. Nevertheless the focus must be seen wider than classification in library science and take into account wider application areas not at least in informatics though the bridge building function and conceptual dimension seem to be its main merit.

### 3. What changes do you foresee in the future that will prove to be the most challenging for ISKO?

Already in the previous chapter when the panel tried to shape KO some statements concerned future challenges and recommendations for the future: the itinerant role of KO (Campbell), the lacking computer orientation, application orientation, economic and social aspects, professionalization (Ohly), clear terminology (Babik, Gnoli), opening to more than library tradition and diversification (Soergel, Gnoli), semantic web challenges (Gnoli). But the subsequent discussion took up more explicit points.

If knowledge organization should be a scientific discipline in its own right, Dahlberg demands to develop it accordingly and start with elaborating its roots, such as Wuester’s work on concepts (10), her contribution on concept definition and concept systems (11), her development of an Information Coding Classification (ICC), as well as the fundamental studies of Ranganathan (1967) on faceted classification.

Ohly demands that KO must be more open to realize that there are new applications, new knowledge sources and quite other applications than library cataloging: virtual knowledge generation, mobile devices, decision making, evaluation indexes. By the way: ISKO UK had in November 2014 a meeting on ‘Knowledge organization goes mobile’(12). Literacy is wanted on KO for users from other communities but as well there is a permanent need for understanding new upcoming techniques and thinking in neighboring fields. Openness to understanding and applicability of neighboring disciplines, specialized areas, and other cultures can be strength-

ened by according tutorials, workshops, and co-operations.

Likewise Lopez-Huertas recommends to invite speakers from other communities to ISKO conferences and to make sure that there is a sufficient number of papers of interest to members of other communities.

And McTavish demands to incorporate better new, different, and upcoming voices, e.g. from students. ISKO should also offer partial conference scholarships to new students. The new ISKO president Tennis in his candidacy statement (2014a) argues similarly for: survey to the membership, convene discussion sessions, working groups. In another contribution (Tennis, 2008) he says: "Our elenchus [rhetoric] is unique, and by acknowledging what is and what it is not, we can see how our work interfaces with myriad research initiatives and the legion of new techniques, tools, and systems of organization". In the discussion round he argued in the same way "to question what is core and what is peripheral" and to have an "open discussion about these issues" instead of relying on 'canons' that define "what is core and what is peripheral" (Tennis, 2014).

Dodebei sees the problem, that with the internet and dynamic approaches we are losing historical traces of knowledge. This leads over to the question of supporting sustainable knowledge for a 'knowledge society'. Whereas the context is unstable we have to find means to preserve a secured constant knowledge pool, especially in the soft sciences (art, history, anthropology, archeology). Compare with this position Buckland (2014), who states: "The tension between the benefits of technology and the limitations imposed by fixity in a changing world provide a central tension in knowledge organization over time". For Dodebei cultural discussions, concerning e.g. the connected societies, have to be included in ISKO topics.

For David the connected world (with the associated functionalities) changes the way knowledge is acquired, represented, managed and exploited. Hence KO has to care about. But mankind should resist the temptation of research [and methods] that are only technologically driven. Though the statements of San Segundo are not published in Green (2014) we can add here her conclusion in 2006 (San Segundo, 2008). "The new organisation of knowledge points to a totally new conception; post-modern epistemology has yet to be articulated. There has been a leap from the invalidity of a general knowledge theory, which culminated in positivist epistemology, to a new digital organisation which is objectively de-

structured and structured from subjectivity, based on semantic networks instead of lexical similarities, within this process of methodological revolution and, with some new material parameters, the forthcoming millennium invades a new organisational form of knowledge in the digital post-modern universe."

In the same way Campbell sees missions for KO in the future (as well as for other disciplines). to negotiate the demands of different cultures instead of enforcing uniformity, and to think more easily and clearly in terms of sustainability. The next international ISKO congress 2016 in Rio de Janeiro will exactly be devoted to such a theme namely 'Knowledge Organization for a Sustainable World'.

Tennis sees a big potential in looking at other neighbor associations and groups, with their own approaches, terminology, means, and purposes, as classification is a problem solving activity not only in very different kinds of libraries but also by everyone trying to organize digital material (2014).

Babik promotes a network approach to knowledge organization, both in its theoretical and conceptual dimensions as well as in practical ISKO activities. Whereas in the past there was an explicit tendency toward automation, globalization and socialization of information and knowledge creation processes, we have now to come back to more human-oriented and sustainable developments.

The question 'Challenges for ISKO/KO' provoked statements concerning its mission and strategy. ISKO has to contribute to sustainability and human aspects in the information task. KO literacy, openness, interdisciplinarity, and network approaches are seen as demands for the future.

#### **4. What is your ideal picture of what the ISKO, resp. the KO, of the future could be? How do we get there?**

Though in the previous question already ideals were formulated (Dahlberg: back to the roots; Ohly, McTavish, Lopez-Huertas: openness; Dodebei, Campbell, Babik: sustainability and human-orientation) the following question on an ideal ISKO brought up further considerations.

For David the current orientation of ISKO should be maintained and reinforced. "To maintain its level of recognition, the community should remain focused on scientific objects rather than technology-dependent issues." Here one could refer to the 10 desiderata of Dahlberg (2011).

Ridenour favors an open access model of publication to provide access for people who may be interested in KO, but are not part of the community. Especially as “KO literature is both difficult to locate and misindexed in databases such as [Library, Information Science & Technology Abstracts] LISTA, usually placed under knowledge management” a forum might be helpful to attract more people and to collaborate with individuals in other research specialties.

Babik says that “ISKO and KO will benefit from the implementation of the idea of information and knowledge society, because this process demands high quality information and knowledge”. The idea of *Knowledge ecology* resp. *Information ecology* (13) holds as well for KO. He distinguishes for ISKO activities three basic levels: international, national and local. Accordingly the ISKO structure should be developed, like it has been done in the Polish chapter.

Ohly thinks of ISKO as a virtual institute where, like in e-science, projects and advice functions are performed virtually with scientists, coming as well from other disciplines. For him it is more important to explain and elaborate the differences, strengths and weakness of special KO approaches in special applications instead of knowing what is the best KO system.

Tennis sees ‘polytely’, the complex problem-solving with multiple goals as a concept for ISKO and KO, what implies for him the open dialogue with other disciplines and professions.

For Soergel ISKO would ideally “develop into a society that covers KO issues in a wide range of applications, with keen attention to common principles, and that attracts people focusing on KO from many communities, serving as a common meeting point for the transfer of basic knowledge and of reusable modules in the development of KO systems. ... ISKO should get involved in formulating information literacy standards ... for deeper understanding of principles of knowledge...” (Green, 2014). One should be aware of the wide range of KO applications, e.g. CYC Ontology, WordNet, Gene Ontology (GO), SnoMed, etc. (14). “This extension of the range should also be pursued for the journal KO. [...] Finally, it would be useful to create a list of associations, conferences, and separate listservs that deal with KO and also repositories for KOS” (Green, 2014).

To sum up, ideally KO and ISKO are seen as an exchange forum, not at least via electronic communication means, that provides access to standards and different approaches in the field of KO. Thus it would meet information

ecological and complex problem solving demands.

## 5. Conclusion

Though the charter of ISKO (1989) mentions “all methods for the organization of knowledge”, “especially the conceptual approaches”, the practice and attraction of ISKO and subsequently of its field KO is mainly restricted to the library and documentation science. KO is stated as a meta science to science, as it is applicable in all science fields and especially has the potential to communicate information between various fields. But the current tendency of mere technical orientation is seen as a threat. Instead semantics and conceptual interoperability should play a bigger part. A common terminology, free access to basic papers, as well as repositories for modules are lacking. Cultural diversity, openness and ploytely as well as high quality information have to be guaranteed by KO, leading to a knowledge ecology. Quite much more application areas must be seen as KO playing grounds than currently perceived in KO.

## Notas

(1) <http://www.isko.org/> (all Internet links refer to 7 Nov. 2014)

(2) <http://www.isko.org/charter.pdf>

(3) c.f. the questions in McIlwaine/Mitchell (2008). 1. Can knowledge organization principles be extended to a broader scope, including hypertexts, multimedia, museum objects, and monuments? 2. Can the two basic approaches, ontological and epistemological, be reconciled? 3. Can any ontological foundation of knowledge organization be identified? 4. Should disciplines continue to be the structural base of knowledge organization? 5. How can viewpoint warrant be respected? 6. How can knowledge organization be adapted to local collection needs? 7. How can knowledge organization deal with changes in knowledge? 8. How can knowledge organization systems represent all the dimensions listed above? 9. How can software and formats be improved to better serve these needs? 10. Who should do knowledge organization: information professionals, authors or readers?

(4) there he explains: “...you cannot classify domains on the basis of theories of knowledge (or other metadisciplines, including genre studies, the sociology of knowledge, etc.)... Epistemology is, however, the best general background... Concepts and semantic relations are not a priori or neutral, but should be examined in relation to their implications for the users they are meant to serve.”

(5) c.f. Hjørland, 2013: “...In order to achieve good consistent indexing, the indexer must have a thorough appreciation of the structure of the subject and the nature of the contribution that the document is making to the advancement of knowledge...”

(6) Sidhom explains in an e-mail discussion, that ‘competitive advantage’ is the core point for the Maghreb world, whereas library and information science is of minor importance. Accordingly under ‘Classification’ as main topics are listed: ‘Knowledge management’ res‘Information man-

agement' (Sidhom, 2014) – another interpretation of these concepts?

(7) "The governance of knowledge seems to be the Scientific Policy most able to creating value with regard of human and its evolution in cultures and civilizations. The duty of good governance is a consideration of the transfer of knowledge related to scientific and technological progress. Intrinsically, this process requires a system of organization and knowledge management by implementing knowledge production and its influence in society. The objective of the ISKO-Maghreb Chapter is to contribute in understanding the factors that organize knowledge and phenomena that affect the information society." <http://iskomaghreb2014.cerist.dz/index.php?lang=en>

(8) <http://www.ergon-verlag.de/en/bibliotheks--informationswissenschaft/index.php>

(9) see: <http://www.isko.org/lit.html>

(10) c.f. DIN 2330 (2013); ISO 704 (2009)

(11) see: Dahlberg, 1979; Dahlberg, 2009

(12) see: [http://www.iskouk.org/events/mobileKO\\_Nov\\_2014.htm](http://www.iskouk.org/events/mobileKO_Nov_2014.htm)

(13) see e.g. Capurro, 2011

(14) more he mentions in Green, 2014

(15) all Internet links refer to 7 Nov. 2014

## Referencias (15)

- Buckland, Michael K. (2014). Knowledge organization and the technology of intellectual work. En: Babik, W. (2014). *Knowledge Organization in the 21st Century: Between Historical Patterns and Future Prospects. Proceedings of the 13th International ISKO Conference 19-22 May 2014 Kraków, Poland*. Würzburg: Ergon, 14-21.
- Capurro, Rafael (2011). Towards an Information Ecology. En: Wormell, I. (1990, updated 2011). *Information Quality. Definitions and Dimensions*. London: Taylor Graham, 122-139. <http://www.capurro.de/nordinf.htm>
- Dahlberg, Ingetraut (2011). How to improve ISKO's standing: ten desiderata for knowledge organization. *Knowl. Org.* 38 (1), 68-74.
- Dahlberg, Ingetraut (2009). Concepts and terms: ISKO's major challenge. *Knowl. Org.* 36 (2/3), 169-177.
- Dahlberg, Ingetraut (1979). On the theory of the concept. En: Neelameghan, A. (1979). *Ordering Systems for Global Information Networks; Proceedings of 3rd International Conference on Classification Research, Bombay, India, January 6-11, 1975*. Bangalore: Sarada Ranganathan Endowment for Library Science, 54-63.
- De Beer, Carel S (2010). The troubadour of knowledge: a knowledge worker for the new knowledge age. En: Gnoli, Claudio; Mazzocchi, Fulvio (2010). *Paradigms and conceptual systems in knowledge organization: Proceedings of the 11th International ISKO Conference*. Würzburg: Ergon, 430-437.
- DIN 2330 (1993). *Begriffe und Benennungen – Allgemeine Grundsätze (concepts and designations – general principles)*. Berlin: Beuth, 1988, updated: 1993, 2013.
- Green, Rebecca (2014). ISKO and Knowledge Organization's 25th Anniversary: The Future of Knowledge Organization and ISKO. Panel Discussion. *Knowl. Org.* 41 (4), 327-331. See also: [http://isko-brasil.org.br/wp-content/uploads/2014/06/relat\\_iskoCrac%C3%B3via2014.pdf](http://isko-brasil.org.br/wp-content/uploads/2014/06/relat_iskoCrac%C3%B3via2014.pdf)
- ISO (1989). *ISKO charter and preamble*. <http://www.isko.org/charter.pdf>
- ISO 704 (2009). *Terminology work - Principles and methods. 2000, updated*. Geneva: International Organization for Standardization.
- Hjørland, Birger (2013). Theories of knowledge organization - Theories of knowledge. *Knowl. Org.* 40 (3), 169-81.
- López-Huertas, María J. (2014). Exploring the boundaries of knowledge organization. Towards future projects. En: Babik, W. (2014): *Knowledge Organization in the 21st Century: Between Historical Patterns and Future Prospects. Proceedings of the 13th International ISKO Conference 19-22 May 2014 Kraków, Poland*. Würzburg: Ergon, 564-571.
- McIlwaine, la C.; Mitchell, Joan S. (2008). Preface to Special Issue. 'What is Knowledge Organization'. *Knowl. Org.* 35 (2/3).
- Ranganathan, Shiyali R. (1967). Prolegomena to Library Classification. *The Five Laws of Library Science*, 1-73. <http://arizona.openrepository.com/arizona/handle/10150/106370>
- Rodríguez-Bárcenas, Gustavo; López-Huertas, María J. (2013). Saaty's analytic hierarchies method for knowledge organization in decision making. *Journal of the American Society for Information Science and Technology*. 64 (7), p 1454-1467.
- San Segundo Manuel, R. (2008). From the invalidity of a general classification theory to a new organization of knowledge for the millennium to come. En: Ohly, H. P.; Netscher, S.; Mitgusch, K. (2008): *Kompatibilität, Medien und Ethik in der Wissensorganisation (Compatibility, media and ethics in knowledge organization). Proceedings of the 10. conference of the German ISKO, Vienna, July 2006*. Ergon: Würzburg, p.12-17.
- Serres, Michel (1991). *Le Tiers-Instruit. Pans*. Editions Francois Bounn. 249 p. (published in English as: *The Troubadour of Knowledge*. University of Michigan Press, 1997)
- Sidhom, Sahbi (2014). e-mail 'Re: CfP: 4th Int. SymISKO Maghreb' sent to the Executive Committee of ISKO, 27th Jan. 2014.
- Sidhom, Sahbi (2014b). *ISKO-Maghreb (Tunisia, Algeria & Morocco) chapter Report May 2014*, sent per e-mail to ISKO Executive Committee 16th May 2014.
- Spinner, Helmut (1994). *Die Wissensordnung: Ein Leitkonzept für die dritte Grundordnung des Informationszeitalters*. Opladen: Leske and Budrich. 276 p.
- Tennis, Joseph T (May 2014). *Canons, Affiliations, and Polytey: The Future of Knowledge Organization and ISKO*. (unpublished presentation for the 13th international ISKO conference, Krakow)
- Tennis, Joseph T. (2014a). *Candidacy Statements: Listen, Deliberate, Act*. <http://www.isko.org/president2014tennis.pdf>
- Tennis, Joseph T. (2008). Epistemology, Theory, and Methodology. *Knowl. Org.* 35 (2/3), 102-112. [http://faculty.washington.edu/jtennis/Publications\\_files/Tennis2008KOEpi.pdf](http://faculty.washington.edu/jtennis/Publications_files/Tennis2008KOEpi.pdf)

---

# HotWheels: catalogação e recuperação da informação em miniaturas de carros colecionáveis

*HotWheels : cataloging and retrieval of information in collectible miniature cars*

---

Rodrigo Eduardo dos Santos (1), Zaira Regina Zafalon (2)

(1) Universidade Federal de São Carlos, Rod. Washington Luís, km 235, São Carlos/SP, Monjolinho, CEP 13565-905. Email: rodrigoedu02@hotmail.com.

(2) Email: zaira@ufscar.br.

## Resumo

A pesquisa centra-se em miniaturas de carros colecionáveis HotWheels tendo em vista sua catalogação, de modo a serem validados os metadados a serem utilizados no *HotWheels Collection*, catálogo online de acesso público desenvolvido com o objetivo de troca de miniaturas entre colecionadores. Buscou-se, como objetivo geral, validar os metadados para miniaturas de carros colecionáveis HotWheels de modo a garantir a recuperação da informação por colecionadores. Com abordagem qualitativa, adotou-se o método exploratório-descritivo e a pesquisa bibliográfica e documental. Analisou-se a teoria e os instrumentos da organização e representação da informação. A coleta de dados concentrou-se nas miniaturas de carros colecionáveis HotWheels e em suas embalagens, em pôsteres da fabricante e em informações disponíveis em site de colecionadores. Procedeu-se a leitura técnica documental, sob a perspectiva do fabricante dos modelos, do colecionador e do bibliotecário, elaborou-se metadados a partir dos elementos identificados, e recorreu-se à análise comparativa dos metadados indicados para a validação. Reafirmou-se, nos resultados, a necessidade de se conhecer as demandas informacionais do público a que se destina a descrição e, mais ainda, o conhecimento técnico especializado do documento e dos instrumentos de representação.

**Palabras clave:** Catalogação. Recuperação da Informação. Carros HotWheels. Catálogos automatizados.

## Abstract

The research focuses on thumbnail HotWheels collectible cars in view of its catalog in order to be validated metadata to be used in *HotWheels Collection*, public access catalog developed with the aim of exchanging miniatures among collectors. It sought, as a general goal, validate metadata to the thumbnail HotWheels collectible cars in order to ensure the retrieval of information by collectors. With a qualitative approach, adopted the exploratory and descriptive method and the bibliographical and documentary research. It analyzed the theory and the tools of organization and representation of information. Data collection focused on the thumbnail collectible HotWheels cars and their packaging, in manufacturer posters and information available on collectors site. We proceeded to document technical reading, from the perspective of the manufacturer of the models, the collector and librarian, was drawn up metadata from the elements identified, and appealed to the comparative analysis of the metadata indicated for validation. Was reaffirmed in the results, the need to meet the informational demands of the public for which it is the description and even more the technical expertise of the document and representation instruments.

**Keywords:** Cataloguing. Information Retrieval. HotWheels cars. Automated catalogs.

## 1. Introdução

Representar carros colecionáveis HotWheels mostrou-se um desafio nesta pesquisa dada a especificidade do objeto e, mais ainda, por considerar o estudo de metadados que atendessem às particularidades da recuperação da informação por colecionadores. A relevância da pesquisa concentra-se em esforços para o desenvolvimento do *HotWheels Collection*, um catálogo online de acesso público.

O ato de colecionar, ou colecionismo, é definido como “o processo de adquirir e possuir coisas de forma ativa, seletiva e apaixonada, que fazem parte de um conjunto de objetos não idênticos e que não são utilizados na sua forma usual” (FARINA; TOLEDO; CORRÊA, 2006). Possuir e colecionar são conceitos maiores que comprar e consumir. Um conjunto de proposições tem sido discutido a respeito de uma teoria sobre o assunto e sua constatação empírica é

explorada em entrevistas com pessoas que se consideram colecionadoras. Pode-se dizer que o colecionador empreende esforços para pesquisar, localizar, negociar, comprar, vender, trocar e guardar um produto sem, necessariamente, consumi-lo.

Quando se inicia uma coleção geralmente não se tem noção do vasto tamanho que ela pode alcançar. Isto porque, em coleções de miniaturas de carro HotWheels, existe uma infinidade de modelos que são lançados todos os anos, o que deixa os colecionadores mais ávidos e com grandes expectativas quanto ao lançamento dos mesmos. No entanto, colecionadores iniciantes, com pouco tato para o desenvolvimento de suas coleções, tem certa desvantagem em relação aos colecionadores mais experientes quando da obtenção de modelos raros ou fora de linha.

Apesar de se poder contar com alguns sites de colecionadores, e de blogs que tratam sobre coleções de miniaturas de carros colecionáveis HotWheels, as informações estão dispersas e não seguem um padrão tanto na definição do conjunto de informações que descrevem os objetos quanto na forma de recuperação. Observou-se que não existe um ambiente digital ou catálogo que contemple, de maneira ampla, informações sobre modelos atuais, fora de comercialização ou de séries especiais. Surge, daí, a motivação para o desenvolvimento da pesquisa: a dificuldade em se obter informações sobre os modelos disponíveis. Destaca-se que, por serem itens de coleção, cada item tem uma história particular, daí a importância de representá-los como itens de um determinado grupo, isto é, dos colecionadores de miniaturas de carros colecionáveis HotWheels.

Inserido nesse cenário e analisando-se com mais cuidado o tema percebeu-se que a funcionalidade de um ambiente digital está atrelada às reais necessidades de seus usuários e aos propósitos de seu acervo, o que faz com que se torne necessária a adoção de padrões. Padrões para representação da informação mostram-se vinculados diretamente às formas de recuperação. Nesse cenário, apresenta-se a questão de pesquisa: quais os metadados que descrevem os carros colecionáveis HotWheels para colecionadores?

Como objetivo geral definiu-se validar os metadados para miniaturas de carros colecionáveis HotWheels de modo a garantir a descrição e a recuperação da informação no *HotWheels Collection*. Para o alcance deste objetivo geral estabelecem-se os seguintes objetivos específicos: a) estudar o processo de catalogação, a definição de metadados e sua relação com a

recuperação da informação; b) proceder análise exploratória inicial para identificar possíveis variações quanto a leitura técnica documental; c) definir os metadados que descrevem as miniaturas de carros colecionáveis HotWheels.

Como contribuição acadêmica ressalta-se a importância dos catálogos como meio de comunicação entre objetos informacionais e usuários e a relevância da adoção de padrões de representação e das formas de apresentação dos resultados aos usuários. A contribuição social é destacada diante das singularidades necessárias ao propósito específico dos colecionadores para que possam adquirir novos modelos, trocar exemplares ou agregar maior valor à sua coleção.

## 2. Aproximações teóricas

A catalogação, enquanto processo de descrição dos mais variados recursos informacionais, é responsável por traçar “mecanismos de comunicação e de mediação entre demanda informacional de usuários e os documentos, por meio de catálogos, resultado dos constructos da representação bibliográfica.” (ZAFALON, 2014).

Muitos foram os que contribuíram com as questões teóricas da catalogação e dos catálogos, dentre os quais Zafalon (2014) destaca: [1] Antonio Panizzi, um dos primeiros a compilar um conjunto de normas, em 1841, com o propósito de reunir e diferenciar matérias similares; [2] Charles Ammi Cutter que, em 1876, apresentou proposta voltada para melhorar as condições de busca do usuário, na qual se podia encontrar o livro a partir do autor, título e assunto; [3] Seymour Lubetzky que, em 1960, propôs um conjunto de regras que visavam facilitar a localização de um item e ao mesmo tempo estabelecer relações entre os itens similares; [4] Elaine Sveonius que, em 2000, indica a necessidade de revisão dos princípios estabelecidos para a catalogação por não considerarem, em seu bojo, o atendimento às necessidades dos usuários. Destacam-se também os princípios nos quais a catalogação deve-se pautar, como os Princípios de Paris, de 1961, e a Declaração dos Princípios Internacionais de Catalogação, publicados pela IFLA, em 2009.

Quando se trata de entender o processo de recuperação de informações em ambiente informacional, seja ele em meio digital ou não, é necessário que se utilize uma forma de representação para que estas informações sejam identificadas e reconhecidas por seus usuários. Segundo Pereira e Santos (1998),

O processo de catalogação pode ser identificado como meio de comunicação, um instrumento de li-

gação entre o usuário e o documento, um processo de representação documentária que desde a antiguidade atua como instrumento de acesso a informação e ao documento a que se utiliza dos instrumentos disponíveis, numa ação que interliga a biblioteconomia e as tecnologias disponíveis, possibilitando uma rápida recuperação e disseminação da informação, proporcionando assim condições para a agilização do conhecimento.

O modo como as informações são organizadas e recuperadas relaciona-se com o desenvolvimento de novas tecnologias, nos quais se privilegia o atendimento às demandas informacionais de seu público. Fica evidente a necessidade, sob esta ótica, de utilizar-se de novas formas e métodos que deem conta de, por meio de estruturas computacionais e de métodos de definição de metadados, contribuir para a produção, a organização, a distribuição, o acesso, o armazenamento, a preservação, o uso e o reuso dos recursos informacionais.

Consideram-se patentes as seguintes premissas que envolvem a catalogação: a) no domínio bibliográfico, os registros descritivos são elaborados a partir do estabelecimento de metadados e de estruturas descritivas; b) há esquemas tradicionais e internacionais de descrição e de visualização de registros descritivos, identificados nos padrões de estrutura de metadados descritivos e nos padrões de conteúdo; c) metadados são estabelecidos com o propósito de promover a identificação e a localização de um documento e estabelecem-se como instrumentos de recuperação da informação; e d) metadados promovem a comunicação e a encontrabilidade entre comunidades e objetos documentais. Dada a relação entre itens documentais e demandas informacionais, o esquema de comunicação proposto por Shannon (1948; 1969) pode ser (re)configurado, conforme Figura 1.



Figura 1. Esquema do processo comunicativo entre recursos informacionais e comunidades por meio de metadados

A definição de metadados e a adoção de estruturas descritivas tornam-se essenciais em processos de recuperação da informação, tanto por humanos quanto por máquinas, por veicularem mensagens codificadas sobre os recursos informacionais de modo a atender e a se nutrir das informações demandadas pelas comunidades. Takahashi (2000), define metadados como “qualquer dado usado para auxiliar na identificação, descrição e localização de informações.

Trata-se, em outras palavras, de dados estruturados que descrevem as características de um recurso de informação.” Grácio (2002), por sua vez, identifica metadados como “um conjunto de dados [...] que descreve o conteúdo de um recurso, possibilitando a um usuário ou a um mecanismo de busca acessar e recuperar esse recurso.”

Em suma, compreende-se que metadados são um conjunto de elementos que, utilizados de forma padronizada, descrevem informações de um recurso informacional, independente do suporte no qual está inscrito, e possibilitam a busca, a recuperação e o acesso ao recurso por um público. Assim, torna-se inerente ao processo de catalogação o estabelecimento de metadados, uma vez que, ao definir catalogação, Siqueira (2003) afirma que é “uma forma de representação sucinta e padronizada de um item documentário, que tem como objetivo ampliar as formas de acesso a um documento facilitando tarefas e os processos de localização de documentos e informações”.

Foulonneau e Riley (2008), Senso Rosa Piñero (2003), Rosetto (2003) e Gilliland-Swetland (1999), classificam os metadados de diferentes formas, quer seja pelas funções que exercem, pelas características que possuem ou, ainda, quanto ao propósito, a comunidade que atende ou a área de domínio. Percebe-se, pela heterogeneidade de recursos informacionais, que, para cada grupo, e de acordo com suas particularidades, pode ser necessário recorrer a vários padrões de metadados para que seja possível garantir a uniformização da descrição dos recursos o que facilita a recuperação e o acesso ao documento.

Nesse estudo, no qual o foco são as miniaturas de carros colecionáveis, a melhor forma de se identificar o que é necessário descrever a fim de atender as necessidades dos seus usuários é, primeiramente, identificar os metadados dos quais os colecionadores se utilizam para descrever, identificar e recuperar seus objetos. Posteriormente, dado o intuito de validar um conjunto de metadados a serem utilizados em uma plataforma de acesso público online, identificada como *HotWheels Collection*, deve-se avaliar uma estrutura de metadados, com o objetivo de que os metadados possam ser reutilizados ou convertidos para outros formatos, caso venha a ser necessário.

Considerando-se também os colecionadores como público-alvo da pesquisa e do *HotWheels Collection*, destaca-se o colecionismo. Desde que o homem começou o processo de socialização nas antigas comunidades, tem sido ob-

servado o comportamento de agrupar ou acumular objetos que lhes eram considerados valiosos. Mesmo que tais objetos não tenham um valor monetário considerável, mas um valor sentimental e muitas vezes inestimável e incalculável, a esse conjunto de objetos guardados define-se coleção. Giacomini Filho (2006) cita vários exemplos de coleções que ilustram o poder que os objetos exercem sobre os colecionadores e mostram o tamanho do mercado que gira em torno deles e indica que colecionar pode ser uma forma de consumo intensa e envolvente.

Belk (1995, p. 67, tradução nossa) define colecionismo como “o processo de, ativamente, seletivamente e apaixonadamente, adquirir e possuir coisas removidas de seu uso ordinário e percebidas como parte de um conjunto não idêntico de objetos e experiências”. Assim, compreende-se que o comportamento do colecionador é diferente do comportamento de uma pessoa que adquire e consome um produto. Farina, Corrêa e Toledo (2006) enfatizam, com exemplos, a diferença entre coleção e consumo de itens, no qual também encaixam-se os colecionadores de miniaturas de carros HotWheels, ao afirmarem que

O ato de colecionar difere do conceito de consumir pelo fato de indicar a formação de um conjunto de coisas, que é a coleção, e o comportamento decorrente do processo de formação desse conjunto. Os objetos que fazem parte do conjunto não são utilizados para o consumo propriamente dito, como, por exemplo, uma coleção de moedas, as quais não são empregadas para compra, ou uma coleção de selos, que não são utilizados para selar cartas para postagem. Pode-se dizer que o consumo de uma coleção é o próprio processo de formação da mesma. O processo implica alto grau de envolvimento e de tempo do colecionador.

As coleções, de um modo geral, são formadas e constituídas em determinados contextos históricos, sociais e, até mesmo, pessoais. Dado que podem ou não ser divulgadas e utilizadas com os mais diversos fins, o centro das coleções está no fato de representar os ideais dos colecionadores, e situa-se entre os limites que estes querem lhes atribuir. Nesse sentido, além de se privilegiar a investigação em torno do ato de colecionar, também se debruça sobre os objetos em si, já que são estes a razão das coleções. O deleite do colecionador reside tanto na singularidade que cada objeto possui quanto na possibilidade de se ter um conjunto de objetos semelhantes.

Para a obtenção de mais um item para a coleção, a busca pelo objeto desejado envolve, também, ações de pesquisa, categorização,

reunião, ordenação e classificação desses objetos. São intrínsecas à coleção, e particulares, ações de representação e organização do acervo.

Conseguir uma determinada miniatura pode ser um verdadeiro desafio, a depender de sua raridade, pois muitos colecionadores irão disputá-la. Certamente, diante do fato de colecionadores não medirem esforços para terem determinado item e exibi-lo, ou contemplá-lo, faz com que se ganhe visibilidade e notoriedade na comunidade em que se insere, uma vez que os objetos colecionados são utilizados como forma de expressão e projeção do colecionador e de sua visão de mundo.

Pearce (2003) define três modos de colecionar: o *coleccionismo de lembranças*, ligado ao desejo autobiográfico do colecionador, construindo suas memórias, apoiando-se nos objetos que selecionou; o *coleccionismo de objetos de fetiche*, no qual o objeto controla o indivíduo e torna-se dominante, obrigando o colecionador, obsessivo, a recolher o máximo de objetos possíveis; e o *coleccionismo sistemático*, caracterizado por objetivos racionais, de construção de conjuntos completos, demonstrando o desejo de se atingir uma coleção fechada.

Independente de qual modo de colecionar se utilize, é fato que tempo, energia e dinheiro são investidos em uma coleção. Assim, é natural que a coleção seja importante na constituição do “eu” do colecionador.

Belk et al. (1988) propuseram algumas características dos colecionadores e suas coleções: começam suas coleções despropositadamente; vício e compulsão fazem parte do colecionismo; a coleção é desenvolvida a partir de critérios afetivos ou cognitivos; o item, quando incorporado à coleção, sai do profano e eleva-se ao sagrado; as coleções servem como extensão de si mesmo; a coleção tende à especialização, uma vez que o colecionador estuda a sua própria coleção e os objetos que virão a compô-la; a distribuição pós-morte da coleção configura-se como problema tanto para o colecionador quanto para sua família; sentimento de desejo e medo de conclusão da coleção. Este tipo de comportamento evidencia que colecionadores procuram por itens muito específicos dentro do universo a que pertencem.

No âmbito dos carros colecionáveis HotWheels destacam-se coleções voltadas a miniaturas de determinada marca de carro, como Ferrari, Ford, Volkswagen, ou mesmo de modelos específicos de uma marca, como a Kombi, da Volkswagen. Retoma-se, então, a relevância da pesquisa na qual se situa o fato de que a cata-

locação dos mesmos pode tanto contribuir para a identificação e a obtenção do item, quanto para a definição da coleção.

Marshall (2005, p. 18) destaca que, para a definição da coleção, utilizam-se critérios:

O colecionista define um critério. Fui exigir do gerente de minha conta uma caneta do banco, que eu cobiçava, muito bonita. Ele prontamente me atendeu e confessou ser ele próprio um colecionador de canetas – mas somente canetas promocionais, com alguma marca corporativa. Já o eng. Marcos Abreu coleciona relógios Omega e canetas da marca Parker 51 (fabricadas em 1947, ano da comemoração dos 51 anos desta marca). Sua coleção de canetas está quase completa, pois ele possui todas, exceto o modelo Empire State e uma variante vermelha, fabricada na Inglaterra, e não são mais produzidas ou comercializadas. Colecionadores de caneta pagam facilmente dezenas de milhares de dólares por exemplares preciosos, em um dos grandes mercados mundiais de colecionáveis, sempre presentes em leilões.

Analisando esse quesito com maior profundidade pode-se estabelecer que, quando um colecionador ou um determinado grupo seleciona critérios para os itens que desejam possuir, estes estão mostrando os pontos principais para localizá-los, ou seja, estão definindo os pontos de acesso. José Mindlin (1997, p. 24 et seq.), bibliófilo notável cuja coleção foi avaliada em 100 milhões de reais e conta com cerca de 17 mil títulos e 40 mil volumes, externa a sua co-moção ao se deparar com um item desejado para a sua coleção, no seu caso um livro raro:

Além do conteúdo, edição, encadernação, diagramação, tipografia, ilustração, ou papel, o livro exerce sobre mim uma atração física. Não me satisfaz ver um livro numa vitrine, sem poder pegá-lo. Minha tese é que a gente deve poder tocar naquilo que gosta, sentir objetos e pessoas. [...] Quando, depois de anos e anos de procura, encontra-se um livro raro, o coração bate mais forte. Sente-se uma emoção grande, mas não se pode deixar que ela transpareça diante do livreiro.

Esse trecho demonstra claramente a paixão do colecionador com o seu objeto de desejo. Assim também é com o colecionador de miniaturas de carros colecionáveis HotWheels ao encontrar numa prateleira um item desejado.

Estava outro dia passeando pelas lojas de um shopping e resolvi visitar uma loja de brinquedos em busca de uma nova miniatura para incrementar a minha coleção de carros HotWheels. Sempre fico alerta aos itens que formam uma de minhas subcoleções: a de carros usados pelo personagem Batman, das revistas em quadrinhos publicadas pela DC Comics. Neste dia, como habitualmente faço, sentei-me no chão da loja e comecei a vasculhar as prateleiras. Não tenho como narrar qual foi a minha alegria ao me deparar com um modelo há

muito tempo desejado. Confesso que nem me importei com o preço tamanha a alegria de encontrá-lo. Um momento de pura felicidade; foi como encontrar um tesouro (SANTOS, 2013).

Outro comportamento observado em bibliófilos é referente ao zelo e à conservação de suas coleções. Assim também são os colecionadores de miniaturas de carros HotWheels, os quais investem em estojos, afixados em paredes, com vários nichos para alocação de carrinhos com a finalidade de protegê-los da ação do tempo, preservando-os. Outros, ainda, chegam a nem tirar suas miniaturas de suas embalagens originais com receio de que danifiquem, o que confere, inclusive, maior valor de mercado em negociações entre colecionadores.

Tendo sido traçados aspectos sobre a representação de recursos informacionais, os quais incidem diretamente na recuperação e obtenção de informações, e um panorama geral sobre coleções, colecionismo e comportamento dos colecionadores, destaca-se que se entende como imprescindível no processo de catalogação o conhecimento do objeto a ser descrito, do público a que se destina a descrição e da linguagem que, por meios de regras e normas, é utilizada na descrição documental. Estes são os requisitos para a compreensão dos atributos das miniaturas de carros colecionáveis HotWheels relevantes na recuperação de informação pelos colecionadores.

### 3. Procedimentos metodológicos

A pesquisa, com abordagem qualitativa, preocupa-se com aspectos da realidade não quantificáveis, com foco no entendimento e explicação das dinâmicas sociais e no “trabalho com o universo de significados, motivos, aspirações, crenças, valores e atitudes, o que corresponde a um espaço mais profundo das relações, dos processos e nos fenômenos que não podem ser reduzidos à operacionalização de variáveis.” (MINAYO, 2001). A pesquisa é de caráter bibliográfico uma vez que o arcabouço teórico se pauta em “[...] informações, dados primordiais, conhecimentos prévios [...]” (BEDIN, 2007). Utiliza-se, também, da pesquisa documental pois “[...] vale-se de materiais que não recebem ainda um tratamento analítico, ou que ainda podem ser reelaborados de acordo com os objetos da pesquisa.” (GIL, 2008). Esse é o enfoque dado na análise dos pôsteres fornecidos pela Mattel e nas próprias embalagens dos carros colecionáveis Hot Wheels. O procedimento adotado é tanto de cunho exploratório quanto descritivo uma vez que mescla a “[...] abordagem do fenômeno pelo levantamento de informações que poderão levar o pesquisador a conhecer

mais a seu respeito”, característica da pesquisa exploratória, e a pesquisa “realizada com o intuito de descrever as características do fenômeno”, o que tipifica a pesquisa descritiva. (GERHARDT, et al, 2009).

Este tópico destina-se a contextualizar o universo das miniaturas de carros colecionáveis HotWheels e à identificação, por meio de uma análise exploratória, das variações de leitura técnica documental dos modelos (pelo fabricante, pelo bibliotecário e pelo colecionador).

### 3.1. Quanto ao objeto

HotWheels é uma marca de miniaturas de carros, em escala 1:64, da Mattel, fabricante de brinquedos americana. É uma das fabricantes mais famosas de automóveis de brinquedo e foi a principal concorrente da Matchbox até 1997, quando a Mattel comprou a Tyco Toys, proprietária da Matchbox, até então.

A Mattel possui licença para fazer modelos em escala de várias marcas de carros, o que permite o uso do design das plantas originais desses. Embora as miniaturas de carros colecionáveis HotWheels tenham sido originalmente destinadas para as crianças, eles se popularizaram por conta dos colecionadores adultos.

Em sua primeira coleção, em 1968, foram feitos 16 modelos. O primeiro modelo produzido foi a réplica de um Camaro, pintado em azul escuro. Dentre esses primeiros, 10 foram baseados em versões customizadas de automóveis da época e 6 baseados em carros reais projetados e construídos para as corridas de pista.

A pintura metálica Spectraflame também marcou o contraste com os modelos pintados em esmalte dos carros da Matchbox. Os acabamentos atraentes foram alcançados dado o polimento do metal para, em seguida, serem revestidos em uma laca de cor clara, que contou com cores exóticas como o magenta e o rosa. O rosa, não muito usado em miniaturas HotWheels, fez com que fossem valorizadas as miniaturas com essa cor.

Em 1995 a Mattel trouxe uma grande mudança para a linha HotWheels e dividiu os modelos em série. Em 1996 teve início a comercialização no Brasil. Em 1998, como marco para os 30 anos da marca, a empresa replicou vários modelos produzidos ao longo de sua história.

Uma das séries que ganhou destaque ao longo dos anos é a Treasure Hunts, lançada em 1995. Também identificada com T-Hunt, esta é uma série composta de 12 carros produzidos a cada ano. A partir de 2011 passou a ser composta

por 15 modelos. Veículos Treasure Hunts são identificáveis por um rótulo na embalagem (figura 2). Os exemplares são decorados e contam com rodas especiais de borracha.



Figura 2. Exemplar da série T-Hunts

Em 2001, a Mattel lançou o portal HotWheels-Collectors.com, um fórum online para unir colecionadores, oferecendo informações sobre miniaturas de edição limitada, próximos lançamentos e eventos, bem como bate-papo e fóruns de comércio.

Existem centenas de páginas web dedicadas aos colecionadores de miniaturas de carros colecionáveis HotWheels. Para a maior parte, este é um hobby relativamente barato, quando comparado com o de colecionar moedas, selos ou bonecas Barbies. Os carros da linha principal custam, atualmente R\$ 7,99 cada (cerca de € 2,50). Depois que os carros não estão mais disponíveis no varejo, o custo pode variar significativamente. Um Volkswagen Beach Bomb, de 1969, item da coleção de Mike Stephenson e Mark Meredith, tem valor estimado em US\$ 125.000,00 (cerca de R\$400.000,00 e €112.000,00). Um exemplar pode ser visualizado na figura 3.



Figura 3. Exemplar do Volkswagen Beach Bomb

Cabe uma digressão sobre a data registrada na base das miniaturas, uma vez que se refere à data de direitos autorais de design e não à data de fabricação. A data de direitos autorais é, geralmente, registrada no ano anterior à produção do carro. A figura 4 indica, como exemplo, as datas da miniatura Baja Beetle: 2010, como data de lançamento, 2009, como data de copyright, e 1983, como data na miniatura.



Figura 4. Exemplos do Baja Beetle

A figura 5, entretanto, indica, para a miniatura Homer, o ano de 2015, como data de lançamento, e 2013 como de copyright e na miniatura.



Figura 5. Exemplos do Homer

### 3.2. Quanto à coleta de dados

A coleta de dados das miniaturas de carros colecionáveis HotWheels ocorreu em pôsteres publicados pela Mattel e enviados pelo Serviço de Atendimento ao Consumidor da empresa; no site HotWheels Br, de um colecionador do Bra-

sil; e nas próprias miniaturas de carros colecionáveis HotWheels e suas embalagens.

Os dados coletados foram analisados sob a ótica da teoria e dos instrumentos da Catalogação, para o qual foi requerida a leitura técnica documental. Sob este horizonte analisou-se os elementos identificados pela Mattel, pelo colecionador e pelo bibliotecário. Considera-se relevante abordar a perspectiva da leitura técnica uma vez que a mesma exerce influência direta na representação documental e, conseqüentemente, na recuperação da informação.

Para a coleta de dados junto ao fabricante identificou-se elementos presentes nos pôsteres. Foram analisados pôsteres dos anos de 2010, 2011, 2012 e 2013 bem como aquelas obtidas via contato com o Serviço de Atendimento ao Consumidor (SAC) da Mattel. Para a definição dos elementos reconhecidos pelo colecionador utilizou-se informações do site HotWheels Br que indicam tanto a identificação quanto as características do modelo. Para a identificação dos elementos por um bibliotecário foram enviadas imagens de 5 embalagens, frente e verso, e solicitou-se que ele procedesse sua avaliação com bases em seus conhecimentos.

a) elementos identificados pela Mattel

A Mattel elabora, anualmente, pôsteres com os modelos da linha principal. Observa-se nos pôsteres uma divisão por séries, cada qual identificada por uma cor. Nos pôsteres há também a reprodução de uma embalagem com a identificação dos elementos (figura 6).



Figura 6. Exemplo de identificação de elementos de descrição no pôster de 2012 da Mattel

Com base nas informações contidas nos pôsteres pode-se afirmar que na parte frontal da embalagem temos: o nome da série, o nome do carrinho, o número do carrinho na série e a cor da série. Observou-se que a cor da série pode mudar a cada ano. Não foi possível identificar

nos pôsteres avaliados informações sobre os dados que constam no verso das embalagens. Desse modo, optou-se por contatar o Serviço de Atendimento ao Cliente (SAC) da Mattel. Em resposta, foram identificados o número de fabricação do modelo, a sub-série do produto e o QR Code.

b) elementos identificados pelo site HotWheels Br

O site HotWheels Br disponibiliza reproduções de embalagens e a identificação dos elementos a partir do ponto de vista de colecionadores. Em relação aos elementos identificados no item anterior pode-se afirmar que os colecionadores disponibilizam mais informações em comparação ao fabricante, o que pode ser observado na figura 7:

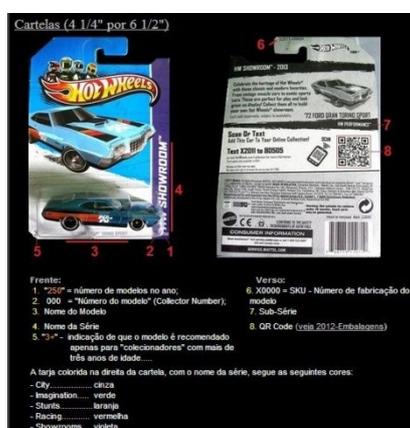


Figura 7. Exemplo de identificação de elementos de descrição de embalagem de 2013 no site HotWheels Br

Com base nas informações disponíveis no site HotWheels Br pode-se afirmar que na parte frontal da embalagem temos: número de modelos no ano, número do modelo, nome do modelo, nome da série e recomendação de faixa etária. No verso estão indicados o número de fabricação do modelo, o nome da sub-série e o QR Code.

c) elementos identificados pelo bibliotecário

A leitura técnica documental do bibliotecário baseou-se na reprodução de 5 embalagens (frente e verso).



Figura 8. Exemplo de identificação de elementos de descrição de embalagem pelo bibliotecário

O profissional identificou, conforme a figura 8, na parte frontal, elementos que descrevem o título, a série e a descrição física, e, no verso da embalagem, ano, local e autor.

#### 4. Análise dos resultados

Este tópico destina-se à análise comparativa dos elementos identificados na leitura técnica documental tendo em vista a validação dos metadados para miniaturas de carros colecionáveis HotWheels de modo a garantir tanto a descrição quanto a recuperação da informação no *HotWheels Collection*.

Realizou-se comparações entre os elementos identificados uma vez que a leitura técnica documental é desenvolvida com o objetivo de extrair os elementos necessários para garantir a representação e a recuperação das informações futuramente requisitadas.

Analisando-se os elementos identificados pelo fabricante, pelo colecionador e pelo bibliotecário identificaram-se diferenças que podem comprometer a representação e a recuperação de informações dos carros colecionáveis HotWheels por colecionadores.

Primeiramente verifica-se diferença na nomenclatura usada para designar a miniatura: modelo, carrinho, veículo e carrinho de brinquedo.

Quanto à série, a Mattel identificou também como segmento.

Embora a cor da série seja destacada com frequência pela Mattel, ela só comparece para o colecionador posteriormente e não é identificada pelo bibliotecário.

O número da série foi identificado tanto pelo fabricante quanto pelo colecionador, entretanto este destaca a quantidade de modelos dentro da série. O nome da série é indicado por ambos, porém, o colecionador também indica a sub-série.

O colecionador também apontou a quantidade de modelos no ano, o número do modelo no ano, a indicação do ano do modelo, a recomendação da faixa etária, o QR Code e a indicação do número de fabricação, elementos que não foram apontados de forma explícita pela fabricante.

Observou-se que a fabricante destacou as cores dos nomes das séries na sua leitura e que o colecionador praticamente não a menciona.

Quanto à leitura técnica do bibliotecário, entende-se que não tenha sido possível proceder uma análise mais acurada dos elementos presentes nas embalagens e que, tampouco, tenha tido condições favoráveis para estabelecer aproximações e significado entre os mesmos. Denota-se que isso possa ser decorrente do fato de não ter tanta familiaridade com miniaturas de carros HotWheels e por não estar habituado a descrever objetos de tal natureza. Percebe-se, porém, que mesmo sem o conhecimento especializado do colecionador, tenha sido possível realizar a leitura do item. Não foram percebidos, entretanto, elementos que denotem as necessidades do público a que se destina a descrição e as informações relevantes para a identificação unívoca do item na coleção, ambos necessários à recuperação da informação com maior precisão e menor revocação. Outra questão que pode ter influenciado a identificação dos elementos pode estar no fato de que, no Brasil, a formação do bibliotecário se dá na graduação, ao passo que, em outros países, seja configurada como sendo de especialização de área específica.

Pondera-se, também, que as embalagens sofrem alterações ao longo do tempo, sendo constantes alguns elementos, tais como o nome do modelo, nome da série e o número de fabricação.

Reconhece-se a incompatibilidade entre os nomes das séries e as cores adotadas para identificá-las ao longo dos anos, o que não nos permite inferir a relevância da cor das séries como

um possível metadado a ser considerado para a descrição e recuperação das miniaturas de carros colecionáveis HotWheels.

## 5. Proposta de metadados para o *HotWheels Collection*

Optou-se, mediante a natureza dos objetos pesquisados e a relevância da adoção de padrões de metadados, por consultar os metadados estabelecidos para a descrição de artefatos tridimensionais presentes no Código de Catalogação Anglo-Americano, 2ª edição, com revisão de 2002 (AACR2r), em específico os capítulos 1, Regras Gerais de Descrição, e 10, Artefatos Tridimensionais e Realia.

De início, justifica-se a escolha do capítulo destinado à descrição de artefatos tridimensionais, dado que:

10.0A1. As regras deste capítulo dizem respeito à descrição de objetos tridimensionais de todos os tipos (menos os abrangidos em capítulos anteriores) como modelos, dioramas, jogos (incluindo quebra-cabeças e simulações), cassetes de braille, esculturas e outras obras de arte tridimensionais, objetos de exposições, máquinas e vestuário. (CÓDIGO..., 2004, p. 10-2).

Com base nesta regra entende-se que a descrição das miniaturas de carros colecionáveis HotWheels encaixa-se neste capítulo por tratar-se de modelos, até mesmo como objetos de exposição, não sendo contemplados em nenhum outro capítulo da AACR2.

Com o intuito de compreender os metadados estabelecidos na AACR2r optou-se pela consulta aos níveis de descrição, indicados na regra 1.0D2 (CÓDIGO..., 2004, p. 1-5). Este nível elenca metadados relacionados:

- ao título: principal, equivalente e outras informações
- à responsabilidade: principal e subsequentes, bem como aquelas relativas à edição e à série
- à designação geral do material: *objeto* ou *brinquedo* ou *modelo*
- à edição
- aos detalhes específicos do material (não utilizada para a descrição de artefatos tridimensionais)
- à publicação: lugar, editor e data
- ao item: extensão, detalhes físicos e dimensões
- à série e à subsérie: título, numeração e ISSN
- às notas
- ao número normalizado

Dado que o processo de representação documental ocorre a partir da descrição de elementos do objeto, destaca-se a orientação presente na regra 10.0B1, a qual indica que a “fonte principal de informação [...] é o próprio objeto, juntamente com qualquer material textual e o contêiner, publicados pelo editor e pelo fabricante do item” (CÓDIGO..., 2004, p. 10-3). Compreende-se, portanto, que para a descrição das miniaturas recorra-se à consulta ao próprio objeto bem como à sua embalagem, uma vez que consultar um ou outro não forneça condições necessárias para uma descrição satisfatória.

Com base nas informações obtidas tanto na fase de coleta de dados quanto naquelas identificadas no Código... (2004), estabelece-se a seguinte discussão sobre os metadados para o *HotWheels Collection*:

- a) título: atribui-se, para o *título principal*, *HotWheels*, dado que é o título impresso em todos as miniaturas e embalagens; para *outras informações sobre o título*, o nome do modelo, como, por exemplo, *Camaro*;
- b) responsabilidade: tendo em vista que a *Mattel* é a responsável pelos objetos, sugere-se adotar o próprio nome para a indicação de responsabilidade principal;
- c) designação geral do material (DGM): embora seja de uso opcional, conforme indicação em 10.1C, decidiu-se usá-la uma vez que se trata de objeto tridimensional. No que se refere ao termo a ser utilizado, e com base nas listas disponíveis em 10.1C1, a qual indica os termos *objeto*, *brinquedo* ou *modelo*, decidiu-se por usar a descrição *modelo* haja vista o fato de remeter à mesma terminologia empregada por colecionadores e pela fabricante, como visto na leitura técnica das embalagens;
- d) edição: não será incluída na proposta uma vez que não houve indícios de informação representativa de edição nos objetos e nas análises;
- e) detalhes específicos do material: não será utilizada por não ser requerida, conforme indicação da regra 10.3A;
- f) publicação: apesar de não constar, nas embalagens, explicitamente indicações do publicador, é possível recorrer à regra 10.4G1, a qual estabelece que “se o nome do editor for desconhecido ou se não houver editor [...] e o lugar e o nome do fabricante forem conhecidos, registre o nome e o lugar [...]” (CÓDIGO..., 2004, p. 10-8). Quanto à data, estabeleceu-se o uso da data indicada na embalagem, concernente ao ano em que o modelo foi lançado no mercado. A data de copyright, indicada no modelo, não será usada, uma vez que, por referir-se à aquisição dos direitos autorais, pode ser usada em outros modelos replicados e relançados;
- g) item: em *extensão* deve ser indicada a quantidade de unidades físicas do artefato seguido do termo que melhor o designa. A regra 10.5B1 indica uma lista de termos possíveis. Porém, apesar de *modelo* parecer apropriado, a mesma regra indica a possibilidade de registrar um nome específico. Assim, optou-se por utilizar o termo *miniatura de carro*, por entender que *modelo* já está em uso em *designação geral do material*. Para a descrição de detalhes físicos recorre-se às regras 10.5C1, que especifica o material de que é feito o objeto, e 10.5C2, para cor. Esta última regra indica o uso do nome da cor, para o caso de uma ou duas cores, *color.*, para objetos multicoloridos, e *p&b*, para preto e branco. Dado o fato de que a cor é um elemento fundamental para a recuperação das informações por colecionadores, o que requer a especificidade na descrição unívoca do objeto optou-se pela indicação de até cinco cores, ao invés de utilizar-se *color*. Para a indicação das dimensões consultou-se as regras 10.5D1 e 10.5D2 para dimensões do objeto e da embalagem;
- h) série e subsérie: para o título da série recorre-se à regra 10.6B1, que indica a consulta às instruções presentes em 1.6, em específico, 1.6B1, para o título principal, 1.6G1, para a numeração, e 1.6H1, para a subsérie;
- i) notas: a indicação de nota é opcional, dado que depende da natureza do item descrito e da finalidade da descrição. Sua redação deve ser concisa e clara, tal qual indicado na orientação 0.26 (CÓDIGO..., 2004). Considerando-se a especificidade do objeto a ser descrito concebe-se o uso de regras referentes à publicação (10.7B9), para a indicação da data de copyright, indicada no modelo; ao item (10.7B10), para a indicação da escala das miniaturas, uma vez que a *Mattel* fabrica outras miniaturas com a logomarca *HotWheels* em diferentes escalas; à série (10.7B12); ao resumo (10.7B17), presente em algumas embalagens; e aos números (10.7B19), para a indicação do número do modelo no ano do lançamento, o número de identificação do modelo e o número de fabricação;
- j) número normalizado: remete à indicação do ISBN e do ISSN. A regra 10.8B1 indica que outros números devem ser indicados em nota, tal qual já foi indicado no item anterior.

Com base nestas discussões realizou-se a validação dos metadados, para o que escolheu-se

20 modelos de miniaturas que apresentavam situações críticas de sucesso da proposta. O estudo sistematizado de três casos, selecionados tendo em vista a especificidade, podem ser observados nas figuras 9 a 11:

a) caso 1: Custom V-8 Vega

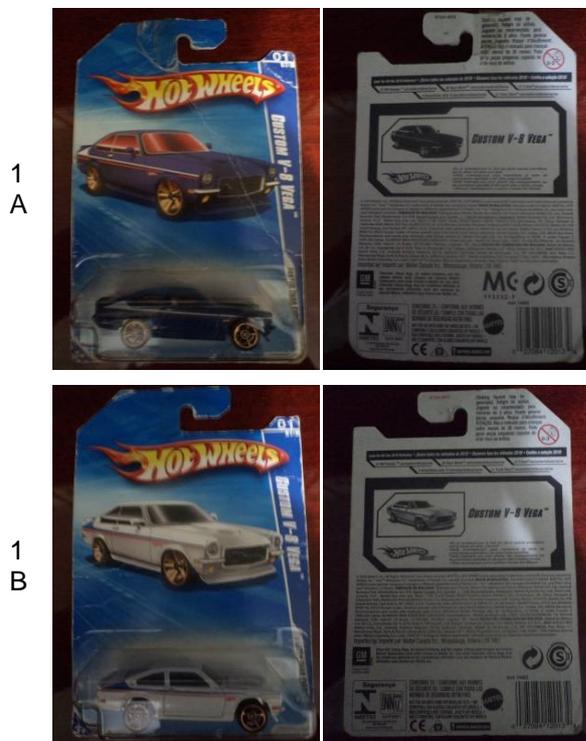


Figura 9. Custom V-8 Vega

O modelo 1A apresenta os seguintes metadados (adotou-se, para demarcar os elementos, a pontuação prescrita na ISBD):

**Título:** HotWheels : Custom V-8 Vega  
**Resp.:** Mattel  
**DGM:** Modelo  
**Publi.:** Malasia : Mattel, c2009  
**Item:** 1 miniatura de carro : ferro, azul e vermelho ; 2 x 7 x 3 cm em embalagem de 17 x 11 x 4 cm  
**Série:** Faster Than Ever ; 1  
**Nota:** Escala 1:64  
**Nota:** Série da Coleção 2010  
**Nota:** Número do modelo no ano de 2010: 127/214  
**Nota:** Número de fabricação do modelo: R7554-B816

Tabela 1. Descrição do modelo 1A

O modelo 1B apresenta os seguintes metadados (adotou-se, para demarcar os elementos, a pontuação prescrita na ISBD):

**Título:** HotWheels : Custom V-8 Vega

**Resp.:** Mattel  
**DGM:** Modelo  
**Publi.:** Malasia : Mattel, c2009  
**Item:** 1 miniatura de carro : ferro, branco e preto ; 2 x 7 x 3 cm em embalagem de 17 x 11 x 4 cm  
**Série:** Faster Than Ever ; 1  
**Nota:** Escala 1:64  
**Nota:** Série da Coleção 2010  
**Nota:** Número do modelo no ano de 2010: 127/214  
**Nota:** Número de fabricação do modelo: R7554-B815

Tabela 2. Descrição do modelo 1B

Observa-se, na descrição de ambos os casos, a presença de metadados idênticos, diferenciando-os pela cor e pelo número de fabricação, evidenciados na descrição.

b) caso 2: '65 Pontiac Boneville

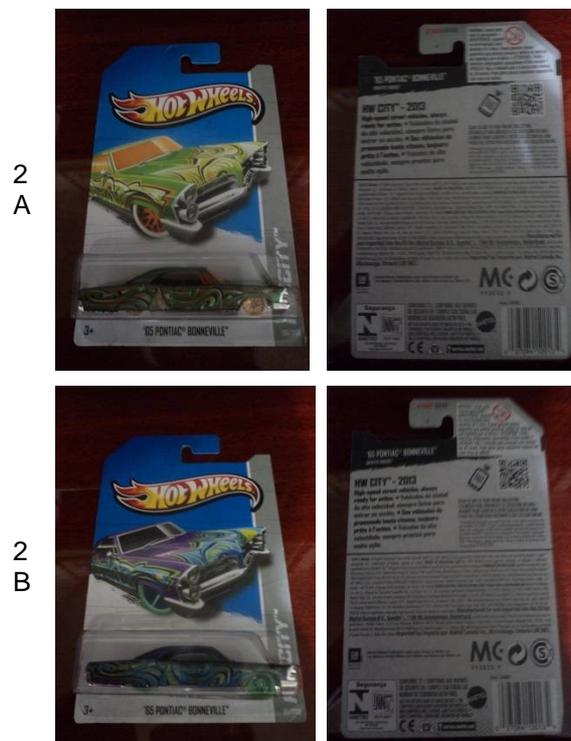


Figura 10. '65 Pontiac Boneville

O modelo 2A apresenta os seguintes metadados (adotou-se, para demarcar os elementos, a pontuação prescrita na ISBD):

**Título:** HotWheels : '65 Pontiac Boneville  
**Resp.:** Mattel  
**DGM:** Modelo  
**Publi.:** Malasia : Mattel, c2012  
**Item:** 1 miniatura de carro : ferro, verde e laranja ; 2 x 9 x 3 cm em embalagem de 17 x 11 x 4 cm  
**Série:** HW City. Graffiti Riders

- Nota: Data de copyright na miniatura c2002  
 Nota: Escala 1:64  
 Nota: Série da Coleção 2013  
 Nota: HW City: veículos de alta velocidade, sempre prontos pra muita ação  
 Nota: Número do modelo no ano de 2013: 35/250  
 Nota: Número de fabricação do modelo: X1881-07A3

Tabela 3. Descrição do modelo 2A

O modelo 2B apresenta os seguintes metadados (adotou-se, para demarcar os elementos, a pontuação prescrita na ISBD):

- Título:* HotWheels : '65 Pontiac Boneville  
*Resp.:* Mattel  
*DGM:* Modelo  
*Publi.:* Malasia : Mattel, c2012  
*Item:* 1 miniatura de carro : ferro, roxo e verde ; 2 x 9 x 3 cm em embalagem de 17 x 11 x 4 cm  
*Série:* HW City. Graffiti Riders  
*Nota:* Data de copyright na miniatura c2002  
*Nota:* Escala 1:64  
*Nota:* Série da Coleção 2013  
*Nota:* HW City: veículos de alta velocidade, sempre prontos pra muita ação  
*Nota:* Número do modelo no ano de 2013: 35/250  
*Nota:* Número de fabricação do modelo: X1687-07A3

Tabela 4. Descrição do modelo 2B

Observa-se, tal qual destacado na descrição anterior, que, em ambos os casos, estão presentes metadados idênticos, diferenciando-se aqueles que indicam a cor e o número de fabricação, evidenciados na descrição.

c) caso 3: Volkswagen Kool Kombi

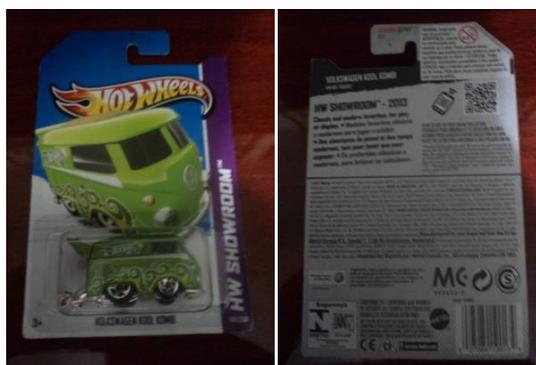
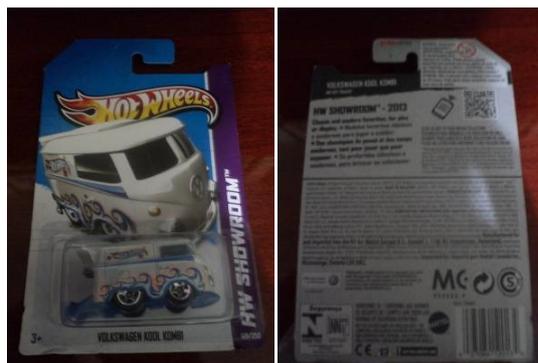
3  
A3  
B

Figura 11. Volkswagen Kool Kombi

O modelo 3A apresenta os seguintes metadados (adotou-se, para demarcar os elementos, a pontuação prescrita na ISBD):

- Título:* HotWheels : Volkswagen Kool Kombi  
*Resp.:* Mattel  
*DGM:* Modelo  
*Publi.:* Malasia : Mattel, c2012  
*Item:* 1 miniatura de carro : ferro, verde ; 3 x 6 x 3 cm em embalagem de 17 x 11 x 4 cm  
*Série:* HW Showroom. HW Hot Trucks  
*Nota:* Data de copyright na miniatura c2012  
*Nota:* Escala 1:64  
*Nota:* Série da Coleção 2013  
*Nota:* HW Showroom: Os preferidos clássicos e modernos, para brincar ou colecionar  
*Nota:* Número do modelo no ano de 2013: 169/250  
*Nota:* Número de fabricação do modelo: X1636-07A3

Tabela 5. Descrição do modelo 3A

O modelo 3B apresenta os seguintes metadados (adotou-se, para demarcar os elementos, a pontuação prescrita na ISBD):

- Título:* HotWheels : Volkswagen Kool Kombi  
*Resp.:* Mattel  
*DGM:* Modelo  
*Publi.:* Malasia : Mattel, c2012  
*Item:* 1 miniatura de carro : ferro, branco e azul ; 3 x 6 x 3 cm em embalagem de 17 x 11 x 4 cm  
*Série:* HW Showroom. HW Hot Trucks  
*Nota:* Data de copyright na miniatura c2012  
*Nota:* Escala 1:64  
*Nota:* Série da Coleção 2013  
*Nota:* HW Showroom: Os preferidos clássicos e modernos, para brincar ou colecionar  
*Nota:* Número do modelo no ano de 2013: 169/250

Nota: Número de fabricação do modelo:  
X1993-07A3

Tabela 6. Descrição do modelo 3B

Observa-se, tal qual destacado na descrição anterior, que, em ambos os casos, estão presentes metadados idênticos, diferenciando-se aqueles que indicam a cor e o número de fabricação, evidenciados na descrição.

A partir da comparação dos resultados obtidos com os metadados definidos para a descrição de miniaturas de carros colecionáveis HotWheels, observa-se que houve validação e pertinência dos mesmos.

Considerou-se, também, para a validação dos metadados indicados, as necessidades de busca dos usuários, marcadas pelo atendimento a questões comuns no universo de colecionadores de miniaturas de carros HotWheels. De um modo geral, um colecionador que está em busca por um determinado modelo de miniatura colecionável HotWheels, o Camaro, por exemplo, realiza perguntas para chegar até o modelo requerido:

- por qual nome o modelo é identificado? Afinal, existem vários modelos de Camaro, e quanto mais específico o modelo, mais rápida é a busca e maior a precisão na recuperação;
- em que ano o modelo foi lançado? Porque um modelo pode ter sido lançado em anos diferentes;
- em qual série ou subsérie o modelo foi lançado? Pois o colecionador pode ter interesse em colecionar miniaturas de um modelo específico de uma série, também específica, que pode se repetir com alguma característica diferente conforme muda a coleção, mudança esta que acontece anualmente;
- quais os modelos de Camaro lançados na cor azul? Uma vez que o colecionador pode, por exemplo, colecionar somente miniaturas de uma cor específica;
- quais os modelos Camaro lançados em 2011? O colecionador pode, por exemplo, só colecionar as miniaturas do modelo Camaro em suas diversas variações.

As formas de busca, como se observa, podem ser as mais variadas possíveis, isso porque, cada colecionador possui uma necessidade específica ao realizar uma busca por um determinado modelo.

Nessa perspectiva, os metadados definidos foram novamente validados, ao mesmo tempo em que atende as necessidades dos colecionadores de miniaturas colecionáveis HotWheels

quanto à representação e recuperação da informação sobre as mesmas.

## 6. Considerações finais

Ao longo desta pesquisa analisou-se tanto aspectos inerentes ao objeto, miniaturas de carros colecionáveis HotWheels, quanto a necessidade de que a representação de recursos informacionais assuma significado diante do público que fará uso de tais recursos.

Destacou-se a contribuição dos conceitos da catalogação para a identificação e elaboração de metadados bem como a evidência que exercem no processo de busca e recuperação da informação. Muito mais do que fazer uso de regras, a catalogação estabelece vínculos entre objetos documentais e usuários, quer sejam humanos ou máquinas, que garantem, a descrição, a recuperação e o acesso aos mais variados recursos informacionais. Desse modo, as premissas foram ratificadas.

Recorrer à análise da leitura técnica documental pelo fabricante, pelo colecionador e pelo bibliotecário fortaleceu a perspectiva de que os metadados devem atender às necessidades de seus usuários e reafirmou como imperioso o conhecimento do objeto a ser descrito, do público a que se destina a descrição e da linguagem utilizada. O segredo para a descrição dos objetos está em considerá-la entrelaçada à forma como são buscados.

A identificação de diferentes leituras do mesmo objeto retoma a afirmação de Saussure (2010, p. 15) de que "[...] é o ponto de vista que cria o objeto". Esse reconhecimento ganhou destaque na pesquisa uma vez que permitiu identificar a visão particular (que pode, de certo modo, ser coletivizada) que fabricante (Mattel), consumidor (coleccionadores) e intermediário (bibliotecário) tem sobre um mesmo objeto.

No conjunto de miniaturas de carros colecionáveis HotWheels, escolhidos como amostra para este estudo, deparou-se com três modelos peculiares que chamaram a atenção. Eles possuíam as mesmas informações em todos os pontos elencados, exceto em relação à cor e ao número de fabricação, como se pode avaliar nos destaques indicados para a descrição das miniaturas indicadas nas imagens 8 a 10.

Considerando-se que o número de fabricação do modelo tenha um destaque unísono diante da cor, haja vista o fato de que se configura como identificador único da miniatura, permite-se fazer uma digressão sobre a cor.

As imagens referentes às figuras 8 a 10 destacam miniaturas com modelo, séries e subséries e ano de lançamento idênticos. Analisa-se, por meio de hipótese, que se um colecionador define como critério de busca o modelo, por exemplo, Volkswagen Kool Kombi, a fim de adquirir mais um item para sua coleção, e supõe-se que já tenha o referido modelo na cor *branca e azul*, o nível de granularidade proposto nos metadados para o *HotWheels Collection* permite-lhe a identificação do modelo em outra cor, *verde*, para remeter ao mesmo exemplo. Destaca-se, assim, a importância de que, tanto o padrão de estrutura de metadados quanto o software para gerenciamento dos dados, assumem para o sucesso da recuperação da informação, uma vez que se o catálogo não estiver modelado para atender tais requisições, de nada adianta ter a previsão de descrição nos metadados. Outro destaque é demarcado pelo fato de que indicar *color.*, para objetos multicoloridos, não atende requisitos de pesquisa por colecionadores. Desse modo, destaca-se a relevância da cor como ponto de acesso essencial para o *HotWheels Collection*.

Considera-se, então, o papel fundamental de o bibliotecário colocar-se no lugar do usuário no processo de catalogação, uma vez que suas necessidades informacionais ganham destaque e autenticam os metadados definidos.

Uma vez que para a publicação do *HotWheels Collection* são requeridos metadados, estruturas descritivas e software de gerenciamento de coleções digitais, iniciaram-se estudos com o Dublin Core, uma vez que é um dos padrões utilizados no Greenstone, software de gerenciamento de acervos digitais, no qual o *HotWheels Collection* está sendo desenvolvido.

Estão previstos no Dublin Core os metadados (DUBLIN..., 2012): *title*, nome pelo qual o recurso é formalmente reconhecido; *creator*, entidade responsável pela criação do recurso; *subject*, tema para o recurso que pode ser representado por meio de palavras-chave ou códigos de classificação, sendo recomendável adotar vocabulário controlado; *description*, consideração sobre o recurso que pode incluir um resumo ou um texto livre; *publisher*, entidade responsável por tornar o recurso disponível; *contributor*, entidade responsável por contribuir para o recurso; *date*, data ou período de tempo associado a um evento ou ciclo de vida do recurso; *type*, natureza ou gênero do recurso, para o qual se recomenda o uso de vocabulário controlado como o DCMITYPE; *format*, formato do arquivo, meio físico ou dimensões do recurso para o qual se recomenda o uso de vocabulário controlado MIME; pode incluir tamanho e duração; *identifi-*

*er*, referência ao recurso, sem ambiguidade, para o qual se recomenda o uso de um sistema de identificação formal; *source*, recurso relacionado do qual o recurso descrito é derivado; *language*, idioma do recurso; *relation*, identificação do recurso relacionado ao recurso descrito para o qual se recomenda um sistema de identificação formal; *coverage*, termo que representa a cobertura espacial ou temporal do recurso, para o qual recomenda-se o uso de vocabulário controlado; *rights*, sobre direitos detidos no e sobre o recurso.

Em uma abordagem ainda inicial, destaca-se que alguns elementos de metadados Dublin Core não seriam requeridos, outros exigiriam a repetibilidade ou, ainda, adaptações (quadro 1).

| Dublin Core | Metadados do HotWheels Collection  |
|-------------|--|
| Title       | Título: HotWheels : modelo   |
| Title       | Série: HW City. Graffiti Riders  |
| Creator     | Resp.: Mattel  |
| Subject     | Não requerido  |
| Description | Item: 1 miniatura de carro : ferro, roxo e verde ; 2 x 9 x 3 cm em embalagem de 17 x 11 x 4 cm |
| Description | Nota: Número do modelo no ano de 2010: 127/214   |
| Description | Nota: Data de copyright na miniatura c2002   |
| Description | Nota: Escala 1:64  |
| Description | Nota: Série da Coleção 2013  |
| Description | Nota: HW City: veículos de alta velocidade, sempre prontos pra muita ação                      |
| Description | Nota: Data de copyright na miniatura c2002   |
| Publisher   | Publ.: Mattel  |
| Contributor | Não requerido  |
| Date        | Publ.: c2012   |
| Type        | DGM: Modelo  |
| Format      | Não requerido  |
| Identifier  | Nota: Número de fabricação do modelo: R7554-B815   |
| Source      | Não requerido  |
| Language    | Não requerido  |
| Relation    | Não requerido  |
| Coverage    | Não requerido  |
| Rights      | Não requerido  |

Quadro 1. Metadados do *HotWheels Collection* em padrão Dublin Core

A pertinência do Dublin Core para a descrição de miniaturas no *HotWheels Collection* pode ser observada nos elementos *title*, *creator*, *description*, *publisher*, *date* e *identifier*. Distinguem-se o *title* e o *description* entre os elementos Dublin Core que exigem repetibilidade na descrição, uma vez que não estão previstos elementos de metadados específicos para série, item e notas. Elementos tais como *subject*, *contributor*, *for-*

*mat*, *source*, *language*, *relation*, *coverage* e *rights*, não se mostraram necessários. Reconhece-se como adaptação a indicação de *modelo*, definido como designação geral do material, no elemento *type*, uma vez que não está indicado no DCMITYPE, que prevê o seguinte vocabulário: *Collection*, *Dataset*, *Event*, *Image*, *Interactive Resource*, *Service*, *Software*, *Sound*, *Text*. Compreende-se, também, que tal elemento destina-se à identificação do tipo de recurso digital, e, nesse caso, dadas as reproduções em imagem digital das miniaturas e das embalagens, seria correto identificar como *Image*.

Embora o Dublin Core tenha se mostrado como suficiente para o HotWheels Collection demarca-se, como horizonte para pesquisas futuras, o intuito de avaliação de outras estruturas que evidenciem a relevância dos metadados identificados.

## Referencias

- BEDIN, F. (2007). Metodologia: O Caminho da ciencia. São Paulo: Edicon.
- BELK, R.W. (1995). Collecting in a Consumer Society. New York: Routledge.
- BELK, R. W.; WALLENDORF, M.; SHERRY, J.; HOLBROOK, M.; ROBERTS, S. (1998). Collectors and Collecting. *Dluth. Advances in Consumer Research*. 15.
- CÓDIGO DE CATALOGAÇÃO ANGLO-AMERICANO 2ª ed. (2004). São Paulo: FEBAB.
- DUBLIN CORE METADATA ELEMENT SET, version 1.1. (2012). Recuperado em 01-06-2101 de <http://dublincore.org/documents/dces/>.
- FARINA, M. C.; TOLEDO, G. L.; CORRÊA, G. B. F. (2006). Colecionismo: uma perspectiva abrangente sobre o comportamento do consumidor. Recuperado em 09-06-2013 de [http://www.ead.fea.usp.br/Semead/9semead/resultado\\_semead/trabalhosPDF/320.pdf](http://www.ead.fea.usp.br/Semead/9semead/resultado_semead/trabalhosPDF/320.pdf).
- FOULONNEAU, M.; RILEY, J. (2008). Metadata for digital resources: Choosing metadata standards for a digital library project. Oxford: Chandos.
- GERHARDT, T. E.; RAMOS, I. C. A.; RIQUINHO, D. L.; SANTOS, D. L. (2009). Estrutura do projeto de pesquisa. In: GERHARDT, E. T. SILVEIRA, D. T. Métodos de pesquisa. Porto Alegre: Editora da UFRGS. 65 – 85.
- GIACOMIN FILHO, G. (2006). O Colecionismo como motivação para a comunicação mercadológica. Recuperado em 09-06-2013 de <http://www.ipcoleccionismo.com.br>.
- GIL A.C. (2008). Como elaborar projetos de pesquisa. São Paulo: Atlas
- GILLILAND-SWETLAND, A. J. (1999). La definición de los metadatos. In: INTRODUCCIÓN a los metadatos: vías a la información digital. [S. I.]: GETTY. 1-9.
- GRÁCIO, J. C. A. (2002). Metadados para a descrição de recursos da Internet: o padrão Dublin Core, aplicações e a questão da interoperabilidade. Marília. 127p.
- HOTWHEELS.BR. (2013). Recuperado em 23-09-2103 em [www.hotwheelsbr.com](http://www.hotwheelsbr.com).
- MARSHALL, F. (2005). Epistemologias históricas do colecionismo. Porto Alegre. 13-23.
- MINAYO, M. C. S. (org.) (2002). Pesquisa Social. Teoria, método e criatividade. Petrópolis: Vozes.
- MINDLIN, J. (1997). Uma vida entre livros: reencontros com o tempo. São Paulo: Edusp, São Paulo: Companhia das Letras.
- PEARCE, S. M. (2003). Collecting reconsidered. In: PEARCE, S. M. (Ed.) Interpreting objects and collection. London: Routledge.
- PEREIRA, A. M., SANTOS, P. L. V. A da C. (1998). O uso estratégico das tecnologias em catalogação. *Cadernos da F.F.C.Marília*. 7. 121-131.
- ROSETTO, M. (2003). Metadados e formatos de metadados em sistemas de informação: caracterização e definição. São Paulo, 95p.
- SANTOS, R. E. (2013). Depoimentos pessoais.
- SENSO, J. A.; ROSA PIÑERO, A. de la. (2003). El concepto de metadato. Algo más que descripción de recursos electrónicos. *Brasília*. 32 (2). 95-106. Recuperado em 14-10-2013 em <http://www.scielo.br/pdf/ci/v32n2/17038.pdf>.
- SIQUEIRA, M. A. (2003). XML na ciência da informação: uma análise do MARC21. Marília. 134p. Recuperado em 19-10-2013 em [http://www.marilia.unesp.br/Home/Pos-Graduacao/CienciadaInformacao/Dissertacoes/siqueira\\_ma\\_me\\_mar.pdf](http://www.marilia.unesp.br/Home/Pos-Graduacao/CienciadaInformacao/Dissertacoes/siqueira_ma_me_mar.pdf).
- TAKAHASHI, T. (Org.) (2000). Sociedade da informação no Brasil: o livro verde. Brasília: Ministério da Ciência e Tecnologia. Recuperado em 29-09-2013 em <http://livroaberto.ibict.br/bitstream/1/434/1/Livro%20Verde.pdf>.
- ZAFALON, Z. R. (2014). Scan for MARC: princípios sintáticos e semânticos de registros bibliográficos aplicados à conversão de dados analógicos para o Formato MARC21 bibliográfico. São Paulo: Ed. UNESP.

**Hito 1: Modelos de organización del conocimiento. Modelaje de sistemas y ciclo de vida de la información**

**Hito 2: Entornos y sistemas para la organización del conocimiento. Diseño, normalización e interoperabilidad de sistemas de información abiertos.**

**Hito 3: Instrumentos para el acceso y transparencia de la información: gestión de documentos, plataformas digitales y archivos**

**Hito 4: Usuarios y usos de la información. Patrones vinculados a la transparencia y el gobierno abierto**

**Hito 5: Gobernanza y Auditoría de información en la organización del conocimiento.**

**Hito 6: Representación del conocimiento: procesos y herramientas en sistemas abiertos y multilingües**

**Hito 7: Epistemología en la organización del conocimiento**

**Hito 8: Igualdad, género y diversidad en la organización del conocimiento**

---

# Bibliotecas eclesiásticas de Castilla y León. Visibilidad y acceso web

*Ecclesiastical libraries from Castilla y León. Visibility and web access*

---

**Esther Carreño Corchete (1), Crispulo Travieso Rodríguez (2)**

(1) Biblioteca Central Capuchinos, C/ Ramón y Cajal, 13. 37002. Salamanca, molucas@usal.es  
(2) Universidad de Salamanca, Departamento de Biblioteconomía y Documentación, C/ Francisco de Vitoria, 6. 37008. Salamanca, ctravieso@usal.es

## Resumen

Existe una tipología de bibliotecas, las eclesiásticas, sobre las que no suele haber suficiente información, incluso entre ámbitos profesionales. Estas bibliotecas han sido, a lo largo de los siglos, un recurso esencial en la formación del clero y se han convertido en guardianas y transmisoras de un valioso e importante legado cultural que llega hasta nuestros días. ¿Cuál es su función? ¿Son meros depósitos de fondo antiguo? ¿Están orientadas al usuario para cubrir sus necesidades de formación e información?

La automatización de bibliotecas, proceso en el que siguen inmersas la mayoría de ellas, no puede entenderse en el momento actual solamente como la automatización de los procesos de adquisición, catalogación y OPAC. Siendo un área de conocimiento profesional, han de incluirse todas las facetas relacionadas con la puesta en marcha de las tecnologías de la información para llevar a cabo los servicios que la biblioteca debe ofrecer en el contexto de la sociedad de la información. Hoy en día, parte de esos servicios están centrados en la web, en el acceso y difusión de los recursos de información bibliográficos a los usuarios. Por ello, se analiza la visibilidad que tienen estas bibliotecas a través de páginas web propias y del acceso directo a sus catálogos en línea. De forma paralela, se amplía el estudio analizando, a través de una encuesta, su grado de automatización, la cooperación plasmada en catálogos colectivos, la participación en proyectos de digitalización de sus fondos, así como su incursión en la web social 2.0.

**Palabras clave:** Bibliotecas eclesiásticas. Acceso web. Automatización de bibliotecas. Catálogos automatizados. Digitalización. Web social 2.0. Biblioteca híbrida. Patrimonio cultural.

## 1. Introducción

Las bibliotecas eclesiásticas son centros privados especializados que cumplen una función imprescindible dentro de la institución a la que pertenecen. Transmisoras de un importante legado cultural, proporcionan los servicios adecuados para poder satisfacer las necesidades de información que requieren sus usuarios du-

## Abstract

Ecclesiastic libraries often go unnoticed, even by professionals in religious fields. Over the 3Mcenturies, these libraries have been an essential resource in the training of clergy and have become both repositories and sources of an invaluable cultural legacy that continues to the present day. Nevertheless, what is the true function of these libraries? Are they mere storage depots for old texts and tomes? Are they able to meet the information and training demands of today's users?

The majority of libraries today are involved in the ongoing process of automation and digitalization, yet this should not be understood as simply automating the processes of acquiring and cataloging their materials. Libraries belong to the sector of professional knowledge, so they must include all the facets of information technology in order to best serve a modern, information-based society. In today's world, one such facet is the use of the internet for access to and distribution of bibliographical resources. For this reason, the visibility and ease of access to library catalogs online will be analyzed. In parallel, the investigation will make use of surveys to determine the level of automation, digitization projects, cooperation and participation between ecclesiastic libraries as well as their inclusion in social networks and "Web 2.0" systems.

**Keywords:** Ecclesiastical libraries. Web access. Online catalogs. Digitalization projects. Social web 2.0. Cultural Heritage. Digital libraries.

rante los procesos de enseñanza, formación, aprendizaje e investigación. Son además, herederas y guardianas de un rico patrimonio bibliográfico; sus salas suelen albergar valiosas e inestimables colecciones de fondo antiguo que celosamente han sabido cuidar y conservar a lo largo de siglos.

El ámbito de las bibliotecas eclesiásticas pasa a menudo desapercibido profesionalmente; ape-

nas existe información sobre ellas y también son escasos los estudios de investigación realizados en este campo. Una de las razones puede ser la dificultad de agruparlas y realizar trabajos de conjunto debido a su variada idiosincrasia y a la diversidad de los centros. La casuística es considerable: están integradas en este sector desde bibliotecas monásticas, parroquiales, de seminarios, etc. Centros celosamente guardados y custodiados, muchos de ellos con un difícil acceso y una apertura limitada al exterior, hasta bibliotecas universitarias o dependientes de centros de estudios católicos, abiertos a la comunidad, con profesionales especializados y servicios actualizados de información orientados a los usuarios a los que van dirigidos.

El valor de estas bibliotecas es indiscutible, tanto a nivel histórico como cultural. Y el empeño de sus responsables por salvaguardar su patrimonio es evidente. En la comunidad de Castilla y León, 61 de ellas están integradas en el programa de inventariado y catalogación de los bienes bibliográficos y documentales establecido por el Ministerio de Cultura, Educación y Deporte, formando parte del Catálogo Colectivo de Patrimonio Bibliográfico Español. El Catálogo se realiza en cumplimiento de la Ley 16/1985 de Patrimonio Histórico español y es elaborado conjuntamente por dicho Ministerio y todas las Comunidades Autónomas.

A partir de los 90 y a raíz de la corriente donde se pusieron en marcha los procesos de automatización bibliotecarios, la mayor parte de estos centros se sumaron a la iniciativa y actualmente casi la totalidad siguen inmersas en procesos de catalogación automatizada o poseen catálogos automatizados. Curiosamente y a pesar de los grandes esfuerzos de presupuesto y de personal empleados, la información de esos catálogos no es visible en la mayor parte de los casos. Este trabajo trata de analizar la visibilidad de estas bibliotecas a través de internet y del acceso a uno de los recursos referenciales más utilizado: el catálogo automatizado en línea.

Con internet, las bibliotecas han sufrido una transformación y una ampliación de sus servicios, de tal modo que hoy en día es indiscutible la presencia de estos centros a través de la red para ofrecer y suministrar la información de sus catálogos, llegando así al mayor número posible de usuarios. A pesar de ser centros privados, muchas de ellas ofrecen servicios de préstamo, consulta e información a usuarios externos sin vinculación con la institución. Por este y otros motivos se hace imprescindible la difusión de estos servicios a través de la red.

Nos encontramos ante un sector donde el grado de especialización y desarrollo tecnológico es muy limitado en muchos casos; a nivel de catalogación automatizada la mayor parte de estas bibliotecas utilizan programas de gestión bibliotecaria. Sin embargo, los intentos de cooperación entre ellas aún no se han materializado, a pesar de los esfuerzos y tentativas desplegadas; se convierte en tarea prioritaria la creación de un catálogo colectivo propio que unifique y dé cobertura a estas bibliotecas.

Este trabajo trata de dar respuesta a tales cuestiones. Cuál es el grado de desarrollo tecnológico de estos centros en relación con el acceso a los servicios bibliográficos en red. Se ha centrado en la visibilidad de páginas web propias de las bibliotecas y en el acceso a los catálogos automatizados en línea. Por otra parte, se intenta ofrecer una visión general del estado actual de estos centros en relación con la catalogación automatizada, la digitalización de sus fondos y los servicios que ofrecen, así como en el uso de la web social para interactuar y ofrecer una participación activa al usuario.

## 2. Marco teórico

### 2.1. Las bibliotecas eclesiásticas

Para delimitar el objeto de estudio de este trabajo es necesario, en primer lugar, precisar el concepto de bibliotecas eclesiásticas. Una definición exhaustiva es la que propone Iguale (1991) en su diccionario sobre patrimonio cultural. Podemos decir que se encuadran en este concepto todas aquellas que pertenecen a la Iglesia Católica. Sin embargo, etimológicamente también podrían considerarse bibliotecas eclesiásticas las adscritas a otras iglesias o a iglesias cristianas no católicas, por ejemplo, las diferentes iglesias protestantes (la Iglesia Evangélica Española, los anglicanos o las Iglesias Bautistas) así como las diferentes iglesias ortodoxas, establecidas alrededor de las comunidades griega y rumana (Casado, 2013).

La Comisión Técnica de Cooperación de Bibliotecas Especializadas (2012), dependiente del Ministerio de Cultura, Educación y Deporte, establece unas directrices al respecto, debido a que las tipologías propuestas por la norma ISO 2789 y el cuestionario UNE presentan dificultades de interpretación cuando se tipifican las bibliotecas especializadas, principalmente en la disparidad de criterios elegidos por las diferentes comunidades autónomas. En este caso, la tipología establecida para las bibliotecas de Instituciones religiosas sería la incluida en la definición del INE: bibliotecas dependientes de

cabildos catedralicios, conventos. etc., es decir, todas las pertenecientes a la Iglesia Católica y sus instituciones. También se incluyen en esta categoría las pertenecientes a otras confesiones ubicadas en sus propios templos, sinagogas o mezquitas y se excluyen las bibliotecas de hospitales gestionados por órdenes religiosas, así como las bibliotecas de centros educativos gestionados por las Iglesias.

No obstante, en el caso de este estudio, el concepto de bibliotecas eclesiásticas queda delimitado a todos los centros cuya titularidad pertenece a la Iglesia Católica por ser la confesión mayoritaria en España, por motivos estadísticos y también de tradición. La localización de las bibliotecas se ha realizado a través de la Guía de las Bibliotecas de la Iglesia (2003), directorio también disponible en versión digital a través de la página web de la ABIE (Asociación de Bibliotecarios de la Iglesia). Esta asociación, creada en 1993, representa a todos los profesionales que trabajan en estas bibliotecas. Su misión se centra en “promover la defensa, conservación y difusión del ingente patrimonio bibliográfico de la Iglesia Católica”, según dicen sus estatutos.

## 2.2. Tipologías de bibliotecas eclesiásticas

Las tipologías utilizadas en el estudio se han basado en el directorio por tipo de biblioteca disponible en la página web de la Asociación de Bibliotecarios de la Iglesia (ABIE). Entre ellas, no sólo se cuenta con las bibliotecas monacales y conventuales, de seminarios, capitulares, diocesanas o adscritas a centros académicos teológicos, sino también con todas aquellas que forman parte de la estructura de las universidades dependientes de entidades eclesiásticas. Asimismo, se omiten aquellas bibliotecas dependientes de centros de enseñanza gestionados por la Iglesia Católica.

En definitiva, distinguiremos las siguientes:

- 1) Capitulares o catedralicias
- 2) Centros de estudio
- 3) Diocesanas / Episcopales
- 4) Especializadas o Centros de documentación
- 5) Monásticas y Conventuales
- 6) Seminarios
- 7) Universitarias
- 8) Otras

### 2.2.1. Bibliotecas capitulares

También llamadas catedralicias, son las bibliotecas de las catedrales y pertenecen a un cabildo, que es el órgano colegiado encargado del gobierno de una catedral o concatedral. Los Cabildos están presididos por un deán y formados por canónigos y dignidades. Las Colegiatas o Iglesias colegiales también poseen cabildo

pero se diferencian de las catedrales porque no son sede de un obispo (Casado, 2013).

Habitualmente las bibliotecas capitulares están ubicadas en la misma catedral o colegiata. En otras ocasiones, sus fondos se localizan externos a la catedral o junto a los archivos catedralicios, porque muchas veces suelen ser bibliotecas auxiliares del archivo. Actualmente, algunos obispados han optado por fusionar las bibliotecas diocesanas con las capitulares para unificar fondos y ahorrar en presupuestos y personal. Es el caso de los archivos y bibliotecas capitulares de Ciudad Rodrigo y Ávila.

Estas bibliotecas suelen estar dirigidas por un miembro del cabildo y, a nivel técnico, atendidas por profesionales seculares, como es el caso del archivo y biblioteca de la Catedral de Salamanca.

### 2.2.2. Bibliotecas de centros de estudio

Son las adscritas a los centros de enseñanza de la Iglesia Católica. Esto es, colegios, seminarios o institutos superiores de teología.

### 2.2.3. Bibliotecas diocesanas y episcopales

Pertenecen a una diócesis o archidiócesis. Históricamente y salvo excepciones, estas bibliotecas propiamente dichas no aparecen hasta la Edad Moderna. Un caso históricamente excepcional es el de la Biblioteca Episcopal de Girona, formada en el siglo XIV con fondos de diversa temática, no sólo religiosa. Estas bibliotecas suelen estar situadas en las dependencias del obispado, en las casas episcopales, de ahí que también se llamen así; se transmiten de obispo a obispo.

Otras bibliotecas episcopales que surgieron en la baja edad Media serían las bibliotecas de uso privado del obispo y estarían claramente diferenciadas de las anteriores. Muchas veces terminarían, por donación y orden expresa del obispo propietario, en las bibliotecas catedralicias o en monasterios (Becedas, 2007).

### 2.2.4. Bibliotecas especializadas o Centros de Documentación

Proporcionan documentación específica relacionada con el ámbito eclesiástico, y recogen y coordinan información técnica y científica generada en una determinada institución religiosa.

### 2.2.5. Bibliotecas monásticas y conventuales

Se incluyen las pertenecientes a monasterios y conventos, ambos con características propias y diferenciadas. En un monasterio se lleva una

vida monástica según una regla concreta (cisterciense, benedictina, etc.) y sus religiosos suelen vivir apartados de la sociedad. Los conventos suelen estar ubicados en las ciudades y acogen a religiosos o religiosas que están en contacto con el mundo (franciscanos, dominicos o jesuitas). El papel de estas bibliotecas ha sido fundamental en la custodia y preservación del patrimonio bibliográfico; gracias a ellas, numerosas obras han sido conservadas y han llegado a nuestros días (Casado, 2013).

#### 2.2.6. Bibliotecas de seminarios

Las bibliotecas de seminarios cumplen la función de ayudar en la formación de los futuros sacerdotes. Actualmente, estas bibliotecas se encuentran en la mayor parte de los casos, desactualizadas y poco atendidas debido a que los escasos seminaristas se nutren de las bibliotecas universitarias para su formación y estudio.

#### 2.2.7. Bibliotecas universitarias

Son las pertenecientes a universidades católicas. Suelen ser centros dotados de importantes fondos para la enseñanza y formación de los estudiantes y profesores, y cuentan con personal profesional especializado. Por su carácter formativo y académico suelen ser centros dotados con las últimas tecnologías y a menudo son pioneras en la implantación de avances tecnológicos. Habitualmente, son una referencia para las demás bibliotecas. La principal biblioteca de esta tipología en esta comunidad es la Biblioteca de la Universidad Pontificia de Salamanca.

#### 2.2.8. Otras

- Bibliotecas parroquiales

Son habitualmente bibliotecas auxiliares de los archivos parroquiales. Aunque existen bibliotecas de este tipo con fondos modernos que dan servicio a sus usuarios, en general, son guardianas de un rico fondo antiguo, que también ha sufrido numerosas pérdidas a causa del mal estado de conservación, la falta de personal especializado y de siglos de saqueos y guerras. Por ello, muchos de sus fondos han pasado a otras bibliotecas, como las públicas o diocesanas, donde se conservan en mejores condiciones (Casado, 2013).

- Bibliotecas de Institutos Seculares y Uniones Pías

Estas instituciones están integradas por laicos que deciden seguir vida consagrada y hacer votos, como los Identes, las Teresianas o las Voluntarias de Don Bosco, todos ellos vinculados a la enseñanza.

- Bibliotecas particulares

Existen bibliotecas privadas, propiedad de personas físicas o jurídicas vinculadas a la Iglesia, o que teniendo otro origen, acaban en manos de la Iglesia por donación, legado testamentario o compraventa.

El Directorio de Bibliotecas Españolas elaborado por el Ministerio de Cultura, Educación y Deporte indica que existen en nuestro país doscientas nueve bibliotecas eclesiológicas entre seminarios, facultades y estudios teológicos, colegios, cabildos y obispados. Según la ABIE y la Guía de Bibliotecas de la Iglesia, serían 250 centros. En comparación, existen 972 de la Administración General del Estado, 744 de la Administración Autonómica, 793 Universitarias, 5.032 Municipales, 3 del Poder Legislativo, 76 del Judicial y 12 de otros organismos del Estado. En lo que se refiere a bibliotecas de instituciones privadas no eclesiológicas (fundaciones, asociaciones y colegios profesionales, etc.) hay casi 800 (Casado, 2013).

#### 2.3. Radiografía de una biblioteca eclesiológica

Las bibliotecas eclesiológicas forman una telaraña desapercibida y poco visible en nuestro país; cada centro tiene su propia identidad y todos ellos en conjunto una serie de características comunes que los identifican dentro de esta tipología de bibliotecas privadas especializadas. Habitualmente están dedicadas a cubrir las necesidades de información de las comunidades religiosas a las que sirven, y también al público en general que requiera sus servicios, en aquellos casos en que su apertura lo contemple.

Si hay una característica que las identifica es la de ser guardianas de excelentes colecciones de fondo antiguo. Asimismo, poseen importantes y extensas colecciones especializadas y generales tanto del ámbito religioso como civil.

Son bibliotecas de un marcado carácter humanista y en ellas se encuentran importantes colecciones de consulta y referencia. Y además de estas, otras colecciones sobre hagiografía, predicación y catequesis, liturgia, espiritualidad, crónicas, etc. Junto a los libros religiosos, también obras de referencia generales; diccionarios bilingües, principalmente de lenguas clásicas; y también amplias colecciones y obras de referencia relativas a literatura, historia, ciencia, derecho y arte.

Asimismo, poseen hemerotecas con un amplio fondo de revistas científicas y de divulgación del ámbito teológico y filosófico, recibidas muchas de ellas por el método tradicional de intercambio científico.

También características son las colecciones de repetidos que albergan sus depósitos, procedentes de donaciones que a menudo reciben de religiosas y religiosos, sacerdotes, profesores, frailes, etc., entre los que se encuentran principalmente manuales de teología, filosofía, obras de patrística, derecho canónico, catecismos, misales, etc.

Como ya se ha dicho, suelen estar adscritas a centros de enseñanza y universidades católicas, seminarios, diócesis, monasterios y conventos, etc. La mayor parte de estos centros están atendidos por los propios religiosos y muchos cuentan con profesionales externos que les ayudan en las tareas técnicas.

Actualmente la sociedad, de forma paulatina, vive un claro proceso de secularización, y estas bibliotecas se enfrentan a una problemática de escasez de usuarios debido a la falta de vocaciones y también al envejecimiento del clero y congregaciones en general. Debido a su orientación religiosa, sus fondos son habitualmente consultados por un público específico, principalmente estudiantes, profesores, religiosas, seminaristas, sacerdotes, frailes, etc., pero el descenso paulatino de sus miembros y el envejecimiento de esta comunidad está repercutiendo directamente en estos centros, en su uso y en el personal que las atiende, que en muchas ocasiones son los propios religiosos. Unido a esto, el mayor problema de estos centros radica en la necesidad de personal técnico especializado, un mayor grado de desarrollo tecnológico y planes concretos y efectivos de cooperación.

#### 2.4. Panorama legislativo

El 80 por 100 del patrimonio artístico y cultural de nuestro país se encuentra relacionado de alguna manera con la iglesia. Pero no es únicamente la gran cantidad de bienes de su titularidad ni la importancia de su riqueza artística lo que hace del patrimonio cultural eclesiástico una cuestión compleja. El problema reside en la confluencia que se da entre los diferentes marcos normativos. Por una parte, la iglesia cuenta con un ordenamiento jurídico propio como sociedad jurídica particular, que está regulada por una normativa que sale directamente del código de Derecho Canónico. Por otra parte, está la normativa civil del estado, que regula las actividades de la sociedad española. Y en un tercer lugar una serie de acuerdos y convenios firmados entre la Iglesia, el Estado y las Comunidades Autónomas (Motilla, 1995).

Según la Pontificia Comisión para los Bienes Culturales de la Iglesia (1994), la misión de las bibliotecas eclesiásticas se centra en tres obje-

vos: conservación, desarrollo y acceso. De ellos, la preservación y conservación del fondo antiguo es el objetivo en el que más inciden la mayoría de estos centros y donde se ha llegado a un mayor desarrollo técnico y de colaboración gracias a los convenios Iglesia-estado y a la ley de patrimonio histórico.

#### 2.5. Bibliotecas eclesiásticas y sociedad de la información

Los acelerados avances tecnológicos que están modificando la vida de los ciudadanos, también se van introduciendo en las instituciones y centros educativos y culturales, entre ellos las bibliotecas eclesiásticas. Las nuevas tecnologías comienzan a producir cambios en los métodos de trabajo y difusión de la información, dotando de nuevos recursos y procesos a las tareas tradicionales. Es la propia sociedad de la información la que demanda una renovación de estas instituciones con el fin de preparar a los ciudadanos para convivir con esas nuevas exigencias. Por la misma razón, estimular la utilización de las tecnologías de la información se está situando en una de las preocupaciones de gobiernos y de organizaciones educativas y culturales internacionales.

El nivel de incorporación de las tecnologías al servicio de las bibliotecas eclesiásticas no es homogéneo. En los estadios más avanzados se encuentran las bibliotecas universitarias católicas y centros de estudios y facultades, que han crecido y evolucionado para atender necesidades concretas de estudio e investigación. En el entorno de la biblioteca católica universitaria la variedad de productos y servicios comerciales basados en las tecnologías de la información, la mayor dotación de recursos económicos, de infraestructuras y de personal, así como la existencia de grandes colecciones de revistas y de un mercado de la información, ha generado muchísimos desarrollos. Casi todos los proyectos tecnológicos punteros pertenecen a este campo y estas bibliotecas son consideradas referentes para las demás.

En cambio, muchas bibliotecas monacales, de seminarios, diocesanas y parroquiales están en precario, tanto por sus recursos como por sus servicios. La mayoría de estas bibliotecas de fondos medianos y pequeños están todavía en situación de atraso, si acaso terminando o continuando la automatización de los catálogos y fuera de la conexión a redes. Si bien es notable el esfuerzo que algunas órdenes de esta comunidad están teniendo en dotar de mejores instalaciones y servicios sus bibliotecas; como es el caso de los dominicos y los jesuitas.

## 2.6. Automatización de bibliotecas

A raíz del nacimiento del formato MARC para el almacenamiento informatizado de registros bibliográficos, los sistemas de automatización de bibliotecas empiezan a consolidarse a finales de la década de los 70. Es entonces cuando se establecen las bases del concepto de sistema integrado, quedando consolidado durante los años 80 este modelo, así como los Catálogos en línea de Acceso Público (Opac). Los hitos señalados por García Melero (1999) en la automatización de bibliotecas son: el surgimiento del formato MARC, el desarrollo de servicios bibliográficos y servicios de recuperación de información y la extensión de la tecnología para la cooperación bibliotecaria.

A partir de los 90, muchas de las bibliotecas eclesíásticas siguen la tendencia general y comienzan los procesos de automatización de sus fondos.

La automatización en estas bibliotecas ha estado centrada en los procesos técnicos bibliotecarios y de tratamiento de datos: catalogación, adquisiciones, circulación y publicaciones periódicas.

La mayor parte de las bibliotecas eclesíásticas utilizan SIGB para gestionar sus fondos y servicios, aunque aún hoy, algunas optan por bases de datos simples en Access o Excel.

### 2.6.1. Software propietario versus software libre

No se debe contemplar el software libre solamente como un programa gratuito, aunque esta sea una de sus características más atractivas. El aspecto más importante de este tipo de software es que proporciona el código fuente, es decir, libertad para ejecutar, copiar, distribuir y mejorar dicho software. Hay diversidad en el uso de software bibliotecario en estos centros. Desde programas propietarios como Liber-Marc o Absysnet hasta las bibliotecas que han optado por programas Open Source como Koha u Openbiblio.

Uno de los programas propietarios más utilizados en este sector eclesíástico es Liber-Marc, de la empresa Cospa & Agilmic. También en algunos centros se ha optado por software libre, por ejemplo Koha, en la biblioteca de la Universidad Pontificia de Salamanca y en el catálogo colectivo de la Orden de los Dominicos.

### 2.6.2. OPACS

Todas las actividades relacionadas con la automatización de bibliotecas se han visto potenciadas y al mismo replanteadas con internet (Jac-

quesson, 1995). El primer elemento visible ha sido el OPAC accesible en internet.

En las bibliotecas eclesíásticas se observa un vacío en este sentido; aunque un razonable número de centros tienen sus catálogos con acceso en línea, no es así en la mayoría de los casos y se convierte en urgente la tarea de poner los catálogos a disposición de los usuarios. La posibilidad de interconexión entre redes locales, el acceso al público de los sistemas de información, la interactividad con el usuario, hacen de la web un contexto lleno de enormes posibilidades.

### 2.6.3. Catálogos colectivos automatizados

El trabajo de automatización cooperativa mediante la creación de redes de bibliotecas responde a la necesidad de afrontar el conflicto que tiene cada biblioteca como entidad aislada. Es imposible que un sólo centro pueda tener disponible toda la información existente; tanto por falta de recursos económicos como por capacidad de proceso.

Teniendo en cuenta el desarrollo de las redes de bibliotecas en nuestro país y a nivel internacional, se observa que en este sector, tanto en la automatización de las bibliotecas como en la prestación de servicios bibliotecarios avanzados, existe un panorama de descoordinación y se hace necesaria una planificación estratégica de trabajo cooperativo.

Los objetivos que han de proponerse serían la colaboración y ayuda mutua en el intercambio de información, la conservación de colecciones, la automatización y la formación del personal.

## 2.7. Proyectos de digitalización en bibliotecas eclesíásticas

Tímidamente y en línea con la tendencia iniciada hace algunos años por universidades y otras instituciones educativas y culturales, algunas de estas bibliotecas han comenzado a participar en proyectos cooperativos de digitalización de fondo antiguo, como el de La Biblioteca Virtual de Patrimonio Bibliográfico, llevado a cabo por el Ministerio de Educación, Cultura y Deporte. En ella se encuentran obras de la biblioteca Diocesana de Zamora y del Archivo capitular de León. La biblioteca digital de Castilla y León también tiene entre sus colaboradores a la diocesana de Zamora y la biblioteca de los pp. Jesuitas de Burgos. También hay otros proyectos, como el que está desarrollando el personal técnico del archivo y biblioteca capitulares de Salamanca. Con un equipo propio, realizan paulatinamente

la digitalización de su fondo para facilitar el uso a los investigadores, llevar un registro automatizado de consultas y mantener la conservación de los documentos originales.

## 2.8. Proyecto web de la biblioteca y Web social 2.0.

### 2.8.1. Proyecto web de la biblioteca

La página web de la biblioteca es uno de los ejes de la política de servicios digitales. Es la carta de presentación de la institución ante sus usuarios y público en general. Gracias a ella la biblioteca se abre al exterior haciendo su contenido accesible a una comunidad y a la sociedad.

En los inicios, las bibliotecas estuvieron presentes en internet mediante el acceso aislado a sus catálogos, y actualmente casi cualquier biblioteca de cierta importancia posee su propia web.

El proyecto web forma parte de la misión de la biblioteca dentro de la institución, de la estrategia de información y servicios de la biblioteca, y no es tan solo una mera presencia institucional en la red. Posee un valor para la organización y no es una actividad secundaria de la biblioteca.

Dado el panorama enormemente competitivo que existe en internet, donde proliferan servicios de información altamente imaginativos y con gran potencialidad de atracción para los usuarios y recuperación de información, las bibliotecas se encuentran con el reto de demostrar la vigencia de su papel en la sociedad digital.

El espacio digital de la biblioteca se articula mediante la oferta al usuario de:

- Información general institucional sobre la biblioteca.
- Diseño y presentación de la información.
- Contenidos.
- Accesibilidad.

### 2.8.1. Web social 2.0.

Es llamativo el escaso número de bibliotecas eclesiósticas que utilizan tecnologías 2.0 en el desarrollo de sus servicios. El uso de estas herramientas conlleva principalmente un cambio de actitud por parte de las bibliotecas. También, la implementación de herramientas y servicios 2.0 implica un estudio y planificación, ya que no todas las herramientas sirven para todas las unidades de información ni para cualquier tipo de usuario. Los casos puntales de la biblioteca de la Universidad Pontificia de Salamanca, el archivo y biblioteca de la catedral y los dominicos también de esta ciudad son casos aislados. Todas las bibliotecas eclesiósticas han de plan-

tearse si quieren llegar a ser lugares abiertos a la comunidad y renovados que ofrecen este tipo de servicios a los usuarios. Según Alonso (2014), la biblioteca no será durante más tiempo un lugar proveedor de información pasiva.

## 3. Justificación y objetivos

De forma general, la automatización de bibliotecas ha sido un duro empeño al que casi la totalidad de las bibliotecas eclesiósticas han dedicado grandes esfuerzos presupuestarios y de personal desde la década de los 90 y durante los últimos años; en la actualidad, la mayor parte de ellas tienen sus colecciones automatizadas y utilizan algún tipo de SIGB para la gestión de sus procesos y servicios.

Sin embargo, no ha ocurrido lo mismo con su visibilidad en internet ni con la disponibilidad de sus catálogos, es decir con el acceso y difusión de la información que proporcionan. En el trabajo diario de catalogación, los profesionales de estas bibliotecas echan de menos catálogos de referencia debidamente normalizados que les ayuden en su tarea, así como la posibilidad de intercambio y descargas de registros para ahorrar tiempo y esfuerzos innecesarios. En este sentido, la única biblioteca de la región que proporciona dichos servicios es la Universidad Pontificia de Salamanca.

Es notable el aislamiento y la falta de planes de cooperación de estos centros, que se traducen en una descoordinación y duplicación de esfuerzos. El proyecto de creación de una red de bibliotecas eclesiósticas que englobe sus catálogos y los ponga a disposición de los usuarios está aún por concretarse y debería ser una iniciativa prioritaria.

En este contexto, las bibliotecas eclesiósticas llevan un camino paralelo al resto de bibliotecas pero bastante más lento en la incorporación de nuevas herramientas tecnológicas y servicios en red. Bien por falta de personal especializado, pues en muchos casos son atendidas por los propios religiosos, a veces con limitada especialización técnica y que deben atender otros trabajos; otras veces, por falta de presupuesto, y en algunos casos, también por una deficiente sensibilidad cultural de los responsables de estos centros.

El objetivo marcado con este trabajo es determinar el nivel de desarrollo tecnológico de estas bibliotecas, traducido en conocer cuál es el grado de automatización de sus catálogos, de su accesibilidad a través de internet y su participación en planes de cooperación y colaboración plasmados en catálogos colectivos. Asimismo,

se determinará la incursión en proyectos de digitalización y en la utilización de herramientas de la web social.

De ese objetivo general se derivan otras cuestiones que constituyen los objetivos específicos:

¿Cuál es el estado actual de la catalogación automatizada en las bibliotecas eclesiásticas de Castilla y León?

- (1) ¿Cuáles son los programas de gestión utilizados en estas bibliotecas?
- (2) ¿Cuántas tienen sus catálogos automatizados en línea?
- (3) ¿Participan en proyectos de digitalización de fondos?
- (4) ¿Utilizan herramientas de la web social 2.0?
- (5) Describir el grado de cooperación entre estos centros a través de sus catálogos colectivos.

## 4. Metodología

### 4.1. Delimitación geográfica

En este estudio se ha realizado una delimitación geográfica que incluye las bibliotecas eclesiásticas ubicadas en la comunidad de Castilla y León. Se han seleccionado un total de 54 bibliotecas. El listado se ha obtenido a través del directorio web de la ABIE.

La investigación se desarrolla en dos partes o líneas paralelas:

- (1) Comprobar la visibilidad y el acceso directo a estas bibliotecas a través de sus páginas web y de sus catálogos automatizados en línea.
- (2) Determinar el nivel de automatización de sus catálogos, la participación de estos centros en proyectos de digitalización y su incursión en la web social mediante una encuesta enviada a las bibliotecas eclesiásticas de Castilla y León seleccionadas para el estudio.

### 4.2. Recogida de datos

En función de los objetivos planteados, se ha optado por el envío de una encuesta como técnica de recogida de datos. Se entiende por encuesta la técnica de investigación que utiliza la consulta o el interrogatorio a un conjunto de individuos como método para recolección de datos (Delgado López-Cózar, 2004). En el sector de las Ciencias Sociales y concretamente en el campo de la Documentación, la encuesta ha sido siempre el principal método de recogida de datos.

Para determinar la primera parte de la investigación, se han realizado búsquedas en internet de todas las bibliotecas incluidas en el estudio. Se han establecido dos situaciones: bibliotecas con página web y catálogo en línea y bibliotecas sin acceso ni información en la web.

La segunda parte de la investigación se ha centrado en el envío de la encuesta a todos los centros participantes.

La encuesta ha sido enviada por email a la mayoría de las bibliotecas y también se han realizado llamadas telefónicas para hablar con los responsables de los centros. Se han recibido 10 Respuestas por email, 24 Respuestas por vía telefónica y 5 por entrevista personal directa y visita a las bibliotecas.

La redacción de la encuesta iba a determinar la naturaleza y alcance de los resultados a obtener. Se pensó en la necesidad de desdoblarse el formulario para que su aplicación fuese más eficaz, de manera que contemplase dos partes independientes: grado de automatización y cooperación de las bibliotecas y web social.

La encuesta se ha enviado a la mayor parte de los centros vía email. Solamente se han recibido por este medio 10 respuestas. Un segundo paso ha consistido en una visita puntual a los siguientes centros: *Biblioteca diocesana y capitular de Ávila*, *Biblioteca diocesana y capitular de Ciudad Rodrigo*, *Biblioteca de los Dominicos de Salamanca*, *Biblioteca del Seminario Diocesano de Salamanca* y *Biblioteca del Teologado Claretiano*.

La mayor parte de los datos se han obtenido vía telefónica, en entrevista directa con los responsables de estos centros. En total se han recogido por esta vía 24 respuestas.

Una vez que se han recogido los datos, el siguiente paso ha sido su ordenación y agrupación para proceder al estudio.

Para determinar la accesibilidad, se han agrupado estadísticamente los resultados en función de las frecuencias absolutas obtenidas por cada grupo de estudio.

La segunda parte de la investigación se ha llevado a cabo con la creación de una base de datos en Excel formada por todas las bibliotecas participantes, y con las respuestas de la encuesta. La base de datos muestra la información de manera esquemática y es mucho más sencillo realizar una valoración de los datos obtenidos. Son, finalmente, 15 apartados que se han estudiado de forma gráfica o textual en función de las respuestas obtenidas.

## 5. Resultados

El primer dato general que arroja este estudio es la distinción entre el número de bibliotecas eclesiásticas que sí ofrecen acceso web a sus catálogos frente al número de aquellas que no dan esa posibilidad. El cotejo de este aspecto nos devuelve (Figura 1) que solo un tercio (18 de los 54 centros analizados) permite acceso a sus catálogos en línea.

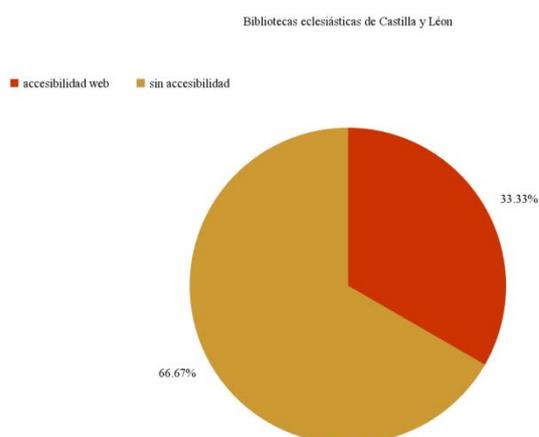


Figura 1. Bibliotecas eclesiásticas según acceso web

Si desglosamos el subconjunto de bibliotecas que dan acceso a sus fondos a través de catálogos en línea disponibles en web, esos 18 centros se distribuyen tal y como presenta la Figura 2, en la que se aparece como el grupo más numeroso (33.3%) las bibliotecas monacales y conventuales.

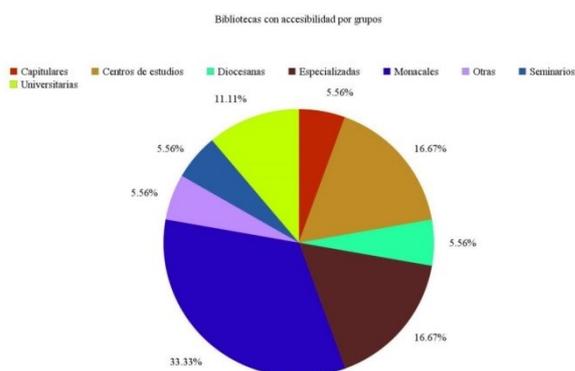


Figura 2. Total de bibliotecas eclesiásticas con acceso web agrupadas por tipología

En todo caso, es más elocuente observar esos datos ponderados en función de la presencia en la muestra de los distintos tipos de bibliotecas, para obtener una idea de cuál es la naturaleza de los centros que están más involucrados en la visibilidad web de sus fondos. Así, se obtiene que proporcionalmente a su inclusión en este

estudio, las tipologías de bibliotecas que de forma más frecuente ofrecen este servicio a todos los usuarios son las especializadas, las universitarias y los centros de documentación y estudios, grupos para los cuales todas las bibliotecas que formaron parte del estudio eran accesibles (Figura 3).

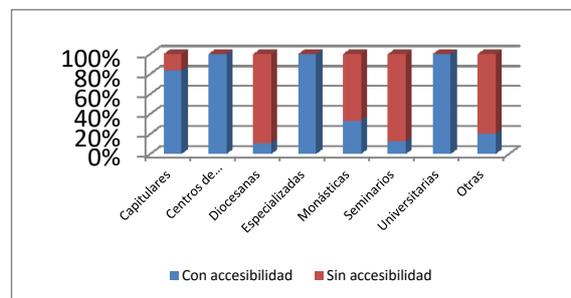


Figura 3. Acceso web de las bibliotecas eclesiásticas en función de su tipología

A continuación se presentan los resultados obtenidos de la encuesta agrupados según las cuestiones a las que se hacía referencia en las preguntas.

- Comienzo de la catalogación automatizada.

La primera pregunta del cuestionario hacía referencia al año en que se abordó la catalogación automatizada de los fondos, paso previo imprescindible para la posterior acceso de los mismos en entorno web. Sobre ella se obtuvo un total de 27 respuestas positivas (justo la mitad de las participantes), mientras que las restantes no habían emprendido esa tarea.

De las que afirmaron haber automatizado la descripción, el año de comienzo se corresponde con la figura 4, en la que se ha agrupado en períodos de cuatro años dicho inicio, para comprobar si existía alguna tendencia en ese sentido. Quizá lo más reseñable sea distinguir que la mayoría de las bibliotecas que respondieron habían empezado con la catalogación entre los años 1989 y 2008; durante esos veinte años, llama la atención el número de bibliotecas (9) que se incorporaron a la catalogación automatizada tardíamente, durante la pasada década, así como el hecho de que solo dos lo han hecho después de 2009.

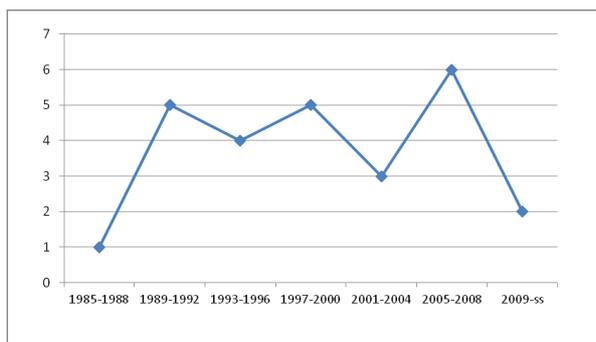


Figura 4. *Bibliotecas según el año de comienzo de la catalogación automatizada*

- Porcentaje de fondos catalogados.

En cuanto a la proporción de fondos que efectivamente ya estaban catalogados, de esas 26 bibliotecas que respondieron afirmativamente a estar inmersas en ese proceso, trece de ellas contaban con más del 90% de su fondo bibliográfico ya descrito, mientras que las 10 restantes declararon contar con más de la mitad del mismo. Los siguientes centros contaban ya con toda su colección catalogada:

Estudio Teológico Agustiniiano, Colegio de San Estanislao de Salamanca, Abadía Cisterciense de San Isidro de Dueñas (Venta de Baños), Abadía de Santo Domingo de Silos (Burgos), Biblioteca Provincial de los Padres Paúles (Salamanca), Agustinos Recoletos (Salamanca), Seminario Mayor de San Froilán (León), Seminario Mayor Diocesano (Salamanca).

- Sistema de gestión de bibliotecas utilizado.

Dentro de este subconjunto de bibliotecas que efectivamente empleaban un software para la gestión automatizada de sus fondos, uno de los aspectos destacable es la significativa presencia de programas no específicos de gestión bibliotecaria, sino de soluciones informáticas genéricas para gestión de bases de datos. Así, nueve bibliotecas organizaban sus fondos en función de herramientas como Access, Excel o FileMaker. De entre las que sí usaban productos específicos para material bibliográfico, los programas más frecuentes fueron por este orden: LIBERMARC, KOHA, WINISIS, ABIES, ABSYSNET, OPENBIBLIO y BIBLIO3000. Además de la variedad de soluciones empleadas, contrasta con la mayor homogeneidad que se da entre otro tipo de bibliotecas, como las públicas o las universitarias, donde la concentración de usuarias de determinados softwares comerciales es

patente. Este hecho puede frenar el desarrollo de catálogos colectivos.

- Inclusión en catálogos colectivos.

Solo once de las bibliotecas encuestadas (20.3%) estaban incluidas en algún catálogo colectivo, dato escaso que está en consonancia con la proporción ya presentada de bibliotecas que contaban con catálogos automatizados (condición previa inexcusable para dar el paso a la difusión de los fondos de forma colectiva). Los catálogos colectivos más citados fueron el Catálogo Colectivo de Patrimonio Bibliográfico Español y RABEL (Red Automatizada de Bibliotecas de Castilla y León).

- Suscripción a bases de datos y acceso externo a las mismas.

El valor explicativo de esta variable viene dada porque la suscripción efectiva a bases de datos externas es un indicador elocuente de la voluntad del centro por constituirse como elemento difusor de información especializada y científica y no solo como agente conservador, así como también puede ser indicativo de los recursos disponibles para este tipo de funciones de las bibliotecas eclesíásticas. Sobre ello, los resultados de la encuesta demuestran que, independientemente de a cuál de esos factores se deba, solo tres de de las bibliotecas consultadas afirmaron poseer acceso a estas fuentes de información. Y como era esperable, dos de ellas eran bibliotecas eclesíásticas del ámbito universitario.

Con respecto al acceso a las mismas, tanto la Universidad Católica Santa Teresa de Ávila como la Universidad Pontificia de Salamanca ofrecían a los usuarios la posibilidad de consulta externa.

- Posibilidad de consultas externas.

Frente al dato de la cuestión anterior, se planteó comprobar la posibilidad de realizar consultas bibliográficas mediante otros medios (correo electrónico o teléfono) a las bibliotecas incluidas en el estudio. En este punto sí se observó una mayor apertura por parte de los centros; de hecho, se obtuvo una respuesta afirmativa en un porcentaje mucho mayor que en otras cuestiones puesto que 31 de las 54 bibliotecas (el 57,4%) sí contaban con esa opción para los usuarios. Ello podría redundar en la idea de un progresivo interés por no aparecer como meros depósitos de información y en un intento por visibilizar su valor como centros documentales. Podríamos interpretar que esos intentos están

más frenados por la falta de recursos que por una negación intencional de esa función.

- Presencia de revistas especializadas.

Además de los fondos bibliográficos relativos a monografías, el dato de la presencia de publicaciones periódicas era relevante en este estudio, pues también puede tomarse como un factor sintomático de estas bibliotecas como especializadas en la información científica y de investigación en los campos de conocimiento en los que están centradas. A este respecto, los datos recogidos en la encuesta demuestran que 31 de las bibliotecas sí cuentan con este tipo de documentación, a la que se presupone un esfuerzo constante de actualización.

Si nos fijamos en la distribución según el número de títulos suscritos (Figura 5), vemos que el grupo más numeroso es el subconjunto de bibliotecas que cuentan con un intervalo de títulos no superior a 100, mientras que solo las dos bibliotecas universitarias ofrecen acceso a un universo mayor del millar de títulos (algo fácilmente explicable dada su consabida función de apoyo a la investigación académica).



Figura 5. Número de publicaciones periódicas

- Publicaciones propias.

Este es otro indicador claro de la iniciativa emprendida por las bibliotecas en su actividad de difusoras de información. Catorce de ellas (lo que representa un 26%) presentaban productos documentales, fundamentalmente revistas especializadas, de edición propia.

- Proyectos de digitalización.

Solo un número escaso de las bibliotecas consultadas (5) respondieron de forma afirmativa a la pregunta sobre si estaban inmersas en algún proyecto de digitalización (y consiguiente difusión en formato electrónico) de sus fondos. De nuevo, parece entreverse que la razón de ello se explica por la insuficiente disposición de recursos humanos y económicos, dado que el potencial de usuarios externos al que este tipo de bibliotecas podría dar servicio es considera-

ble, teniendo en cuenta el valor de los materiales que atesoran y las restricciones de consulta a las que usualmente están sujetos.

- Uso de herramientas de web social.

En consonancia con el resto de resultados de las cuestiones referidas a la inmersión tecnológica, el uso de herramientas 2.0 fue también exiguo (4 de las 54 bibliotecas). En este caso, y dado el carácter libre de este tipo de servicios, no cabría achacarlo a la falta de recursos, sino que nos atrevemos a aventurar a que la justificación esté más relacionada con un desconocimiento de las ventajas de comunicación y visibilidad que este tipo de recursos pueden reportar.

## 6. Conclusiones y recomendaciones

Las bibliotecas eclesíásticas han custodiado y conservado durante siglos un rico patrimonio cultural, documental y bibliográfico. Sus fondos abarcan un amplio arco cronológico que va desde el siglo XIII al XXI. Poseen fondos de consulta imprescindibles para un conocimiento integral de la historia, la cultura, las artes y las ciencias. Estas bibliotecas necesitan poner al servicio de los investigadores y público en general sus fondos, contribuyendo con ello al conocimiento y difusión de la cultura no sólo eclesíástica sino también civil.

Este trabajo muestra que es problemático el escaso grado de especialización de estos centros y las lagunas que existen tanto a nivel tecnológico como en la dotación de profesionales que sepan situar a estos centros al nivel actual de desarrollo tecnológico y digital. Es necesaria la puesta en marcha, tanto de forma individual como asociada, de nuevas iniciativas y planes de cooperación para que muchas de ellas no corran el peligro de quedar como simples depósitos bibliográficos desactualizados e incluso desaparecer. Las bibliotecas eclesíásticas necesitan seguir siendo en la sociedad un importante factor de creación, conservación y difusión de conocimiento, formación y cultura.

Para ello, es necesario llevar a cabo diversas iniciativas:

La primera y más importante es la planificación de un proyecto serio de cooperación entre estos centros. Es necesario un esfuerzo serio de organización que posibilite definitivamente la creación de un catálogo colectivo único. Esta tarea sería viable, desde un punto de vista jurídico, a través de la ABIE como asociación coordinadora junto a la Comisión para los Bienes Culturales de la Iglesia, dependiente de la Conferencia Episcopal Española. Y una vez esta-

blecidas las bases jurídicas, la biblioteca de la Universidad Pontificia de Salamanca como centro universitario dotado de medios adecuados y un suficiente desarrollo tecnológico, el proyecto del catálogo sería viable. Esta biblioteca universitaria ha sido la primera en España en cambiar de un sistema propietario, "Innopac Millenium" a software libre, concretamente al programa Koha. Sería recomendable que muchas de estas bibliotecas, que aún funcionan con bases de datos sencillas y programas básicos, optaran por cambiar a este programa gratuito para concretar definitivamente el proyecto, tan necesario como urgente, de un catálogo colectivo que englobe a todas estas bibliotecas.

Una cooperación de doble vertiente, intentando por otra parte la búsqueda de acuerdos de colaboración con instituciones y redes de bibliotecas del ámbito público, académico y científico. Los fondos de estas bibliotecas podrían llegar a un mayor número de usuarios potenciales que se beneficiarían de sus servicios, contribuyendo así a una mayor proyección social y servicio a los usuarios. En el campo de las humanidades y en concreto en literatura, historia, psicología, etc. muchas de ellas poseen importantes colecciones que pueden ser consultadas por especialistas y alumnos tanto de centros católicos como públicos.

Otra de las prioridades absolutamente importantes en estos centros es la necesidad de trabajar con personal técnico especializado. Es imprescindible que los procesos estén normalizados, que haya personal suficientemente formado para situar a estos centros al nivel que requiere hoy día la sociedad de la información. Por ello se hace también imprescindible realizar un esfuerzo económico para incorporar nuevos y modernos medios técnicos.

De forma individual, los responsables de estos centros necesitan dotar a sus bibliotecas de una infraestructura adecuada que posibilite la conservación de su rico patrimonio documental y bibliográfico y facilite la tarea investigadora.

Asimismo, colaborar con otras instituciones eclesíásticas y civiles, tanto políticas como culturales y financieras, para contribuir conjuntamente a la promoción de la investigación y de la cultura.

Estas bibliotecas a pesar de ser privadas y especializadas deberían mantener un acceso público generalizado debido a su importante valor cultural y humanista. No solamente por recomendaciones de la IFLA en su manifiesto sobre servicios bibliotecarios y de información en Internet. También la Comisión Pontificia para la conservación de los bienes culturales eclesiásti-

cos tiene como uno de sus objetivos principales el acceso y difusión a las colecciones eclesíásticas; por lo tanto, es urgente y necesario que estas bibliotecas comiencen a desarrollar nuevas políticas de desarrollo en este sentido para actualizarse y crear nuevos servicios al usuario.

## Referencias

- Acuerdo entre el Estado Español y la Santa Sede sobre asuntos jurídicos (1979). Recuperado 12-6-2014, de [http://www.vatican.va/roman\\_curia/secretariat\\_state/archivio/documents/rc\\_seg-st\\_19790103\\_santa-sede-spagna\\_sp.html](http://www.vatican.va/roman_curia/secretariat_state/archivio/documents/rc_seg-st_19790103_santa-sede-spagna_sp.html)
- Alonso Arévalo, J.; Cordon García, J. A.; Gómez Díaz, R.; García-Delgado Jiménez, B. (2014). Uso y aplicación de herramientas 2.0 en los servicios, producción, organización y difusión de la información en la biblioteca universitaria. *Investigación bibliotecológica*. 28 (64), 51-74. Recuperado 15-1-2015, de <https://dl.dropboxusercontent.com/u/29684975/Usos%20y%20aplicacion%20de%20herramientas%202.0.pdf>
- Archivos y bibliotecas eclesíásticas de Castilla y León*. Valladolid: Junta de C. y L., Consejería de Cultura y Bienestar Social, 1989.
- Barrio Gozalo, Maximiliano. Los archivos eclesíásticos, diocesanos y de las órdenes religiosas, y los investigadores. En: Ribot García, Luis Antonio (coord.). (2003) *El libro, las bibliotecas y los archivos en España a comienzos del Tercer Milenio*. Madrid: Sociedad Estatal España Nuevo Milenio, 485-504.
- Becedas, Margarita. *Las bibliotecas históricas de Castilla y León*. Valladolid: Junta de Castilla y León, Consejería de Cultura y Turismo, 2007.
- Carrión Gutiérrez, Alejandro (1999). RABEL, un instrumento de cooperación. *Educación y Biblioteca*. 101, 96-99. Recuperado 15-9-2014, de [http://gredos.usal.es/jspui/bitstream/10366/115428/1/EB\\_11\\_N101\\_P96-99.pdf](http://gredos.usal.es/jspui/bitstream/10366/115428/1/EB_11_N101_P96-99.pdf)
- Casado, Antonio (2013). Otras tipologías bibliotecarias. Las bibliotecas eclesíásticas. Recuperado 18-6-2014, de <http://www.biblogtecarios.es/antoniocasado/otras-tipologias-bibliotecarias-las-bibliotecas-eclesiasticas/>
- Comisión técnica de cooperación de bibliotecas especializadas (2012). Tipología de bibliotecas especializadas. Recuperado 6-4-2014, de [http://www.mecd.gob.es/cultura-mecd/dms/mecd/cultura-mecd/areas-cultura/bibliotecas/mc/consejocb/comisiones-tecnicas-de-cooperacion/especializadas/Tipologia\\_de\\_bibliotecasdefinitivo2abril2012.pdf](http://www.mecd.gob.es/cultura-mecd/dms/mecd/cultura-mecd/areas-cultura/bibliotecas/mc/consejocb/comisiones-tecnicas-de-cooperacion/especializadas/Tipologia_de_bibliotecasdefinitivo2abril2012.pdf)
- Cordón García, José Antonio; Alonso Arévalo, Julio; Gómez Díaz, Raquel; López Lucas, Jesús. *Las Nuevas fuentes de información: información y búsqueda documental en el contexto de la web 2.0*. Madrid: Pirámide, 2012.
- Curros Ares, M. A. Las bibliotecas de las órdenes y congregaciones. Organización y conservación. *Patrimonio cultural de los religiosos españoles: Primeras jornadas sobre el patrimonio histórico-cultural de los religiosos españoles*. Madrid, CONFER, 1992. 37-48.
- Directrices para planificar la digitalización de colecciones de libros impresos antiguos y manuscritos. Elaboradas por la Sección de Libros Raros y Manuscritos de la IFLA (2014). Recuperado 13-1-2005, de [http://www.ifla.org/files/assets/rare-books-and-manuscripts/rbms-guidelines/directrices\\_para\\_planificar\\_la\\_digitalizacion\\_de\\_colecciones\\_de\\_libros\\_antiguos\\_impresos\\_y\\_manuscritos.pdf](http://www.ifla.org/files/assets/rare-books-and-manuscripts/rbms-guidelines/directrices_para_planificar_la_digitalizacion_de_colecciones_de_libros_antiguos_impresos_y_manuscritos.pdf)

- García Madrid, A.; Manzano García, M. I. Recomendaciones y experiencias para una política de normalización y crecimiento de las bibliotecas y de los archivos de la Iglesia, hechas desde una biblioteca universitaria. *Revista Patrimonio Cultural*, 2002, 35, 25-31.
- García Melero, L.A. y García Camarero, E. (1999). *Automatización de bibliotecas*. Madrid: Arco.
- González Herrera, Guadalupe (2010). *Software libre vs. software propietario: una evaluación de sistemas Janium vs. Koha*. Recuperado 24-1-2015, de <http://eprints.rclis.org/14590/1/tesis.pdf>
- Guía de las bibliotecas de la Iglesia* (2003). Comisión Episcopal para el Patrimonio Cultural de la Iglesia. Madrid: EDICE.
- Hevia Ballina, A. La Iglesia y la cultura: bibliotecas eclesiales. Génesis y desarrollo. *Revista Patrimonio cultural*, 2001, 33, 99-114.
- Iguale Borau, Damián (1991). *Diccionario del Patrimonio Cultural de la Iglesia*. Madrid: Encuentro, 208-209.
- Jacquesson A. (1995). L'informatisation des bibliothèques: historique, stratégie et perspectives. Paris: Cercle de la Librairie.
- Marcos Mora, M. C. El acceso a los catálogos en línea a través de internet. *Revista general de información y documentación*. 1998, 8 (2) 85-118.
- Margaix-Arnal, D. (2008). El OPAC 2.0. Puerta de acceso a los contenidos de la biblioteca. IV Congreso Nacional de Bibliotecas Públicas (A Coruña, 24-26 de Septiembre de 2008). Madrid; Ministerio de Cultura, 237-246.
- Marín Martínez, Tomás (1972). Bibliotecas eclesiales. En Aldea Vaquero, Quintín, Marín Martínez, Tomás y Vives Gatell, José (Dir.). *Diccionario de Historia Eclesial de España*. Madrid: CSIC Instituto Enrique Flórez, 250-261
- Martínez Arellano, F. F.; Ruiz Vaca, J. O. Las nuevas características de los catálogos de las bibliotecas. En Martínez Comeche, Juan Antonio y Marcos Recio, Juan Carlos (Coords.). *Información y documentación: investigación y futuro en red*. 2011, 383-388.
- Melo Alves, M. F. y Quiroa Herrera, M.L. (2007). Análisis y evaluación de sitios Web de bibliotecas nacionales: los casos de Brasil y de Portugal. *Revista Española de Documentación Científica*. 30, 2, 199-217.
- Miguel Alonso, Aurora (1992). La biblioteca de los Reales Estudios de San Isidro, Madrid, Editorial de la Universidad Complutense.
- Mateu, F. Los catálogos de las bibliotecas y archivos eclesiales de España. *Hispania Sacra*, 1948.
- Morillo, M<sup>a</sup> Jesús (2010). Opacs 2.0. Recuperado 3-12-2014, de [http://biblioteca.cchs.csic.es/docs/formacion/herramientas\\_web2\\_bibliotecarios/OPACs%2020.pdf](http://biblioteca.cchs.csic.es/docs/formacion/herramientas_web2_bibliotecarios/OPACs%2020.pdf)
- Motilla, A. (1995). Régimen jurídico de los bienes histórico-artísticos de la Iglesia Católica. Madrid: Eurolex, Madrid.
- Pontificia Comisión para los Bienes Culturales de la Iglesia (1994). *Le biblioteche ecclesiastiche nella missione della chiesa*. Recuperado 1-7-2014, de [http://www.vatican.va/roman\\_curia/pontifical\\_commissions/pcchc/documents/rc\\_com\\_pcchc\\_19940319\\_biblioteche-ecclesiastiche\\_it.html](http://www.vatican.va/roman_curia/pontifical_commissions/pcchc/documents/rc_com_pcchc_19940319_biblioteche-ecclesiastiche_it.html)
- Rica Barriga, J. E. Identidad, misión y objetivos de las Bibliotecas Diocesanas y Eclesiales como centros dinamizadores de cultura cristiana en el entorno en el que se ubican. *Patrimonio cultural* 2007, 2, 17-38.
- Travieso Rodríguez, C.; Frías Montoya, J. A. La visualización de la información bibliográfica en los catálogos en línea y en entorno web: tendencias de investigación. En Travieso Rodríguez, Crispulo y Frías Montoya, José Antonio (Coords.). *Tendencias de investigación en organización del conocimiento = Trends in knowledge organization research*. Salamanca: Universidad de Salamanca, 2003, 455-462.
- Viñas Torner, Vicente. La conservación del patrimonio bibliográfico y documental de la Iglesia. *Revista Patrimonio Cultural*. 2001, 33.

---

# Modelagem de dados para representação descritiva no domínio da História da Arte: uma proposta para a construção de um banco de dados

*Data modeling for descriptive representation in the field of History of Art: a proposal to build a database*

---

Célia da Consolação Dias (1), Cíntia de Azevedo Lourenço (2)

(1) Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Brasil, celiadias@gmail.com

(2) cintia.eci.ufmg@gmail.com

## Resumen

O trabalho apresenta um projeto para organização e tratamento da informação de um acervo de fotos de pinturas de tetos e paredes de edifícios antigos. O objetivo principal é apresentar a modelagem, propor uma estrutura para a criação de uma base de dados, contendo imagens e informações sobre as pinturas de falsa arquitetura existentes em tetos e paredes em edifícios históricos no século XVIII. O projeto aconteceu em duas fases, onde foi desenvolvido um protótipo de uma base de dados com as fotos, uma interface dessa base com a identificação e descrição dos metadados necessários para identificar e para disponibilizar o acervo na web para a criação de forma de uma biblioteca digital. Como procedimentos metodológicos destacam-se a análise da ficha descritiva das fotos que existiam em papel, o estudo para a definição estrutura da base de dados e, ao mesmo tempo da busca por outras bases de dados de sistemas de documentação museológica; o levantamento dos metadados descritivos necessários para a representação dessa informação; a descrição dos metadados para facilitar a recuperação da informação e por último foi feita a realização da modelagem conceitual com o uso da ferramenta MER (Modelagem entidade – relacionamento) e estrutura de metadados que possibilitou a criação da interface para uma biblioteca digital foi proposta em Dublin Core. Além dessa estrutura foi apresentada, também a descrição e o padrão de cada metadado utilizado para a organização das informações sobre o acervo e para favorecer a sua recuperação.

**Palabras clave:** Modelagem de dados, História da Arte, representação descritiva, Dublin Core.

## 1. Introdução

A organização e recuperação da informação, na sociedade atual, tem se tornado cada vez mais desafiante, visto a quantidade de informações disponíveis na internet.

Nesse universo, as mais diferentes áreas do saber humano buscam visibilidade científica, artística e cultural, disponibilizando conteúdos e informações na rede. Essa realidade tem trazido

## Abstract

The paper presents a project for the organization and processing of information of a photo collection of paintings of ceilings and walls of old buildings. The main objective is to present the modeling, proposing a framework for the creation of a database containing images and information about existing architecture fake paintings on ceilings and walls in historic buildings in the eighteenth century. The project took place in two phases, where it was developed a prototype of a database with photos, an interface that base with the identification and description of the metadata required to identify and to provide the web assets for creating form of a digital library. As methodological procedures include the analysis of the fact sheet of the photos that exist on paper, the study for the structure of the database setting, while the search for other databases of museum documentation systems; the survey of descriptive metadata needed for the representation of such information; the description of metadata to facilitate the retrieval of information was taken and finally the achievement of conceptual modeling using the MER tool (modeling entity - Relationship) and metadata structure which enabled interface for the creation of a digital library was proposed in Dublin Core. Besides this structure was presented also the description and the pattern of each metadata used for the organization of information on the acquis and to promote to its recovery

**Keywords:** Data modeling, art history, descriptive representation, Dublin Core

os bibliotecários cada vez mais para fora das bibliotecas, com o desafio de transmitir, adequar, adaptar e inovar toda sua gama de conhecimentos e técnicas desenvolvidos e aperfeiçoados desde os primórdios da humanidade para organização e recuperação do conhecimento produzido pelo homem.

O presente estudo vai ao encontro desses novos desafios da biblioteconomia e da ciência da

informação. Neste sentido, este trabalho está sendo realizado com uma equipe multidisciplinar, através de um projeto que visa a disponibilizar informações, documentos acumulados e estudos sobre a pintura de tetos em Portugal, Itália e Brasil a partir da realização de inventários realizados por um grupo de pesquisadores em História da Arte.

Este projeto que está sendo desenvolvido pelas autoras deste trabalho foi concebido em 3 fases que compreendem: 1) A etapa de organização das imagens das pinturas de tetos das igrejas e os documentos relacionados que, se iniciou com a estruturação da base de dados de imagens; 2) Apresentação de uma proposta de interface, em Dublin Core, para a construção de uma biblioteca digital. Essa biblioteca tem como objetivo reunir, disponibilizar e compartilhar o acervo de documentos relacionados às imagens dos tetos das igrejas barrocas; 3) A construção de um instrumento que permita a padronização das informações da base de imagens e facilite a recuperação da informações do acervo organizado pela comunidade científica.

O objetivo principal deste trabalho é propor uma estrutura para a criação de uma base de dados, contendo imagens e informações sobre as pinturas de falsa arquitetura existentes em tetos e paredes em edifícios históricos no século XVIII. Pretende-se com essa estrutura organizar esse inventário para facilitar e ampliar as pesquisas sobre esse tema, tão importante para historiadores e artistas.

A ideia é aplicar os conhecimentos de representação da informação, combinados com as possibilidades tecnológicas disponíveis para assim, facilitar e melhorar a recuperação das informações contidas nesse inventário.

## 2. Bases teóricas

### 2.1. A organização do acervo

A coleta de informações sobre as coleções é uma atividade resultante do processo de pesquisa, etapa importante para a constituição do processo de documentação que registra os dados sobre determinado objeto. Neste estudo o objeto em questão é representado pelas pinturas de tetos de falsa arquitetura cujos documentos foram coletados com o objetivo de fazer um inventário deste acervo que posteriormente será organizado para o compartilhamento das informações desta coleção.

Por se tratar de um acervo de fotos das pinturas em questão, a estrutura do banco de dados proposto, procurou seguir as diretrizes das mais recentes pesquisas em catalogação, que vêm

se ampliando e acrescentando aos seus processos a modelagem conceitual Entidade-Relacionamento, como nos estudos dos FRBR (Functional Requirements of Bibliographic Records). De acordo com Fusco (2010, p. 20), essa modelagem das estruturas bibliográficas através do desenvolvimento de modelos conceituais, pode servir de base para a construção de esquemas de banco de dados.

Já Ferrez (1994), em seu estudo sobre os sistemas de documentação museológica afirma que a documentação museológica é “um sistema de recuperação de informação capaz de transformar as coleções dos museus de fontes de informação em fontes de pesquisa científica”. Neste sentido, transformar as coleções em registro de informação que serão armazenados em um banco de dados e, posteriormente possam ser fontes de informação passíveis de recuperação e compartilhamento do conhecimento potencial é um dos objetivos dos sistemas de informação museológica.

Para cumprir os objetivos dos sistemas de recuperação de informação apontados por Ferrez (1994), destacam-se dois aspectos que estão em conformidade com os estudos realizados neste projeto. São eles: “armazenar informações sobre os objetos, individualmente e completar estas informações através da literatura e de outras fontes documentais existentes, e do registro fotográfico”.

A organização do acervo do inventário da pintura de tetos em Portugal no século XVIII e no Brasil é fundamental para possibilitar a recuperação de informação pelos pesquisadores. Além disto, essa organização vai possibilitar a formação de redes de informação com o objetivo de reunir uma gama de estudos realizados ou em andamento, encorajar a realização de novas pesquisas sobre as pinturas de teto, temática pouco estudada no Brasil e, ao mesmo tempo, facilitar o compartilhamento, via web, de informações com a comunidade científica.

O acervo inicial que subsidiou a elaboração do inventário dos tetos no período é formado, em sua maioria por fotografias e imagens digitais de tetos de igrejas de Portugal e Brasil e os registros do levantamento documental realizado pelo grupo de pesquisa *Perspectiva Pictorum*, registrado no CNPq. A pintura de falsa arquitetura em Minas Gerais, o objeto de estudo deste grupo de pesquisa, está incluída na categoria de pintura de perspectiva. Neste sentido, a intenção do grupo de pesquisa é resgatar o universo dos tratados estudando ainda a transmigração da pintura setecentista desde a Europa, até a América Portuguesa de modo a reconhecer

neste gênero pictórico uma rede de difusão do saber perspéctico. Para complementar este acervo foram identificados além dos registros das imagens coletadas, artigos, teses e dissertações, produtos dos estudos realizados por alunos da pós-graduação e por pesquisadores do referido grupo de pesquisa, bem como outros materiais relacionados aos estudos sobre a pintura de falsa arquitetura em Minas Gerais, assim como todos os documentos e registros realizados durante o inventário.

Preocupados em tornar este acervo passível de recuperação para a geração de novas pesquisas e ao mesmo tempo compartilhar as informações existentes é que se pensou na organização do acervo dos documentos relacionados ao inventário das imagens de teto das igrejas barrocas.

## 2.2. Modelagem dos dados

Em sua pesquisa de doutorado, o Professor Magno Moraes Mello, do Departamento de história da UFMG, inventariou pinturas em tetos e paredes de edifícios históricos no Brasil, em Portugal e na Itália, coletando fotos, informações sobre as pinturas e bibliografias.

Analisando as fichas preenchidas sobre cada uma dessas pinturas fotografadas e os edifícios onde se encontram, tentou-se chegar a uma estrutura que permitisse uma padronização dos dados inseridos na base de dados, além de uma melhor customização e usabilidade que facilitasse a inserção dos dados.

Buscou-se assim, inicialmente trabalhar a nomenclatura dos campos, pois essa base de dados conteria informações do Brasil e de outros países, havendo necessidade de se utilizar termos que fossem inteligíveis em qualquer parte do mundo.

Chegado a um consenso inicial, foi realizada uma modelagem de dados, para um mapeamento conceitual mais consistente das entidades centrais dessa base de dados e como estas se relacionavam entre si. Para essa modelagem, optou-se pela “modelagem entidade-relacionamento” também conhecida pela sigla MER (CHEN, 1990), por ser de fácil utilização por pesquisadores da área de ciências humanas e sociais.

Outro motivo para a escolha do MER para essa modelagem de dados é o uso dessa técnica, nos estudos para o delineamento do novo código de Catalogação RDA, lançado em 2013, que foi desenvolvido com base na modelagem conceitual do ISBD (International Standard Book Description), pelo JSC (Joint Steering Committee).

Essa modelagem resultou no modelo FRBR (Functional Requirements of Bibliographic Records), base conceitual do RDA (Resource Description Access).

O objetivo da MER é identificar as entidades, ou seja, os núcleos principais de informação da base de dados, entidades secundárias vinculadas a essas entidades bases e a maneira como elas se relacionam.

Essa modelagem facilita a visualização de cada parte das informações da base de dados e, permite a identificação clara dos atributos, ou seja, dos campos necessários para a inserção das informações, além da elaboração de um dicionário de dados para eliminar ambiguidades e polissemias dos termos utilizados para nomear esses campos.

Na FIG. 1, pode-se observar a modelagem inicial. Para o inventário de pinturas em tetos e paredes, identificou-se 3 Entidades Nucleares: Iconografia, Edifício e Referências.

Na Entidade Iconografia, estão vinculadas as informações sobre a pintura propriamente dita. Nesse universo, identificou-se também 4 entidades secundárias, importantes para a padronização da entrada dos dados na base de dados: Suporte, Autorias, Material e Técnica.

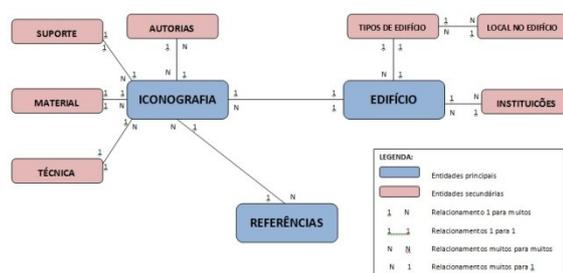


Figura 1. Modelagem entidade-relacionamento. Fonte: elaborado pela equipe do projeto

Essas entidades secundárias consistem basicamente em listas padronizadas, pré-definidas com informações que podem se repetir em várias pinturas da base de dados. Maiores informações serão apresentadas quando tratarmos dos atributos das entidades e do dicionário de dados.

Na entidade Edifício, serão registradas as informações sobre os prédios onde essas pinturas estão localizadas. Aqui também foram identificadas mais 3 entidades secundárias: Tipos de edifícios, Local no edifício e Instituições.

E por fim, na Entidade Referências, estarão as informações relacionadas à documentação e a bibliografias acerca de estudos realizados sobre a pintura em questão. Além dessas informações

a modelagem apresentada na FIG. 1, ainda nos informa sobre como essas entidades se relacionam entre si através das cardinalidades:

- 1 para N – um para muitos
- 1 para 1 – um para um
- N para N – muitos para muitos

Por exemplo, se observarmos a relação da entidade Iconografia com a entidade Edifício, poderemos observar que uma Iconografia só poderá existir em um único edifício, mas em contra partida, um Edifício poderá conter várias Iconografias.

Feita a modelagem, o passo seguinte é identificar os atributos, ou seja, os campos necessários para que cada entidade seja descrita e/ou identificada, e um Dicionário de dados que explica que tipo de informação cada atributo conterá:

| ICONOGRAFIA                | EDIFÍCIO                 |
|----------------------------|--------------------------|
| Título                     | Nome                     |
| Legenda                    | Tipo de edifício         |
| Criador                    | Local do edifício        |
| Data de nascimento e morte | País                     |
| Data de atividade          | Estado                   |
| Criador atribuído: (S/N)   | Município                |
| Atribuidor                 | Distrito                 |
| Suporte                    | Bairro                   |
| Material                   | Freguesia                |
| Técnica                    | Concelho                 |
| Dimensão                   | Tombamento (S/N)         |
| Fotógrafo                  | Instituição Tombadora    |
|                            | Data do tombamento       |
|                            | Restauração (S/N)        |
|                            | Instituição restauradora |
|                            | Data da restauração      |

#### a. Atributos das entidades nucleares

| AUTORIAS                   | TÉCNICA |
|----------------------------|---------|
| Nome                       | Nome    |
| Data de nascimento e morte |         |
| Data de atividade          |         |

| SUPORTE | TIPO DE EDIFÍCIO |
|---------|------------------|
| Nome    | Nome             |

| MATERIAL | LOCAL NO EDIFÍCIO |
|----------|-------------------|
| Nome     | Nome              |

#### INSTITUIÇÕES

Nome  
b. Atributos das entidades secundárias

#### ENTIDADES NUCLEARES

| ICONOGRAFIA | Descrição   |
|-------------|---|
| Título      | Nome ou descrição da imagem pintada.                              |
| Legenda     | Texto inserido da própria pintura em latim.                       |
| Criador     | Nome do pintor a quem a autoria da pintura foi atestada por fonte |

|                            |   |
|----------------------------|---|
| Data de nascimento e morte | Período em que o pintor viveu.  |
| Data de atividade          | Período em que o pintor trabalhou.  |
| Criador atribuído (S/N)    | Caso a pintura tenha sido atribuída posteriormente por estudos, selecionar S (Sim). Dessa forma o nome da pessoa que atribuiu essa autoria será solicitado.   |
| Atribuidor                 | Nome do pesquisador ou estudioso que atribuiu a um pintor uma obra de autoria não documentada.  |
| Suporte                    | Nome da estrutura predial onde a pintura foi realizada. Pode ser: parede, abóbada, cúpula, teto plano, caixotão, ou outro identificado posteriormente.  |
| Material                   | Tipo de material do suporte da pintura. Pode ser: alvenaria, madeira, ou outro identificado posteriormente.   |
| Técnica                    | Nome da técnica de pintura utilizada na pintura. Pode ser: óleo, têmpera, estuque, afresco, ou outra identificada posteriormente.   |
| Dimensão                   | Tamanho em metros quadrados, exato ou aproximado, da pintura.   |
| Fotógrafo                  | Nome do fotógrafo que fez a foto da pintura.  |
| EDIFÍCIO                   | Descrição   |
| Nome                       | Nome pelo qual o edifício onde a pintura se encontra é conhecido.   |
| Tipo de edifício           | Tipo de edifício. Pode ser: igreja, palácio, casa particular, sepulcro, museu, ou outro identificado posteriormente.  |
| Local no edifício          | Cada edifício se divide em cômodos internos. Aqui identifica-se em qual cômodo a pintura se encontra. Pode ser: nave, sacristia, capela mor, capela lateral, sala do trono, biblioteca, sala de jantar, sala de estar, quarto, ou outro identificado posteriormente.  |
| País                       | Localização patrimonial:<br>Como o endereço dos edifícios no Brasil e em Portugal é personalizado, aqui se fará a opção pelo país onde se encontra o edifício. Feito a escolha os campos de endereçamento específicos de cada país surgirão para serem preenchidos. Posteriormente outros países podem ser incluídos. |
| Estado                     | Bairro  |
| Município                  | Freguesia   |
| Distrito                   | Concelho  |
| Tombamento (S/N)           | Caso o edifício onde a pintura se encontra tenha sido tombado, escolha a opção S (Sim). Dessa forma será solicitado o nome da instituição tombadora e a data do tombamento.   |
| Instituição Tom-           | Instituição responsável pelo tom-   |

|                          |  |
|--------------------------|--|
| badora                   | bamento do edifício.   |
| Data do tombamento       | Data do tombamento.  |
| Restauração (S/N)        | Caso o edifício e/ou pintura tenha sido restaurado, escolha a opção S (Sim). Dessa forma será solicitado o nome da instituição restauradora e a data da restauração. |
| Instituição restauradora | Nome da instituição responsável pela restauração.  |
| Data da restauração      | Data da restauração.   |

| REFERÊNCIAS      | Descrição   |
|------------------|---|
| Fonte documental | Documento que atesta a autoria da pintura a um pintor. Pode ser inserida a referência onde esse documento foi coletado, ou pode ser inserida uma imagem em PDF desse documento. |
| Bibliografia     | Lista de referências de estudos realizados sobre essa pintura.  |
| Foto da pintura  | Imagem da pintura feita pelo fotógrafo.   |
| Gravura fonte    | Gravura de onde a imagem pintada foi copiada.   |

#### ENTIDADES SECUNDÁRIAS

| AUTORIAS                   | Descrição  |
|----------------------------|--|
| Nome                       | Nome de pessoa relacionada à pintura. Pode ser o criador, o atribuidor, o fotógrafo.   |
| Data de nascimento e morte | Período em que o pintor viveu. Preencher somente no caso da pessoa ser o criador da pintura. Quando inserido o nome do criador, essa informação será inserida automaticamente na descrição da pintura.     |
| Data de atividade          | Período em que o pintor trabalhou. Preencher somente no caso da pessoa ser o criador da pintura. Quando inserido o nome do criador, essa informação será inserida automaticamente na descrição da pintura. |

| SUPORTE | Descrição  |
|---------|--|
| Nome    | Nome da estrutura predial onde a pintura foi realizada. Pode ser: parede, abóbada, cúpula, teto plano, caixotão, ou outro identificado posteriormente. |

| MATERIAL | Descrição  |
|----------|--|
| Nome     | Nome da estrutura predial onde a pintura foi realizada. Pode ser: parede, abóbada, cúpula, teto plano, caixotão, ou outro identificado posteriormente. |

| TÉCNICA | Descrição   |
|---------|---|
| Nome    | Nome da técnica de pintura utilizada na pintura. Pode ser: óleo, têmpera, estuque, afresco, ou outra identificada posteriormente. |

| TIPO DE EDIFÍCIO | Descrição                                 |
|------------------|---|
| Nome             | Tipo de edifício. Pode ser: igreja, palá- |

cio, casa particular, sepulcro, museu, ou outro identificado posteriormente.

| LOCAL NO EDIFÍCIO | Descrição  |
|-------------------|--|
| Nome              | Cada edifício se divide em cômodos internos. Aqui identifica-se em qual cômodo a pintura se encontra. Pode ser: nave, sacristia, capela mor, capela lateral, sala do trono, biblioteca, sala de jantar, sala de estar, quarto, ou outro identificado posteriormente. |

| INSTITUIÇÕES | Descrição   |
|--------------|---|
| Nome         | Nome de instituições vinculadas à imagem, podem ser responsáveis pelo tombamento do edifício ou pela restauração. |

#### c. Dicionário de dados

#### 2.3. O padrão de metadados Dublin Core- DC

O projeto da base de imagens sobre pinturas em tetos tem a meta de disponibilizá-la na internet para que pesquisadores e interessados no tema possam ter acesso a essas informações compiladas.

Dessa forma, identificou-se a necessidade de um padrão específico para a descrição dessas imagens em uma base de dados digital. Nesse sentido, a área da biblioteconomia vem estudando e desenvolvendo padrões de metadados que possam ajudar na construção bibliotecas digitais, para as mais diversas áreas e temáticas.

Dentre os padrões de metadados que tem sido estudados e desenvolvidos, o padrão Dublin Core tem apresentado uma proposta bem flexível para a organização e tratamento da informação em ambiente digital.

O Dublin Core é uma recomendação do Dublin Core Metadata Initiative, e vem sendo estudado, atualizado e ampliado com base nas aplicações desenvolvidas a partir dos seus 15 elementos nucleares.

A proposta do DC é que existem 15 elementos básicos para a descrição de objetos digitais, aplicáveis a maioria das situações informacionais. Como esses 15 elementos não são suficientes para descrever especificidades, o DC propõe a utilização de qualificadores vinculados a cada um desses 15 elementos. Os 15 elementos do DC são:

Title - título do objeto

Creator - responsáveis pelo conteúdo intelectual do objeto

Subject - tópico relacionado ao objeto descrito

Description - contém uma descrição textual do objeto

Publisher - agente responsável por tornar o objeto disponível

Contributor - outros "autores" do conteúdo intelectual do objeto

Date - data de publicação

Type - tipo do objeto

Format - formato de dado do objeto

Identifier - identifica o recurso de forma única

Source - objetos dos quais o objeto descrito é derivado

Language - idioma relativo ao conteúdo intelectual do objeto

Relation - indica um tipo de relacionamento com outros objetos

Coverage - localização espacial e duração temporal do objeto

Rights - contém referencia ou direitos de propriedade

Um projeto que decida utilizar o DC, pode tanto utilizar os qualificadores já julgados, validados e autorizados pelo *DCMI Usage Committee*, como podem especificar novos qualificadores adequados às suas necessidades informacionais.

Segundo Dublin (2013), para que os qualificadores especificados por um projeto possam ser validados e incorporados à lista oficial de qualificadores do DC, é importante que se observe duas questões específicas. Os qualificadores devem se enquadrar em uma das duas grandes classes de qualificadores:

Refinamento de elemento - são qualificadores de refinam, ou especificam variáveis possíveis de cada elemento nuclear do DC, como por exemplo, o elemento *Date*, pode usar os qualificadores *created*, *valid*, *available*, *issued*, *modified*, entre outros. Dessa forma teremos data de criação, data de validação, data de disponibilização, data de emissão e data de modificação de um objeto informacional.

Esquema de codificação - são qualificadores que identificam esquemas que ajudam na identificação de um valor específico do elemento do DC. Podem ser vocabulários controlados, regras gramaticais ou uma norma que padroniza o preenchimento do campo.

Além disso, cada qualificador utilizado precisa conter as propriedades básicas que irão definir cada um deles que são:

Name - Símbolo único que identifica o qualificador.

Label - rótulo legível por pessoas que nomeiam o qualificador.

Definition - uma declaração que representa o conceito e a natureza essencial do qualificador.

Comment - Informações adicionais sobre o qualificador (opcional)

See also - um link com mais informações sobre o qualificador (opcional)

Além da proposta de qualificadores, um projeto específico pode também mesclar vários elementos ou etiquetas, oriundos de vários padrões de metadados distintos, uma das grandes possibilidades atualmente das linguagens de marcação da *web*, através da garantia de interoperabilidade entre sistemas diversos.

### 3. Aspectos metodológicos e norteadores

A base contendo informações sobre as fotos do inventário do prof. Magno, já existia e vinha sendo alimentada de forma não padronizada por diversos bolsistas. Assim, continha muitas inconsistências e registros inseridos com informações despadronizadas o que tornava a recuperação da informação ineficiente.

Para a melhoria da representação descritiva e temática da base de dados e sua posterior conversão em uma biblioteca digital, foi realizada uma análise da ficha descritiva das fotos que existem em papel, da estrutura da base de dados já existente e, ao mesmo tempo, a busca por outras bases de dados de sistemas de documentação museológica. Ressalta-se um ponto muito relevante neste projeto que é o caráter inédito de organização e estruturação das informações do acervo de documentos referentes à pintura de perspectiva, bem como a construção do banco de dados com todos os conhecimentos gerados a partir da elaboração do inventário deste acervo.

Feito esse levantamento dos metadados descritivos necessários para a representação dessa informação, foi realizada a modelagem conceitual com o uso da ferramenta MER. Através dessa modelagem, delineou-se uma estrutura descritiva mais adequada, tanto em termos tecnológicos, quanto em relação à catalogação dessas fotos.

Com a modelagem a base de dados pode ser reorganizada e padronizada. Após a modelagem, iniciaram-se as atividades de validação dos dados e dos aspectos de usabilidade dessa base. Esta atividade foi realizada em dois momentos diferentes: 1) com o especialista do domínio trabalhado e que representa o usuário técnico e 2) com alunos do curso de museologia representando os usuários não especializa-

dos do domínio, mas com interesse específico no assunto.

A atividade de validação foi operacionalizada com o especialista do domínio para a verificação e validação do modelo conceitual com o objetivo de confirmar cada conceito, o seu grau de representatividade e compreensão, bem como avaliar as entidades e os relacionamentos especificados. Estas atividades foram realizadas em 6 encontros agendados com o professor Magno Mello, com as autoras do trabalho e com o analista de sistema durante os meses de agosto a setembro para as discussões de validação do modelo conceitual e, posteriormente para a implementação e para os ajustes técnicos de desenvolvimento do banco de dados.

A validação, também ocorreu com 40 alunos do 3º período da disciplina documentação museológica do curso de museologia com o objetivo de avaliar o grau de dificuldade da entrada de dados no sistema e a compreensão dos campos do formulário para facilitar ao máximo a entrada de dados, posteriormente. Esta atividade foi realizada durante o horário de aula da referida disciplina, na biblioteca da ECI/UFMG, onde cada aluno usou um computador com acesso a internet para proceder à entrada de dados. A validação da entrada de dados contou com a presença da professora responsável pela disciplina e da equipe de desenvolvimento deste projeto. Esta equipe foi responsável por disponibilizar a ficha de dados padrão usada por todos os envolvidos nesta etapa. Tal ficha contém informações sobre uma pintura de teto para a entrada de dados pelos alunos. Após esta entrada de dados os alunos acessaram e preencheram o formulário eletrônico (Avaliação da etapa de Validação do formulário de entrada de dados - *Perspectiva Pictorum Digital*). Este formulário foi elaborado com o objetivo de coletar as impressões dos alunos durante a entrada de dados e possui as seguintes características: foi construído no *Google Docs* para facilitar o acesso de todos, contém 10 perguntas abertas e fechadas que abordam basicamente sobre a organização do formulário de entrada de dados, as dúvidas e dificuldades no preenchimento, bem como sobre as questões de funcionalidades do formulário.

#### 4. Fase 1: estruturação da base de dados de imagens

Nessa primeira fase, o objetivo é desenvolver uma base de dados para organização do inventário já existente, em uma estrutura que facilite a alimentação das informações na base de dados. As FIG.2 e 3 apresentam o *layout* das telas de

entrada de dados, que foram criadas pela equipe responsável para a alimentação.

#### 5. Fase 2: Interface em Dublin Core para uma biblioteca digital

Inicialmente, é importante ressaltar que a escolha do Dublin Core como padrão de metadados para a estruturação de uma interface da base de imagens para ser disponibilizada em forma de uma biblioteca digital, se deu porque este padrão faz parte marcante nas pesquisas da área da biblioteconomia. Além disso, é uma proposta flexível que permite personalizações necessárias a aplicações em acervos específicos.

Após a modelagem de dados, os atributos identificados para a base de dados em questão, foram utilizados para a construção de uma estrutura completa de DC que pudesse atender às necessidades dos objetos de informação a serem organizados e tratados.

Fez-se uma relação com os 15 elementos básicos do DC e os campos descritivos identificados na modelagem de dados. Posteriormente identificaram-se os qualificadores necessários para especificar cada um desses elementos e se fosse o caso, quais os esquemas de codificação necessários para cada elemento descritivo, conforme indicado nos QUADROS 1 (anexo 1) e 2.

The image shows a web form titled "Iconografia" with a navigation bar at the top containing tabs: "Iconografia", "Edifício", "Referências", "Autoria", and "Confirmar". The "Iconografia" tab is active. The form contains the following fields:

- Título
- Legenda
- Criador
- Data de nascimento e Morte
- Data de Atividade
- Autoria Atribuída? (radio buttons for Sim and Não)
- Nome do Atribuidor
- Suporte
- Material
- Técnica
- Dimensão
- Fotógrafo

Figura 2 – Layout da tela 1 de entrada de dados

Figura 3 – Layout da tela 2 de entrada de dados

| Campo do DC | Qualificador recomendado | Qualificador proposto   | Esquema proposto  |
|-------------|--------------------------|---|---|
| Title       | -----                    | Legend  | -----   |
| Creator     | -----                    | -----   | -----   |
| Contributor | -----                    | AttributedFor<br>Photographer<br>RegistrationInstitution<br>RestoringInstitution  | -----   |
| Date        | -----                    | BirthDeath<br>ActivityDate<br>Registration<br>Restoration   | -----   |
| Format      | Medium<br>Extent         | Material<br>Technique   | Tipo de suporte<br>Material do suporte<br>Técnica de pintura<br>Tamanho em m <sup>2</sup> |
| Source      | -----                    | DocumentalSource<br>EngravingSource   | -----   |
| Relation    | isReferencedBy           | -----   | -----   |
| Coverage    | -----                    | SpatialBuilding<br>BuildingType<br>PlaceBuilding<br>SpatialCountry<br>SpatialState<br>SpatialMunicipalDistrict<br>SpatialDistrict<br>SpatialNeighborhood<br>SpatialClientele<br>SpatialConcelho | Tipo de edifício<br>Local no edifício   |

Quadro 2 - Estrutura de DC proposta

### 5.1. Propriedades dos qualificadores propostos

Para que os qualificadores aqui propostos para a utilização do DC na biblioteca digital de imagens de falsa arquitetura, possa vir a ser valida-

da e reconhecida pelo grupo responsável pela iniciativa Dublin Core, é importante que as propriedades de cada qualificador proposto fiquem bem definidas. Dessa forma, essas propriedades foram delineadas no QUADRO 3.

## 6. Considerações Finais

O projeto aqui apresentado está na implementação da fase 1, etapa fundamental para auxiliar na organização do conjunto de documentos relacionados ao inventário das pinturas de tetos de edifícios do século XVIII. Os próximos passos compreendem um esforço subsequente de reflexão sobre a importância do processo de padronização e tratamento de informação, com a finalidade de promover a recuperação da informação e, a geração de conhecimentos pelos estudiosos da área.

Para a construção da base de dados, o grupo de pesquisa conta com alunos bolsistas do curso de Sistema de Informação da UFMG. Com a estruturação e alimentação dessa base em andamento, a fase 2, relativa à interface de uma biblioteca digital, já poderá ser iniciada. A perspectiva é de que as duas primeiras fases sejam concluídas quase que simultaneamente.

Com parte do inventário organizado e disponibilizado na *web*, outras informações serão alimentadas e a fase 3 poderá ser iniciada. Com o desenvolvimento de um tesouro para a área de pinturas em tetos e paredes, espera-se que a recuperação dessas informações, tanto na base de dados, quanto da biblioteca digital possa ser melhorada, tornando-se o mais eficiente possível.

## Referências

- Chen, P (1990). *Modelagem de dados: a abordagem entidade-relacionamento para projeto lógico*. São Paulo : McGraw Hill.
- Dublin Core Metadata Initiative (2011). *Dublin Core Qualifiers (SUPERSEDED, see DCMI Metadata Terms [documents/dcmi-terms/]*. Recuperado 09-19-2013 de <http://dublincore.org/documents/2000/07/11/dcmes-qualifiers>
- Dublin Core Metadata Initiative (2011). Elementos do núcleo metadata “*Dublin Core*”, versão 1.1: descrição de referência. Recuperado 08-22-2003 de <http://bnd.bn.pt/ed/dcmes/dcmes11-20000518.htm>
- Ferrez, Helena Dodd (1994). Documentação museológica: teoria para uma boa prática. *Cadernos de ensaios: estudos de museologia*. Riode Janeiro. IPHAN, 1994, p.65-74.
- Fusco, E. (2010). *Modelos conceituais de dados como parte do processo da catalogação*: perspectiva de uso dos FRBR no desenvolvimento de catálogos bibliográficos digitais. Tese (Doutorado em Ciência da Informação) – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília.

Yassuda, Sílvia Nathaly (2009) . *Documentação museológica: uma reflexão sobre o tratamento descritivo do objeto no Museu Paulista*. 123 f. Dissertação (Mestrado em Ciência da Informação) – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, 2009.

---

# OntoSmart: proposta de um modelo de recuperação de informação baseado em ontologia

*OntoSmart: proposal for an ontology-based information retrieval model*

---

**Edberto Ferneda (1), Guilherme Ataíde Dias (2)**

(1) Universidade Estadual Paulista – UNESP-Marília. Departamento de Ciência da Informação.  
Av. Hygino Muzzi Filho, 737 - 17.525-000 – Marília - SP, ferneda@marilia.unesp.br

(2) Universidade Federal da Paraíba – UFPB. Departamento de Ciência da Informação  
Campus I - Jardim Cidade Universitária - 58.051-900 – João Pessoa - PB, guilhermeataide@gmail.com

## Resumo

Um sistema de recuperação de informação é um ambiente linguístico mediador da comunicação entre um estoque de informação e seus requisitantes. Sua eficiência depende de um controle adequado da representação dos documentos e das requisições de seus usuários. No contexto da Ciência da Informação o tesouro se consolidou como uma ferramenta bastante eficiente na representação da informação. Porém, as ontologias surgem como uma nova tecnologia que auxiliam na representação e na organização da informação e do conhecimento. Esta pesquisa tem por objetivo desenvolver um modelo de recuperação de informação baseado em ontologia que utiliza como estrutura formal o Modelo Espaço Vetorial. Os vetores que representam os documentos são criados durante o processo de indexação automática. A partir de um conjunto inicial de termos extraídos dos documentos, procede-se uma inferência na ontologia com a finalidade de enriquecer a indexação. De forma semelhante, a expressão de busca de um usuário é também representada por um vetor, criado a partir de um processo de expansão de consulta por meio de inferências em uma ontologia. Utilizando o modelo proposto, deu-se início ao desenvolvimento de um sistema de recuperação de informação denominado OntoSmart, cujos resultados parciais apontam para um aumento significativo na precisão da recuperação.

**Palavras-chave:** Recuperação de Informação; Ontologia; Modelo Espaço Vetorial; Indexação Automática; Expansão de Consulta.

## 1. Introdução

Um sistema de recuperação de informação é um ambiente linguístico cuja eficiência depende de um controle adequado da representação dos itens de informação e das requisições de seus usuários. Insere-se como um elemento mediador da comunicação entre um estoque de informação e os seus potenciais requisitantes.

Em seu papel de mediador de um processo comunicativo, é tarefa do sistema definir um código ou uma linguagem comum entre emissor

## Abstract

An information retrieval system is a linguistic environment that mediates the communication between a stock of information and its requesters. Their efficiency depends on an adequate control of representation of the documents and of the requests of its users. In the context of Information Science, thesaurus has established as a very efficient tool in the representation of information. However, ontologies are emerging as a new technology that assist the representation and organization of information and knowledge. This research aims to develop an ontology-based information retrieval model that uses the formal structure of the Vector Space Model. The vectors representing the documents are created during the automatic indexing process. From an initial set of terms extracted from the documents there shall be an inference in the ontology in order to enrich the indexing. Similarly, a user's query is also represented by a vector created from a query expansion process through inferences in ontology. Using the proposed model, it was initiated the development of an information retrieval system called OntoSmart, whose partial results point to a significant increase in accuracy of retrieval.

**Keywords:** Information retrieval; Ontology; Vector Space Model; Automatic Indexing; Query Expansion.

e receptor, entre os conteúdos informacionais dos documentos e as requisições dos usuários. Na Ciência da Informação, tradicionalmente as linguagens documentárias são utilizadas como um elemento comutador entre a informação e o usuário que a necessita. Para Fujita (2004), as linguagens documentárias são um conjunto controlado de termos que visam representar os assuntos tratados pelos documentos e utilizados na fase de indexação e busca. Essas linguagens promovem a convergência entre a lingua-

gem do indexador e a linguagem do usuário de um sistema de recuperação informação.

Tálamo, Lara e Kobashi (1992, p.197) apontam que:

As Linguagens Documentárias são tradicionalmente consideradas instrumentos de controle terminológico que atuam em dois níveis: a) na representação da informação obtida pela análise e síntese de textos; b) na formulação de equações de busca da informação.

Na década de 1970, Salton (1972) já propunha métodos de construção de tesouros para serem utilizados em sistemas de recuperação de informação. Segundo Salton e McGill (1983, p.75), apresentado por meio de uma interface adequada, um tesouro poderia ajudar o usuário a elaborar suas buscas, ao mesmo tempo em que o familiariza com o vocabulário utilizado pelo sistema.

A partir da década de 1990 o termo “ontologia” começa a ser frequentemente referenciado na área da Ciência da Computação. O tema tomou notoriedade ainda maior e se expandiu para outras áreas com o surgimento do projeto da Web Semântica, na qual as ontologias aparecem como parte de destaque na sua estrutura.

Muitos trabalhos tratam das diferenças e semelhanças entre tesouros e ontologias (Codina; Pedraza-Jiménez, 2011; Kless; Milton, 2011; Sales; Café, 2009; Jiménez, 2004;). Dentre as semelhanças, pode-se destacar que: (1) ambos têm como objetivo representar e compartilhar os conceitos ou o vocabulário de um domínio a fim de possibilitar uma comunicação eficiente; (2) as suas estruturas básicas são hierárquicas, agrupando termos ou conceitos em categorias e subcategorias (classes e subclasses); (3) ambas podem ser utilizadas para catalogar ou organizar recursos informacionais.

Na Ciência da Computação, a recuperação de informação baseada em ontologia (*ontology-based information retrieval*) já é um campo de pesquisa consolidado, com um grande número dissertações e teses defendidas em diversos países. Tais trabalhos abordam uma diversificada gama de propostas e abordagens para a utilização de ontologias no processo de recuperação de informação.

O sistema CIRI (AIRIO *et al*, 2004) utiliza ontologias na indexação de documentos, criação e expansão de consultas. Inicialmente o usuário escolhe a ontologia relacionada ao seu interesse de busca e seleciona os termos em uma representação hierárquica e visual dos conceitos da ontologia escolhida. A partir de um conjunto inicial de termos escolhidos pelo usuário, o

sistema expande automaticamente a consulta, considerando os relacionamentos entre os conceitos da ontologia.

O sistema OnAIR (PAZ-TRILLO; WASSERMANN; BRAGA, 2005) é um sistema de recuperação de trechos de vídeos a partir de consultas em linguagem natural. Foi testado utilizando um conjunto de entrevistas com a artista brasileira Ana Teixeira. Para esse objetivo foi desenvolvida uma ontologia sobre arte contemporânea.

Os trechos de vídeo são indexados por meio de palavras-chave atribuídas por um especialista do domínio e por palavras contidas na transcrição do vídeo. A partir das consultas em texto livre, o sistema extrai termos relevantes e elimina palavras de pouca importância semântica. Para cada termo é atribuído um peso em função da frequência no *corpus* e de sua ocorrência na ontologia. A expansão das consultas é feita com a utilização dos conceitos e das relações da ontologia.

O sistema OntoSeek (Guarino; Masolo; Vetere, 1999) é um sistema de recuperação de informação baseado na descrição de produtos disponíveis em páginas amarelas e catálogos *on-line*. A descrição dos produtos e as consultas dos usuários são representadas por meio de grafos conceituais derivados de ontologias. Assim, o problema de recuperação de informação se reduz à equiparação (*matching*) de grafos. Os nós e arcos de um grafo que representa uma consulta são comparados aos nós e arcos de um grafo que representa um produto.

O sistema OWLIR (FININ *et al*, 2005) recupera documentos textuais contendo marcações semânticas provenientes do próprio texto e de uma ontologia. Tais marcações auxiliam no processo de indexação dos documentos, melhorando o desempenho da recuperação de informação.

Esse sistema utiliza uma ontologia sobre eventos de uma universidade e foi aplicado sobre um *corpus* de páginas de anúncios de eventos desta mesma universidade. Inicialmente são extraídos termos das páginas visando identificar os tipos de eventos tratados na coleção. O sistema, então, anota as páginas utilizando informações extraídas dos textos, acrescidas do conhecimento inferido na ontologia. Em seguida é realizada a indexação dos documentos anotados. A ontologia é utilizada também na expansão das consultas dos usuários.

O sistema FROM (Pereira; Ricarte; Gomide, 2006) implementa o modelo ontológico relacional *fuzzy* para recuperação de informação textual. O sistema faz a expansão da consulta con-

siderando as relações existentes em uma ontologia de domínio composta por categorias e palavras-chaves. As categorias denotam os conceitos mais gerais e as palavras-chaves denotam conceitos mais específicos. Uma consulta do usuário pode ser composta apenas por palavras-chaves, por categorias ou por ambas. A expansão da consulta é feita pela adição de novas categorias e palavras-chaves, em função das conexões existentes na ontologia. A similaridade dos documentos em relação à consulta é calculada por meio de operações *fuzzy*, e são recuperados os documentos que apresentarem similaridade acima de um determinado valor.

Este trabalho propõe um modelo de recuperação de informação baseado em ontologia, denominado OntoSmart, no qual as ontologias, vistas como vocabulários controlados, são utilizadas como ferramentas de padronização do vocabulário tanto das representações dos documentos como das buscas dos usuários. Utiliza como alicerce teórico e prático o Modelo Espaço Vetorial (Salton, 1968), que permite incorporar diversos métodos e técnicas desenvolvidas ao longo de mais de quatro décadas de pesquisas nesse modelo. O OntoSmart possui muitas características semelhantes aos sistemas citados, porém se distingue por uma abordagem relativamente simples e intuitiva.

## 2. Ontologia

Até a última década do século XX, ontologia era considerada primariamente uma disciplina da Filosofia. Atualmente, as ontologias têm sido utilizadas de diferentes maneiras em diversas áreas, e vêm ocupando cada vez mais a atenção de estudiosos da Ciência da Informação e da Ciência da Computação, tendo em vista a possibilidade de melhorar significativamente a representação de um domínio do conhecimento.

O tema tem despertado o interesse de inúmeros pesquisadores da área, principalmente após a criação de fóruns temáticos tal como a série de conferências FOIS (*Formal Ontology and Information Systems*), em meados da década de 1990. Porém, somente a partir de 2001 é que se observa uma grande quantidade de trabalhos relacionados ao tema (Guizzardi, 2005, p.56).

Uma definição de ontologia muito citada é a de Gruber (1993), que descreve uma ontologia como uma “especificação formal explícita de uma conceitualização compartilhada”. Por *formal* entende-se que esta especificação seja expressa num formato legível por computadores; *explícita* significa que os conceitos, as propriedades, as relações, as funções, as restrições e os axiomas devem estar formalmente

definidos e passíveis de serem manipulados por computadores. Entende-se por *conceitualização* que tal representação seja referente a algum modelo abstrato de algum fenômeno do mundo real. Por *compartilhada*, compreende-se que esse conhecimento seja consensual (Gruber, 1995; Fensel, 2001; Borst, 1997).

Uma ontologia pode ser considerada um vocabulário de representação, geralmente especializado em algum domínio ou assunto, qualificado por conceitualizações de tipos de objetos e suas relações no mundo. Em outras palavras, é um corpo de conhecimento que descreve algum domínio, utilizando um vocabulário de representação (Chandrasekaran; Josephson; Benjamin, 1999).

Segundo Jacob (2003, p.19):

Ontologias são categorias de coisas que existem ou podem existir em um determinado domínio particular, produzindo um catálogo onde existem as relações entre os tipos e até os subtipos do domínio, provendo um entendimento comum e compartilhado do conhecimento de um domínio que pode ser comunicado entre pessoas e programas de aplicação.

Uschold (1996) ressalta a necessidade de explicitação dos relacionamentos entre os termos de uma ontologia:

Uma ontologia pode possuir uma variedade de formas, mas necessariamente incluirá um vocabulário de termos e alguma especificação de seus significados. Isto inclui definições e uma indicação de como conceitos estão inter-relacionados, o que impõe uma estrutura no domínio e restringe as possíveis interpretações dos termos.

Uma ontologia define os conceitos usados em uma determinada área de conhecimento, padronizando seus significados. Pode ser usada por pessoas, bases de dados e aplicações que precisam compartilhar informações e conceitos de um domínio (Daconta; Obrst; Smith, 2003, p.167). Ramalho (2010, p.38) apresenta resumidamente, os componentes de uma ontologia:

- **Classes e Subclasses:** As classes e subclasses de uma ontologia agrupam um conjunto de elementos, “coisas”, do “mundo real”, que são representadas e categorizadas de acordo com suas similaridades, levando-se em consideração um domínio concreto. Os elementos podem representar coisas físicas ou conceituais, desde objetos inanimados até teorias científicas ou correntes teóricas;

- **Propriedades Descritivas:** Descrevem as características, adjetivos e/ou qualidades das classes;
- **Propriedades Relacionais:** Trata-se dos relacionamentos entre classes pertencentes ou não a uma mesma hierarquia, descrevendo e rotulando os tipos de relações existentes no domínio representado;
- **Regras e Axiomas:** Enunciados lógicos que possibilitam impor condições como tipos de valores aceitos, descrevendo formalmente as regras da ontologia e possibilitando a realização de inferências automáticas a partir de informações que não necessariamente foram explicitadas no domínio, mas que podem estar implícitas na estrutura da ontologia;
- **Instâncias:** Indicam os valores das classes e subclasses, constituindo uma representação de objetos ou indivíduos pertencentes ao domínio modelado, de acordo com as características das classes, relacionamentos e restrições definidas;
- **Valores:** Atribuem valores concretos às propriedades descritivas, indicando os formatos e tipos de valores aceitos em cada classe.

A construção de uma ontologia pode ser pensada como uma união de peças que formam uma estrutura completa. Classes e subclasses definem um “esqueleto” na forma de uma hierarquia que pode ser expressa na forma de uma árvore ou de um grafo, complementada por propriedades descritivas, propriedades relacionais, regras e axiomas. A sua abrangência (domínio) deve ser previamente definida e estabelece uma área do conhecimento ou uma parte do mundo que se pretende tratar.

Toda classe é caracterizada por seus atributos ou propriedades. Uma subclasse herda as características (atributos) da classe pai. Uma instância é a materialização de uma classe e representa um conceito ou uma entidade do mundo real. Quando uma classe é instanciada, cada um dos seus atributos pode então receber valores que irão individualizar aquele conceito ou entidade. É possível estabelecer regras que impõem restrições e limites às classes e atributos, e que se refletem nas suas instâncias.

Uma ontologia é, enfim, uma estrutura conceitual que visa representar formalmente os conceitos e suas relações, regras e restrições lógicas de um determinado domínio.

Na Ciência da Informação, segundo Soergel (1999) e Vickery (1997), o termo ontologia começou a ser utilizado no final da década de 1990, principalmente por pesquisadores da área

de Organização do Conhecimento. Nessa época, os instrumentos e métodos de classificação passaram a despertar um maior interesse de pesquisadores da comunidade de Ciência da Computação, devido principalmente à necessidade de desenvolvimento de instrumentos de organização da informação no ambiente Web.

A Organização do Conhecimento vem se consolidando como um importante campo de investigação da Ciência da Informação a partir da fundação da *International Society for Knowledge Organization* (ISKO), em 1989, quando as principais ações para a consolidação da área foram adotadas.

Para Esteban Navarro (1996) a Organização do Conhecimento é a disciplina da Ciência da Informação que se dedica ao estudo dos fundamentos teóricos do tratamento e recuperação da informação, avaliando o uso de instrumentos lógico-linguísticos para controlar os processos de representação, classificação, ordenação e armazenamento do conteúdo informativo dos documentos com a finalidade de permitir sua recuperação e disseminação.

Segundo Ramalho (2010, p.37):

Entre os instrumentos de representação tradicionalmente utilizados na área de Ciência da Informação, os tesouros apresentam-se como os que possuem maior aproximação com as ontologias, devido ao fato de ambos os instrumentos serem constituídos por meio de linguagens de estruturas combinatórias, de caráter especializado, representando termos e conceitos organizados a partir de tipos de relacionamentos.

Ao longo dos últimos anos, inúmeros estudos comparativos entre ontologias e tesouros têm constatado que, apesar de possuírem características comuns, tais instrumentos caracterizam-se como diferentes modelos de representação do conhecimento. Enquanto os tesouros são desenvolvidos como ferramentas de auxílio para os usuários na busca de informações, as ontologias têm como principal objetivo descrever formalmente os recursos informacionais para possibilitar a realização de inferências automáticas.

O autor acrescenta ainda que:

As ontologias proporcionam liberdade para representar tipos de relacionamentos que não seriam possíveis em outros modelos de representação, podendo ser concebidas a partir de diversas técnicas de organização do conhecimento (Ramalho, 2010, p.37).

Outro fator determinante para a distinção das ontologias e os modelos de representação tradicionalmente utilizados na Ciência da Informação é a própria natureza dos relacionamentos. Para possibilitar a realização de inferências automáticas é necessário que as relações expressas na

ontologia sejam formalmente descritas, rotuladas e categorizadas. As ontologias proporcionam liberdade para representar tipos de relacionamentos que não seriam possíveis em outros modelos de representação, podendo ser concebidas a partir de diversas técnicas de organização do conhecimento, cabendo aos desenvolvedores importantes decisões no momento da modelagem.

As ontologias se colocam como um novo instrumento a ser incorporado ao arsenal teórico e prático da Ciência da Informação. A aprendizagem de novos conceitos e novos recursos oferecidos pelas ontologias é um desafio para os profissionais da informação, mas que pode ser facilmente enfrentado utilizando toda bagagem teórica acumulada durante a história da Ciência da Informação.

### 3. Ontologias na Recuperação de Informação

Existem dois elementos linguísticos que afetam diretamente na eficiência de um sistema de recuperação de informação: a representação dos documentos e a representação da expressão de busca. As ontologias se inserem no processo de recuperação de informação com o objetivo de prover um maior nível semântico das representações dos documentos e das necessidades de informação dos usuários.

A representação dos documentos de um *corpus* é feita por meio da indexação, que visa descrever o conteúdo informacional de cada documento por meio de um conjunto de termos extraído do texto do próprio documento ou selecionado de um elemento auxiliar de padronização terminológica. As ontologias podem desempenhar um papel importante no processo de indexação por meio da disponibilização de uma estrutura conceitual e terminológica contextualizada em determinado domínio de conhecimento.

A representação adequada das necessidades de informação dos usuários é também um fator determinante para a eficiência de um sistema de recuperação de informação. A tradução da necessidade de informação em uma expressão de busca envolve elementos difíceis de serem formalizados. Um usuário não familiarizado com a terminologia de uma área do conhecimento ou de um determinado assunto de seu interesse tenderá a expressar sua necessidade de informação utilizando termos muito genéricos ou coloquiais, o que pode resultar na recuperação de um número excessivo de documentos não relevantes. A utilização de uma ontologia no processo de especificação de buscas permite derivar novos termos e agregá-los automática-

mente à expressão de busca inicial do usuário, em um processo conhecido como “expansão de consulta” (*query expansion*).

Segundo Ferneda (2012, p.20):

Um modelo de recuperação de informação é a especificação formal de três elementos: a representação dos documentos, a representação da expressão de busca e como esses dois elementos serão comparados, a função de busca. A eficiência de um sistema de recuperação de informação está diretamente ligada ao modelo que utiliza, influenciando diretamente em seu modo de operação.

No modelo OntoSmart utiliza-se o Modelo Espaço Vetorial como estrutura básica para a representação dos documentos e das expressões de busca, assim como diversas técnicas derivadas das pesquisas nesse modelo.

### 4. Modelo Espaço Vetorial

Como mencionado anteriormente, o modelo OntoSmart tem como base o Modelo Espaço Vetorial. No Modelo Vetorial (Salton, 1968) um documento é representado por um vetor no qual cada elemento determina o peso ou a importância do respectivo termo na representação do conteúdo informacional do documento. Cada elemento do vetor é normalizado de forma a assumir valor entre zero (0) e um (1).

Uma expressão de busca (consulta) é também representada por um vetor numérico onde cada elemento representa o grau de relevância do respectivo termo na necessidade de informação do usuário.

A utilização de uma mesma representação tanto para os documentos como para as expressões de busca permite calcular o grau de similaridade entre uma determinada busca e cada um dos documentos do *corpus*. Em um espaço vetorial contendo  $N$  dimensões, a similaridade (*sim*) entre um documento  $d_i$  e uma expressão de busca  $q$  é calculada utilizando a seguinte fórmula:

$$sim(d_i, q) = \frac{\vec{d}_i \bullet \vec{q}}{|\vec{d}_i| \times |\vec{q}|} = \frac{\sum_{k=1}^N (w_{k,i} \times w_{k,q})}{\sqrt{\sum_{k=1}^N w_{k,i}^2} \times \sqrt{\sum_{k=1}^N w_{k,q}^2}}$$

onde  $w_{k,i}$  é o peso do  $k$ -ésimo elemento do vetor que representa o documento  $d_i$  e  $w_{k,q}$  é o peso do  $k$ -ésimo elemento do vetor da expressão de busca  $q$ .

Os valores da similaridade entre uma expressão de busca e cada um dos documentos do *corpus* são utilizados no ordenamento dos documentos resultantes. Esse ordenamento (*ranking*) permi-

te agregar a um sistema alguns parâmetros que permitem restringir o resultado a um número máximo de documentos ou determinar um limite mínimo para o valor da similaridade dos documentos resultantes de uma determinada busca.

## 5. Indexação Automática baseada em Ontologia

A indexação de um documento visa representar o seu conteúdo informacional por meio de um conjunto de termos com o objetivo de sintetizar o seu conteúdo, ressaltando o que lhe é essencial. Um termo de indexação é constituído de uma ou mais palavras cujo significado remete pretensamente a um conceito único, não ambíguo. Os termos de indexação servem também como pontos de acesso mediante os quais o documento é localizado e recuperado em um sistema de informação.

Embora a prática da indexação possa ser regulada por políticas e princípios institucionais, o processo de indexação realizada por seres humanos (manual) é dependente de critérios subjetivos e pessoais relacionados à formação e à experiência do indexador. Assim, o tempo despendido e a qualidade da indexação ficam fortemente atrelados a fatores não controláveis, o que pode afetar o custo desse processo.

As dificuldades inerentes à indexação manual e a grande quantidade de documentos derivados da explosão informacional após a Segunda Guerra, justificaram estudos que buscavam soluções alternativas para auxiliar o indexador no exercício de sua atividade. As primeiras pesquisas em indexação automática aconteceram no final dos anos de 1950, época de rápido desenvolvimento da Ciência da Computação. O surgimento dos microcomputadores na década de 1980 e da Web nos anos de 1990 fez com que o nível de interesse nas pesquisas sobre indexação automática permanecesse praticamente constante até os dias de hoje.

Os argumentos contra a indexação automatizada estão centrados na capacidade inerente do ser humano em tratar com a linguagem. Para um ser humano as palavras deixam de ser meros dados vazios de significado e tornam-se formas de representação mental de elementos do conhecimento. Assim, um indexador humano, utilizando o seu conhecimento e sua bagagem cultural, pode reconhecer os diferentes significados de uma palavra ou frase em seus diferentes contextos. Tais significados, convertidos em novos termos de indexação, proporcionam uma melhoria na representação dos documentos de um *corpus*, melhorando, por conse-

guinte, a eficiência e a eficácia do processo de recuperação de informação.

Particularmente, as ontologias abrem novas perspectivas para as pesquisas em indexação automática, pois oferecem uma estrutura conceitual e terminológica restrita a um determinado domínio e originalmente representadas em linguagens processáveis por computador, o que permite a sua utilização nos mais variados processos computacionais.

A utilização de uma base ontológica possibilita uma abordagem mais rica para a indexação, pois permite oferecer algum tipo de análise semântica. Essa análise pode ser efetuada a partir dos textos dos documentos, onde são identificados e selecionados termos que possam ser mapeados para os conceitos de uma determinada ontologia. Esse mapeamento permite padronizar o vocabulário e restringir o campo semântico dos termos, contextualizando-os ao domínio da ontologia, solucionando assim possíveis ambigüidades.

## 6. Expansão de Consultas baseada em Ontologia

Embora importante para uma recuperação eficiente, a especificação da busca (consulta) é dependente do usuário, com toda a variabilidade e imprecisão inerente ao ser humano. Além disso, geralmente as buscas dos usuários são expressas por meio de um número reduzido de termos, não permitindo uma interpretação exata e inequívoca da necessidade de informação do usuário.

A importância e as dificuldades do processo de especificação de buscas fizeram surgir na área de Recuperação de Informação um campo de pesquisa em expansão de consulta. Expansão de consulta (*query expansion*) é o termo utilizado para referenciar os métodos e processos que visam melhorar a eficiência da recuperação de informação baseados no pressuposto de que as consultas definidas pelos usuários muitas vezes não refletem suas reais necessidades de informação. O objetivo principal é adicionar novos termos à consulta inicialmente formulada pelo usuário a fim de melhorar os resultados obtidos.

Uma ontologia pode ser utilizada na expansão de consultas por meio da inserção de novos termos de busca derivados dos relacionamentos entre os seus conceitos. Além disso, a partir de uma interface adequada, as ontologias podem servir também como ferramentas para a seleção dos termos que irão compor a consulta do usuário. Isso permite que uma pessoa leiga em um determinado assunto consiga realizar consultas

pertinentes, ao mesmo tempo em que se familiariza com a terminologia do domínio de interesse.

Utilizando-se conceitos de indexação baseada em ontologia e de expansão de consulta baseada em ontologia, foi desenvolvido um modelo e um sistema de recuperação denominado OntoSmart.

## 7. O modelo OntoSmart

O modelo OntoSmart utiliza a estrutura formal do Modelo Espaço Vetorial, associado ao uso de ontologias como ferramenta de normalização da terminologia durante o processo de criação dos vetores representativos dos documentos e das buscas. A seguir serão definidos alguns conceitos básicos do modelo.

### 7.1. Distância semântica

Uma ontologia  $O = (C, R)$  é composta por um conjunto de conceitos  $C = \{c_1, c_2, \dots, c_n\}$  interconectados por um conjunto de relacionamentos em  $R = \{r_1, r_2, \dots, r_n\}$ . A distância semântica (**ds**) entre dois conceitos de uma ontologia ( $c_1$  e  $c_2$ ) é igual ao número de relacionamentos existentes no menor caminho entre  $c_1$  e  $c_2$ .

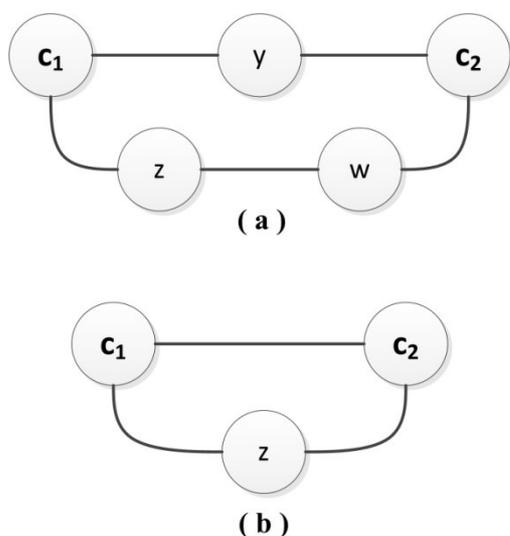


Figura 1. Ilustração do conceito de distância semântica (**ds**)

Na Figura 1a, o menor caminho entre  $c_1$  e  $c_2$  é passando pelo conceito  $y$  e dois relacionamentos ( $c_1, y$ ) e ( $y, c_2$ ). Portanto,  $ds(c_1, c_2) = 2$ . Na Figura 1b  $ds(c_1, c_2) = 1$ , pois os conceitos  $c_1$  e  $c_2$  são adjacentes, separados por um único relacionamento.

O valor de **ds** entre um conceito e ele próprio é igual a zero. Assim, por exemplo,  $ds(c_1, c_1) = 0$  e  $ds(c_2, c_2) = 0$ .

### 7.2. Valor semântico

Tomando-se como referência um conceito  $c$  de uma ontologia, pode-se inferir que exista uma progressiva degradação do nível semântico dos conceitos a ele relacionados à medida que a distância semântica (**ds**) vá aumentando.

Dado um conceito  $c$  de uma ontologia, o valor semântico (**vs**) de cada conceito ( $c_i$ ) é calculado da seguinte forma:

$$vs(c_i, c) = 1 - [ds(c_i, c) \times p]$$

onde  $p$  é um parâmetro numérico, entre 0 e 1, que define a diferença dos valores de **vs** a cada distância **ds** de um dado conceito  $c_i$  em relação a  $c$ .

Quanto maior a distância semântica (**ds**) do conceito central  $c$ , menor será o valor semântico (**vs**) de um dado conceito da ontologia. O parâmetro  $p$  define a diferença dos valores de **vs** a cada valor de **ds**. A Figura 2 apresenta uma ilustração da aplicação dos conceitos de distância semântica (**ds**) e valor semântico (**vs**), considerando o parâmetro  $p = 0.2$ .

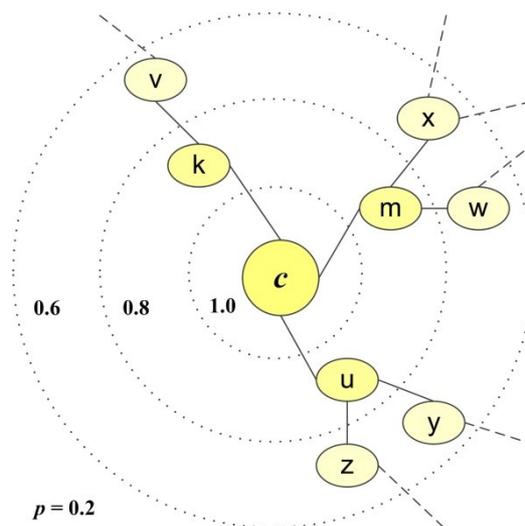


Figura 2. Ilustração do conceito de valor semântico (**vs**)

A definição de um conceito central em uma ontologia faz surgir diversos níveis ou “camadas” concêntricas, onde cada camada é definida pela distância semântica (**ds**) em relação ao conceito central. Os conceitos de uma mesma camada recebem o mesmo valor semântico (**vs**). O conceito central possui **vs** igual a 1 e os demais conceitos terão **vs** menores, de acordo com a

camada que ocupam e com o valor do parâmetro  $p$ . Considerando  $c$  o conceito central e  $p=0.2$ , os conceitos da Figura 2 terão os seguintes valores:

| Conceito        | $ds$ | $vs$ |
|-----------------|------|------|
| $c$             | 0    | 1.0  |
| $k, m, u$       | 1    | 0.8  |
| $v, x, w, y, z$ | 2    | 0.6  |

Na Figura 2 foram apresentadas apenas três camadas de uma ontologia genérica. Com parâmetro  $p$  igual a 0.2, cada camada corresponde a um valor decrescente de  $vs$ , variando de 1 a 0.6. Considerando que o valor de  $vs$  não pode ser negativo e que uma ontologia pode ter um grande número de conceitos, é necessário definir um parâmetro  $k$  que limite o número de camadas a serem consideradas no cálculo de  $vs$ .

Os valores dos parâmetros  $p$  e  $k$  são interdependentes. Não faz sentido, por exemplo,  $p = 0.2$  e  $k = 8$ , pois isso acarretaria valores negativos de  $vs$ . Portanto, o valor máximo que o parâmetro  $p$  pode assumir ( $p_{max}$ ) é igual  $1/k$ . De maneira formal temos:

$$p_{max} = \frac{1}{k}$$

No exemplo da Figura 2 o valor de  $k$  é igual a três ( $k=3$ ). Portanto, o valor máximo que o parâmetro  $p$  pode assumir é 0.33 ( $p_{max}=0.33$ ).

### 7.3. Indexação dos documentos

No OntoSmart um documento é representado por um único vetor numérico no qual cada elemento representa a importância (peso) do respectivo termo na representação do documento. Porém, diferentemente do Modelo Vetorial, no modelo OntoSmart os pesos são calculados por meio da ontologia associada ao documento. A indexação de cada documento é realizada em duas fases: extração de termos e expansão dos índices.

Inicialmente será extraído do documento um conjunto de termos que represente o seu conteúdo informacional. Para cada termo é atribuído um valor numérico (peso) que expressa a relevância do respectivo termo na representação do conteúdo informacional do documento. A extração de termos e o cálculo de seus pesos são realizados por meio de algum método extração automática. Ao final desse processo poderia se obter, por exemplo, os seguintes termos de indexação e seus respectivos pesos:



|               |             |
|---------------|-------------|
| <b>Leão</b>   | <b>0.9</b>  |
| <b>Girafa</b> | <b>0.85</b> |
| <b>Zebra</b>  | <b>0.8</b>  |
| Macaco        | 0.5         |
| Floresta      | 0.4         |
| Savana        | 0.35        |

Nesse exemplo foram extraídos do documento seis termos, mas considerados apenas aqueles com peso igual ou superior à um determinado valor, no caso 0.8. Dessa forma, o documento será representado por três termos: “Leão”, “Girafa” e “Zebra”, que serão utilizados para realizar inferências em uma ontologia.

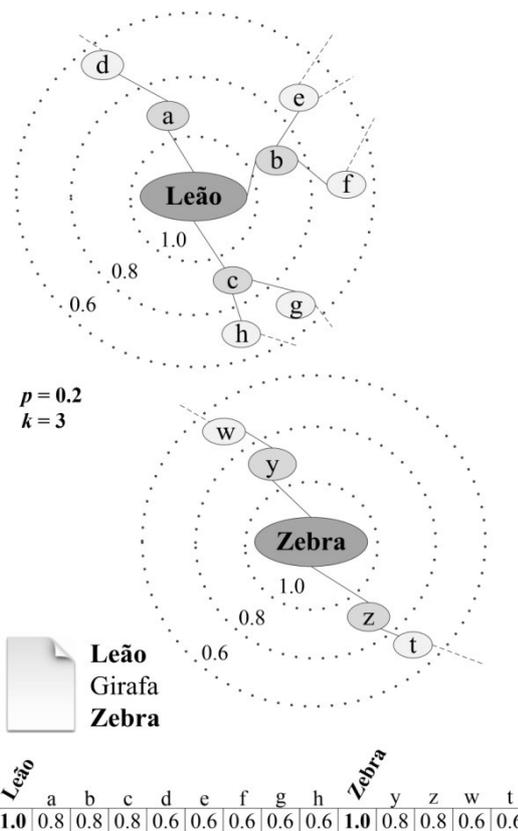


Figura 3 - Representação vetorial de um documento utilizando ontologia

Na Figura 3, pode-se verificar que apenas os termos “Leão” e “Zebra” têm relação com conceitos da ontologia. Assim, esses termos farão parte do vetor que representa o documento, com valor semântico ( $vs$ ) igual a 1 (um). Os demais termos que irão compor o vetor do documento serão derivados desses dois termos coincidentes, por meio suas relações.

Tomando-se “Leão” como conceito central, deriva-se os demais termos de indexação, observando o valor de  $vs$  para cada camada da ontologia. A diferença dos valores de  $vs$  em cada

camada é dado pelo parâmetro  $p$ , que no exemplo possui valor igual a 0.2. Assim, os termos **a**, **b** e **c** receberão o valor 0.8. Os termos **d**, **e**, **f**, **g** e **h** terão valor igual a 0.6.

Considerando agora “Zebra” como o conceito central, os conceitos **y** e **z** serão considerados termos de indexação do documento, ambos com  $vs$  igual a 0.8. Os conceitos **w** e **t** terão  $vs$  igual a 0.6.

O termo “Girafa” foi descartado por não estar representado por um conceito da ontologia. Porém, há de se considerar que esse termo foi extraído do texto do documento por método estatístico que lhe atribuiu um peso relativamente alto. Esses termos serão armazenados em um tipo de repositório, formando um conjunto de potenciais conceitos a serem inseridos na ontologia.

#### 7.4. Expressão de busca

Uma expressão de busca é representada por um único vetor numérico no qual cada elemento corresponde à importância do respectivo termo para a descrição da necessidade de informação do usuário.

Antes da execução da busca, o usuário deve selecionar a ontologia do domínio ao qual se refere a sua necessidade de informação. Os parâmetros  $k$  e  $p$  devem estar previamente definidos antes da expansão da sua busca.

Os termos definidos pelo usuário na sua expressão de busca serão utilizados como conceitos centrais da ontologia associada à essa busca. A ontologia terá duas funções: (1) expandir o conjunto inicial de termos, acrescentando novos termos; (2) atribuir pesos a cada um dos termos. Essas funções tomam como base a distância dos termos inicialmente definidos na busca e que se encontram diretamente representados na ontologia.

Considere uma expressão de busca na qual o usuário utilizou dois termos: “Leão” e “Girafa”. Fazendo-se uma pesquisa na ontologia, verifica-se que apenas o primeiro termo está representado na ontologia. Assim, no vetor que representará esta busca apenas o termo “Leão” estará presente, com peso igual a 1 (um).

Tomando-se “Leão” como conceito central e considerando os parâmetros  $p=0.2$  e  $k=2$ , derivam-se os termos **a**, **b** e **c**, que farão parte da expressão de busca expandida. Tais termos receberão o valor 0.8, como exemplificado na Figura 4.

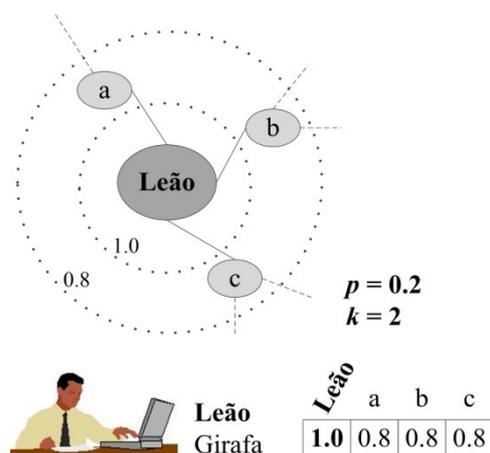


Figura 4. Representação de uma expressão de busca

O termo “Girafa”, que não está presente na ontologia, da mesma forma que nos documentos, será armazenado em um tipo de repositório. Se esse termo for frequentemente utilizado pelos usuários ele pode vir a ser integrado na ontologia, de acordo com critérios previamente determinados em um sistema.

## 8. Considerações Finais

No modelo de recuperação proposto, elementos linguísticos que formam uma ontologia são considerados termos de um vocabulário de domínio, utilizado como ferramenta de padronização terminológica das representações dos documentos e das buscas em um sistema de recuperação de informação. Tais representações utilizam como base formal o Modelo Espaço Vetorial, que fornece uma base matemática consistente e consolidada.

Uma vantagem evidente do modelo de recuperação proposto é a delimitação explícita do contexto no qual o processo de recuperação de informação é realizado. No OntoSmart o domínio de um documento é explicitamente definido por meio de sua relação com uma determinada ontologia. Os documentos são indexados utilizando o vocabulário de domínio definido pelos conceitos dessa ontologia. Por sua vez, antes de expressar sua necessidade de informação o usuário define o seu domínio de interesse por meio da seleção de uma ontologia, que será utilizada para agregar novos termos à expressão de busca inicialmente formulada por ele. O Modelo Vetorial fornece a estrutura formal de representação tanto para os documentos como para as buscas, o que permite fornecer como resultado uma lista de documentos ordenados pelo grau de similaridade/relevância.

Percebe-se que o modelo OntoSmart possui uma simetria na forma de se construir os veto-

res que representam os documentos e os vetores das expressões de busca. Tal característica permitiu uma redução significativa do tempo e do esforço de desenvolvimento do sistema OntoSmart, pois permitiu racionalizar um considerável compartilhamento de código fonte.

O sistema OntoSmart, uma implementação do modelo de mesmo nome, ainda está em desenvolvimento, mas já apresenta resultados bastante expressivos no que se refere à precisão dos resultados de uma busca. Alguns desses resultados foram apresentados por FERNEDA e DIAS (2013).

## 9. Agradecimentos

Este trabalho é parte do resultado de pesquisa de pós-doutorado financiada pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) por meio do projeto PROCAD-NF-099/2009.

## Referências

- AIRIO, E.; JÄRVELIN, K.; SAATSI, P.; KEKÄLÄINEN, J.; SUOMELA, S. CIRI – an ontology-based query interface for text retrieval. In: HYVÖNEN, E.; KAUPPINEN T.; SALMINEN, M.; VILJANEN, L.; ALA-SIURU, P. (Eds) Proceedings of the 11th Finnish Artificial Intelligence Conference, 2004.
- BORST, W.N. Construction of Engineering Ontologies for Knowledge Sharing and Reuse. 1997. Tese (Doutorado). Centre for Telematics for Information Technology, University of Twente, Enschede, 1997.
- CHANDRASEKARAN, B.; JOSEPHSON, J.R.; BENJAMINS, V. R. 1999. What are ontologies, and why do we need them? IEEE Intelligent Systems, v.14, n.1, 1999.
- CODINA, L.; PEDRAZA-JIMÉNEZ, R. Tesoros y Ontologías en Sistemas de Información Documental. El profesional de la Información, v.20, n.5, 2011.
- DACONTA, M.C.; OBRST, L.J.; SMITH, K.T. The Semantic Web: a guide to the Future of XML, Web Services, and Knowledge Management. Indianápolis: Wiley Publishing, 2003.
- MEADOW, C.T.; BOYCE, B.R.; KRAFT, D.H.; BARRY, C. (2007). Text Information Retrieval System. 3rd ed. London UK: Elsevier, 2007.
- CINTRA, A. M. M. (Org.) (2002). Para entender as linguagens documentárias. 2.ed. São Paulo: Polis, 2002.
- DACONTA, M.C.; OBRST, L.J.; SMITH, K.T. The Semantic Web: a guide to the Future of XML, Web Services, and Knowledge Management. Indianápolis: Wiley Publishing, 2003.
- DING, Y.; FOO, S. Ontology research and development. Part 1- a review of ontology generation. Journal of Information Science, v.28, n. 2, 2002.
- ESTEBAN NAVARRO, M.A. El marco disciplinar de los lenguajes documentales: la Organización del Conocimiento y las ciencias sociales. Scire, Zaragoza, v.2, n.1, 1996.
- FININ, T.; MAYFIELD, J.; JOSHI, A.; COST, R.S.; FINK, C. Information retrieval and the semantic web. In: Proceedings of the 38th Annual Hawaii International Conference on System Sciences (HICSS'05). Washington: IEEE Computer Society, 2005.
- FENSEL, D. Ontologies: a silver bullet for knowledge management and electronic commerce. Springer, 2001.
- FERNEDA, E. Introdução aos Modelos Computacionais de Recuperação de Informação. Rio de Janeiro: Ciência Moderna, 2012.
- FERNEDA, E.; DIAS, G.A. Um Método de Expansão Automática de Consulta Baseada em Ontologia. In: XIV Encontro Nacional de Pesquisa em Ciência da Informação (ENANCIB). Florianópolis-SC, 2013.
- FUJITA M. S. L. (2004). A leitura Documentária na Perspectiva de suas Variáveis: leitor-texto-contexto. DataGramaZero: Revista de Ciência da Informação, Rio de Janeiro, v.5, n.4, ago. 2004.
- GRUBER, T. A Translation Approach to Portable Ontology Specifications. Knowledge Acquisition, v.6, n.2, 1993.
- GRUBER, T. Toward Principles for the Design of Ontologies Used for Knowledge Sharing. International Journal Human-Computer Studies v.43, n.5-6, 1995.
- GUARINO, N.; MASOLO, C.; VETERE, G. Ontoseek: Content-based access to the web. IEEE Intelligent Systems, v.14, n.3, 1999.
- GUIZZARDI, G. Ontological Foundations for Structural Conceptual Models. The Netherlands: Universal Press, 2005.
- JACOB, E. K. Ontologies and the semantic web. Bulletin of the American Society for Information Science and Technology, apr./may,, 2003.
- JIMÉNEZ, A.G. Instrumentos de Representación del Conocimiento: tesauros versus ontologías. Anales de Documentación, n.7. Universidad de Murcia, 2004.
- KLESS, D.; MILTON, S. Comparison of thesauri and ontologies from a semiotic perspective. In: Proceedings of the Sixth Australasian Ontology Workshop. Conferences in Research and Practice in Information Technology. Advances in Ontologies. Adelaide, Australia: Australian Computer Society, 2010.
- PAZ-TRILLO, C.; WASSERMANN, R.; BRAGA, P.P. An information retrieval application using ontologies. Journal of the Brazilian Computer Society, v.11, n.2, 2005.
- PEREIRA, R.; RICARTE, I.; GOMIDE, F. Fuzzy relational ontological model in information search systems. In: SANCHEZ, Elie (Ed.). Fuzzy Logic and The Semantic Web, p.395–412, Elsevier B.V.: Amsterdam, 2006.
- QIN, J.; PALING, S. Converting a controlled vocabulary into an ontology: the case of GEM. Information Research, v.6, n2, 2000-01.
- RAMALHO, R.A.S. Desenvolvimento e utilização de ontologias em Bibliotecas Digitais: uma proposta de aplicação. Tese (Doutorado em Ciências da Informação) – Universidade Estadual Paulista, 2010.
- SALES, R.; CAFÉ, L. Semelhanças e Diferenças entre Tesoros e Ontologias. DataGramaZero, Rio de Janeiro, v.9, n.4, ago. 2008.
- SALES, R.; CAFÉ, L. Diferenças entre tesauros e ontologias. Perspectivas em Ciência da Informação. v.14, n.1, 2009.
- SALTON, G. Automatic information organization and retrieval. New York: McGraw-Hill Book Company, 1968.
- SALTON, G. Experiments in Automatic Thesaurus Construction for Information Retrieval. In: FREIMAN, C. V.; GRIFFITH, J.E.; ROSENFELD, J.L. (eds.) Information Processing 71: Proceedings of IFIP Congress 71, v.1. North-Holland, 1972.
- SALTON, G.; MCGILL, J.M. Introduction to Modern Information Retrieval. New York, McGraw-Hill, 1983.
- SOERGEL, D. The rise of ontologies or the reinvention of classification. Journal of the American Society for Information Science. v. 50, n. 12, 1999.

TÁLAMO, M.F.G.M.; LARA, M.L.G.; KOBASHI, N.Y. Contribuição da terminologia para a elaboração de tesouros. *Ciência da Informação*, v.21, n.3, 1992.

USCHOLD, M. Building Ontologies: Towards a Unified Methodology. proceedings of Expert Systems. In: Annual Conference of the British Computer Society Specialist Group on Expert Systems, 16, p.16-18 December, 1996, Cambridge. *Anais...*,Cambridge, 1996

VICKERY, B. C. Ontologies. *Journal of Information Science*. v.23, n.4, 1997

---

# Estudo da comunicabilidade da interface e do comportamento de busca da informação no Portal de Periódicos da CAPES

*Study of communicability of interface and information seeking behavior in the CAPES Journal Portal*

---

**Gracielle Mendonça Rodrigues Gomes (1), Beatriz Valadares Cendón (2), Raquel Oliveira Prates (3)**

- (1) Doutoranda do Programa de Pós-Graduação em Ciência da Informação da Universidade Federal de Minas Gerais, Avenida Presidente Antônio Carlos, 6627 – Pampulha – Minas Gerais – Brasil – Cep: 31270-901, [graciellemendonca@yahoo.com.br](mailto:graciellemendonca@yahoo.com.br).
- (2) Professora Titular do Programa de Pós-Graduação em Ciência da Informação da Universidade Federal de Minas Gerais, [cendon@eci.ufmg.br](mailto:cendon@eci.ufmg.br)
- (3) Professora Associada do Programa de Pós-Graduação em Ciência da Computação da Universidade Federal de Minas Gerais, [rprates@dcc.ufmg.br](mailto:rprates@dcc.ufmg.br)

## Resumo

Desde o seu lançamento, o Portal de Periódicos da CAPES passou por modificações em sua interface, *layout* e recursos de busca num esforço constante para facilitar a navegação e melhorar os serviços oferecidos aos usuários. Este estudo almejou integrar as áreas da Ciência da Informação e da Ciência da Computação para analisar o processo de busca da informação no Portal de Periódicos da CAPES, identificar quais problemas na interface e no sistema de busca que impedem ou dificultam a interação e a busca de informação, e propor possíveis soluções para aperfeiçoar este sistema. A investigação caracteriza-se como um estudo de caráter exploratório e descritivo, sob a abordagem metodológica qualitativa. Teve como sujeitos da pesquisa alunos de doutorado das áreas de Ciências Humanas, Linguística, Letras e Artes, Ciências Biológicas, Ciências Exatas e bibliotecários da Universidade Federal de Minas Gerais (UFMG). Na coleta de dados, aplicou-se um método de avaliação realizado por especialistas, denominado Método de Inspeção Semiótica (MIS), e um método de avaliação com usuários, que integra o Método de Explicitação do Discurso Subjacente (MEDS) com fundamentos de modelos que representam a busca de informação em SRIs. A partir dos resultados, foi possível identificar os problemas existentes; observar táticas, julgamentos e mudanças de estratégias realizadas durante um episódio de busca; conhecer o que o usuário pensa a respeito da interface e o sistema de busca, propor recomendações para melhoria do sistema, além de contribuir para o desenvolvimento da área de Recuperação da Informação Interativa.

**Palavras chave:** Portal de Periódicos da CAPES. Sistemas de Recuperação da Informação - Avaliação e Uso. Comportamento de Busca de Informação em SRI. Satisfação do Usuário. Recuperação da Infor-

mação Interativa. Interação Humano-Computador. Método de Inspeção Semiótica.

## Abstract

Since its launch, the CAPES Portal of E-Journals underwent modifications in its interface, layout and search capabilities in a constant effort to facilitate navigation and improve the services offered to users. This study integrated knowledge from the areas of information science and computer science to analyze the information search process in the CAPES Portal of E-Journals, identify problems in the interface and the search system that prevent or hinder the interaction and the search for information, and propose possible solutions to improve this system. This research is characterized as an exploratory and descriptive study, under a qualitative approach. The research subjects were doctoral students in the areas of Humanities, Linguistics, Literature and Arts, Life Sciences, Exact Sciences and librarians of the Federal University of Minas Gerais. Data collection applied the Semiotics Inspection Method, an evaluation method performed by specialists, and an evaluation method by users, which includes the Underlying Discourse Unveiling Method (UDUM) grounded on information seeking models. From the results, it was possible to identify existing problems; observe tactics, judgments and changes in the strategies used during an episode of search; know what users think about the interface and the search system, make recommendations to improve the system and contribute to the development of the area of Interactive Information Retrieval.

**Keywords:** CAPES Portal of E-Journals. Information Retrieval Systems – Evaluation and Use. Information Search Behaviour. User Satisfaction. Interactive Information Retrieval. Human-Computer Interaction. Semiotic Inspection Method.

## 1. Introdução

O interesse em compreender as experiências dos usuários na busca de informação em Sistemas de Recuperação da Informação (SRI), particularmente no ambiente *Web*, implica em avaliar a sua interface, portanto a comunidade de pesquisa na área de Recuperação de Informação (RI) tem realizado o esforço de avaliar as interfaces de busca de informação para compreender o comportamento dos usuários e tornar mais fácil a elaboração de estratégias de pesquisas permitindo tanto aos usuários novatos quanto aos experientes realizarem uma pesquisa mais rápida e eficiente.

Para avaliar e aperfeiçoar SRIs, em conformidade com as exigências dos usuários, é necessário compreender como estes utilizam e interagem com o sistema. Além disso, é relevante mensurar se os resultados alcançados são satisfatórios, bem como identificar o grau de satisfação com os serviços encontrados e quais recursos o sistema oferece para facilitar a localização da informação desejada e caso a encontre, se a mesma é eficiente, útil e confiável.

No Brasil, o Portal de Periódicos da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) foi desenvolvido para disponibilizar às instituições de ensino e pesquisa, a produção científica nacional e internacional, através do acesso remoto. De acordo com o site eletrônico da CAPES (2015), “o Portal de Periódicos atende às demandas dos setores acadêmico, produtivo e governamental e propicia o aumento da produção científica nacional e o crescimento da inserção científica brasileira no exterior. É, portanto, uma ferramenta fundamental para fomento, avaliação e regulação dos cursos de Pós-Graduação e desenvolvimento da pesquisa científica no Brasil.”

O Portal de Periódicos da CAPES procura constantemente otimizar sua interface, ampliar seus serviços de busca e incorporar novas ferramentas para facilitar a navegação, além de fazer investimentos para modernizar a infraestrutura, aumentar a visibilidade e a qualidade dos serviços oferecidos. Em virtude desse contexto, colocou-se como pergunta de pesquisa: como o Portal de Periódicos da CAPES tem oferecido suporte às atividades de interação, busca e recuperação da informação de seus usuários?

## 2. Objetivos

Para investigar a questão, esta pesquisa propôs atender o objetivo geral e os objetivos específicos a seguir.

### 2.1. Objetivo geral

Avaliar a qualidade da interface e do sistema de busca do Portal de Periódicos da CAPES, e a experiência dos usuários durante os processos de interação, busca e recuperação da informação com este sistema.

### 2.2. Objetivos específicos

- a) Integrar áreas da Ciência da Informação e da Ciência da Computação para alcançar uma análise mais completa do processo de busca e recuperação da informação no Portal de Periódicos da CAPES;
- b) Identificar quais problemas existem na interface e no sistema de busca que impedem ou dificultam o seu uso;
- c) Avaliar a experiência dos usuários durante os processos de interação, busca e recuperação da informação;
- d) Apresentar soluções para aperfeiçoar o Portal de Periódicos da CAPES;
- e) Melhorar a compreensão dos processos de interação do usuário com, com vistas a contribuir para o desenvolvimento teórico da Ciência da Informação na área de Recuperação da Informação Interativa.

## 3. Fundamentação Teórica

Neste tópico, são apresentadas as bases teóricas que alicerçaram a pesquisa.

### 3.1. Comportamento de Busca e Recuperação da Informação das Comunidades Científicas

As atividades de pesquisa científica têm sido beneficiadas pela infraestrutura existente para o uso das tecnologias da informação e comunicação (TICs) e estas foram incorporadas intensamente às várias atividades do meio acadêmico, para aumentar a velocidade da disseminação das pesquisas, o intercâmbio das ideias, a colaboração entre pesquisadores, o desenvolvimento das pesquisas, a geração de conhecimentos e ampliação do acesso aos diversos recursos de informação criados favorecendo a expansão das comunidades científicas.

Diversas fontes para a busca de informações estão disponibilizadas *online*. Lopes e Silva (2007) declaram que as obras de referência (como, por exemplo, bibliografias, índices e resumos) estão sendo substituídas por bases de dados digitais. Paralelamente surgiram as bibliotecas virtuais, que facilitam o processo de busca e o acesso à informação pelos cientistas, pois disponibilizam documentos eletrônicos para

consulta em tela ou mediante impressão e os repositórios digitais que operam com a tecnologia e a filosofia dos arquivos abertos (*open archives*), e estão em processo de consolidação e reconhecimento como fontes confiáveis de informação científica.

O processo de busca da informação de forma *online* garante a rapidez na obtenção de resultados, elimina barreiras de espaço e tempo, diminui a dependência de intermediários, sejam eles instituições ou pessoas. Portanto, os profissionais da informação precisam trabalhar no sentido de aumentar a disponibilização de recursos *online* para pesquisa, como também têm que investir esforços no treinamento para tornar os usuários mais autônomos, pois, na atualidade, os usuários dependem muito mais de serviços de organização e acesso à informação e de interfaces amigáveis que facilitem a busca, a localização e a recuperação, do que dos próprios recursos humanos e físicos existentes nas unidades de informação, de acordo com Lopes e Silva (2007).

Conforme Abels, Liebscher e Denman (1996), os fatores que impedem ou incentivam a adoção das TICs pelos pesquisadores estão relacionados ao sistema, referindo-se à acessibilidade (proximidade, experiência, facilidade de uso e disponibilidade de equipamentos); aos fatores pessoais e profissionais (questões de comportamento na busca da informação, nas quais estão incluídas as diferenças entre áreas do conhecimento, tarefas desenvolvidas e utilidades percebidas); e aos fatores institucionais, relacionados à infraestrutura, no que se refere aos investimentos e à manutenção dos equipamentos e redes, bem como à realização de treinamentos.

Portanto, com o avanço das TICs, várias mudanças estão ocorrendo na comunicação científica permitindo serem cada vez mais eficientes e rápidas, além de vencer barreiras geográficas, hierárquicas e financeiras. Novos formatos e canais de comunicação se tornaram disponíveis, possibilitando novas possibilidades de comunicação entre os cientistas. O entendimento do comportamento e das necessidades informacionais dos cientistas neste contexto e o modo como as tecnologias da informação impactam em suas atividades permite avaliar e promover serviços de informação mais adequados para atender suas demandas atuais e o desenvolvimento da ciência.

### 3.2. Interação Humano-Computador fundamentada na Engenharia Semiótica

A área de Interação Humano-Computador (IHC) tem como uma de suas linhas de investigação a apreciação das qualidades de uso dos sistemas. Isso porque, conforme mencionado por Rogers, Sharp e Preece (2012), o sucesso dos sistemas computacionais depende fortemente do modo como os usuários se sentem ao utilizá-los e, para que isto ocorra, é necessário desenvolver tecnologias que realmente atendam as necessidades dos usuários, considerando suas diversidades e limitações.

De acordo com Barbosa e Silva (2010), para que o apoio computacional de um software seja satisfatório para o usuário, independente de suas condições físicas e/ou psicológicas, é desejável que contemple propriedades de qualidade de uso como: comunicabilidade, acessibilidade e usabilidade adequadas às novas formas de interação e de relações sociais.

A Engenharia Semiótica tem seu foco na avaliação da qualidade de uso denominada comunicabilidade, ou seja, na sua capacidade de transmitir aos usuários as decisões e princípios que guiaram seu design de forma eficiente e capaz, conforme cita Prates, Souza e Barbosa (2000). Neste sentido, quando o usuário não é capaz de entender a comunicação do projetista, ocorre então uma ruptura da comunicação, declara Souza (2005).

Prates e Barbosa (2007), afirmam que a Engenharia Semiótica é uma teoria explicativa de IHC que nos permite entender os fenômenos envolvidos no design, uso e avaliação de um sistema interativo. Do ponto de vista da Engenharia Semiótica, a interface de um sistema interativo é vista como uma comunicação indireta e unidirecional entre o projetista do sistema e o usuário final, ver Figura 1.



Figura 1. Metacomunicação designer-usuário e comunicação usuário-sistema

Segundo Souza (2005), o objetivo dessa comunicação é transmitir aos usuários a sua visão

sobre a quem o sistema se destina; que problemas ele pode resolver e como interagir com ele. Tal comunicação recebe o nome de metacomunicação e a esta metamensagem é transmitida do projetista para o usuário a partir do momento que o usuário interage com o sistema. Peirce (1992) explica que a metamensagem transmitida é composta por signos, que são definidos como qualquer coisa que representa algo para alguém.

Na Engenharia Semiótica, estes signos são classificados como signos metalinguísticos, estáticos e dinâmicos. Souza e Leitão (2009) explicam que são considerados *signos metalinguísticos*, aqueles signos que se referem a outros signos na interface do sistema, por exemplo, a documentação e a ajuda do sistema. Os *signos estáticos* são aqueles que expressam o estado do sistema e podem ser interpretados independentemente das relações causais ou temporais, por exemplo, o estado de botões, os elementos de interação utilizados na interface e opções selecionadas. Os *signos dinâmicos* são signos que só podem ser percebidos a partir do momento que há uma interação com o sistema, estes por sua vez, representam o comportamento do sistema e relacionam-se aos aspectos temporais e causais da interface, por exemplo, um botão Enviar tornar-se habilitado após a digitação de um texto no campo do e-mail, o impacto de se selecionar um determinado valor para um atributo ou a ação ocasionada pelo evento de disparo de um botão.

De acordo com Souza *et al.* (2006), um dos métodos qualitativos e interpretativos propostos pela Engenharia Semiótica para a apreciação da comunicabilidade, é o Método de Inspeção Semiótica (MIS), onde um avaliador com experiência em IHC e Engenharia Semiótica, inspeciona o sistema com o objetivo de identificar as potenciais rupturas de comunicação na emissão da metamensagem entre projetista-usuário.

Segundo Prates e Barbosa (2007), há outro método de avaliação baseado na Engenharia Semiótica diferente do MIS, denominado Método de Avaliação da Comunicabilidade (MAC). Este método que envolve a participação dos usuários em um ambiente controlado no qual um especialista acompanha e posteriormente analisa a interação do usuário com o sistema. A principal diferença do ponto de vista teórico entre o MIS e o MAC é que enquanto o MIS foca em como a mensagem é enviada pelo projetista através da interface, o MAC enfatiza em como esta mensagem está sendo recebida e entendida pelo usuário.

Além do método de avaliação MIS e MAC, em 2010, foi desenvolvido na pesquisa de Oliveira (2010) um novo método baseado também na Teoria da Engenharia Semiótica denominado Método de Inspeção Semiótica Intermediado (MISI). O MISI também tem como foco avaliar a capacidade de um sistema transmitir ao usuário os princípios de interação e decisões que o guiaram, isto é, a qualidade de comunicabilidade.

Para obter a análise, de acordo com Oliveira, Luz e Prates (2008), o MISI combina passos do MIS, para definir o roteiro da interação, com alguns passos do Método de Explicitação de Discurso Subjacente (MEDS), segundo Nicolacida Costa, Leitão e Romão-Dias (2004), que é um método de entrevista semiestruturado. Por ser baseado no MIS e no MEDS, ambos são métodos qualitativos e interpretativos, o MISI também é qualitativo e interpretativo.

O MEDS é um método qualitativo de pesquisa originalmente utilizado nas ciências humanas e sociais. Em IHC, o seu principal objetivo é tornar visível às preferências, dificuldades, desejos, dentre outros aspectos; importantes para o desenvolvimento de sistemas interativos. Assim, sua maior utilidade para a área de IHC é a possibilidade de captar o que não é perceptível por outros métodos. O MEDS prevê um roteiro semiestruturado de perguntas abertas, ou seja, não é necessário ter perguntas prontas, mas sim um roteiro, para garantir que os mesmos pontos serão explorados com todos os entrevistados. O avaliador pode aprofundar mais ou menos, de acordo com o que é dito pelo entrevistado, explica Oliveira (2010).

### 3.3. Integração das áreas de Recuperação da Informação, *Information Search Behaviour* e Interação Humano-Computador

Os pesquisadores da Ciência da Informação estão procurando demonstrar a importância da sinergia entre as áreas de Recuperação da Informação (RI), Comportamento de Busca da Informação e Interação Humano-Computador (IHC), ou pelo menos entre duas destas áreas para criar sistemas altamente interativos, alcançar avanços substanciais nas capacidades de busca de informação e navegação, criar interfaces que suportam adequadamente a formulação de estratégias e a reformulação de buscas, tornar mais fácil para os usuários recuperar a informação que precisam e auxiliar na compreensão do conjunto de informações localizadas, conforme MARCHIONINI (2006); WILDEMURTH (2006); COSTA (2008), KESHAVARZ (2008); MAMOON, EL-BAKRY E SALAMA

(2013); JOSEPH, DEBOWSKI E GOLDSCHMIDT (2013).

Ingwersen e Jarvelin (2005) analisaram os objetivos das investigações sobre Comportamento de Busca de Informação, Recuperação da Informação e Sistemas de Informação e observaram pouca sobreposição entre as três disciplinas, pois seguiram suas próprias agendas de pesquisa. Ainda concluíram que a pesquisa sobre o comportamento de busca de informação tem sido bem sucedida até agora no desenvolvimento de uma compreensão teórica do processo de busca e em fornecer uma compreensão empírica dos fenômenos subjacentes, mas ficou aquém no apoio a gestão da informação e no *design* de sistemas. A razão para isso está no foco da pesquisa, que geralmente exclui o sistema. Em contraste, quando o SRI está no centro da investigação e a interação é o foco da pesquisa, os usuários estão envolvidos, consegue-se identificar fatores básicos de interação, como por exemplo, o tempo de conclusão e número de documentos relevantes recuperados, mas falha em capturar a relação da qualidade da informação com o seu uso ou o processo que desencadeou a atividade de pesquisa.

O campo da RI tem dado pouca atenção às questões de interface e, por outro lado, a área de IHC, embora tenha amadurecido como disciplina, ainda não realizou a integração com a pesquisa na área da RI. Esperava-se que com as interfaces baseadas na *Web* haveria uma melhora na busca de informação para os usuários finais, entretanto, os estudos realizados sobre a usabilidade desses sistemas foram suficientes para demonstrar que as dificuldades fundamentais de recuperação da informação ainda permanecem (AHMED; MCKNIGHT; OPPENHEIM, 2009).

De acordo com Kelly e Sugimoto (2013), a combinação entre os campos da Recuperação de Informação, Comportamento de Busca de Informação e Interação Humano-Computador para formar uma especialidade única de pesquisa que visa capacitar as pessoas a explorar, resolver e gerenciar seus problemas de informações através das interações com os sistemas de informação é denominada de Recuperação de Informação Interativa (RII).

Esta área tem sido difundida através de várias instâncias práticas e acadêmicas. Os principais periódicos de RII são *Journal of the American Society for Information Science and Technology*, *Information Processing & Management*, *Journal of Documentation*, *Journal of Information Retrieval*, and *ACM Transactions on Information Systems*, todos os artigos regularmente presentes

na pesquisa de alta qualidade em RII assim como fazem os principais periódicos em Interação Humano-Computador, incluindo *ACM Transactions on Computer-Human Interaction*, *Interacting with Computers* e, em menor extensão, *Human-Computer Interaction* (RUTHVEN, 2008, p. 46).

De acordo com Belkin (2010), é bastante provável que o estado de conhecimento do usuário de informação se altere no decurso do episódio de busca de informação, levando novas ideias sobre o que pode ser útil e até como a pessoa poderia compreender o problema ou a tarefa que o levou a buscar informações. A perspectiva de avaliação do paradigma tradicional de RI não leva em conta estes tipos de alterações, pois não responde a natureza da RII, em termos dos tipos de metas que a busca de informação pressupõe, em termos de sua incapacidade de avaliar os episódios inteiros de busca de informação e em termos da sua incapacidade para levar em conta as alterações do pesquisador que são inerentes à busca de informação interativa.

As demandas por abordagens que considerem os aspectos humanos em recuperação da informação têm sido difundidas devido a predominantemente a três mudanças realizadas na literatura, apontadas por Robertson e Hancock-Beaulieu (1992) e Ingwersen (1992), que buscaram novas interpretações para caracterizar as necessidades informacionais: a) perspectiva da relevância, ou seja, a relevância de um objeto informacional deve ser medida em relação ao potencial para atender uma necessidade informacional; b) perspectiva cognitiva, no qual compreende que uma lacuna no conhecimento do usuário pode levá-lo a interagir com um SRI e de acordo com a gravidade desta lacuna, ele pode não conseguir elaborar satisfatoriamente uma consulta no sistema; e a c) perspectiva interativa, em que deve ser considerado não apenas a relação consulta-documentos, mas também as interações do usuário com o sistema e os documentos, além do comportamento das reformulações da consulta (MARTINS, 2009, p.42-43).

Conforme Ruthven (2008), a área de RII abarca a investigação relacionada com o estudo e o suporte aos diversos usuários finais para recuperar e acessarem a informação através de sistemas. A RII é formada pela pesquisa sobre: (a) a busca e o comportamento de busca de informação e (b) o desenvolvimento de novos métodos de interação com os recursos eletrônicos. Ambas as abordagens são importantes, enquanto a pesquisa sobre o comportamento de busca de informação fornece um panorama

sobre as decisões dos usuários envolvidas no processo e permite contextualizar a maior parte do trabalho em RII, a pesquisa sobre os métodos de interação com sistemas de busca promove novos entendimentos sobre os métodos adequados para facilitar ao usuário acessar a informação. Desta maneira, é importante perceber que o sucesso e a satisfação do usuário durante a busca no sistema não depende, inevitavelmente, apenas do que o sistema interativo oferece ou sobre como incentiva os pesquisadores a empregarem esses recursos, mas também depende de como o sistema suporta as estratégias pessoais do usuário e o conduz a entender como o sistema funciona. (RUTHVEN, 2008, p. 44).

A área de RII estabeleceu um novo rumo dentro da longa tradição de RI, ao longo das duas últimas décadas, que introduz o usuário em seu centro e coloca novos desafios para avaliação do sistema. O desempenho dos SRIs pode ser melhorado através do uso de informações sobre todo o processo interativo de pesquisa. No entanto, esta abordagem até agora foi só inicialmente explorada e possui grande potencial de pesquisa para o futuro (BIERING; GWIZDA; BELKIN, 2010, p. 673).

Portanto, o potencial de uma boa pesquisa nesta área não vem só a partir do conhecimento técnico no desenvolvimento de sistemas interativos, mas também a partir do conhecimento sobre o comportamento de busca de informação das pessoas e enquanto interagem com os SRIs. Atualmente, os pesquisadores que se interessam pelos estudos dos SRIs têm privilegiado a perspectiva do usuário, concentrando em suas necessidades e comportamentos de informação para projetar interfaces de busca que suportem ricas interações.

#### 4. Metodologia

Esta investigação trata-se de uma pesquisa qualitativa de caráter exploratório-descritivo e o objeto de estudo foi o Portal de Periódicos da CAPES, no contexto da Universidade Federal de Minas Gerais, no Brasil.

A abordagem adotada para a seleção dos participantes foi a não probabilística e a intencional. Foram selecionados 12 alunos de doutorado sendo 03 alunos de doutorado das áreas de Ciências Humanas, 03 alunos de doutorado das áreas Linguística, Letras e Artes, 03 alunos de doutorado das áreas de Ciências Biológicas, e 03 alunos de doutorado das áreas de Ciências Exatas e Engenharias e, ainda, 03 bibliotecários do Sistema de Bibliotecas da UFMG, totalizando 15 voluntários. De acordo com Prates e Barbosa

(2003), normalmente, nos estudos de avaliação de interface é recomendado à participação de 05 a 08 usuários. Quando o sistema é destinado a mais de um perfil de usuário, então se deve fazer uma avaliação envolvendo pelo menos 03 usuários de cada perfil avaliado.

Apesar do Portal de Periódicos da CAPES ter como objetivos atender docentes, pesquisadores e estudantes, fizeram parte desta investigação apenas alunos de doutorado, pois seu trabalho requer a dedicação de mais tempo para realização de pesquisas e precisam ter acesso, constantemente, a grande quantidade de informação científica de qualidade. Além disso, a intenção foi compor uma amostra de perfil semelhante, mas de áreas de conhecimento e características individuais diferentes. Já a escolha dos bibliotecários como sujeitos desta pesquisa teve como finalidade incluir a interpretação de especialistas de informação e de profissionais que recebem as demandas dos usuários e vivenciam as dificuldades que eles encontram para uso do Portal CAPES.

Para coleta de dados, combinou-se um método de inspeção de interface realizado por especialistas com um método de avaliação realizados com os usuários para conhecer seus interesses, objetivos e desafios com relação à interação e busca de informação no Portal de Periódicos da CAPES e, assim ter uma avaliação sobre dois pontos de vista.

##### 4.1. Método de Inspeção Semiótica

Para avaliar a interface do Portal de Periódicos da CAPES, o MIS foi aplicado por 02 especialistas de IHC e Engenharia Semiótica, sendo um dos avaliadores a própria pesquisadora, com o intuito de identificar a repetição de interpretações e enriquecer o relatório com visões distintas. As atividades relacionadas com o método de inspeção semiótica envolvem cinco etapas:

- Passo 1: Inspeção dos signos metalinguísticos
- Passo 2: Inspeção dos signos estáticos
- Passo 3: Inspeção dos signos dinâmicos
- Passo 4: Consolidação e contraste
- Passo 5: Apreciação da qualidade da metacomunicação

Nos passos 1 ao 3, o especialista deve ser capaz de reconstruir a metamensagem proposta pelo projetista a partir da análise de cada um dos signos inspecionados. No passo 4, é feito o contraste entre as metamensagens geradas, com o objetivo de rever intencionalmente as metamensagens para encontrar inconsistências e ambiguidades. Finalmente, no último passo do

método, é feita a apreciação final por parte do avaliador acerca de sua inspeção.

Através deste método foram identificados quais obstáculos na interface impedem ou dificultam os usuários de compreender o funcionamento do Portal de Periódicos da CAPES e proposto possíveis soluções para aperfeiçoar estes problemas. A partir da avaliação dos signos metalinguísticos, estáticos e dinâmicos, os problemas encontrados foram classificados de acordo com os níveis de gravidade elaborados por Nielsen (1994), ver (Quadro 1, no Anexo 1).

#### 4.2. Método de Avaliação com Usuários

Nesta pesquisa, para adaptar ao contexto e as especificidades do objeto do estudo, foi proposta a aplicação do Método de Explicitação do Discurso Subjacente, orientado por pressupostos de alguns modelos que representam a busca de informação em SRIs. Baseiam-se nestes modelos, portanto, os roteiros de entrevista e as tarefas de interação propostas, além da definição das diferentes ações realizadas pelos usuários durante a busca no Portal de Periódicos da CAPES, a serem observadas.

Foram considerados 07 dos 08 fatores que afetam o processo de busca do Quadro Conceitual sobre Recuperação Bibliográfica de Fidel e Soregel (1983) identificados como: *contexto, usuário, consulta, base de dados, sistema de busca, processo de busca e resultados da busca*. Foi incluído mais 01 fator para atender as especificidades desta pesquisa denominado *interface*. Além disso, foram observadas as mudanças que ocorrem no processo de busca de informação, as intenções interativas e as estratégias em que os usuários se envolvem dentro de um episódio em busca de informações. Vários modelos analisam a ocorrência destes movimentos dentro de um episódio de busca, tais como: o Modelo Berry-Picking de Bates (1989), o Modelo de Interação Episódica de Belkin *et al.* (1995), o Modelo de *Feedback* de Spink (1997), o Modelo de Spink e Wilson (1999) e o Modelo Situacional de Recuperação Interativa de Xie (2000).

Os passos necessários para aplicação deste método são:

- Passo 1 - Delineamento do escopo: consiste em delimitar a parte do sistema a ser inspecionada.
- Passo 2 - Recrutamento dos participantes: após a definição do escopo, selecionar os participantes para entrevista.
- Passo 3 - Preparação para a coleta de dados: elaborar o material a ser utilizado durante e inspeção do sistema que inclui *script* de apre-

sentação, termo de consentimento, roteiros de entrevista, tarefas de interação e formulários de acompanhamento, ver apêndice A.

- Passo 4 - Realização de teste-piloto: para testar a adequação das tarefas, roteiros de entrevista e equipamentos.
- Passo 5 - Coleta de dados: nesta etapa é recomendada a gravação do áudio e da interação (sincronizados) ou o vídeo (sendo capaz de distinguir elementos apontados na tela).
- Passo 6 - Preparação para análise dos dados: consiste na transcrição de todas as entrevistas e descrição de vídeos.
- Passo 7 - Análise dos dados: a avaliador faz a análise intra-sujeito, com o objetivo de realizar a análise individualmente e, em seguida, o avaliador faz a análise inter-sujeito, a partir da análise entre os participantes.
- Passo 8 - Interpretação dos dados: o avaliador faz uma análise das categorias que surgiram na avaliação.

Após a execução de 02 testes-piloto, o convite para participação foi feito através de e-mail, telefone e contato face a face. Os participantes realizaram a avaliação durante 03 semanas na sala de teste do Departamento de Ciência da Computação da Universidade Federal de Minas Gerais. Dois avaliadores estavam presentes. Um avaliador estava disponível para instruir, explicar as tarefas e entrevistar os participantes e o outro avaliador estava presente para observar e fazer anotações. As sessões duraram de 45 a 60 minutos, foram gravadas e posteriormente analisadas.

Fariam parte da amostra 15 voluntários, porém durante a realização da coleta de dados, o site do Portal de Periódicos da CAPES apresentou uma nova interface com diversas mudanças e já haviam sido executados testes com 14 voluntários na interface anterior, assim, decidiu-se finalizar os testes. Portanto, a amostra foi formada por 03 alunos de doutorado das áreas de Ciências Humanas, 02 alunos de doutorado das áreas de Linguística, Letras e Artes, 03 alunos de doutorado de Ciências Biológicas e 03 alunos de doutorado de Ciências Exatas e Engenharias, da UFMG, além de 03 bibliotecários do Sistema de Bibliotecas da UFMG, totalizando 14 voluntários.

Antes de iniciar as tarefas de interação, foi realizada a entrevista pré-teste para fazer o levantamento do perfil dos participantes, prática corrente de pesquisa e conhecimentos sobre o Portal de Periódicos da CAPES. Logo após, foi solicitado ao participante realizar 03 tarefas de busca e após o término de cada tarefa foram

realizadas várias perguntas referentes à interface, ferramentas de busca, construção da consulta, qualidade dos resultados, estratégias de busca e mudanças realizadas durante o episódio. O roteiro de entrevista foi adaptado com novas perguntas de acordo com o comportamento do voluntário. Por fim, após o término das tarefas foi realizada também uma entrevista para verificar questões como facilidades e dificuldades encontradas na execução das tarefas, expectativas e sugestões para aprimorar o sistema.

Durante a realização das tarefas de interação, foi aplicada a técnica “*think-aloud protocol*”, ou seja, os usuários são requisitados a expressarem seus pensamentos e ações durante a execução de tarefas. A gravação do áudio e da tela digital foi executada pelo *software* Monrae que capturou os movimentos do cursor na tela, à adição ou remoção de texto, desempenho do sistema, a voz do participante e a voz do avaliador. Também foi posicionada uma câmera de vídeo atrás do participante para gravar os movimentos e gestos dos usuários direcionados à tela.

Através da aplicação deste método foram alcançados os objetivos de integrar perspectivas da Ciência da Informação e da Interação Humano-Computador da Ciência da Computação para alcançar uma análise mais completa da interação, busca e recuperação da informação no Portal de Periódicos da CAPES; identificar quais problemas existem na interface e no sistema de busca que impedem ou dificultam o uso; propor soluções para aperfeiçoar este sistema e, melhorar a compreensão dos processos de interação do usuário com SRIs, com vistas a contribuir para o desenvolvimento teórico da Ciência da Informação na área de Recuperação da Informação Interativa.

## 5. Resultados

A seguir, apresenta-se o relato da avaliação da comunicabilidade do Portal de Periódicos da CAPES com o julgamento dos avaliadores sobre os problemas de comunicabilidade encontrados a partir da aplicação do Método de Inspeção Semiótica e os resultados do Método de Avaliação com Usuários.

### 5.1. Apreciação global da comunicabilidade da metacomunicação do Portal de Periódicos da CAPES

A metamensagem disponibilizada pelos projetistas através de signos metalinguísticos na ajuda do sistema e na própria interface do sistema para apoiar eventuais rupturas de comunicação

são insuficientes e ineficientes. Muitas vezes, o usuário necessita aprender através de tentativa e erro. O emprego de informações convenientes para compreender as diversas aplicações e comportamentos do sistema, tal como: a escolha dos critérios adotados para ordenar e apresentar os resultados, a influência do uso de estratégias de buscas, a necessidade de digitar o maior número de palavras corretas para localizar o item desejado, entre outras funções, ajudariam o usuário a assimilar melhor as reações do sistema e atingir os seus objetivos. Estes problemas poderiam ser solucionados através da aplicação de instruções na tela e o uso de signos do tipo *tooltips* seria eficiente para dar explicações sobre botões que executam funções importantes.

Os projetistas lançam mão principalmente dos signos estáticos para comunicar as funções para os usuários, porém existem incoerências. Como por exemplo, os botões com a função de realizar buscas, na maioria das vezes, apresentam-se como “Buscar”, mas em algumas janelas expressam-se com outros nomes como “Bases de Dados” ou “Ir para a Busca”. A adoção de um padrão tornaria o sistema mais fácil de ser compreendido pelo usuário.

O tipo de organização adotada também pode confundir o usuário, como apresentar três modos de pesquisa dentro da busca avançada repetindo alguns tipos de busca. No entanto, a reunião destes modos num único simplificaria o sistema, ou quando a busca avançada é subdividida em três buscas sendo que uma delas foi também designada por busca avançada tornando a metamensagem inconsistente. Ainda, algumas funções como, por exemplo, a ordenação dos resultados poderia ficar mais visível na tela para ajudar ao usuário compreender, rapidamente, como os resultados foram ou podem ser apresentados.

Existem partes na interface que desmotivam o usuário a explorar e utilizar suas funcionalidades. Como tal, a janela para “Busca das Bases de Dados”, no qual após apresentar uma lista de bases de dados, exige que o usuário selecione as bases de dados desejadas e clique no botão “Ir para a busca” na página inferior da tela. Caso o usuário, compreenda que este botão é para fazer uma busca e clicá-lo, o sistema ainda direciona o usuário novamente para a janela da busca avançada, apresentando os mesmos resultados da busca inicial e, em seguida, é necessário o usuário também compreender que precisa clicar no botão “Buscar” da busca avançada. Neste caso, a metamensagem não é clara, pois ele não consegue entender o que deve fazer em seguida. Portanto, seria mais

fácil usar o sistema, se o usuário não precisasse interagir com diversas camadas do projeto de interface e, também, que este apresentasse uma lógica mais coerente e intuitiva.

A identificação de excessos e duplicações de funções exibidas na tela como o uso de dois ou mais botões e ícones com função igual na mesma tela, ou funções como a aba “Detalhes” que tem como objetivo dar uma visão detalhada do registro, porém apresenta praticamente todas as informações já disponibilizadas abaixo do título do documento. Isto pode aumentar a dificuldade do usuário operar o sistema, portanto, é importante diminuir a quantidade de informações disponibilizadas na tela.

O principal estilo de interação utilizado é a manipulação direta. Praticamente todos os itens da interface podem ser explorados através de cliques do mouse e preenchimento de campos de pesquisa. O Portal disponibiliza vários signos dinâmicos, porém a maioria deles são ativações de janelas, no entanto, seria útil disponibilizar para os usuários mais recursos como apresentação de informações sobre um elemento da interface ao ser sobreposto pelo cursor do mouse, associação causal entre a seleção de um item de menu e a exibição do diálogo, o surgimento de dica para entrada de dados durante o preenchimento de um formulário, ativação e desativação de um comando.

O algoritmo de busca do Portal não permite recuperar os documentos, se o usuário digitar termos errados ou não digitar todos os termos do periódico, artigo ou base de dados desejada, e nem sempre, os primeiros resultados apresentados são os mais relevantes dentre o conjunto de documentos recuperados.

O Portal de Periódicos integrou o programa *Library Links* do Google Acadêmico que é uma ferramenta que permite identificar, nos resultados de pesquisa, os documentos disponíveis no Portal. Assim, muitos usuários preferem pesquisar no Google, pois recuperam os documentos desejados com muito mais facilidade e eficiência. Entretanto, esta prática pode impedir os usuários de identificar e recuperar outras fontes importantes relacionadas aos temas desejados.

O sistema não conta com um *design* adequado para a busca avançada. Por exemplo, as buscas avançadas do Portal de Periódicos da CAPES não disponibilizam ou apenas oferecem poucas opções de campos para combinações de termos impedindo ou limitando a possibilidade de recuperar mais registros relevantes. Poderia oferecer mais campos de buscas e possibilidades de elaborar estratégias de buscas.

Os projetistas poderiam lançar mão de instrumentos, no Buscar Assunto, Buscar Periódico e Buscar Base, como vocabulários controlados ou cabeçalhos de assunto para auxiliar os usuários na seleção de termos, possibilitar o usuário retornar para telas anteriores, caso o usuário tente desfazer uma ação, retirar funções desnecessárias como a ferramenta “Ir para a página” na lista de resultados ou o *link* “Visualizar” no Buscar Base, e remover *links* de ajuda que não são utilizáveis ou trazem pouco benefício.

Para usuários experientes com buscas em bases de dados e dispostos a explorar a interface do Portal e que tem consciência da qualidade do conteúdo disponibilizado no Portal CAPES, a met mensagem apesar de insuficiente será possível de ser transmitida. No entanto, o custo desta estratégia é que se o usuário não tiver este perfil poderá não perceber através da interface as várias possibilidades de uso, não usá-las ou até mesmo achar que o sistema não o atende.

## 5.2. Resultados do Método de Avaliação com Usuários

Durante a aplicação deste método observa-se que não havia variáveis importantes que permitissem generalizar ou distinguir o comportamento dos participantes de acordo com a área do conhecimento no qual estão vinculados. Porém, foi possível perceber que havia diferenças e semelhanças entre o desempenho e o comportamento dos usuários como a capacidade de empregar as estratégias de busca, o grau de conhecimento sobre o tema pesquisado e fontes disponíveis, a habilidade para manipular o sistema, a aptidão para formular a consulta, o domínio da língua inglesa e a persistência combinada com a capacidade de definir o que se quer buscar.

Assim, os resultados desta avaliação foram agrupados através de faixas de experiências dos participantes com a interface e o sistema de busca do Portal de Periódicos da CAPES, ver (Quadro 2, no Anexo 1). Primeiramente, são apresentadas as análises referentes ao grupo de bibliotecários, pois estes são especialistas em construção de estratégias de busca e recuperação da informação e, em seguida, são explicitados os resultados de acordo com os 04 níveis nos quais os alunos de doutorados foram enquadrados.

Apesar deste método não ter sido realizado em contexto real de uso e a limitação do tempo para execução dos testes, houve a preocupação de abarcar as novas interpretações para caracterizar as necessidades informacionais dos usu-

ários, apontadas por Robertson e Hancock-Beaulieu (1992), como: a) perspectiva da relevância, a satisfação do usuário com a informação recuperada; b) perspectiva cognitiva, que o grau de conhecimento do usuário pode levá-lo a alcançar sucesso ou obter fracasso ao interagir com SRIs; e a c) perspectiva interativa, em que considera as interações dos usuários com o sistema e os documentos, além do comportamento das reformulações da consulta.

Deste modo, foi considerado adequado para esta pesquisa basear em sete dos oito fatores que afetam o processo de busca apresentados pelo Quadro Conceitual de Recuperação Bibliográfica de Fidel e Soergel (1983) e observar as mudanças que ocorrem no processo de busca de informação, as intenções interativas e as estratégias que os usuários se envolvem dentro de um episódio em busca de informações para compreender melhor a natureza da recuperação de informação interativa. Estas mudanças foram analisadas em vários modelos que tentam representar a busca de informação em SRIs. E ainda, para atender as características do sistema avaliado, o fator *usuário intermediário* foi substituído pelo fator denominado *interface*.

Estes fatores são abordados durante as três etapas da avaliação: 1) contexto – tipo de grupo ou área do conhecimento; 2) usuário – características e experiências da pessoa que necessita de informação e como inicia a busca; 3) consulta – como foi construída a consulta apresentada pelo usuário; 4) banco de dados – cobertura e listas de referências; 5) sistema de busca – suporte, capacidade de busca, formato de apresentação de resultados, filtros de refinamento, etc.; 6) interface – capacidade para suportar a interação com o usuário, diminuir os erros cometidos, reduzir os custos com treinamentos, aumentar o sucesso para obter resultados satisfatórios, etc.; 7) processo de busca – seleção das ferramentas de busca, formulação da consulta, mudanças que ocorrem no processo de busca de informação, etc.; e 8) resultados da busca – quantidade e qualidade dos resultados recuperados.

No momento da identificação das partes dos textos transcritos sobre as barreiras e os problemas encontrados pelos participantes para uso do Portal de Periódicos da CAPES foram observados as seguintes categorias de análise: interface, conteúdo, sistema de busca e comportamento do usuário.

A partir das categorias de análise, foram identificados 141 problemas durante a avaliação das ferramentas Buscar Assunto, Buscar Periódico e do Buscar Base que podem dificultar ou impedir

os usuários de interagirem com o sistema de modo satisfatório.

## 6. Conclusões

Para construção metodológica desta pesquisa, inicialmente, foi selecionada a mais recente das abordagens da Interação Humano-Computador conhecida como Engenharia Semiótica. Esta abordagem possibilitou através do Método de Inspeção Semiótica inspecionar minuciosamente as perguntas frequentes da Aba Suporte, o Buscar Assunto, o Buscar Periódico e o Buscar Base do Portal de Periódicos da CAPES para identificar as potenciais rupturas de comunicação na emissão da metamensagem projetista-usuário. Através da análise dos signos metalinguísticos, estáticos e dinâmicos foi possível identificar 52 problemas.

A partir da etapa de apreciação global da qualidade da metacomunicação foi possível apresentar a síntese geral dos problemas e as respectivas soluções para aperfeiçoar a interface do Portal de Periódicos da CAPES:

- a) *As informações disponibilizadas são insuficientes para apoiar eventuais rupturas de comunicação:* estes problemas poderiam ser solucionados através da aplicação de instruções e uso de signos do tipo *tooltips* para dar orientações aos usuários sobre recursos e botões que executam funções importantes;
- b) *Há incoerência e falta padronização para comunicar algumas funções:* a adoção de um padrão para os botões e outros recursos tornaria o sistema mais fácil de ser compreendido e utilizado pelo usuário.
- c) *tipo de organização adotada dificulta a compreensão e pode confundir o usuário:* evitar a repetição de campos de busca, diminuir a quantidade de informações, reunir as opções de busca de avançada e melhorar a visibilidade de funções importantes e a apresentação dos resultados podem tornar mais simples o uso do sistema;
- d) *Existem partes na interface que desmotivam o usuário a explorar e utilizar suas funcionalidades:* seria mais fácil usar o sistema, se possibilitasse ao usuário retornar para telas anteriores, caso tente desfazer uma ação; se o usuário não precisasse interagir com muitas camadas do projeto de interface e também que este apresentasse uma lógica mais coerente e intuitiva;
- e) *A identificação de excessos e duplicações de funções exibidas na tela:* é importante diminuir a quantidade de informações disponibilizadas na tela, pois os usuários não conse-

quem prestar atenção em todas as informações e pode aumentar o esforço cognitivo;

- f) *Praticamente todos os itens da interface podem ser explorados através de cliques do mouse e preenchimento de campos de pesquisa*: seria útil disponibilizar para os usuários mais recursos como apresentação de informações sobre um elemento da interface ao ser sobreposto pelo cursor do mouse, associação causal entre a seleção de um item de menu e a exibição do diálogo, o surgimento de dica para entrada de dados durante o preenchimento de um formulário, ativação e desativação de um comando.
- g) *algoritmo de busca do Portal não permite recuperar documentos se o usuário digitar termos errados ou não digitar todos os termos do periódico, artigo ou base de dados desejada*: o algoritmo de busca poderia lidar melhor com as consultas mais longas e complexas que tem recebido dos usuários e também corrigir e sugerir palavras e expressões de busca.
- h) *sistema não conta com um projeto adequado para a busca avançada*: poderia oferecer mais campos de buscas e possibilidades de elaborar estratégias de buscas.
- i) *Falta apoio aos usuários para seleção dos termos das buscas*: os projetistas poderiam lançar mão de ferramentas de auxílio como vocabulários controlados ou cabeçalhos de assunto;
- j) *Disponibilizam recursos que possuem pouca utilidade ou não funcionam*: retirar funções desnecessárias como a ferramenta “Ir para a página” na lista de resultados ou o link “Visualizar” no Buscar Base, e remover links de ajuda que não são utilizáveis ou trazem pouco benefício.

Além disso, com base nos resultados alcançados através da aplicação do Método de Inspeção Semiótica foi possível concluir que para os usuários experientes com buscas em SRIs, dispostos a explorar a interface do Portal e que tem consciência da qualidade do conteúdo disponibilizado, a qualidade da comunicação do projetista para o usuário apesar de insuficiente será possível de ser transmitida. No entanto, o custo desta estratégia é que se o usuário não tiver este perfil poderá não perceber através da interface as várias possibilidades de uso, não usá-las ou até mesmo achar que o sistema não o atende.

Para executar a segunda parte da metodologia, foi proposto à aplicação do Método de Explicitação do Discurso Subjacente (MEDS) orientado pelos pressupostos de modelos que represen-

tam a busca de informação em SRIs baseados na área de Recuperação da Informação Interativa para avaliar a interação dos usuários com a interface e o sistema de busca do Portal de Periódicos da CAPES. Também, adaptaram-se os fatores envolvidos no processo de busca de informação identificados pelo Quadro Conceitual de Recuperação Bibliográfica de Fidel e Soergel (1983) para atender aos objetivos e as especificidades do objeto de estudo desta pesquisa.

A partir dos resultados desta avaliação constatou-se a necessidade de que os SRIs suportem a interação de usuários com diferentes capacidades e demandas. Os usuários novatos e inexperientes precisam interagir com telas mais simples, ter acesso rápido a ajuda do sistema, de recursos para correção e sugestão de palavras, ter facilidade para mover-se entre as aplicações e de mais apoio para entrada, combinação e seleção dos termos de busca. Por outro lado, os usuários mais experientes demandam opções avançadas para construção das estratégias de busca, possibilidade de clicar em vínculos de hipertexto num documento, diversos campos para especificação da busca, gerenciar os históricos de busca, opções mais sofisticadas para seleção dos resultados e mover-se rapidamente entre as aplicações.

Deste modo, foi possível concluir que os projetistas destes sistemas têm o desafio de compatibilizar recursos que atendam usuários com perfis e demandas diferentes para aumentar a popularidade e o uso, além de serem suficientemente flexíveis para permitir aos usuários adaptar o processo de busca de informações às suas próprias necessidades.

Ademais foram considerados 07 fatores envolvidos no processo de busca de informação do usuário de Portal de Periódicos da CAPES do Quadro Conceitual de Recuperação Bibliográfica de Fidel e Soergel (1983). A partir da avaliação com usuários foi permitido estimar a capacidade destes fatores para representar os padrões de busca de informação dos participantes e certificar a competência deste modelo para orientar as avaliações de SRIs em ambientes digitais.

Também foi evidenciado que o sucesso e a satisfação do usuário durante a busca no sistema não depende apenas do que o sistema interativo oferece ou sobre como incentiva a empregarem esses recursos, mas também depende de como o sistema suporta as estratégias pessoais do usuário e o conduz para interagir com as várias camadas do projeto e de componentes como estrutura de banco de dados, sistema de busca e objetos de informação.

Na avaliação com os usuários, empenhou-se em observar as razões para os usuários realizarem sucessivas mudanças de estratégias de busca durante um episódio de busca. Dentre as várias estratégias utilizadas, realizar a primeira busca somente para explorar o sistema, incluir novos termos de busca, mudar de termos de busca e alterar da busca simples para a busca avançada do Buscar Assunto, Buscar Periódico e Buscar Base foram às estratégias mais utilizadas pelos participantes.

Ao reconhecer estas necessidades dos usuários, os desenvolvedores dos SRIs podem disponibilizar mais recursos que os auxiliem a realizarem facilmente estas mudanças sem sentirem desmotivados ou desistirem de explorar as funcionalidades do sistema. Além de corroborar com os resultados de investigações recentes que revelam a limitação da pesquisa no campo de RI tradicional no qual pressupõe que a busca de informações é estática e que as pessoas formulam um único tipo de consulta durante o processo de busca.

Durante as análises dos resultados da avaliação dos usuários foi possível explorar as motivações e dificuldades encontradas para uso do Portal de Periódicos da CAPES relacionadas a quatro elementos: interface, conteúdo, sistema de busca e características do comportamento dos usuários. Portanto, julgou-se que há quatro elementos importantes de serem considerados no planejamento, avaliação e desenvolvimento de SRIs, e que estes elementos interagem e tem influência entre si.

Ademais esta pesquisa contribuiu ainda para três áreas distintas: Recuperação da Informação (RI), Comportamento de Busca da Informação em Sistemas de Recuperação da Informação, também denominada *Information Search Behaviour*, e Interação Humano-Computador. Para a área de RI realiza a contribuição á medida que combina métodos para proporcionar melhor compreensão das necessidades dos usuários durante o processo de busca de informação. Para a área de *Information Search Behaviour*, a colaboração se dá por apresentar os diversos modelos relacionados com a busca em SRIs, aplicar um destes modelos e tratar de um campo que atualmente tem sido pouco explorado no contexto nacional. Para a área de IHC e, em especial, teoria da Engenharia Semiótica, a contribuição se dá pela combinação de métodos que permitem a avaliação da comunicabilidade adaptadas ao contexto de SRIs. Ainda, favorece ao evidenciar que a interseção entre estas áreas formou uma especialidade de pesquisa denominada Recuperação da Informação Interativa (RII) que tem como objetivo capacitar indi-

víduos a entender e a solucionar suas necessidades de informação através das interações com os sistemas de recuperação da informação.

Esta pesquisa apresentou limitações por não ter sido realizada em contexto real de uso. Apesar do Portal de Periódicos da CAPES, a interação e os usuários estarem no foco desta investigação, a seleção da amostra foi oportunista e o contexto de avaliação não foi real. Esta descontextualização não permite capturar a relação da qualidade da informação com o seu uso ou o processo que desencadeou a atividade de pesquisa.

A restrição de tempo para execução dos testes não possibilitou identificar o nível macro dos componentes da busca e recuperação como as dimensões das tarefas que influenciam na forma como os usuários interagem com os SRIs e integrar outras camadas, como por exemplo, as estruturas de conhecimento do usuário, o contexto social e organizacional que são subjacentes a experiência humana no que diz respeito à busca de informação.

Devido à necessidade de constituir uma amostra com poucos usuários, não foram incluídos docentes, alunos de mestrado e graduação para participarem da avaliação e também permitir que as avaliações fossem tão semelhantes quanto possível de modo que a causalidade pudesse ser estudada e, e em seguida, fosse proposto um novo modelo de busca de informação integrando as diversas variáveis estudadas e manifestadas durante os testes com os participantes.

Apesar de todo esforço para aperfeiçoar os SRIs e melhorar a sua capacidade de apresentar informações que atendam às necessidades dos usuários, reconhece-se que o comportamento intelectual do cientista muda enormemente á medida que ele progride em sua pesquisa. A preferência por este ou aquele tipo de fonte de informação ou SRIs pode também variar de acordo com as necessidades e essas, por sua vez, variam dependendo do estágio da pesquisa.

Reconhece-se também que não é fácil mesmo unir os estudos comportamentais, contextuais ou cognitivos com decisões de *design*, devido aos diversos fatores envolvidos na interação humano-computador para apoiar o aperfeiçoamento de SRIs, e encontrar um único modelo de busca de informação ou metodologia para avaliação que consiga explicar todas as questões do processo de busca de informação. No entanto, estes resultados apresentados pelos avaliadores e usuários sobre a interface e sistema de busca do Portal de Periódicos da CAPES po-

dem ser considerados para aperfeiçoá-lo, melhorar o seu sucesso e uso e diminuir o investimento com treinamentos para os usuários.

## Referências

- Abels, E. G.; Liebscher, P.; Denman, D. (1996). Factors that influence the use of electronic network by science and engineering faculty at small institutions. part I. Queries. *Journal of the American Society for Information Science*, 47:2, 146-158.
- Ahmed, S. M. Z.; Mcknigh, C.; Oppenheim, C. (2009). A review of research on human-computer interfaces for online information retrieval systems. *The Electronic Library*, 27:1, 96-116.
- Barbosa, S. D. F.; Silva, B. S. (2010). *Interação humano-computador*. Rio de Janeiro: Elsevier.
- Bates, M. J. (1989). The design of browsing and Berry picking techniques for the online search interface. *Online Information Review*, 13:5, 407-424.
- Belkin, N. J. et al. (1995). Cases, scripts, and information-seeking strategies: on the design of interactive information retrieval systems. *Expert Systems with Applications*, 9:3, 379-395.
- Belkin, N. J. (2010). On the evaluation of interactive information retrieval systems. In: Larsen, B. (Ed.). (2010). *The Janus faced scholar: a fest schrift in honor of Peter Ingwersen*. Copenhagen: Det Informationsvidenskabelige Akademi, 13-21.
- Bierig, R.; Cole, M.; Belkin, N. (2010). A data analysis and modelling framework for the evaluation of interactive information retrieval. In: Gurrin, C. (Ed.). (2010). *Advances Information Retrieval: 32<sup>nd</sup> European Conference on IR Research, ECIR 2010*. Berlin: Springer-Verlag, 673-674.
- Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) (2015). *Portal de Periódicos da CAPES*. Recuperado 26-05-2015 de [http://www.periodicos.capes.gov.br/index.php?option=com\\_phome&mn=68](http://www.periodicos.capes.gov.br/index.php?option=com_phome&mn=68)
- Costa, L. F. (2008). *A usabilidade do Portal de periódicos da CAPES*. 2008. 238 f. Dissertação (Mestrado em Ciência da Informação) – Universidade Federal da Paraíba, João Pessoa.
- Fidel, R.; Soergel, D. (1983). Factors affecting online bibliographic retrieval: a conceptual framework for research. *Journal of the American Society for Information Science*, 34:3, 163-180.
- Ingwersen, P. (1992). *Information retrieval interaction*. London: Taylor Graham.
- Ingwersen, P., Jarvelin, K. (2005). *The turn: integration of information seeking and retrieval in context*. Netherlands: Springer.
- Joseph, P.; Debowski, S.; Goldschmidt, P. (2013). Models of information search: a comparative analysis. *Information Research*, 18:3.
- Kelly, D.; Sugimoto, C. R. (2013). A systematic review of interactive information retrieval evaluation studies, 1967-2006. *Journal of American Society for Information Science and Technology*, 64: 4, 745-770.
- Keshavarz, H. (2008). Human information behaviour and design, development and evaluation of information retrieval systems, *Program: electronic library and information systems*, 42:4, 391-401.
- Lopes, M. I.; Silva, E. L. (2007). A internet e a busca da informação em comunidades científicas: um estudo focado nos pesquisadores da UFSC. *Perspectivas em Ciência da Informação*, 12:3, 21-40.
- Mamoon, M. H.; El-Bakri, H. M.; Salama, A. A. (2013). Interactive visualization of retrieved information. *International Journal of Knowledge Engineering and Research*, 2:4, 217-231.
- Marchionini, G. (2006). Towards human-computer information retrieval. *Bulletin of the American Society for Information Science and Technology*, 32, 20-22.
- Martins, D. S. (2009). *Uma abordagem para recuperação de informações sensível ao contexto usando a retroalimentação implícita de relevância*. São Carlos, SP, 111 p. Dissertação (Mestrado em Ciência da Computação) – Universidade Federal de São Carlos.
- Nielsen, J. (1994). Heuristic evaluation. In: Mack, R.; Nielsen, J. (Ed.). (1994). *Usability inspection methods*. New York, NY: John Wiley & Sons.
- Oliveira, E. R. (2010). *Investigação sobre a aplicabilidade dos métodos de avaliação de comunicabilidade ao domínio educacional*. Belo Horizonte, 226 p. Dissertação (Mestrado em Ciência da Informação) – Instituto de Ciências Exatas, Universidade Federal de Minas Gerais, Belo Horizonte.
- Oliveira, E. R.; Luz, L. C. S.; Prates, R. O. (2008). Aplicação semi-estruturada do método de inspeção semiótica: Estudo de caso para domínio educacional. In: Brazilian Symposium on Human Factors in Computing Systems, 8., 2008. *Proceedings...* Porto Alegre, RS: Sociedade Brasileira de Computação, 50-59.
- Peirce, C. S. (1992). *The essential Peirce*. Indiana University Press.
- Prates, R. O.; Barbosa, S. D. J. (2007). Avaliação de interfaces de usuários: conceitos e métodos. In: T. Kowaltowski; K. K. Breitman. (Org.). (2007). *Atualizações em Informática*. Porto Alegre: Editora PUC-Rio, 263-326.
- Prates, R. O.; Barbosa, S. D. J. (2003). Introdução à teoria e prática da Interação Humano-Computador fundamentada na Engenharia Semiótica. In: *Atualizações em Informática*.
- Robertson, S. E.; Hancock-Bealieu, M. M. (1992). On the evaluation of IR systems. *Information Processing & Management*, 28:4, 457-466.
- Rogers, Y.; Sharp, H.; Preece, J. (2012). *Interaction design: beyond human-computer interaction*. 3<sup>rd</sup>. Wiley.
- Ruthven, I. (2008). Interactive information retrieval. *Annual Review of Information Science and Technology*, 42:1, 43-91.
- Souza, C. S. (2005). *The semiotic engineering of human-computer interaction*. Cambridge, MA: The MIT Press.
- Souza, C. S. et al. (2006). The Semiotic Inspection Method. In: IHC'06: Brazilian Symposium on Human Factors in Computing Systems, 7., 2006. *Proceedings...* New York, USA: ACM, 148-157.
- Souza, C. S.; Leitão, C. F. (2009). *Semiotic engineering methods for scientific research in HCI*. Morgan and Claypool Publishers.
- Spink, A. (1997) Information science: a third feedback framework. *Journal of the American Society for Information Science*, 48:8, 728-740.
- Spink, A.; Wilson, T. D. (1999). Toward a theoretical framework for information retrieval (IR) evaluation in an information seeking context. In: MIRA 99: Evaluation Frameworks for Multimedia Information Retrieval Applications, April 14-16, 1999. *Proceedings...* University of Glasgow: Scotland.

## Anexo 1

| NÍVEL | GRAVIDADE          | DESCRIÇÃO  |
|-------|--------------------|--|
| 0     | Não Concordo       | Não concordo que isto seja um problema (este valor pode resultar da avaliação de um especialista sobre um problema apontado por outro especialista). |
| 1     | Problema Cosmético | Não precisa ser consertado a menos que haja tempo extra no projeto.  |
| 2     | Problema Pequeno   | O conserto deste problema é desejável, mas deve receber baixa prioridade.  |
| 3     | Problema Grande    | Importante de ser consertado; deve receber alta prioridade.  |
| 4     | Catastrófico       | É imperativo consertar este problema antes do lançamento do produto.   |

QUADRO 1 NÍVEIS DE GRAVIDADE PARA CLASSIFICAÇÃO DOS PROBLEMAS ENCONTRADOS

| CATEGORIAS               | DESCRIÇÃO  | PARTICIPANTES  |
|--------------------------|--|--|
| <b>Bibliotecários</b>    | Têm experiência no atendimento das demandas informacionais dos usuários, foram treinados e possuem domínio para construção de estratégias de busca elaboradas.   | Participante 1, Participante 2, Participante 3                   |
| <b>Usuário nível I</b>   | Precisam aprender como executar tarefas de recuperação da informação, com mais rapidez e facilidade. Encontram muitas dificuldades para usar a interface e o sistema de busca do Portal de Periódicos da CAPES | Participante 6, Participante 11                                  |
| <b>Usuário nível II</b>  | Executam algumas estratégias de recuperação de informação com facilidade obtém sucesso em algumas das buscas realizadas no Portal, tentam aprender a usar o sistema durante a interação.                       | Participante 5, Participante 9, Participante 12, Participante 13 |
| <b>Usuário nível III</b> | Sabem aplicar várias estratégias de busca, são tidos como experientes, embora limitem a variedade de funções que utilizam no sistema. Conhecem as fontes de informação relacionadas à sua área de pesquisa.    | Participante 4, Participante 7, Participante 8                   |
| <b>Usuário nível IV</b>  | São pesquisadores experientes, utilizam regularmente sistemas de recuperação da informação, bases de dados e motores de busca, têm domínio sobre tarefas e recursos de busca.                                  | Participante 10, Participante 14                                 |

QUADRO 2 CATEGORIAS USADAS PARA ENQUADRAR AS FAIXAS DE EXPERIÊNCIAS DOS PARTICIPANTES

---

# Elaboração de um vocabulário taxonómico sobre História Medieval portuguesa: problemas e desafios

*Preparation of a taxonomic vocabulary of Portuguese Medieval History: problems and challenges*

---

Filipa Medeiros (1), José Antonio Moreiro González (2)

(1) Centro Interdisciplinar de História, Culturas e Sociedades da Universidade de Évora, Palácio do Vimi-oso, Largo do Marquês de Marialva, nº 8, Apartado 94, 7800-809 Évora, PORTUGAL, fmcm@uevora.pt

(2) Universidad Carlos III de Madrid, Departamento de Biblioteconomía y Documentación, Edificio López Aranguren, Despacho 15.1.63, Calle Madrid, 126, 28903 Getafe (Madrid) España, jamore@bib.uc3m.es

## Resumo

Apresenta-se uma taxonomia de domínio sobre História Medieval Portuguesa (séculos XII-XV), cuja elaboração se enquadra no âmbito de uma tese de doutoramento de índole interdisciplinar. A elaboração deste vocabulário taxonómico procura colmatar a necessidade de indexar recursos digitais, em particular os contidos em bases de dados especializadas, consideradas hoje veículos fundamentais para a disseminação e distribuição da produção científica em acesso aberto. Por conseguinte, os principais objetivos da taxonomia são: verificar os métodos já comprovados com êxito em estudos de caso para projetar, desenvolver e manter taxonomias de domínio; organizar e representar informação da especialidade, de forma a recuperá-la de forma consistente e coerente; refletir sobre as potencialidades deste tipo de vocabulário controlado, nomeadamente no que respeita à organização de recursos digitais. Para a construção da taxonomia optou-se, de forma geral, pelo método descendente, isto é, partiu-se do genérico para o específico (top-down). Apresentam-se as 17 categorias que compõem a taxonomia. Como conclusões, destacam-se as questões relacionadas com a extensão e profundidade das categorias suas constituintes. Por fim, reflete-se sobre o desenvolvimento futuro da taxonomia, em especial no que concerne ao seu tratamento ontológico.

**Palavras-chave:** Taxonomias; Sistemas de organização do conhecimento; Representação da informação; História Medieval Portuguesa; Interdisciplinaridade.

## 1. Introdução

Este trabalho apresenta uma taxonomia de domínio destinada a organizar e a representar informação relativa à História Medieval portuguesa entre os séculos XII e XV e que já se encontra em linha sob a forma de lista estruturada. Trata-se, portanto, de um trabalho aplicado (Medeiros, 2014) - fruto de uma investigação de doutoramento desenvolvida na Universidade de Évora

## Abstract

We are presenting a domain taxonomy of Portuguese Medieval History (12<sup>th</sup>-15<sup>th</sup> centuries) that was prepared within the scope of a doctoral research work and based on an interdisciplinary approach. The preparation of this taxonomic vocabulary seeks to fill a gap in terms of the representation of information of digital resources, namely the ones contained in specialized databases, which are currently seen as crucial vehicles for disseminating and sharing open-access scientific knowledge. So, its main goals are: to test methods that have already proven successful in various case studies to design, develop and maintain domain taxonomies; to organize and represent specialized sources of information, in order to allow recovering it in a consistent and uniform way; to reflect on the potential of this type of controlled vocabulary, particularly for the organization of digital resources. In order to build this taxonomy we used a top-down, or general-to-specific, method. We present and categorize the 17 categories that compose the taxonomy. With regard to the conclusions, there is an emphasis on issues related to the size and depth of the different categories. Finally, we reflect on the future developments of this taxonomy, especially on its ontological development.

**Keywords:** Taxonomies; KOS, representation of information; Portuguese Medieval History, Interdisciplinarity

em colaboração com a Universidade Carlos III de Madrid - que utilizou os métodos já testados por especialistas para construir taxonomias.

A elaboração de uma taxonomia no âmbito da História Medieval constitui, como adiante se verá, uma opção pioneira e como tal sem parâmetros passíveis de comparação.

Não obstante as dificuldades e as opções, sempre discutíveis, subjacentes à sua elaboração, a

taxonomia constitui um instrumento de inegável interesse e de apoio aos investigadores, em especial numa sociedade marcada pelo crescimento exponencial da informação disponível em acesso aberto.

Ela deve ser entendida como um instrumento em permanente construção e avaliação, mas clara nos propósitos que se propõe a atingir.

## 2. Objetivos

Assim, esta comunicação apresenta três objetivos fundamentais:

- a) Testar os métodos já comprovados com êxito em estudos de caso para projetar, desenvolver e manter taxonomias de domínio;
- b) Organizar e representar informação da especialidade, tendo em vista uma recuperação da mesma com consistência e uniformidade;
- c) Refletir sobre as potencialidades deste tipo de vocabulário controlado - particularmente em termos da organização de recursos digitais e a possibilidade de navegação pelos mesmos - aqui aplicadas ao universo dos estudos medievais portugueses. Mais especificamente, a criação desta taxonomia procura colmatar uma lacuna existente no plano das unidades de informação portuguesas especializadas em História Medieval, muitas delas incorporadas nas universidades e respetivos centros de investigação. Na verdade, é premente o tratamento diferenciado da informação da especialidade, designadamente no que diz respeito à indexação do universo temático da História medieval portuguesa.

Para além disso, a elaboração deste vocabulário controlado procura, sobretudo, responder à necessidade de indexação de recursos digitais, nomeadamente os contidos em bases de dados especializadas, considerados hoje veículos fundamentais para a disseminação e partilha da produção científica em acesso aberto. Veja-se, a título exemplificativo, algumas bases de dados já produzidas pelo Instituto de Estudos Medievais da Faculdade de Ciências Sociais e Humanas da Universidade Nova de Lisboa (IEM - FCSH/UNL) e disponíveis em: <http://iem.fcsh.unl.pt/section.aspx?kind=noticia&id=49>

## 3. Metodologia

Para a construção e desenvolvimento da taxonomia utilizou-se o método descendente ou do geral para o particular, top-down (Conway e Slinger, 2002; Cumming, 2003; Gilchrist, 2003; Jagerman, 2006; Zhonghong e outros, 2006; Moreira González, 2006 e 2011), o mais utilizado na

criação deste tipo de vocabulário controlado. Este método compõe-se das seguintes etapas:

- a) Compilar o conhecimento;
- b) Reduzir os sinónimos e proceder à escolha dos termos preferenciais;
- c) Preparar uma estrutura sistemática prévia;
- d) Elaborar uma primeira versão do esquema;
- e) Completar a taxonomia com todos os seus termos;
- f) Avaliar o seu funcionamento;
- g) Publicar uma primeira versão.

Apresenta-se, de seguida, o desenvolvimento específico de cada uma das etapas da construção da taxonomia de História Medieval Portuguesa, de acordo com o método descendente:

### 3.1. Compilação do conhecimento

Nesta fase, recorreu-se às fontes de informação especializadas, em particular dicionários, léxicos e vocabulários relacionados com a Idade Média.

Não foram encontradas quaisquer produções de origem portuguesa destas tipologias documentais, pelo que se reuniram contribuições estrangeiras de referência no âmbito da História Medieval geral e que poderiam enriquecer a taxonomia, particularmente no que se relaciona com a constituição das suas categorias gerais, quer do ponto de vista da forma, quer do conteúdo.

Chegou-se a um total de 11 fontes de informação estrangeiras. No que respeita às fontes de informação portuguesas, na impossibilidade de se dispor desse tipo de fontes em língua nacional, recorreu-se a histórias gerais e a manuais universitários de referência no âmbito dos estudos medievais portugueses, em concreto a nove fontes de informação (Serrão e Marques, 1987 e 1996; Marques, 1988; Tavares, 1990; Coelho, 1991; Tavares, 1992; Branco e Costa, 1992; Serrão, 1994; Moreno, 1995; Mattoso, 1997; Ramos e outros, 2009).

### 3.2. Redução dos sinónimos e escolha dos termos preferenciais

Este ponto corresponde à segunda etapa de construção da taxonomia, no qual se expõem as questões atinentes ao controlo formal e semântico do vocabulário, de acordo com o seguinte esquema de análise:

- Normalização formal: controlo morfológico (língua, género e número) e controlo sintático (termos simples e termos compostos);

- Normalização semântica: controlo das ambiguidades via significante (homógrafos ou polissemos) e via significado (palavras sinónimas e quase-sinónimas), através do estabelecimento de relações semânticas.

Uma vez que a língua de partida da taxonomia é a portuguesa, seguiu-se, naturalmente, as orientações dos normativos nacionais, em concreto a Norma Portuguesa 4036 (pontos 6 e 7) e o manual *Siporbase: sistema de indexação em português* (secção 4, Terminologia) (NP 4036, 1992: 8-22; Portugal, 1998).

O controlo morfológico e sintático não se revelou muito complexo, tendo-se seguido as regras da língua portuguesa assim como as normas nacionais. O mesmo não ocorreu com o controlo semântico de alguns termos.

Por conseguinte, foi dado especial cuidado aos homógrafos ou polissemos, isto é, às palavras constituídas pela mesma cadeia de caracteres mas que têm significados diferentes. No entanto, estas palavras não se revelaram uma dificuldade efetiva, pois estamos perante um vocabulário controlado altamente especializado, no qual este tipo de situações é substancialmente menor se compararmos com outros vocabulários de cobertura temática de maior abrangência. Nestes casos, recorreremos ao uso dos designados «qualificadores» para contextualizar e fixar o âmbito semântico (significado) do termo e que, de resto, são parte integrante do termo de indexação, formando, assim, um termo composto:

Ex. Jantar (tributo senhorial)

Jantar (refeição)

Ao nível do controlo semântico via significado, ao reunirmos uma primeira amostra de termos da especialidade, efetuámos, à partida, a eliminação de sinónimos e de quase-sinónimos, elegendo termos preferenciais, destinados a representar univocamente um só conceito:

Ex. Cantigas de escárnio e de maldizer [termo preferencial]

Ex. Canções de escárnio e de maldizer [termo não preferencial]

A ordenação dos termos na estrutura hierárquica foi efetuada por ordem alfabética, exceto na categoria «Acontecimentos», «Personalidades» e «Reinados», nas quais os termos foram inseridos por ordem cronológica. Só deste modo, se conseguiu organizá-los de forma diacrónica, possibilitando ao utilizador navegar na estrutura taxonómica de acordo com a sucessão natural dos acontecimentos, das personalidades e dos reinados, respetivamente:

Ex. 1128, Batalha de São Mamede  
Ex. 1350-1405, Leonor Teles  
Ex. 1385-1433, Reinado de D. João I

### 3.3. Processo de construção da taxonomia

Após a normalização formal e semântica do vocabulário, conforme os critérios e procedimentos supramencionados, efetuou-se o registo alfabético dos termos compilados, distribuindo-os pelas categorias que compuseram a estrutura sistemática prévia da taxonomia, correspondendo esta operação à terceira etapa.

Seguidamente, e numa quarta etapa, elaborou-se uma primeira versão do esquema taxonómico, definindo a posição hierárquica dos termos dentro das suas respetivas categorias e ordenando, por conseguinte, cada termo específico sob o termo genérico correspondente. Por sua vez, no interior de cada categoria, os conceitos foram organizados em classes, que podem ser de dois tipos: cadeias e renques. As primeiras são séries verticais de conceitos e as segundas séries horizontais de conceitos, podendo ser ambas do tipo genérico (é tipo de?) ou partitivo (é parte de? é todo de?) (Campos e Gomes, 2007: 6).

Estes procedimentos permitiram avançar para a quinta etapa e elaborar a superestrutura classificatória da taxonomia e completá-la com todos os seus termos. A sua versão final é constituída por 17 categorias, a saber:

- (1) Administração
- (2) Áreas da História
- (3) Ciências e técnicas historiográficas
- (4) Coroa
- (5) Cronologia
- (6) Cultura
- (7) Economia
- (8) Estruturação social
- (9) Fontes
- (10) Geografia
- (11) Guerra
- (12) Personalidades
- (13) Recursos de informação
- (14) Reinados
- (15) Religião
- (16) Sociedade
- (17) Vida privada

Numa sexta etapa, avaliou-se a taxonomia de acordo com as recomendações da ANSI/NISO Z39.19-2005, pois quer a NP 4036, quer a ISO 25964 são omissas relativamente a esta questão. Deste modo, optou-se pela avaliação heurística da versão final da taxonomia, pois pensa-se ser aquela que mais se adequa ao vocabulário controlado construído, que é uma taxonomia de domínio, e que, enquanto tal, pressupõe a inter-

venção e monitorização efetivas e permanentes por parte de especialistas no processo de elaboração da mesma, como, de resto, ocorreu (ANSI/NISO Z39.19, 2005: 95-96).

Desta forma, foi solicitada a colaboração formal a dois especialistas, um da área da História medieval portuguesa e outro da área das CID, especializado em sistemas de organização do conhecimento e cuja avaliação se apresenta mais adiante.

Na sétima e última etapa, publicou-se uma primeira versão da taxonomia. Os termos recolhidos e normalizados foram, na sua totalidade, introduzidos num *software* próprio de gestão de taxonomias denominado *Knowledge Manager* (KM), comercializado pela empresa espanhola *The Reuse Company*, que colabora com o Departamento de Biblioteconomia e Documentação da Universidade Carlos III de Madrid. A versão utilizada é uma versão para *download* de teste (edição 5.0.0.). O KM permite a criação de diferentes tipos de relatórios que podem ser exportados em vários formatos, possibilitando a introdução dos termos para outras aplicações.

Para a versão final da taxonomia, preferiu-se a apresentação alfabética, pois é aquela que possibilita navegação pela estrutura semântica do vocabulário taxonómico, permitindo a visão geral das suas categorias, assim como das suas respetivas cadeias e renques, tal como demonstra a figura 1 (anexo 1).

#### 4. Resultados

A taxonomia conta atualmente com 2987 termos, em permanente revisão e atualização.

A versão integral da taxonomia encontra-se disponível *online* em: <http://www.cidehus.uevora.pt/Recursos/Taxonomia-de-Historia-Medieval-Portuguesa>

Tal como se depreende pelo gráfico 1 (anexo 2), a taxonomia de História Medieval Portuguesa é constituída por 17 categorias, com diferentes níveis de profundidade. Assim, a categoria de maior extensão é a «Cronologia», que conta com 570 termos. Ao invés, a categoria «Ciências e técnicas historiográficas» reúne somente 9 termos.

De igual modo em posição de destaque, encontram-se as categorias «Personalidades», «Economia», «Guerra», «Cultura» e «Sociedade» com 523, 486, 265, 244 e 231 termos, respetivamente.

Numa hierarquia intermédia, listam-se as categorias «Administração», «Estruturação social»,

«Religião» e «Fontes», reunindo, na devida ordem, 142, 132, 114 e 93 termos.

Por fim, e com um menor número de termos, identificam-se as categorias «Geografia», «Coroa», «Vida privada», «Recursos de informação», «Áreas da História», «Reinados» e «Ciências e técnicas historiográficas», que agregam 47, 40, 38, 22, 17, 14 e 9 termos, respetivamente.

É notória a existência de diferentes níveis de profundidade das categorias do vocabulário taxonómico em análise. Esta situação deve-se ao facto das categorias deterem diferentes escopos de análise. Outro dos fatores relaciona-se com a distribuição desigual das fontes disponíveis para cada uma das categorias e a forma como os diferentes temas se encontram representados na produção científica universitária sobre História Medieval.

No que respeita à avaliação da taxonomia, em termos gerais, solicitou-se ao medievalista a avaliação semântica da taxonomia (atualidade e comunicabilidade dos termos) e ao especialista da área da representação da informação a avaliação formal (morfológica e sintática) da mesma. Para tal, foram elaboradas duas grelhas de análise, de acordo com as recomendações da ANSI/NISO Z39.19-2005. Os resultados de ambas as avaliações foram bastante positivos e funcionaram de forma complementar, tal como é inerente a qualquer investigação de índole interdisciplinar e que, no contexto em observação, se assume como fundamental pela necessidade de validação terminológica conjunta.

As tabelas I e II apresentam os critérios de avaliação delineados e a respetiva pontuação atribuída pelos especialistas a cada um deles (cf. anexo 3).

#### 5. Conclusões

Em jeito de conclusão tecem-se algumas considerações fundamentais:

Em primeiro lugar, importa salientar que a taxonomia construída, enquanto uma investigação aplicada, testou com sucesso os métodos já experimentados por especialistas para desenhar, desenvolver e manter vocabulários controlados desta natureza.

Em segundo lugar, há que ter em linha de conta que embora este vocabulário taxonómico seja representativo de um país e de um momento histórico em particular, mesmo assim poderá funcionar como ponto de partida para a construção de outros vocabulários controlados com o

mesmo escopo temático, mas relativos a outras geografias, com o mesmo recorte cronológico.

Sublinha-se, por fim, o caráter pioneiro deste vocabulário taxonómico no panorama dos estudos medievais portugueses. A sua elaboração, exatamente pela ausência de estruturas comparativas ao nível nacional e internacional, constituiu um enorme repto, só alcançado pelo trabalho interdisciplinar entre medievalistas e profissionais I&D. Por isso mesmo, reforça-se a ideia de que a taxonomia elaborada representa, acima de tudo, uma primeira tentativa de categorização terminológica neste domínio disciplinar. Enquanto tal, ela é discutível, encontrando-se os seus termos em constante evolução formal e semântica e, de igual modo, a serem testados atualmente em bases de dados especializadas.

## 6. Desenvolvimentos futuros

No que respeita aos desenvolvimentos futuros, deve procurar-se gerir o conhecimento contido na taxonomia com um maior alcance, isto é, para além daquele contido na organização hierárquico-classificatória do seu vocabulário. Pretende-se, assim, orientar o vocabulário taxonómico para uma esquematização que permita ligar os assuntos da investigação que representa a outros vocabulários existentes e com os quais detenha afinidades. Para alcançar esta interoperabilidade, terá que se estabelecer esquemas SKOS que codifiquem a taxonomia em XML e que permitam a sua migração no âmbito da web semântica (Daconta e outros, 2003).

Portanto, este vocabulário taxonómico é, ainda, fundamentalmente terminológico, não detendo, por isso, uma natureza ontológica. Por conseguinte, e numa fase mais avançada, este será, sem dúvida, o grande desafio futuro que se coloca ao vocabulário taxonómico aqui apresentado.

## Referências

Referências bibliográficas sobre História Medieval

Nota: Na impossibilidade de aqui apresentar todas as referências bibliográficas que constituíram as nossas fontes de informação para a obtenção de dados, optámos por incluir apenas aquelas que dizem respeito a pontos de situação da historiografia medieval portuguesa e, também, algumas obras de referência da especialidade que serviram de base ao estabelecimento das temáticas, das cronologias e de determinados conceitos.

Branco, Maria João Violante; Costa, Adelaide Millán da (1992). *História da Idade Média*: caderno de apoio. Lisboa: Universidade Aberta. ISBN 972-674-370-2.

Coelho, Maria Helena da Cruz (1991). *História medieval de Portugal: guia de estudo*. Porto: Universidade Portuguesa. ISBN 972-9354-07-3.

Marques, A. H. de Oliveira Marques (1988). *Guia do estudante de História medieval portuguesa*. 3.<sup>a</sup> ed. Lisboa: Estampa.

Mattoso, José (dir.) (1997). *História de Portugal: a monarquia feudal (1096-1480)*. Lisboa: Estampa, vol. 2. ISBN 972-33-1263-8.

Moreno, Humberto Baquero (1995). *História de Portugal medieval: político e institucional*. Colab. Maria da Conceição Falcão Ferreira, Luís Carlos do Amaral, Luís Miguel Duarte. Lisboa: Universidade Aberta, 2 vols. ISBN 972-674-132-7 (vol. 1). ISBN 972-674-135-1 (vol. 2).

Ramos, Rui; Sousa, Bernardo Vasconcelos e; Monteiro, Nuno Gonçalo (2009). *História de Portugal*. Lisboa: A Esfera dos Livros. ISBN 978-989-626-139-9.

Serrão, Joaquim Veríssimo (1994). *História de Portugal: Estado, pátria, nação*. 5.<sup>a</sup> ed. Lisboa: Verbo, vol. 1. ISBN 972-22-0271-5.

Serrão, Joel; Marques, A. H. de Oliveira (dir.) (1987). *Nova História de Portugal: Portugal na crise dos séculos XIV e XV*. Coord. A. H. de Oliveira Marques. Lisboa: Estampa, vol. 4.

Serrão, Joel; Marques, A. H. de Oliveira (dir.) (1996). *Nova História de Portugal: Portugal em definição de fronteiras (1096-1325): do Condado Portucalense à crise do século XIV*. Coord. Maria Helena da Cruz Coelho, Armando Luís de Carvalho Homem. Lisboa: Presença, 1996. ISBN 972-23-2039-4, vol. 3.

Tavares, Maria José Ferro (1990). *Sociedades e culturas portuguesas*. Lisboa: Universidade Aberta, 2 vols. ISBN 972-674-030-4 (vol. 1). ISBN 972-674-049-5 (vol. 2).

Tavares, Maria José Ferro (1992). *História de Portugal medieval: economia e sociedade*. Lisboa: Universidade Aberta. ISBN 972-674-100-9.

Referências bibliográficas sobre representação da informação:

ANSI/NISO Z39.19-2005. Guidelines for the construction, format and management of monolingual controlled vocabularies. NISO Press: Bethesda, MD. ISBN 1-880-124-65-3.

Campos, Maria Luiza; Gomes, Hagar Espanha. Taxonomia e classificação: a categorização como princípio. *Enancib: Encontro Nacional de Pesquisa em Ciência da Informação*. [Electronic Version]. Salvador, 8, 2007. Retrieved 22 sep. 2010, from <http://www.enancib.ppgci.ufba.br/artigos/GT2--101.pdf>

Conway, Susan; Sligar, Char (2002). *Unlocking knowledge assets*. Redmont: Microsoft Press. ISBN 0-7356-1463-6.

Cumming, Maewin (2003). Tomatoes are not the only fruit: a rough guide to taxonomies, thesauri, ontologies and the like. *Record Management Bulletin*, nº 113, 1-4.

Daconta, M.; Obrst, L. J.; Smith, K. T. (2003). The semantic web. A guide to the future of XML, web services and knowledge management. Indianapolis: Wiley.

Gilchrist, Alan (2003). Thesauri, taxonomies and ontologies: an etymological note. *Journal of documentation*, vol. 59, (1) 7-18. ISSN 0022-0418.

Jagerman, Evert (2006). *Creating, maintaining and applying taxonomies*. Zoetermeer: E. Jagerman. ISBN 90-811-274-1-1.

Medeiros, Filipa (2014). *A historiografia medieval portuguesa na viragem do milénio: análise bibliométrica (2000-2010) e representação taxonómica* [Electronic Version]. Retrieved 22 oct. 2014, from <https://dspace.uevora.pt/rdpc/handle/10174/11220?mode=full>

Moreiro González, José Antonio (2006). La representación y recuperación de los contenidos digitales: de los tesauros conceptuales a las folksonomías. In: Tramullas, J. (coord.). *Tendencias en Documentación Digital*. Gijón: TREA: 81-109.

- Moreiro González, José Antonio (2011). *Linguagens documentárias e vocabulários semânticos para a web: elementos conceituais*. Salvador: EDUFBA. ISBN 978-85-232-0824-0.
- NP 4036 (1992). Documentação. *Tesaurus monolingues: directivas para a sua construção e desenvolvimento*. Lisboa: Instituto Português da Qualidade.
- Portugal. Biblioteca Nacional. Área de Indexação e Classificação (1998). *SIPORBASE: sistema de indexação em português: manual*. 3.<sup>a</sup> ed. rev. e aum. Lisboa: Biblioteca Nacional. ISBN 972-565-154-5.
- Zhonghong, Wang; Chaudhry, Abdus Sattar; Khoo, Christopher (2006). Potencial and prospects of taxonomies for content organization. *Knowledge organization: international journal devoted to concept theory, classification, indexing and knowledge representation*, vol. 33 (3), 160-170. ISSN 0943-7444.

## Anexo 1

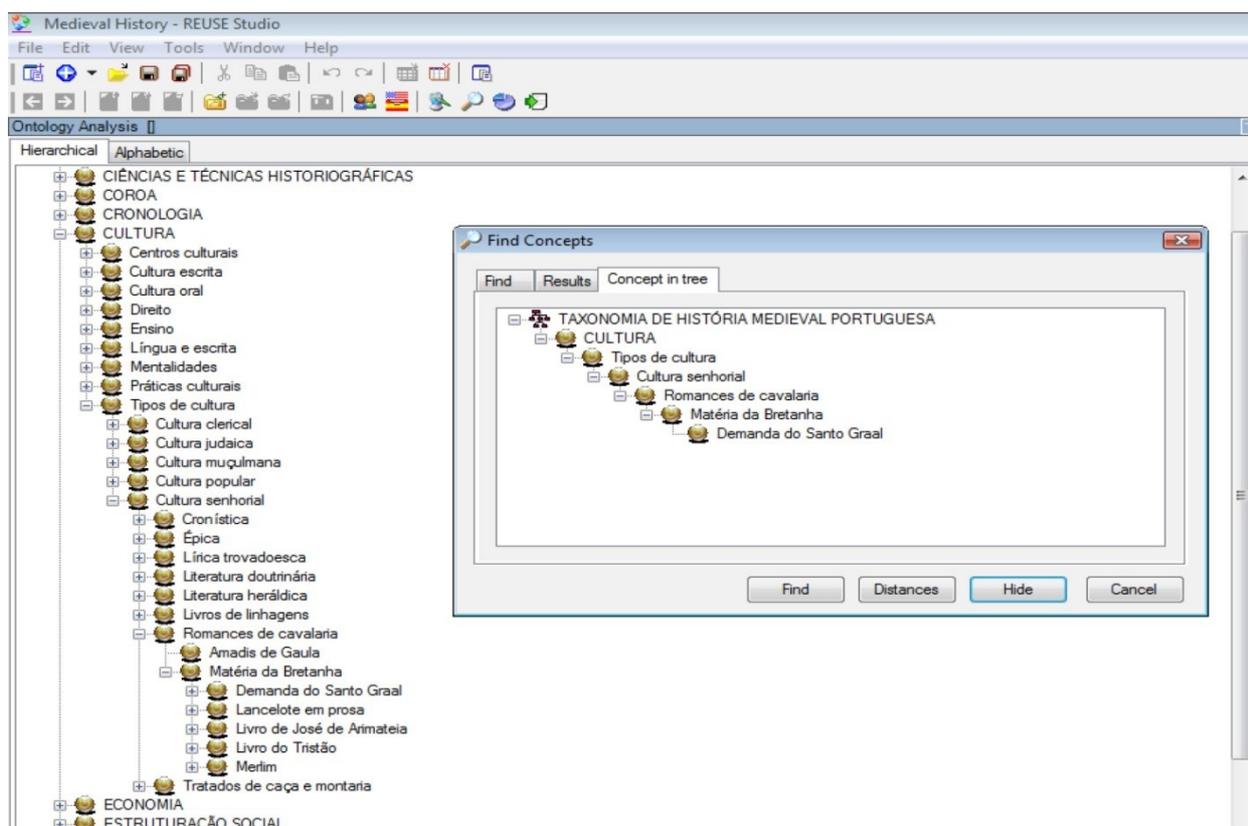


Figura 1: Visão geral da taxonomia de História Medieval portuguesa no software KM

## Anexo 2

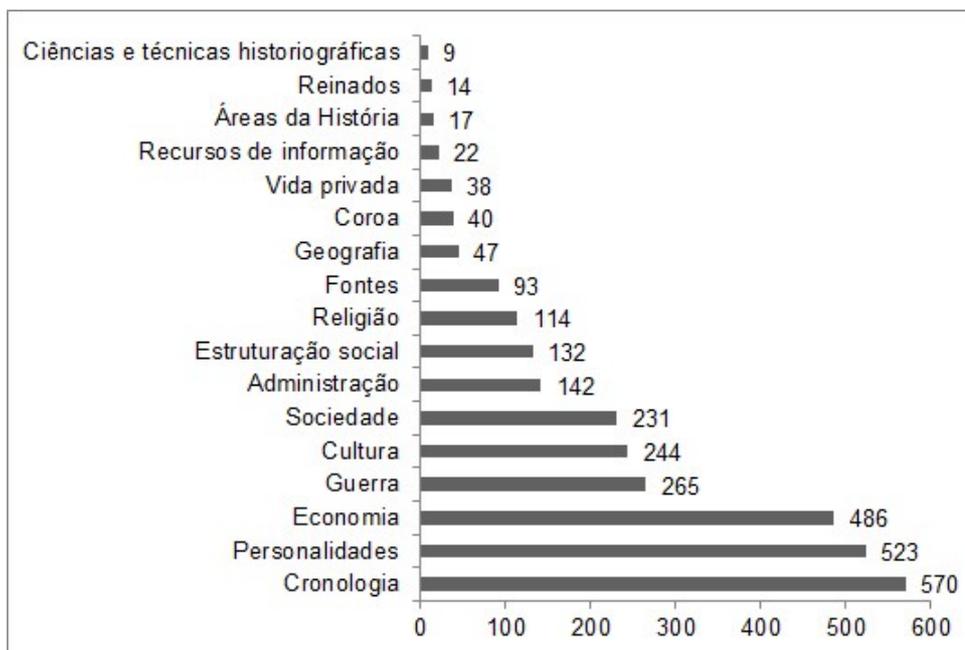


Gráfico 1: Distribuição do nº de termos pelas categorias da taxonomia de História Medieval Portuguesa

## Anexo 3

| <b>CRITÉRIOS DE AVALIAÇÃO</b>  | <b>PONTUAÇÃO</b> |
|--|------------------|
| 1. Fontes de informação gerais e específicas utilizadas  | Muito bom        |
| 2. Uniformidade ao nível da representação das várias áreas temáticas incluídas na taxonomia  | Muito bom        |
| 3. Pertinência semântica dos termos face ao vocabulário utilizado pelos utilizadores (atualidade dos termos)   | Muito bom        |
| 4. Adequação da estrutura hierárquica  | Muito bom        |
| 5. Eficiência do sistema de navegação  | Muito bom        |
| 6. Eficácia do sistema de pesquisa (recuperação e opções de pesquisa)  | Muito bom        |
| 7. Observações/Sugestões:<br>Solicitações de pareceres específicos a especialistas oriundos dos vários domínios temáticos abrangidos pela taxonomia. |                  |

Tabela I: Grelha de avaliação (medievalista)

| <b>CRITÉRIOS</b>  | <b>PONTUAÇÃO</b> |
|---|------------------|
| 1. Fontes de referência e procedimentos utilizados na normalização do vocabulário | Muito bom        |
| 2. Consistência morfológica e sintática dos termos                                | Muito bom        |
| 3. Correção concetual ao nível do estabelecimento das relações hierárquicas       | Muito bom        |
| 4. Eficiência do sistema de navegação   | Muito bom        |
| 5. Eficácia do sistema de pesquisa (recuperação e opções de pesquisa)             | Bom              |
| 6. Observações/Sugestões:<br>Diversificação e refinamento do leque de pesquisas.  |                  |

Tabela II: Grelha de avaliação (especialista em KOS)

---

# Apropriação do Conhecimento na Web de Dados - Gestão de Recursos Computacionais em Sistemas Abertos

*Web de Dados e os Sistemas Abertos- Desafios na Gestão e Organização do Conhecimentos em Bibliotecas Digitais*

---

**Francisco Carlos Paletta**

Universidade de São Paulo, Av. Prof. Lúcio Martins Rodrigues, 443 - São Paulo, Brasil, CEP 05508-020  
– São Paulo, fcpaletta@usp.br

## Resumo

Como resultado deste estudo espera-se uma contribuição para o entendimento dos recursos da tecnologia colaborativa utilizados em ambientes informacionais digitais. Com base nos recursos identificados e coletados em bibliotecas e repositórios digitais deseja-se observar como são aplicados os recursos da tecnologia colaborativa no contexto da Web Semântica, também chamada de Web de Dados. A inserção dessas tecnologias apresenta-se como inovação que devem estar vinculadas à tradição e a missão das bibliotecas e dos repositórios de informação na gestão e organização do conhecimento em sistemas abertos. Avaliar a flexibilidade das estruturas computacionais, sua atratividade e dinâmica na qual o usuário torna-se agente na construção de seu ambiente, demandando recursos de customização e personalização na criação de Serviços de Informação inovadores que permitam à Biblioteca continuar a ocupar papel relevante na produção de novos conhecimentos por parte de seus Usuários.

**Palavras-Chave:** Web de Dados. Sistemas Abertos. Recursos Computacionais. Biblioteca Digital. Gestão do Ciclo de Vida de TI. Organização do Conhecimento. Usuário da Informação.

## 1. Introdução

A Web é uma rede cujos conteúdos estão interligados através de documentos de hipertexto. Seu estudo é possível por um processo de análise e coleta sucessiva das páginas, a partir de um conjunto de sites previamente conhecidos. Essa busca é feita de forma automática por um programa de computador normalmente chamado de “crawler” – programa que visita Web Sites e lê suas páginas e qualquer outra informação com objetivo de criar entradas e indexação para as ferramentas de busca. Nem toda a Web está interligada, contudo, embora a maior parte dela esteja: há “ilhas” de tamanhos variados sem ligação com o restante da rede. Isso significa que o conjunto inicial de sítios a partir dos quais

## Abstract

As a result of this study, it is expected to contribute to the understanding of collaborative technology features used in digital information environments. Based on resources identified and collected in digital libraries and repositories, it is intended to observe how the resources of collaborative technology are applied in the context of the Semantic Web also called Data Web. The integration of these technologies is presented as an innovation that have to be linked to tradition and to the mission of the information units in open systems management and knowledge organization. Evaluate the computational structures flexibility, their attractiveness and dynamic in which the user becomes an agent in the environment construction, requiring a custom creation of innovative information services, enabling the Library to maintain its relevant role in the production of new knowledge by their users.

**Keywords:** Web of Data. Open Systems. Computational Resources. Digital library. Management of IT Lifecycle. Production and Organization Knowledge. Information User.

a pesquisa é feita influencia o resultado, e encontrar o conjunto adequado, geralmente o mais completo possível, é um passo importante. O primeiro princípio da Web, proposto pelo W3C Brasil (World Wide Web Consortium), afirma que “o principal valor da Web é o social. Mais do que tecnológico, este é um ambiente de comunicação humana, de transações comerciais, de oportunidades para compartilhar conhecimentos e, para ser um ambiente universal, deve estar disponível para todas as pessoas, independentemente dos equipamentos e softwares que utilizem, principalmente da cultura em que se inserem da localização geográfica, das habilidades físicas ou mentais, das condições sócio econômicas ou de instrução. A universalidade

da Web só pode ser garantida e aprofundada com um modelo de governança democrático e pluralista que tenha foco no acesso por todos e na sua própria evolução tecnológica. (.gov.br, 2010).

A temática da inclusão digital vem sendo tratada, desde os anos 90, como a necessidade de permitir o acesso a computadores e ferramentas de TICs (Tecnologia da Informação e Comunicação). No entanto, essa perspectiva em explicar muito com muito pouco da inclusão digital vem sendo substituída, no novo milênio, por proposições da inclusão que tratam não apenas do acesso às ferramentas Digitais, mas também dos usos e apropriações dos conteúdos disponibilizados na WEB pelos internautas. Promover a inclusão digital e, por conseqüência, a inclusão social não significa apenas promover as ferramentas, mas possibilitar seu uso de forma crítica, estimulando o aperfeiçoamento das potencialidades informativas e cognitivas e também, as atividades cidadãs. Na sociedade em rede identificam-se duas “ondas”. Na primeira delas, a questão central girava em torno da necessidade da inclusão digital. Já na segunda, vivida atualmente, se evidencia a presença marcante de uma geração de nativos Digitais. Assim a preocupação deixa de ser o aprendizado de ferramentas básicas de navegação na WEB e desloca-se, mais especificamente, para diferentes formas de apropriação e de produção do conhecimento na Internet. Este novo foco traz ao centro do debate questões de inclusão social através da inclusão digital e das práticas sociais e educacionais vigentes nas culturas conectadas (Passarelli, 2012).

Com o presente estudo objetiva-se verificar a apropriação que as Bibliotecas e os Repositórios Digitais têm feito das tecnologias colaborativas em ambientes de informação digital e sistemas abertos bem como a gestão da infraestrutura de seus recursos computacionais.

## 2. Usuário da Informação e Competência Informacional

No novo contexto mundial definido pela globalização e pela mudança tecnológica, o conhecimento tornou-se principal riqueza das nações, das empresas e das pessoas, podendo também vir a constituir o principal fator de desigualdade. A Sociedade da Informação é a pedra angular das Sociedades do Conhecimento. O conceito de “Sociedade da Informação” está relacionado à idéia da “inovação tecnológica”, enquanto o conceito de “Sociedade do Conhecimento” inclui uma dimensão de transformação social, cultural, econômica, política e institucional, assim como

uma perspectiva mais pluralista e de desenvolvimento. O conceito de “sociedades do conhecimento” expressa a complexidade e o dinamismo das mudanças que estão ocorrendo. O conhecimento em questão não só é importante para o crescimento econômico, mas também para fortalecer e desenvolver todos os setores da sociedade (Matias, 2005).

Competência Informacional ou Information Literacy esta ligada a necessidade de se desenvolver nos indivíduos aptidões sobre habilidades e competências relacionadas ao acesso, uso e disseminação da informação, objetivando fazer uso desta de forma ética e eficiente, para que o ser humano através de seu intelecto e processo cognitivo possa produzir novo conhecimento (Dudziak, 2001). A expressão Information Literacy tem suas origens no surgimento da sociedade da informação, caracterizada pelo rápido crescimento da informação disponibilizada e as mudanças ocasionadas pela tecnologia usada para gerar, disseminar, acessar e usar a informação (Melo, 2007).

A noção de fácil acesso à informação por meio da tecnologia a partir do estabelecimento de redes, Internet e as telecomunicações, criou uma noção errônea do imperativo tecnológico como resposta às deficiências de comunicação e educacionais da sociedade. O conhecimento e o uso deste ferramental tecnológico são essenciais nos dias de hoje, porém é preciso considerar que por si só a tecnologia não leva à comunicação e à educação. A competência informacional esta fortemente relacionada ao processo de interiorização de conhecimentos, habilidades e valores ligados à informação e ao aprendizado (Dudziak, 2001). Em um contexto mais prático converge para um conjunto de habilidades necessárias para localizar, interpretar, analisar, sintetizar, avaliar e comunicar a informação em diferentes ferramentas e suportes (Dudziak, 2003).

O comportamento de busca e uso de informação são modelados pelo estilo cognitivo do indivíduo e por fatores que geram o encontro do usuário com os sistemas de informação ou as conseqüências de tal confronto. É preciso atentar para o fato de que não é mais possível se limitar à tarefa de localizar fontes de informação, não levando em consideração as tarefas de interpretação, formulação e aprendizagem envolvidas no processo de busca e acesso a informação. A velocidade de acesso a informação no universo digital, requer serviços que estejam centrados no significado da busca mais do que meramente na localização da fonte.

Nessa perspectiva, os usuários da informação não podem ser vistos apenas como um dos integrantes do sistema, mas como a “razão de ser” do Serviço de Informação. Sistemas de Informação organizados nessa perspectiva tradicional concentram-se prioritariamente na aquisição e administração de grandes coleções de materiais. Assumiu-se, durante décadas, que as atividades técnicas dos sistemas eram o seu ponto estratégico. Considerava-se que os usuários utilizavam o sistema exatamente da maneira como estes tinham sido projetados. Não se imaginava indagar, aos sistemas, questões imprescindíveis sobre a identidade e propósitos principais de seus usuários. Como a informação era considerada como algo existente fora das pessoas e passível de ser transferida de uma para outra, parecia ser possível que eficiência e sucesso das operações de um sistema pudessem ser medidos em função do número de fontes de informações recuperadas pelo sistema versus o que realmente foi de interesse do usuário. Isso, na realidade, coloca novamente o usuário como um processador imperfeito da informação, pois já é sabido que nem todas as pessoas se interessam pelas mesmas fontes indicadas. Resulta desse procedimento que hoje em dia se conhece muita coisa sobre planejamento, aquisição, organização, controle e desenvolvimento de coleções, mas muito pouco sobre como as pessoas fazem uso dos sistemas ou para que fins e como a informação que é a matéria prima dos sistemas está sendo utilizada (Ferreira, 1995).

Considerando as sete faces da Information Literacy: Tecnologia da Informação, Fontes de Informação, Processo de Informação, Controle da Informação, Construção do Conhecimento, Extensão do Conhecimento e Inteligência (Bruce, 2003), podemos afirmar que neste contexto, as Bibliotecas são vistas como modelo de ambiente informacional e como espaço de aprendizagem. Os bibliotecários são educadores, ativamente envolvidos com os processos de ensino aprendido. Sua crença se baseia no aprendizado independente, auto-orientado e no aprendizado baseado em recursos informacionais (Passarelli, 2008).

O ser humano esta constantemente em buscas que permitam renovar os seus conhecimentos. Neste contexto, observa-se o surgimento uma nova forma de interatividade entre usuário e a Internet: a Web Semântica - uma Web Inteligente com toda sua informação organizada de forma que não somente seres humanos possam entendê-la, mas principalmente máquinas.

A construção de uma Internet mais inteligente caminha fortemente na direção de produzir uma revolução no universo digital da organização da informação e do conhecimento. Com o uso de novas tecnologias é imperativo o uso das Tecnologias da Informação e das Comunicações em tornar os processos de busca de informação e a geração de novo conhecimento mais ágeis. É neste ponto que surge um novo usuário da informação com novas demandas por recursos computacionais e novas capacidades em produzir novos conhecimentos.

Literacia Informacional tem constituído, nos últimos anos, novo campo de pesquisa que interessa à educação, à ciência da informação e às ciências cognitivas. O desafio esta em, num primeiro momento, aprender a utilização básica dos recursos tecnológicos - literacia digital - e a seguir apropriar-se dos mesmos para gerar novos conhecimentos - literacia informacional. Entretanto esta apresentação reducionista de ambas as formas de literacia gestadas no bojo da sociedade em rede não dá conta de delinear as profundas rupturas imersas nas mesmas.

No contexto da literacia informacional os desafios são múltiplos e multifacetados abarcando pelo menos: a dificuldade das Bibliotecas com suas coleções caríssimas em concorrer com acervos Digitais, as Bibliotecas virtuais, e os motores de busca oferecendo informações de forma imediata e a custo zero. A desconstrução do conceito de autor individual e a emergência dos coletivos digitais; a nova valoração das fontes de informação bem como as novas competências que os Profissionais da Informação precisam desenvolver para atuar como gestores da informação em ambientes em rede (IFLA, 1997).

Os novos contornos da sociedade em rede têm constituído objeto de estudo e pesquisa para profissionais das mais diversas áreas de atuação. A importância das redes sociais como fenômeno emergente dos coletivos digitais como autores e consumidores da informação e produtores de conhecimento introduz profundas rupturas nos papéis tradicionais de editores e Bibliotecários, de tal forma que atualmente é difícil diferenciar se um conteúdo esta editado por uma editora ou se é uma auto-edição na WEB 2.0. Estudos indicam que a produção de conteúdo na WEB em 2007 concentrava-se nas mãos de uma população mais madura e com maior grau de literacia informacional. Entretanto novas pesquisas precisam ser realizadas, agora, para verificar se a tendência se confirma ou se dissipa (Sousa, 2010).

### 3. Produção e Organização do Conhecimento na WEB de Dados

Dado que os saberes científicos e as inovações tecnológicas estão desigualmente repartidos entre países ricos e pobres, por níveis educacionais e faixas etárias, a problemática da diversidade cultural e os estudos sobre ela devem fazer parte da consideração teórica, da investigação empírica e do planeamento de políticas neste campo. Também é necessário delimitar o alcance da posição oposta, que afirma, a partir da antropologia, que todas as sociedades, em todas as épocas, foram sociedades do conhecimento, ou seja, que todo o grupo humano dispôs de um conjunto de saberes apropriado ao seu contexto e aos seus desafios históricos (Choo, 2003).

A organização do conhecimento liga os três processos de uso estratégico da informação - a criação de significado, a construção do conhecimento e a tomada de decisões - num ciclo contínuo de aprendizagem e adaptação que podemos chamar de ciclo do conhecimento. Entre os elementos mais importantes que influenciam o uso da informação estão as atitudes do indivíduo em relação à informação e a sua busca, atitudes essas que são fruto da educação, do treinamento, da experiência passada, das preferências pessoais. O risco aqui é de uma super simplificação, de ver o usuário da informação como alguém que quer extrair informações específicas e definitivas no menor tempo possível, ou como alguém disposto a investir esforço para buscar e explorar bases de dados. A verdade é que os indivíduos oscilam continuamente entre extrair e explorar, e que o uso da informação é um processo confuso, desordenado, sujeito aos caprichos da natureza humana, como qualquer outra atividade (Canclini, 2009).

A informação sempre foi uma importante fonte de poder, responsável por controlar e administrar povos. Observamos que a tecnologia da informação objetiva a integração das operações das organizações e entre as organizações, com ganhos de agilidade e redução de custos operacionais. Integrando o fluxo de informações, os novos sistemas assumem certas tarefas, eliminam controles nas entradas e saídas de processos, diminuindo a duração dos ciclos operacionais. Embora a escolha dos sistemas computacionais seja fundamental para a geração de vantagem competitiva, exercer a capacidade de escolha entre sistemas ainda é pouco. Independentemente do sistema, as operações podem ser melhoradas a partir de um processo de reengenharia, através do mapeamento dos processos, da avaliação crítica, do redesenho e da

implementação do fluxo da informação (Sousa, 2010).

Atualmente existe consenso para caracterizar três gerações da Web. A Web 1.0 para designar a primeira geração comercial da Internet com conteúdos de baixa interatividade. A geração Web 2.0 atualmente entre nós, caracterizada por redes sociais e folksonomias - sites onde os usuários agregam valor a conteúdos com valor pessoal. A Web de Dados foi também denominada Web Semântica por Tim Berners-Lee ao final dos anos 90, para denominar uma Web com maior capacidade de busca e auto-reconhecimento dos conteúdos por meio de metadados com descrições ligados aos conteúdos originais (Passarelli, 2008).

A Internet é provavelmente a mais sofisticada tecnologia de informação e comunicação atualmente disponível para a sociedade, em função da sua forma de organização e de seus impactos nas esferas tecnológicas, social, econômica e política. Ela é também a infraestrutura necessária para uma de suas maiores e mais conhecida aplicação: a Web. Principal responsável pela popularização da Internet, a ponto de hoje ser confundida com esta. Internet e Web são, portanto, conceitos distintos. A Web pode ser definida, como a parte da Internet acessada por meio de navegadores, ou browsers. O impacto do uso da Internet e da Web na sociedade, nos indivíduos e nas organizações tornou-se objeto de pesquisa, extrapolando o campo especializado da computação aplicada, e atingindo áreas de estudos organizacionais e sociológicos. Por ser essencialmente dinâmica e sem fronteiras, tanto do ponto de vista físico como virtual, é importante que seja conhecida em detalhes, tanto para assegurar sua livre transformação quanto para permitir sua disponibilidade, confiabilidade e acessibilidade por todos.

A Web Semântica ou Web de dados é a terceira geração da Internet. Esta nova geração prevê que os conteúdos online estarão organizados de forma semântica, muito mais personalizada para cada usuário, sites e aplicações inteligentes e publicidade baseada nas pesquisas e nos comportamentos de busca, acesso e apropriação na informação no universo digital. Esta nova Web também pode ser chamada de a "Web Inteligente" (Paletta, 2014).

As constantes mudanças nas formas de organização tradicional da informação presentes em ambientes informacionais digitais são reflexos da incorporação das Tecnologias de Informação e Comunicação (TICs), como no caso das Bibliotecas e Repositórios Digitais, que armazenam, preservam, disseminam e permite o acesso a

produção intelectual da comunidade científica e acadêmica, visando contribuir para o aumento da visibilidade e do valor da instituição ao agregar recursos que possibilitam o processo de construção do conhecimento, a partir da participação colaborativa aplicada em diferentes ambientes.

A Web 2.0 segue uma filosofia com princípios de leitura e escrita de natureza participativa, em que cada usuário pode intervir diretamente na escolha e introdução de dados no âmbito de cada site; é cooperativa, uma vez que compartilha ideias, preferências, informações e conhecimento; é interativa, na medida em que, através de toda a gama de recursos multimídia, é possível um diálogo simultâneo com os usuários; é democrática, pois sob essa filosofia existe liberdade de expressão, de pensamento, e, sobretudo, de trânsito de informações, independentemente dos interesses de cada um. É também socio-técnica, pois, através de todas as suas características, é possível um intercâmbio de culturas, religiões, etnias e outros.

No contexto da Biblioteca 2.0 a maior parte dos pesquisadores concordaria que muito do que as Bibliotecas aprovaram na primeira revolução da Web são estáticos. Por exemplo, catálogos online de acesso público (OPAC) exigem que os usuários busquem a informação. Embora muitos estejam iniciando a incorporar técnicas da Web 2.0 relativas à pesquisa de dados. Do mesmo modo, a primeira geração de Biblioteca online foi elaborada através de textos tutoriais estáticos e que não respondiam às necessidades dos usuários, nem permitiam que interagissem uns com os outros. As Bibliotecas, porém, tem começado a evoluir para uma estrutura mais interativa, meios de comunicação social e rico em tutoriais, programação e animação com o uso de banco de dados mais sofisticados.

A Web 2.0 nas Bibliotecas pode ser uma ferramenta que possibilite a gênese de uma base de conhecimento a partir da inteligência coletiva, como também ferramenta para a gestão do conhecimento que facilite, de maneira interativa, a descoberta dos mesmos. Passamos de uma Biblioteca para o usuário para uma Biblioteca com o usuário. Em este integrar e complementar estes dois paradigmas, devemos também considerar, conhecer, explorar e avaliar as novas ferramentas de comunicação, organização, participação e construção coletiva do conhecimento que estão disponíveis na Web.

São aplicações, na sua maioria intuitiva, gratuita e que respondem as novas necessidades de informação e participação da comunidade. Deste modo, fica clara a transformação, ou seja, a

nova abordagem da relação entre informação e o conhecimento no contexto das Bibliotecas e dos Repositórios, ao contrário dos ambientes tradicionais que só permitia ao usuário uma única forma de se relacionar com o conteúdo armazenado fisicamente.

Atualmente, não só o usuário participa como também adiciona conteúdos que ao serem compartilhados em outros recursos colaborativos passam a receber comentários dos membros vinculados a sua comunidade agregando valor à informação compartilhada, além de criar vários caminhos para a localização do próprio recurso. Assim, temos uma nova relação e interação na qual aumenta o nível sofisticado de atitudes sociais eletrônicas (Observatório Ciber-sociedad, 2008).

Na Web 1.0 conecta informação. O usuário desempenha o papel de espectador, o conteúdo é pouco interativo. A Web 2.0 conecta pessoas. O foco esta na construção coletiva do conhecimento. A essência da Web Social é permitir que os usuários não sejam mais apenas espectadores, e sim que eles se tornem colaboradores. A Web Semântica ou Web de Dados conecta conhecimento. Programas serão capazes de interpretar nossas preferências e guiar nossa navegação pela Web. Esta ligada a um conjunto de tecnologias com formas mais eficientes para ajudar os computadores a organizar e analisar a informação disponível na rede. É uma das grandes propostas para o futuro da Internet, pois será ela que definitivamente organizará todas as informações que estejam na Internet. Fará com que todos os aplicativos baseados na Web sejam Open-Source e viabilizar uma grande interatividade em diversas áreas da Web.

O universo da informação digital se expande no contexto do que hoje chamamos de Big Data onde a busca por informações fica cada vez mais difícil uma vez que o universo digital esta constituído por dados não estruturados que precisam ser organizados, acessados, apropriados e então utilizados na produção de novos conhecimentos. A Web de Dados virá com o objetivo de organizar estas informações para que os usuários tenham mais facilidade na busca da informação. A busca Semântica como é chamada os buscadores da Web de Dados, organizam informações por assuntos determinados, conectando conhecimentos. A busca Semântica da Web de Dados dividirá os resultados em categorias para que a busca da informação seja mais rápida e organizada.

A Figura 1 relaciona o Grau de Conectividade da Informação com o Grau de Conectividade Social de acordo com a evolução tecnológica:

Web - Conectando Informação, Web Social - Conectando Pessoas, Web Semântica – Conectando Conhecimento e a Metaweb – Conectando Inteligência.

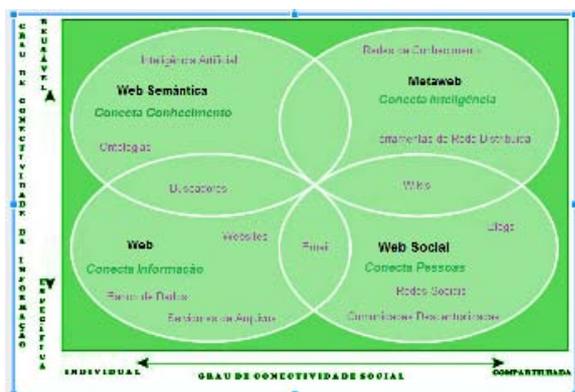


Figura 1. *Evolução da WEB.* Fonte: <https://swedinbalchik.wordpress.com/category/studies-work/algc/> – Adaptação: Ligia Capobianco

Biblioteca 3.0 refere-se a Bibliotecas utilizando tecnologias como a Web Semântica, a computação em nuvem, dispositivos móveis, e reimaginando o uso de tecnologias estabelecidas, tais como pesquisa federada (RFID), para facilitar a disseminação do conteúdo gerado pelo usuário e a colaboração para promover e fazer coleções de Bibliotecas e Unidades de Informação acessíveis. O resultado final da Biblioteca 3.0 é a expansão da "Biblioteca sem fronteiras", onde as coleções podem ser feitas facilmente disponíveis aos usuários da Biblioteca, independentemente da sua localização física. Biblioteca 3.0 é um complemento virtual para os espaços físicos da Biblioteca (Belling, 2010).

#### 4. Gestão do Ciclo de Vida da Tecnologia da Informação em Biblioteca Digital

A inovação tecnológica constitui uma ferramenta essencial para aumentar a produtividade e a competitividade das organizações, assim como para impulsionar o desenvolvimento econômico da sociedade. O desenvolvimento não deriva de um mero crescimento das atividades econômicas existentes, mas reside fundamentalmente em um processo qualitativo de transformação da estrutura produtiva no sentido de incorporar novos produtos e processos e agregar valor à produção por meio da intensificação do uso da informação e do conhecimento.

A Universidade atua como organismo gerador, transmissor e receptor de conhecimentos e a

Biblioteca através dos Serviços de Informação, torna-se consciente de sua função intermediadora realizando os processos documentários e preservando a informação para sua próxima transformação em conhecimento em uma espiral de evolução científica e tecnológica. Neste contexto a Universidade tem como foco a socialização dos saberes e a Biblioteca atua como instrumento de socialização. As funções básicas da Biblioteca derivam dessa dinâmica social que, em um movimento circular, fornece insumos para sua própria continuidade. Dentro dessa dinâmica, visualizamos as funções de:

- **Armazenagem do Conhecimento:** desenvolvimento de coleções, memória da produção científica e tecnológica, preservação e conservação;
- **Organização do Conhecimento:** qualidade de tratamento temático e descritivo que favoreça o intercâmbio de registros entre Bibliotecas e sua recuperação;
- **Acesso ao Conhecimento:** a exigência de informação transcende o valor, o lugar e a forma e necessita de acesso. Por isso devemos pensar não só em fornecer a informação, mas possibilitar o acesso simultâneo de todos.

Essas três funções estão presentes em toda a evolução do processo de socialização do conhecimento realizado pela Universidade ao longo dos tempos, mesmo considerando a permanente mudança dos formatos documentários para registro do conhecimento e seu modo de acesso. A Biblioteca insere-se em um contexto universitário cujos objetivos maiores são o desenvolvimento educacional, social, político e econômico da sociedade humana (Fujita, 2005).

Uma solução integrada de gestão de ativos combina as disciplinas de gerenciamento de recursos e de serviços da Biblioteca Digital em uma única arquitetura baseada na WEB, ajudando a unir departamentos e processos diferentes. Ao gerenciar ativamente todo o ciclo de vida dos recursos, a solução permite a Biblioteca Digital eliminar custos desnecessários de SW e HW, a gerenciar proativamente contratos com fornecedores e alinhar os recursos dos serviços com os modelos de Governança da Tecnologia da Informação com objetivo de assegurar a otimização dos investimentos em TI (Turban, 2006). Os benefícios incluem:

- Monitorar a configuração, as versões implementadas, as informações históricas dos recursos de TI, uso do software e hardware para realocação e negociação de contratos

- Assegurar a disponibilidade dos recursos através do gerenciamento de incidentes

Com base neste cenário, torna-se fundamental analisar os fatores críticos que devem ser considerados pelas Bibliotecas Digitais na gestão do ciclo de vida de seus recursos de tecnologia da informação. O gerenciamento da infraestrutura de TI torna-se cada vez mais caro e complexo. Estudos indicam que mais que 50% de todos os gastos de TI são alocados para configurar, atualizar, migrar e gerenciar recursos (Paletta, 2008). A Figura 2 apresenta as principais dimensões do Gerenciamento do Ciclo de Vida de Tecnologia da Informação.



Figura 2. Gerenciamento do Ciclo de Vida de TI  
Fonte: Altiris Inc

A Figura 3 ilustra a modularidade para o desenvolvimento da infra-estrutura de TI necessária para a implantação de uma solução de gestão de ativos:



Figura 3. Modularidade da Infraestrutura Computacional - Fonte: Altiris Inc

O gerenciamento de clientes e equipamentos móveis permite que os administradores implementem, gerenciem e solucionem problemas de sistemas a partir de qualquer lugar. Os benefícios incluem:

- Gerenciamento consolidado de desktops, notebooks e handhelds
- Implementação do SO e migração de personalidade do PC com intervenção zero
- Inventário abrangente de HW e SW com geração de relatórios pela Internet
- Avaliação das vulnerabilidades do sistema com distribuição de SW e gerenciamento e patches de segurança em tempo real
- Gerenciamento de estados através dos recursos de auto correção e reversão de aplicativos

O gerenciamento de servidores oferece as funções de implementação, gerenciamento e monitoramento a partir de um console centralizado, reduzindo os custos totais de infraestrutura. Os benefícios incluem:

- Melhorar a confiabilidade e a estabilidade de servidores, minimizando paralisações da biblioteca digital e melhorando a satisfação do usuário
- Automatizar o gerenciamento das operações de TI para responder rapidamente à mudanças nas necessidades da biblioteca digital
- Monitorar o desempenho, restaurar a operação e minimizar os patches de segurança, de modo a assegurar a continuidade da operação e a sua disponibilidade

As soluções de gerenciamento do ciclo de vida de TI estão organizadas em três níveis ao longo de um modelo de maturidade conforme as necessidades de recursos computacionais.



Figura 4. Infraestrutura Computacional - Fonte: Altiris Inc

As aplicações estão evoluindo em direção a soluções abrangentes de gerenciamento de TI que utilizem um único repositório e uma única

interface, reduzindo radicalmente os custos e a complexidade do gerenciamento de seus recursos computacionais, incluindo desktops, notebooks, handhelds e dispositivos de redes. É fundamental automatizar, simplificar e integrar suas funções de gerenciamento de TI a partir de um único console baseado na Web.

## 5. Objetivos - Metodologia - Resultados - Conclusões - Recomendações

A seguir apresentamos os aspectos relacionados aos objetivos, à metodologia utilizada, a estratégia de coleta de dados, tratamento e análise dos dados, introduzindo os tópicos relativos aos objetivos, métodos e instrumentos de pesquisa. Esta prospecção tem caráter experimental e aplicado, em razão principalmente da atualidade e emergência do tema.

### 5.1. Objetivos

Há atualmente a necessidade de novas propostas para abordagens gerenciais e estratégicas associadas à utilização da Tecnologia da Informação na apropriação e geração de conhecimento no contexto da WEB. Neste Projeto de Pesquisa, nossa proposta é identificar e discutir as tendências nesse campo, em especial interesse aquelas que influenciam o contexto da chamada Biblioteca Digital, de modo que, cada vez mais, a tecnologia possa proporcionar a criação e obtenção de valor ao usuário da informação. Espera-se que o desenvolvimento dessa área leve a um maior entendimento dos recursos da tecnologia colaborativa utilizados em ambientes informacionais digitais. Em qualquer caso, a inserção dessas tecnologias consiste em inovação que deve estar vinculada à tradição e à missão das Bibliotecas e dos repositórios de informação.

Com o presente estudo objetiva-se verificar a apropriação que as Bibliotecas e os Repositórios Digitais têm feito das tecnologias colaborativas em ambientes de informação digital e sistemas abertos bem como a gestão da infraestrutura de seus recursos computacionais na configuração tecnológica dos Serviços de Informação orientados ao Usuário da Informação na Web.

### 5.2. Metodologia

O pesquisador utiliza a metodologia científica para a obtenção de novos conhecimentos no campo da realidade social. A elaboração e o desenvolvimento da pesquisa necessitam, para que os resultados sejam satisfatórios, estarem baseadas em planejamento cuidadoso, refle-

xões conceituais sólidas e alicerçados em conhecimentos existentes.

Para Gil (1999, p.42), a pesquisa deve ser pragmática, um “processo formal e sistemático de desenvolvimento do método científico. O objetivo fundamental da pesquisa é descobrir respostas para problemas mediante o emprego de procedimentos científicos”. Tendo em vista o objeto deste trabalho conforme descrito inicialmente, é importante, neste instante, estabelecer a relação entre a metodologia científica e a utilização desta para estudar os fenômenos sociais.

Entende-se por Metodologia Científica um conjunto de etapas ordenadamente dispostas que o pesquisador deve executar na investigação de um fenômeno. Inclui a escolha do tema, o planejamento da investigação, o desenvolvimento metodológico, a coleta e a tabulação de dados, a análise dos resultados, a elaboração das conclusões e a divulgação de resultados.

Método Científico é o conjunto de processos ou operações mentais que se devem empregar na investigação. É a linha de raciocínio adotada no processo de pesquisa. Os métodos que fornecem as bases lógicas à investigação são: dedutivo, indutivo, hipotético-dedutivo, dialético e fenomenológico (Lakatos; Marconi 1992).

Partindo dessas concepções, é possível então apresentar a metodologia utilizada neste estudo. O Projeto de Pesquisa esta dividido em três partes:

- Estudo dos Modelos de Busca e Comportamento Informacional do Usuário da Informação na Web
- Infraestrutura Tecnológica e Recursos Computacionais das Bibliotecas Digitais
- Ferramentas de Busca, Acesso, Apropriação, e Uso da Informação na WEB

A pesquisa será estabelecida a partir de referenciais teóricos e do estudo de melhores práticas na gestão de recursos computacionais em Unidades de Informação e Bibliotecas Digitais. Será desenvolvido um conjunto de ferramentas de estudo com foco no Estudo de Usuários da Informação em ambiente Web, a partir do qual será possível, com base no paradigma teórico oriundo da revisão da literatura, estruturar o método de coleta de dados que verificarão na prática as proposições referenciais.

O estudo tem por escopo confrontar a utilização os modelos de estudo de usuários da informação conhecidos com os resultados de pesquisas científicas para buscar conhecer e entender as

formas de acesso, apropriação e uso da informação em ambiente Web.

A população ou universo da pesquisa incluiu as Bibliotecas da Universidade de São Paulo que possuem as mesmas características definidas para este estudo: Bibliotecas Digitais. A amostra, a parte da população ou do universo selecionada de acordo com o plano preestabelecido, estará limitada ao âmbito do Sistema Integrado de Bibliotecas da Universidade de São Paulo.

O problema da pesquisa é, para Kerlinger (1980), “uma questão que mostra uma situação necessitada de discussão, investigação, decisão ou solução”. Problema é uma questão a que a pesquisa pretende responder. Todo o processo de pesquisa irá girar em torno da busca da solução.

A pergunta que esta pesquisa se propõe a estudar e responder é: quais os modelos de acesso, apropriação e uso da informação utilizada pelos usuários da informação baseados na Web? Os resultados desta pesquisa poderão servir de base para a orientação de novos Serviços ao Usuário propostos pela Biblioteca e/ou Centro de Informação no cenário fortemente dependente da Web.

Os critérios de escolha e a estratégia de coleta e análise dos dados convergem para o instrumento de pesquisa, que esta em desenvolvimento a partir do equacionamento das dimensões e fronteiras investigadas na pesquisa de campo e que serão, para ser devidamente perscrutadas, particionadas em três enfoques primordiais, subdivididos de acordo com a conveniência e importância do assunto.



Figura 5. Dimensões Consideradas no Projeto de Pesquisa

### 5.3. Resultados

Como resultado deste estudo espera-se uma contribuição para o entendimento dos recursos da tecnologia colaborativa utilizados em ambientes informacionais digitais. Com base nos recursos identificados e coletados em Bibliotecas e repositórios digitais deseja-se observar como são aplicados os recursos da tecnologia colaborativa no contexto da Web Semântica ou Web de Dados. A inserção dessas tecnologias apresenta-se como inovação que devem estar

vinculadas à tradição e a missão das Bibliotecas e dos repositórios de informação. Avaliar a flexibilidade das estruturas computacionais, sua atratividade e dinâmica na qual o usuário torna-se agente na construção de seu ambiente, demandando recursos de customização e personalização na criação de Serviços de Informação inovadores e que permitam que a Biblioteca continue a ocupar papel relevante na produção de novos conhecimentos.

Neste trabalho apresentamos os resultados da Fase II do Projeto de Pesquisa: Infraestrutura Tecnológica e Recursos Computacionais das Bibliotecas Digitais.

A gestão consolidada do ambiente de trabalho exige que as Bibliotecas Digitais adotem uma abordagem holística orientada a pessoas, processos, resultados e tecnologia em todo o ambiente de computação. Ela também exige que as Bibliotecas trabalhem com fornecedores de TI que possam analisar suas necessidades operacionais, que assessorem a implementação e o gerenciamento e suporte contínuos das soluções implementadas. Os desafios básicos que as Bibliotecas Digitais enfrentam dentro dos ambientes computacionais incluem:

**Redução de Custos** - Os ambientes de atendimento ao usuário estão mudando rapidamente para locais de pesquisa móveis, globais e virtuais, diversificados culturalmente, que são onerosos para manter e suportar. Através da consolidação de hardware, dos aplicativos e processos de suporte dentro de seus ambientes de trabalho, as Bibliotecas Digitais podem gerenciar e reduzir os custos de TI, ao mesmo tempo em que priorizam a satisfação do usuário e o retorno no investimento.

**Aumento da Produtividade dos Profissionais da Informação** - Para realizar este objetivo, as Bibliotecas Digitais estão buscando maneiras de aumentar a colaboração e o trabalho de equipe, através da criação de um ambiente de trabalho sem fronteiras, confiável e seguro, proporcionando a conexão e acesso à informação a qualquer hora de qualquer lugar.

**Redução da complexidade da TI** - A falta de padronização dentro do ambiente computacional pode aumentar o tempo e os custos necessários para gerenciar e suportar este ambiente. Ao mesmo tempo, à medida que os ambientes de computação se tornam mais complexos, o nível de conhecimento e especialização necessários para oferecer suporte a eles aumenta.

As ferramentas de gestão do ciclo de vida de TI permitem a padronização da plataforma de hardware; redução de dispositivos redundantes;

simplifica e automatiza os processos computacionais; além de gerenciar as funções de suporte e construir a flexibilidade e estabilidade que permite a criação das condições dinâmicas da gestão da informação digital (Laudon, 2007).

A maior despesa da propriedade de recursos de TI não reside na compra inicial do hardware e software, mas sim na complexidade de implementar e de manter estes dispositivos. A fim de reduzir esses custos, as Bibliotecas Digitais devem investir em software de gerenciamento de sistemas para melhorar a confiabilidade e a disponibilidade do hardware e do software através de todas as fases do ciclo de vida de um recurso.

A crescente complexidade dos dispositivos tecnológicos tem incentivados os gestores de TI a buscarem meios de melhorar a eficiência na operação visando reduzir custos, estar de acordo com os aspectos reguladores de Governança da Tecnologia da Informação, e responder às constantes exigências das Bibliotecas Digitais por uma melhor resposta as demandas geradas pelo usuário. Esses fatores têm sido um impulsor para que os gestores de TI procurem formas eficientes de ter o controle de tudo o que existe em sua rede de dados. Apresentamos a seguir os imperativos para que uma biblioteca digital alcance a excelência operacional e maximize o seu desempenho (Turban, 2005).

- Desenvolver relacionamento entre serviço oferecido e demanda por informação
- Entregar e implantar novos sistemas baseados nas demandas do usuário
- Construir e administrar a infraestrutura computacional e garantir acessibilidade
- Capacitar à organização no uso dos recursos computacionais
- Administrar parcerias e colaboração em redes de dados abertos
- Desenvolver soluções Web de alto desempenho e disponibilidade
- Alinhamento estratégico: infraestrutura computacional e valor percebido pelo usuário

Ao avaliar uma ferramenta de Gestão do Ciclo de Vida da Tecnologia da Informação apresentamos na Figura 6 as recomendações das Melhores Práticas de Gestão de TI com maior impacto organizacional, administrativo e financeiro para a Unidade de Informação.

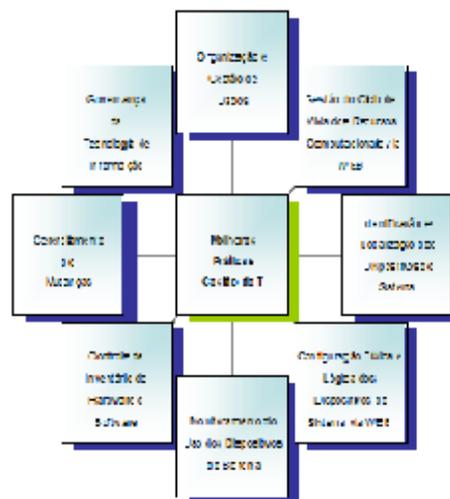


Figura 6. Melhores Práticas Gestão de Recursos de TI via Web

As soluções de gestão do ciclo de vida dos recursos computacionais devem permitir um tratamento adequado às complexidades associadas ao gerenciamento dos recursos de TI. Os sistemas devem ser modulares, permitindo a definição de uma estrutura tecnológica compatível com as necessidades computacionais da Unidade de Informação (Weill, 2006).

#### 5.4 Conclusões

Para muitas Bibliotecas Digitais, a crescente disponibilização das tecnologias tem demonstrado uma ambigüidade em seu gerenciamento. No aspecto positivo, estas novas tecnologias têm ajudado a aumentar a produtividade dos profissionais da informação, aprimorar o processo de tomada de decisão e acentuar a satisfação do usuário da informação. Porém, a gestão e o suporte destes ambientes heterogêneos e complexos - repletos de diferentes PCs, desktops e portáteis, dispositivos móveis e sem fios, impressoras, redes e aplicativos - comprovadamente têm se revelado difíceis e dispendiosos para os departamentos de Tecnologia da Informação. Neste contexto torna-se relevante avaliar os principais desafios que as Bibliotecas Digitais terão que enfrentar com relação ao gerenciamento do ciclo de vida de suas tecnologias, consolidação e simplificação de seus processos dentro de seus ambientes computacionais, com objetivo de aumentar a produtividade e construir ambientes ágeis que permitam às Bibliotecas responder as demandas da gestão da informação digital.

## 5.5 Recomendações

Gestores de Bibliotecas Digitais estão cada vez mais envolvidos em maior ou menor grau, no desenvolvimento, controle e monitoração dos ativos tecnológicos de sua organização. As constantes pressões por manter os investimentos de TI eficientes fazem que seja prioritário administrar estes ativos de duas formas: como função do departamento de TI bem como parte de um processo integral da biblioteca digital.

## Referências

- Belling, A., et al. (2010). Exploring Library 3.0 and beyond. Disponível em: [http://www.libraries.vic.gov.au/downloads/20102011\\_Shared\\_Leadership\\_Program\\_Presentation\\_Day\\_exploring\\_library\\_3.pdf](http://www.libraries.vic.gov.au/downloads/20102011_Shared_Leadership_Program_Presentation_Day_exploring_library_3.pdf) Acesso em: 25 jun. 2015.
- Bruce, C. (2003).. Seven Faces of Information Literacy. Disponível em: <http://www.bestlibrary.org/digital/files/bruce.pdf> Acesso em: 26 mar. 2013.
- Canclini, N. G. (2009). Diferentes, desiguais e desconectados. Rio de Janeiro: Editora. UFRJ.
- Choo, C. W. (2003). A organização do conhecimento: como as organizações usam a informação para criar significado, construir conhecimento e tomar decisões. São Paulo: Editora Senac São Paulo.
- Dimensões e características da WEB brasileira: um estudo do .gov.br.2010. (2010). Disponível em: <https://www.ufmg.br/proex/cpinfo/educacao/docs/01b.pdf>
- Dudziak, E. A. A Information Literacy e o papel educacional das bibliotecas. 2001. Dissertação (Mestrado) - Escola de Comunicações e Artes, Universidade de São Paulo, São Paulo.
- Dudziak, E. A. (2003).Information literacy: princípios, filosofia e prática. Ciência da Informação, Brasília, v.32, n.1, p. 23-35, jan./abr.
- Ferreira, S. M. S. P. ( 1995). Novos paradigmas e novos usuários da informação. Ciência da Informação, Brasília, v.25, n.2, p. 217-223, maio/ago.
- Fujita, M., S. L. (2005). A Biblioteca digital no contexto da gestão de bibliotecas universitárias: análise de aspectos conceituais e evolutivos para a organização da informação. In: Encontro Nacional de Ciência da Informação. Salvador. Anais... Salvador: Editora da Universidade Federal da Bahia - EDUFBA.
- IFLA. (1997). Intelligent Library Buildings. Retrieved July 15, 2015, from <http://www.ifla.org/publications/intelligent-library-buildings>
- Kerlinger, F.N. (1980). Metodologia de pesquisa em ciências sociais. São Paulo : EPU.
- Lakatos, E.M., & Marconi, M.A.(1992). Metodologia do trabalho científico: procedimentos básicos , pesquisa bibliográfica, projeto e relatório, publicações e trabalhos científicos. 4 ed.São Paulo: Atlas.
- Laudon, K., Laudon, J. P. (2007). Management Information Systems: managingthe digital firm. New Jersey: Person.
- Matías Clavero, G. (2005). La Estrategia de Lisboa sobre la sociedad del conocimiento: la nueva economía. Información Comercial Española, ICE: Revista de economía.
- Melo, A. V. C.,Araujo, E. A.( 2007). Competência Informacional e gestão do conhecimento: uma relação necessária no contexto da sociedade da informação. Perspectivas em Ciência da Informação, Belo Horizonte, v.12, n.2, p.185-201, maio/ago.
- Observatorio para a Cibersociedad. Ambientes informacionais: As bibliotecas digitais e os repositórios institucionais no contexto da web 2.0. (2009). Retrieved July 13, 2015, from <http://www.cibersociedad.net/congres2009/po/coms/ambientes-informacionais-as-bibliotecas-digitais-e-os-repositorios-institucionais-no-contexto-da-web-20/1056/>
- Paletta, Francisco Carlos, & Dias Vieira Junior, Nilson. (2008). Information technology and communication and best practices in it lifecycle management. *Journal of technology management & innovation*, 3(4), 80-94. Recuperado em 12 de julio de 2015, de [http://www.scielo.cl/scielo.php?script=sci\\_arttext&pid=S0718-27242008000200007&lng=es&tlng=en.10.4067/S0718-27242008000200007](http://www.scielo.cl/scielo.php?script=sci_arttext&pid=S0718-27242008000200007&lng=es&tlng=en.10.4067/S0718-27242008000200007).
- Paletta, F., & Mucheroni, M. (2014). O Desenvolvimento da WEB 3.0: Linked Data e DBPEDIA. PRISMA.COM, (25). Retrieved July 15, 13, from file:///C:/Users/Francisco/Downloads/3047-10898-1-PB (5).pdf
- Passarelli, B. (2008). Do Mundaneum à WEB Semântica: discussão sobre a revolução nos conceitos de autor e autoridade das fontes de informação. DataGramaZero - Revista de Ciência da Informação - v.9 n.5, P.1-13.
- Passarelli, B., JUNQUEIRA, A.H. (2012). Gerações Interativas Brasil. Crianças e Adolescentes Diante das Telas. São Paulo: Escola do Futuro/USP.
- Sousa, R. S. C., Nascimento. B. S. (2010). Competências Informacionais: uma análise focada no currículo e na produção docente dos cursos de Biblioteconomia e gestão da informação. Revista ACB: Biblioteconomia em Santa Catarina, Florianópolis, v.15, n.2, p.130-150, jul./dez.
- Turban, E., et al. (2006). Administração de tecnologia da informação: teoria e prática. Rio deJaneiro: Elsevier, 2005.
- Weill, Peter; ROSS, Jeanne W. Governança de TI, Tecnologia da Informação. São Paulo:M. Books do Brasil Editora.

## Anexo 1

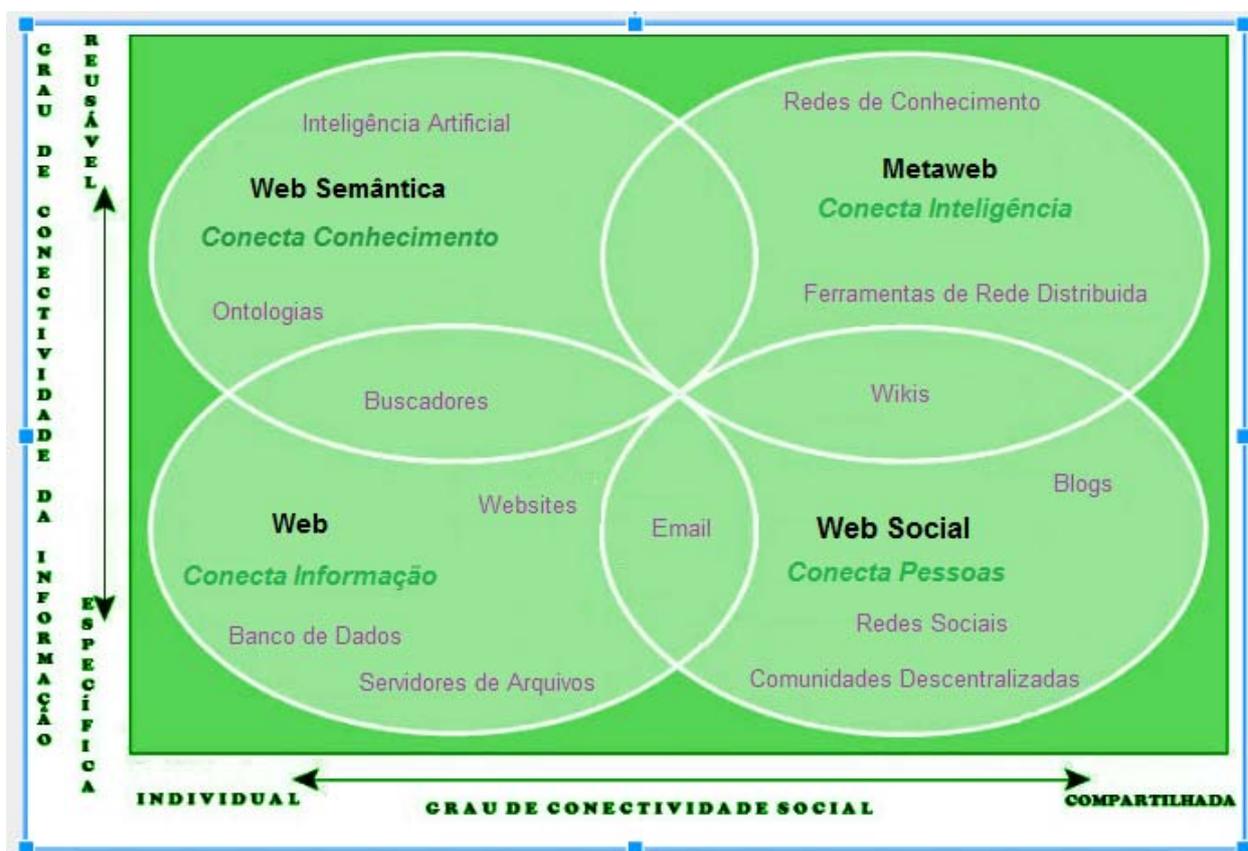


Figura 1. Evolução da WEB. Fonte: <https://swedinbalchik.wordpress.com/category/studies-work/algc/> – Adaptação: Ligia Capobianco

## Anexo 2



Figura 2. Gerenciamento do Ciclo de Vida de TI - Fonte: Altiris Inc

## Anexo 3



Figura 3. Modularidade da Infraestrutura Computacional - Fonte: Altiris Inc

## Anexo 4



Figura 4. Infraestrutura Computacional - Fonte: Altiris Inc

## Anexo 5



Figura 5. Dimensões Consideradas no Projeto de Pesquisa

## Anexo 6



Figura 6. Melhores Práticas Gestão de Recursos de TI via Web

---

# Organización del conocimiento en entornos wiki: una experiencia de organización de información sobre lecturas académicas

*Knowledge organization in wiki environments: an experience of organizing information about academic readings*

---

Jesús Tramullas (1), Ana I. Sánchez Casabón (2), Piedad Garrido Picazo (3)

(1) Depto., Ciencias de la Documentación, Univ. de Zaragoza, C/ Pedro Cerbuna 12, 50009 Zaragoza, tramullas@unizar.es

(2) asanchez@unizar.es

(3) Depto. Informática e Ingeniería de Sistemas, Univ. de Zaragoza, Esc. Univ. Politécnica, Ciudad Escolar s/n, 44003 Teruel, piedad@unizar.es

## Resumen

Este trabajo revisa el comportamiento informacional de una comunidad de estudiantes universitarios durante el desarrollo de una actividad de aprendizaje con un wiki. Mediante el estudio de caso, analiza los datos disponibles en el wiki, e identifica pautas de creación y organización de contenidos. El estudio del wiki se realiza también dentro del marco de gestión de información propuesto por Rowley. Las conclusiones permiten afirmar que los estudiantes aplican el principio de economía de esfuerzo en su comportamiento informacional, orientados por los requerimientos de evaluación de la actividad, y que la propuesta de Rowley no es adecuada para analizar y evaluar los procesos educativos mediados tecnológicamente.

**Palabras clave:** wiki, educación, comportamiento informacional.

## Abstract

This paper reviews the informational behavior of a community of university students during the development of a learning activity with a wiki. Through a case study, analyzes the data available on the wiki, and identifies patterns of creating and organizing content. The wiki study is also done within the information management framework proposed by Rowley. The findings support the conclusion that students apply the principle of economy of effort in their informational behavior, guided by the assessment requirements of that activity, and Rowley's proposal is not suitable for analyzing and evaluating educational processes technologically mediated.

**Keywords:** wiki, education, information behavior.

## 1. Planteamiento

Las herramientas wiki se popularizaron en la década de 2000 dentro del panorama general del web 2.0. Su facilidad de uso y la posibilidad de adecuarlas a numerosos contextos y requerimientos resultaron claves en la aparición de gran número de proyectos e iniciativas que usaban esta aproximación como soporte para sus productos y servicios de información digital. Sin embargo, la evolución del contexto digital y de las herramientas disponibles, así como de las pautas informacionales de los usuarios, han llevado a reflexionar sobre el planteamiento y la utilización de los wikis en entornos de actividad más específicos.

Tapscott y Williams (2006) definieron los procesos colaborativos en la creación de conocimiento que se desarrollaban en el entorno wiki. Koblas (2006) y Tramullas (2008) establecieron

los usos básicos de los wikis en la práctica de la gestión de información en las organizaciones. La utilización de los wikis, y por extensión, de *Wikipedia*, ha sido una constante en la literatura sobre investigación educativa durante la última década (Schoereder, 2009; Konieczny, 2012). También se ha señalado su uso como plataforma para desarrollos de buenas prácticas en entornos organizativos y educativos (Tramullas y Sánchez Casabón, 2014).

En este contexto, en la docencia de varias asignaturas que se imparten en el grado en Información y Documentación de la Universidad de Zaragoza se han utilizado plataformas wiki como instrumentos para desarrollar procesos de enseñanza y aprendizaje (Tramullas, Garrido, Esteban y Sánchez Casabón, 2008). Estos procesos corresponden a tareas de creación y organización de conocimiento relacionado con literatura científica, por una parte, y a la adquisi-

ción de las habilidades técnicas de procesamiento de información con herramientas wiki, por otra.

Para alcanzar un adecuado ajuste entre los objetivos de aprendizaje y las herramientas empleadas, es necesario comprender los procesos que sigue el estudiante, el contexto en el cual se llevan a cabo, y los factores que intervienen. Este trabajo plantea un estudio de estas características, tomando como referencia el marco de trabajo para la gestión de información propuesto por Rowley (1998).

## 2. Objetivos y metodología

El proceso educativo puede tratarse como un proceso cognitivo en el cual los procedimientos de gestión de información desempeñan un papel nuclear. Este procedimiento se encuentra enmarcado en un contexto más amplio, en el que intervienen factores sociales, económicos y personales. Por lo tanto, al abordar su estudio resulta necesario adoptar un enfoque que contemple esta pluralidad.

La propuesta de Rowley (1998) para estudiar la gestión de información ha sido ampliamente referenciada como un marco válido para analizar la gestión de información en contextos complejos. A esto hay que añadir que en este tipo de procesos educativos mediados tecnológicamente, el comportamiento informacional de los usuarios desempeña una papel fundamental, en cuanto es parte y base de los procesos cognitivos (Tramullas y Garrido, 2011a).

De acuerdo con lo anterior, este trabajo pretende estudiar el comportamiento informacional de los estudiantes en un entorno wiki, atendiendo a los siguientes objetivos específicos: 1) identificar y analizar las pautas de producción y organización de la información en un entorno wiki por parte de una comunidad de usuarios; 2) determinar las posibles causas y motivaciones de la organización de la información adoptada por los usuarios; y 3) establecer si el estudio del comportamiento informacional en entornos wiki puede abordarse desde el marco de gestión de información propuesto por Rowley en 1998.

Para ello se ha adoptado una metodología de descripción y análisis de caso, que revisa el objeto de estudio, las herramientas disponibles, el desarrollo del proceso y los resultados obtenidos, y contextualiza la actividad en el entorno educativo en el cual se lleva a cabo.

## 3. Estudio de caso

### 3.1. La actividad de enseñanza y aprendizaje en su contexto

Esta actividad de aprendizaje se ha llevado a cabo anualmente desde el curso académico 2008-2009 hasta el año 2013. Dada la reforma educativa de la universidad española en este periodo, han sido varias las asignaturas en las que se ha realizado la actividad objeto de estudio (*Sistemas electrónicos de tratamiento documental, Aplicaciones para recursos de información digital y Gestión de contenidos*), todas ellas correspondientes al tercer o cuarto año de la titulación. Los estudiantes elaboran un proyecto personal de lectura, revisión y valoración de un trabajo de literatura científica sobre aspectos de las tecnologías de la información en información y documentación. A tal fin se les provee de una guía de la actividad, en la que se recogen los objetivos y método básico para la ejecución de la tarea, un listado de revistas científicas de referencia, y los criterios de evaluación.

La actividad se lleva a cabo sobre una herramienta wiki, instalada y configurada especialmente a tal fin (Anexo 1, figuras 1 y 2). Los estudiantes deben crear un redactar una página en el wiki para la lectura que han revisado (Anexo 1, figura 3), así como incorporar a la misma referencias bibliográficas complementarias, y usar categorías propias del wiki. La herramienta implementada es *MediaWiki* (1), desarrollada por *Wikimedia Foundation*, con licencia de software libre (Tramullas, 2006). Los contenidos aportados por los estudiantes se publican con licencia *GNU Free Documentation License* ver. 1.3 (2).

Los objetivos educativos de la actividad son dos, y se corresponden con las dos facetas principales sobre las que se quiere desarrollar las competencias: 1) crear y estructurar contenidos informativos hipertextuales en entorno digital; y 2) aprender a usar, de manera autónoma, una herramienta wiki.

### 3.2. Requerimientos para el estudiante

La guía de la actividad señala al estudiante los criterios de evaluación, haciendo referencia al tamaño mínimo del contenido, estructura del mismo, uso de bibliografía complementaria, organización dentro del wiki, y categorización. El estudiante dispone de un mes para llevar a cabo la tarea.

### 3.3. Análisis de resultados

Los resultados obtenidos del análisis de los contenidos aportados por los estudiantes durante los años de ejecución de la actividad (véase Anexo 1 tabla 1) permiten hacer las siguientes afirmaciones: 1) los estudiantes aportan un trabajo sobre una única lectura; 2) usan una media de tres categorías de clasificación; 3) aunque enlazan su trabajo desde el índice general de referencia, no revisan el resto de trabajos de los otros estudiantes para localizar y poder enlazar el suyo con otros relacionados temáticamente; y 4) la media de referencias bibliográficas que usan para documentar su trabajo personal en el wiki ha fluctuado, durante los diferentes cursos, entre el 2,9 y el 3,8 de media.

El proceso de contrastar estos resultados con los requerimientos de evaluación establecidos en la guía de desarrollo de la actividad, de la que disponen previamente los estudiantes, permite afirmar que existe un paralelismo entre los requisitos previamente establecidos y la actividad que llevan a cabo. Como los criterios y requerimientos de evaluación establecen el nivel mínimo a partir del cual se pueda considerar superada satisfactoriamente la actividad, los estudiantes tienden a ajustar su comportamiento informacional para cumplir con esos mínimos que les aseguran la consecución del objetivo académico.

Se detectan, además, tres cuestiones de importancia relacionadas con otras competencias de gestión de información que los estudiantes han debido adquirir previamente, si se revisa la estructura curricular de la titulación. En primer lugar, la falta de consistencia y la existencia de errores en la redacción de las referencias bibliográficas. En segundo lugar, el uso de las categorías de clasificación es completamente personal y contextual, sin uso de esquemas de clasificación ya construidos, ni de reutilización de categorías usadas por otros estudiantes (Anexo 1, figura 4). Por último, la falta de exploración del propio espacio wiki de la actividad: sólo en tres casos los estudiantes han enlazado sus referencias complementarias con resúmenes de los originales, elaborados por otros estudiantes, y que ya se encontraban disponibles en el mismo espacio.

Estos tres elementos indican que los estudiantes parecen descontextualizar y aislar por completo la actividad que están desarrollando, del resto de actividades que llevan o han llevado a cabo durante el estudio de la titulación. También demuestra que su concepción de la actividad es a corto plazo, y no es considerada como una actividad formativa, sino como un ejercicio obli-

gatorio que debe ejecutarse con el menor coste posible de recursos.

En lo que concierne al uso técnico del wiki, la media de ediciones por página oscila entre 4,2 y 5,6 de media lo que indica que los estudiantes redactan el contenido externamente, y luego proceden a trasladarlo a la plataforma (Anexo 1, figura 5). Esta actuación acelera el proceso de finalización de la actividad, pero limita el aprendizaje de las funcionalidades técnicas de la herramienta. De la misma forma, el comportamiento indicado en párrafos anteriores hace que los estudiantes no desarrollen las estructuras hipertextuales típicas de los wiki para organizar la información.

## 4. El wiki de lecturas en el marco de trabajo de gestión de información de Rowley

La aproximación positivista de Rowley, ya citada, propone que la gestión de información es una disciplina basada en la práctica, y que debe abordarse desde diferentes perspectivas, cuyo elemento común sería el procesamiento de la información.

El marco de trabajo se estructura en cuatro niveles, del más externo al más interno, correspondientes a: 1) entorno de información, es el ambiente que rodea al contexto de la información, con aspectos tecnológico, político, social, regulatorio... 2) contexto de información, en el que se encuentran los sistemas de información, la organización, educación... 3) sistemas de información, diseñados para introducir, almacenar y recuperar información; y 4) recuperación de información, la interfaz individual de un sistema, para satisfacer requerimientos de información de los usuarios (Rowley, 1988: 363-364).

Los cuatro niveles se agrupan, a su vez, en dos subdisciplinas: 1) la microinformática, que corresponde al individuo, su uso de la información, y la utilización de los sistemas diseñados para ello; y 2) la macroinformática, que se refiere a la relación entre la sociedad y la información (ibídem, 365-366). A continuación se caracteriza el wiki de lecturas en los cuatro niveles propuestos.

### 4.1. Entorno de información

El wiki de lecturas ha sido implantado como una herramienta de aprendizaje para unas competencias específicas, en el marco del Grado en Información y Documentación de la Universidad de Zaragoza. La actividad que se lleva a cabo con el wiki se encuentra regulada por la memo-

ria del citado Grado. Además, la actividad requiere de un conocimiento y práctica adquiridos en otras materias del Grado.

#### 4.2. Contexto de información

El wiki es un recurso de información que se usa dentro de un proceso formativo más amplio. Es el resultado de seleccionar la herramienta más adecuada para la formación de profesionales, pero no se han tenido en cuenta específicamente las necesidades de información de éstos.

#### 4.3. Sistemas de información

El sistema de información se ha resuelto tecnológicamente mediante la implementación de una herramienta wiki. Se han utilizado las interfaces de usuario estándares, sin diseños a medida. No ha habido un diseño específico de sistemas de información, ya que los estudiantes debían conformarlo por sí mismos.

#### 4.4. Recuperación de información

Los estudiantes, como usuarios y diseñadores, llevan a cabo su propia propuesta de organización y clasificación de la información. La misma es limitada, por el principio de economía de esfuerzo. Es el nivel en el cual los estudiantes desempeñan más claramente tareas de gestión de información.

Middleton (2007) ha aplicado la propuesta de Rowley para analizar varios servicios de información científica y técnica de Australia, usando una metodología de descripción de caso. Concluye que el marco es de aplicación, pero que necesita una mayor elaboración teórica y práctica, proponiendo unir los niveles de entorno y contexto, y renombrar el nivel de recuperación como de procesamiento.

El resultado de la aplicación de la propuesta de Rowley al wiki de lecturas permite corroborar las afirmaciones de Middleton. Realmente, se trata de un marco de trabajo genérico, que no ofrece técnicas específicas de análisis, y que parece especialmente orientado a estudiar las actividades de gestores de información, en un contexto organizacional, antes que las de usuarios. Los niveles de entorno y contexto resultan realmente difusos, y no se analiza la actividad de los usuarios con la suficiente profundidad. Tampoco ofrece una propuesta de evaluación adecuada a las características que identifican a los entornos actuales para aprendizaje en línea.

## 5. Conclusiones

El análisis y revisión de los resultados obtenidos durante el estudio de caso permite afirmar que se han cubierto los objetivos establecidos para este trabajo. Se han identificado las pautas de creación y organización de información hipertextual por parte de una comunidad de estudiantes, en el contexto de una actividad de aprendizaje.

El comportamiento informacional subyacente es coincidente con el principio de economía de recursos que los estudiantes aplican en actividades mediadas tecnológicamente (Tramullas y Garrido, 2011b). Al disponer previamente de los criterios que establecen los mínimos exigibles para el producto de información que se les demanda, su comportamiento tiende a excluir el esfuerzo que supone el aprendizaje personal de la herramienta wiki. En paralelo con lo anterior, la organización de la información se limita a seguir esquemas básicos de contenido y clasificación.

El trabajo también pretendía establecer si la propuesta de Rowley era válida para su aplicación en entornos de aprendizaje de este tipo. Esta hipótesis debe rechazarse, ya que no ofrece una propuesta de aplicación lo suficientemente específica y detallada. Una aproximación basada en el estudio del comportamiento informacional de los estudiantes sigue siendo la más adecuada para analizar las pautas de creación, organización y consumo de información de los estudiantes.

## Notas

- (1) <http://www.mediawiki.org>
- (2) <http://www.gnu.org/licenses/fdl-1.3.html>

## Referencias

- Head, A. J., & Eisenberg, M. B. (2010). How today's college students use Wikipedia for course-related research. *First Monday*, 15(3) doi: 10.5210/fm.v15i3.2830
- Koblas, J. (2006). *Wikis: Tools for Information Work and Collaboration*. London: Chandos Pub.
- Konieczny, P. (2012). Wikis and wikipedia as a teaching tool: Five years later. *First Monday*, 17. doi:10.5210/fm.v0i0.3583
- Rowley, J. (1998). Towards a framework for information management. *International Journal of Information Management*, 18(5), 359–369.
- Schroeder, B. (2009). Within the wiki: Best practices for educators. *AACE Journal*, 17(3), 181-197.
- Tapscott, D., Williams, A.D. (2006): *Wikinomics: How mass collaboration changes anything*. Portfolio.
- Tramullas, J. (2006): Wikis: MediaWiki. En: Tramullas J. y Garrido, P. (coords.) *Software libre para servicios de información digital*. Madrid: Pearson Prentice Hall, 71-92.
- Tramullas J., Garrido, P. (2011a). Comportamiento informacional y estudios de usuario. En Cacheda, F., Fernán-

dez-Luna, J.M., Huete, J. (coords.) *Recuperación de Información. Un enfoque práctico y multidisciplinar*. Madrid: Ra-Ma, 457-473.

Tramullas, J., Garrido, P. (2011b). Weblog publishing behaviour of librarianship and information science students: a case study. *Information Research*, 16(1), paper 456. <http://www.informationr.net/ir/16-1/paper456.html>

Tramullas, J., Garrido, P., Esteban, M.A., Sánchez, A., (2008). Azumi: servidor de información para la enseñanza y el aprendizaje de informática documental mediante herramientas de software libre. En: Frías, J.A., Travieso, C. (eds.) *Formación, investigación y mercado laboral en información y documentación en España y Portugal*. Salamanca: Univ. de Salamanca, 359-370.

Tramullas, J., Sánchez Casabón, Ana I., (2014). Pautas para la implementación de wikis para el desarrollo colaborativo de sistemas de buenas prácticas En: *Estudios de Información, Documentación y archivos. Homenaje a la profesora Pilar Gay Molins*. Zaragoza: Prensas Universitarias de Zaragoza, 299-312.

## Anexo 1

| Curso académico | N.º páginas publicadas | Enlaces internos (media) | Categorías (media) | Referencias bibliográficas | Ediciones por página (media) |
|-----------------|------------------------|--------------------------|--------------------|----------------------------|------------------------------|
| 2008-2009       | 20                     | 0,0                      | 3,0                | 2,9                        | 5,1                          |
| 2009-2010       | 22                     | 0,0                      | 3,0                | 3,0                        | 4,8                          |
| 2010-2011       | 18                     | 0,0                      | 3,2                | 3,2                        | 5,6                          |
| 2011-2012       | 12                     | 0,3                      | 3,1                | 3,8                        | 4,5                          |
| 2012-2013       | 8                      | 0,0                      | 3,0                | 3,2                        | 4,2                          |

Tabla 1. Datos numéricos básicos

[página](#) [discusión](#) [ver código fuente](#) [historial](#) Regístrate/Entrar

## Portada

**Bienvenido al wiki de lecturas.**

Este es el wiki de lecturas de los estudiantes de la asignatura **Sistemas Electrónicos de Tratamiento Documental**.

**Contenido** [ocultar]

- 1 Revistas
- 2 Categorías
- 3 Para empezar
- 4 Ejemplo

### Revistas

La revistas que van a utilizarse como fuente son las siguientes:

- Revista Española de Documentación Científica
- El Profesional de la Información
- Journal of the American Society for Information Science and Technology
- Journal of Academic Librarianship
- Journal of Librarianship and Information Science

### Categorías

Se generarán las [categorías](#) automáticamente desde la indización de artículos.

### Para empezar

- [Guía de usuario](#)
- [PMF sobre MediaWiki](#)

### Ejemplo

Una página con un [ejemplo de resumen de un artículo](#) que puede tomarse como punto de partida para trabajar.

Esta página fue modificada por última vez el 17:43, 13 ene 2009. Esta página ha sido visitada 4.738 veces. El contenido está disponible bajo los términos de la [GNU Free Documentation License 1.2](#). [Política de protección de datos](#) [Acerca de Lecturas](#) [Aviso legal](#)

Figura 1. Wiki de lecturas

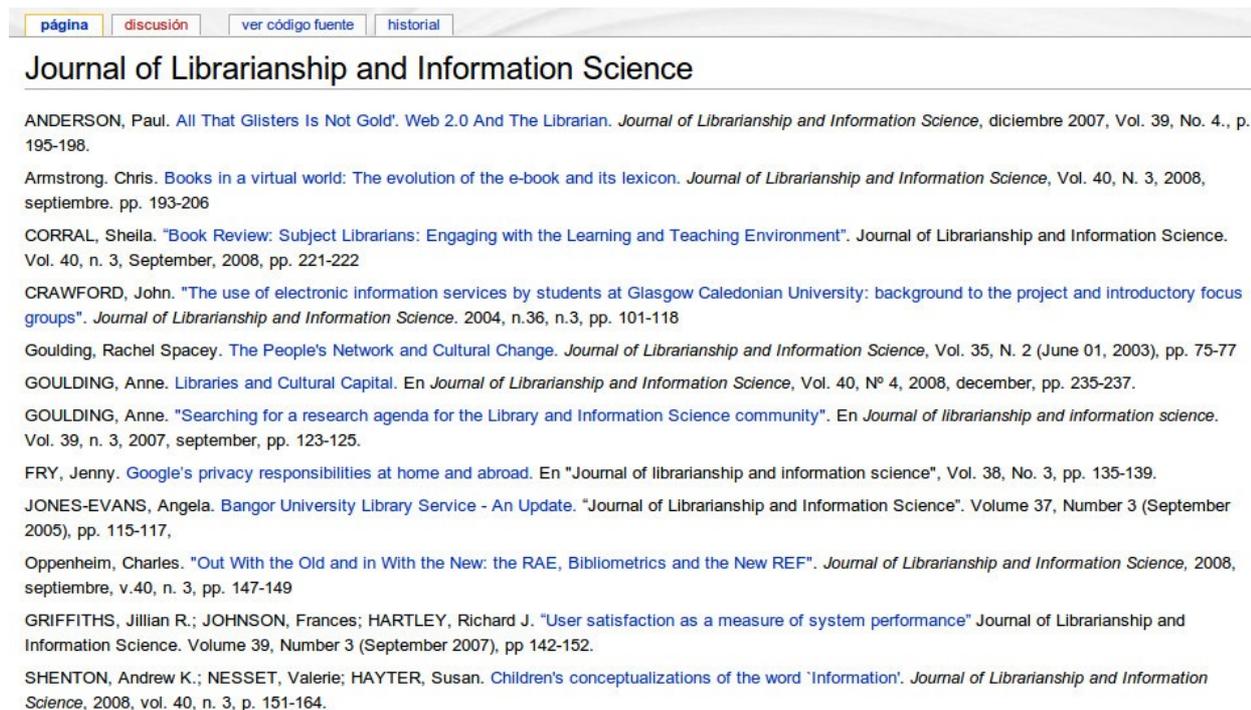


Figura 2. Ejemplo de listado de referencias analizadas

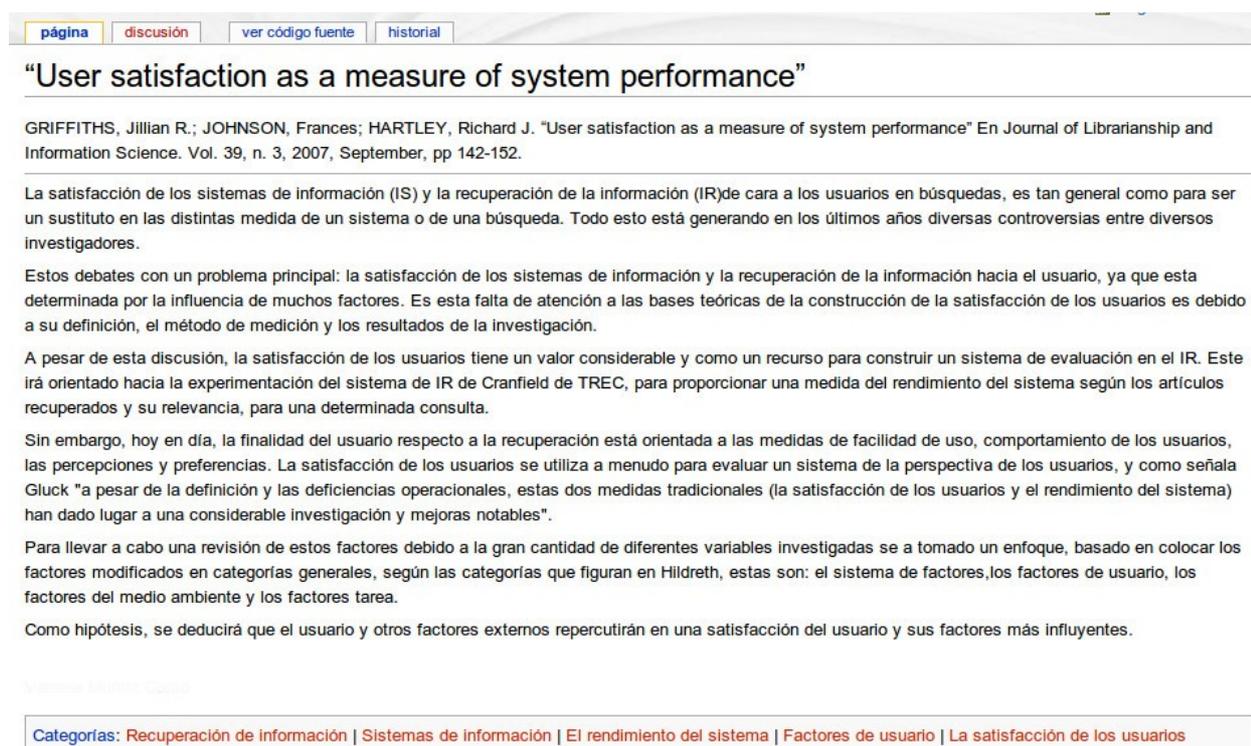


Figura 3. Ejemplo de página redactada

especial

## Categorías

Las siguientes categorías contienen páginas o fiche

Categorías \_\_\_\_\_

Mostrar categorías que empiecen por:

(primeras | últimas) Ver (50 previas) (50 siguientes) |

- 11-M (1 artículo)
- ARTEMIS Digital Library (1 artículo)
- Acceso (1 artículo)
- Acceso abierto (1 artículo)
- Acceso público (1 artículo)
- Adoctrinamiento Ideologico (1 artículo)
- Alfabetización informacional (2 artículos)
- Ambiente de estudio (1 artículo)
- Ambito Científico (1 artículo)
- América Latina (1 artículo)
- Análisis cuantitativo (1 artículo)
- Análisis de Co-enlaces (1 artículo)
- Archivo (1 artículo)
- Archivo digital (1 artículo)
- Archivos (1 artículo)
- Arquitectura-información (1 artículo)
- Association of Research Libraries (1 artículo)
- Audiolibros (1 artículo)
- Autenticación (1 artículo)
- Avances informáticos (1 artículo)
- Bases de Datos Electrónicas (1 artículo)
- Bibliografía (1 artículo)
- Bibliometría (1 artículo)
- Biblioteca (2 artículos)
- Biblioteca 2.0 (2 artículos)
- Biblioteca Académica (3 artículos)
- Biblioteca Digital (1 artículo)
- Biblioteca Unversitaria (1 artículo)
- Biblioteca digital (1 artículo)

Fig. 4. Listado de categorías

página discusión ver código fuente historial

## Historial de revisiones para «“User satisfaction as a measure of system performance”»

Ver los registros de esta página

(Últimas | Primeras) Ver (50 siguientes) (50 anteriores) (20 | 50 | 100 | 250 | 500).

Seleccione diferencias: marca los selectores de las versiones a comparar y pulsa *enter* o el botón de abajo.

Leyenda: (act) = diferencias con la versión actual, (prev) = diferencias con la versión previa, M = edición menor

Comparar versiones seleccionadas

- (act) (prev)  09:58 19 ene 2009 Vanesa (Discutir | contribuciones) (2.666 bytes)
- (act) (prev)  17:26 28 nov 2008 Vanesa (Discutir | contribuciones) (2.442 bytes)
- (act) (prev)  17:26 28 nov 2008 Vanesa (Discutir | contribuciones) (2.441 bytes)
- (act) (prev)  17:24 28 nov 2008 Vanesa (Discutir | contribuciones) (2.454 bytes) (Página nueva: La satisfacción de los sistemas de información (IS) y la recuperación de la información (IR)de cara a los usuarios en búsquedas, es tan general como para ser un sustituto en las ...)

Comparar versiones seleccionadas

(Últimas | Primeras) Ver (50 siguientes) (50 anteriores) (20 | 50 | 100 | 250 | 500).

Fig. 5. Ejemplo de histórico de ediciones

---

# Sistema de informação aberto como recurso de organização do conhecimento e de mediação da aprendizagem

*Open information system as feature of knowledge organizing and learning mediation*

---

**Aida Varela Varela (1), Maria Giovanna Guedes Farias (2), Marilene Lobo Abreu Barbosa (3)**

(1) Instituto de Ciência da Informação, Universidade Federal da Bahia, Av. Reitor Miguel Calmon (parte superior), s/n – Campus do Canela, 40110-100, Salvador-BA, Brasil, varela@ufba.br

(2) Departamento de Ciências da Informação, Universidade Federal do Ceará, Av. da Universidade, 2853 – Benfica, Fortaleza-CE, 60020-181, Brasil, giovannaguedes@ufc.br

(3) Instituto de Ciência da Informação, Universidade Federal da Bahia, marilene@ufba.br

## Resumo

Bibliotecas, arquivos e centros de documentação podem ser considerados sistemas de informação de cunho social, cuja função é a de desenvolver processos de gestão, de modo a viabilizar a difusão da informação para o usuário, objetivando ampliar a base de conhecimento do indivíduo e, conseqüentemente, da sociedade. Assim, esses sistemas, são planejados para operacionalizar o fluxo informacional e possibilitar o acesso e uso da informação pelo usuário, mediante a integração de métodos, pessoas e equipamentos. Nesta comunicação objetiva-se mostrar de que forma um sistema de informação aberto, que tem como finalidade a comunicação e mediação da informação, pode oferecer ferramentas facilitadoras para o aprendizado. Para alcançar tais objetivos, procedeu-se a uma revisão de literatura sobre a temática em foco, buscando estabelecer pressupostos teórico-conceituais, subsidiados pelas diferentes visões de autores nacionais e internacionais. Conclui-se que a construção de um sistema aberto de informação – pautado na partilha do conhecimento, na responsabilidade social da Ciência da Informação, no processo comunicativo, mediacional, interativo e dialógico entre o profissional da informação e o usuário – viabiliza a disseminação do conhecimento e facilita o aprendizado.

**Palavras-chave:** Sistema aberto de informação. Organização do conhecimento. Aprendizagem. Usuário e profissional da informação.

## 1. Introdução

As mudanças constantes nas concepções sobre a natureza do próprio homem e sobre os fenômenos que ele procura compreender caracterizam-se como a metamorfose do conhecimento. Retoma-se, assim, a discussão sobre o sujeito cognoscente, ou seja, o sujeito que constrói o conhecimento.

## Abstract

Libraries, archives and documentation centers can be considered socially oriented information systems whose function is to develop management processes in order to facilitate the dissemination of information to the user, aiming to expand the individual's knowledge base and consequently of the society. Thus, these systems are designed to operationalize the information flow and enable access and use of information by the user, by integrating methods, people and equipment. In this paper aims to show how an open information system, which aims at communication and mediation of information, can provide enabling tools for learning. To achieve these goals, this article proceeded to a review of literature on the subject in focus, seeking to establish theoretical and conceptual assumptions, subsidized by the dialogue with national and international authors. It is concluded that the construction of an open information system - founded on the sharing of knowledge, social responsibility of the Information Science, in the communicative process, mediational, interactive and dialogical with professional information - enables the dissemination of knowledge and facilitates learning of users.

**Keywords:** Open information system. Knowledge organization. Learning. Users. Information professional.

Outrossim, o avanço dos estudos sobre os processos cognitivos comprovam que não há um modo único para a organização do conhecimento. Diferentes formas de ler e compreender o mundo dão origem a diferentes esquemas de ordenação. Há uma pluralidade de paradigmas de conhecimento, que são contextualizados do ponto de vista social, histórico, científico e cultural o que implica uma operacionalização cognoscente objetiva e subjetiva.

Nesta comunicação, analisa-se se o funcionamento de um sistema de informação aberto, cuja finalidade é a comunicação e a mediação da informação, pode oferecer ferramentas facilitadoras para o aprendizado. Para tal, apresentam-se modelos de processamento de informação e de aprendizagem, a exemplo do de Wilson e de Gagné, bem como o modelo social da CI, os quais podem fornecer elementos para a montagem de experiências pedagógicas – processos dialógicos -, que poderão ser aplicados para usuários de informação em contextos variados.

Nesses processos, os integrantes do sistema aberto de informação – entrada, processamento, saída e retroalimentação - serão operacionalizados. A entrada se configura na necessidade informacional dos indivíduos, bem como a análise dos conhecimentos prévios que o sujeito aprendiz detém. É no processamento que ocorre a transformação dos elementos de entrada em contato com os novos conhecimentos provenientes do meio ambiente. A saída está relacionada com a análise do conhecimento obtido pelos usuários após participarem de experiências pedagógicas. Para isso, podem-se estabelecer níveis de codificação de novos conhecimentos e a trajetória desenvolvida a partir do modelo básico de aprendizagem, a exemplo da teoria do processamento de informação de Gagné. A retroalimentação tem como base o paradigma social da Ciência da Informação, que pressupõe a atuação para e com os sujeitos, em uma ação compartilhada, participativa, baseada na responsabilidade social do profissional da informação.

## 2. Modelos de processamento da informação

As questões relacionadas aos processos de busca e acesso à informação não são tarefas simples, isto porque estão na dependência do desempenho dos usuários da informação e da competência de quem organiza e disponibiliza a informação. Estes processos articulam-se entre o desenvolvimento psicológico humano e o desenvolvimento do conhecimento, que, por sua vez, transcende à questão mais ampla que diz respeito à interação entre conhecimento, indivíduo e sociedade, o que configura a transformação e a usabilidade do conhecimento. Aqui, ressalta-se que as práticas profissionais, consideradas práticas sociais, fundamentam-se em crenças, valores, representações sociais e teorias particulares que as sustentam. Essas práticas incluem concepções sobre a relação homem-conhecimento.

Segundo Wilson (1981, p. 13), ao se construir esses diagramas e ao interpretá-los, objetiva-se sugerir que “[...] uma análise pode ser utilizada como ponto de partida para uma pesquisa baseada numa visão holística dos usuários de informação.” Dessa forma, essa visão significa ser percebida não apenas como um indivíduo impulsionado a buscar informações para fins cognitivos, “[...] mas como viver e trabalhar em ambientes sociais que criam suas próprias motivações para buscar informações, para ajudar a satisfazer grandes necessidades afetivas.”

Para fornecer um quadro geral mais eficaz para a reflexão sobre o comportamento de busca de informação, Wilson publicou em 1996 uma revisão sobre o modelo de 1981, com base na pesquisa de uma variedade de outros campos da CI, “[...] incluindo a tomada de decisões, psicologia, inovação, comunicação em saúde e pesquisa do consumidor.” (Wilson, 1999, p. 256). Segundo esse autor (*ibidem*, p. 36), essa revisão foi necessária, pois, percebeu-se a necessidade de incluir uma etapa entre, “[...] o que deve ser chamado a pessoa-em-contexto e a decisão de buscar informações [...]”. Além disso, Wilson (1996, p. 36) sentiu que era preciso elaborar o conceito de variáveis intervenientes ou barreiras, para especificar “[...] o fato de que as características da fonte de informação podem constituir uma barreira para o comportamento de busca de informação ou processamento de informações, e que as variáveis pessoais podem ser psicológicas ou demográficas”, como se observa na Figura 1, no anexo A.

Essa pesquisa promovida por Wilson (1996, p. 36) também proporcionou, ao autor, visualizar a “[...] necessidade de mais um estágio intermediário (ou mecanismo de ativação) entre a determinação da necessidade e o início de ação para satisfazer a necessidade.” De acordo com ele, a proposta é a:

[...] teoria do risco/recompensa, a teoria da aprendizagem social e do conceito de autoeficácia, como conceitos de intervenção. Dada a relação de autoeficácia, é discutível se o conceito deve aparecer aqui ou como parte do esforço/elemento de enfrentamento, mas, dada a sua relação com a teoria da aprendizagem social, a localização mostrada parece pelo menos apropriada. Finalmente, o modelo precisa de extensão, para incluir o processamento e uso de informações, que são os estágios, além da busca de informação e que fornecem o elo de volta para a situação de necessidade da pessoa-em-contexto. (Wilson, 1996, p. 36, tradução dos autores).

Além das teorias acima citadas, envolvidas na proposta de Wilson, para explicar a busca de informação, há ainda, nessa revisão do modelo,

a teoria do estresse/enfrentamento, que oferece possibilidades de explicar porque algumas necessidades não invocam o comportamento de busca por informações (*ibidem*). Como se pode observar, essa versão expandida do modelo de Wilson, de 1981, foi criada, como explica o autor, em 1996, para ser aplicada ao comportamento informacional pleno, ao invés de referir-se, exclusivamente, a um comportamento de busca de informação. Entretanto, como o próprio autor (*ibidem*, p. 251) explica esse tipo de modelo tem suas limitações, o fato é “[...] que ele faz pouco mais do que fornecer um mapa da área e chamar a atenção para as lacunas na pesquisa”, não ofertando sugestões de “[...] fatores causais do comportamento da informação e, conseqüentemente, não indica diretamente hipóteses a serem testadas.” O autor tentou mostrar de forma mais completa como ocorre o comportamento informacional de uma pessoa considerando toda a complexidade em torno do contexto.

Para contribuir para a formação e aprendizagem, é preciso intervir, mas antes de se chegar à efetiva intervenção, existem outros níveis de mediação. Kuhlthau (1993), por exemplo, indica níveis de mediação direcionados e adaptados para bibliotecários, os quais podem ser ampliados para outros profissionais da informação, que assumem papel de mediador.

- (a) **organizador**: é aquele que atua apenas na operacionalização do sistema;
- (b) **localizador**: é aquele, cuja ação concentra-se na localização de informações ou na resposta de questionamentos do usuário; trata-se de uma intervenção pontual;
- (c) **identificador**: é aquele que indica fontes de informação, visando satisfazer as necessidades do usuário;
- (d) **conselheiro**: é aquele que procura entender o problema, para então recomendar fontes pertinentes, do geral ao específico;
- (e) **tutor**: é aquele que percebe a real necessidade do usuário, estabelece um diálogo, interage, recomenda, encoraja, elabora estratégias e avança para o momento da resolução do processo, contribuindo na construção do conhecimento do usuário.

Os cinco níveis de Kuhlthau demonstram que o mediador para estabelecer uma relação de intervenção com o usuário, no processo de organização do saber, recorre ao seu repertório intelectual, social, ideológico e afetivo, como explica Varela (2007, p. 83), acrescentando que o mediado ao receber a informação, transforma-a “[...] por meio de seu próprio repertório, constituído por seus conhecimentos prévios, por suas

predisposições com relação ao conteúdo e ao próprio mediador.” Obviamente, que as mensagens entre esses dois sujeitos transitam por “[...] caminhos multiformes, interligados, gerando movimentos em diferentes condições, uma vez que as fontes se alteram ciclicamente e em espiral numa composição de infinitas variáveis.” (*idem*, p. 83).

Neste contexto, conforme Farias (2014), ao tratar de codificação, é preciso refletir a respeito do processamento de informação, e para isso, utiliza-se a perspectiva da teoria da aprendizagem do neo-behaviorista Robert Gagné, que entende a aprendizagem como um processo, que modifica nosso comportamento de forma rápida e de modo mais ou menos permanente. Gagné (1980, p. 5) desenvolve estudos sobre a representação do que ocorre internamente no indivíduo, elaborando um constructo da possível trajetória interna da aprendizagem, que pode ser verificada por uma mudança comportamental externa, relativamente persistente, o que significa dizer que um observador externo pode “[...] reconhecer que houve aprendizagem quando observa a ocorrência da mudança comportamental e também a permanência desta mudança. Inferindo destas observações um novo ‘estado permanente’, atingido pelo aluno.”

Gagné (*ibidem*, p.14), nos pressupostos da teoria de processamento de informação esclarece como ocorre o fenômeno da aprendizagem nos sujeitos:

[...] aqueles que realizam determinados tipos de transformação de “insumos” para “exsumos” [produto, output], numa forma um tanto análoga às operações de um computador. Por exemplo, quando um estudante está numa situação de aprendizagem, a estimulação física dos seus olhos, ouvidos e outros sentidos é transformada em certas “mensagens” neurais. Por sua vez, as mensagens neurais sofrem outras transformações no sistema nervoso, de tal forma que possam ser armazenadas e, mais tarde, lembradas. A informação lembrada é novamente transformada em ainda outros tipos de “mensagens”, as quais controlam a ação dos músculos. O resultado é um discurso ou outros tipos de movimento que indicam que o desempenho foi aprendido. Estas formas diferentes de transformação são chamadas processos de aprendizagem, e se constituem naquilo que ocorre “dentro da cabeça do estudante”.

Este fenômeno, considerado na sua totalidade, com suas características e seu *modus operandi* constituem a essência da teoria de processamento da aprendizagem de Gagné, e está representado na Figura 2, no anexo B, a partir do qual, segundo esse autor (*ibidem*, p. 16) pode-se obter, “[...] um ‘quadro mental’ de tal modo que se possa seguir o fluxo de informação e

captar a ideia de que este é processado (transformado) de várias maneiras à medida que passa de uma estrutura para outra.”

Com este modelo Gagné (1980) objetivou apresentar as estruturas que pressupõe existirem no sistema nervoso central do aluno, supondo que elas sejam redes neurais. Os componentes definidos pelo autor são:

- (a) **ambiente** – a informação proveniente do ambiente influencia os seus receptores e entra no sistema nervoso, através de um registrador sensorial;
- (b) **receptores** - esta é a estrutura responsável pela recepção inicial de objetos e eventos que o aluno vê, ouve ou sente, o que envolve a atenção e a percepção;
- (c) **registrador sensorial** – momento em que o filtro perceptual seleciona a informação de interesse e codifica-a e esta adquire a forma de uma representação padronizada;
- (d) **memória de curto prazo** - ao entrar neste tipo de memória, a informação se reforçada, por estímulos, adquire um formato conceitual, embora sua permanência seja de pouco tempo. Entretanto, se processada novamente por ensaio interno será preservada na memória de breve duração por período mais longo;
- (e) **memória de longo prazo** – o processo de repetição da informação nova, em novos contextos, desempenha um papel de reforço e, se a informação for apreendida, é de novo transformada e entra na memória de longa duração, onde é armazenada para posterior rememoração;
- (f) **gerador de respostas** - quando recuperada, a informação, proveniente tanto da memória de breve duração quanto da memória de longa duração, passa por um gerador de resposta, que tem a função de transformar a informação, para retornar ao mundo externo, em signos ou em ação;
- (g) **executores** – a “mensagem” neural, proveniente da estrutura, ativa os executores físicos, levando à externalização, por meio de signos e ações. O que acontece neste momento é o que capacita o observador externo a perceber se a estimulação inicial promoveu o efeito esperado, ou seja, se a “informação foi processada” e se ocorreu a aprendizagem;
- (h) **controle executivo e expectativas** – os sinais vindos destas duas estruturas, como estereótipos, ideias preconcebidas, radicalismos etc., supostamente ativam e modificam o fluxo de informação. Por exemplo, o estudante tem uma perspectiva do que ele será capaz de fazer, uma vez que ele tenha aprendido, e sua convicção, por sua vez, pode modificar a percepção do estímulo externo, sua codificação na memória e transformação da informação em ação. Os processos de controle originados na estrutura de controle executivo podem determinar como a informação é codificada, quando entra na memória de longa duração e como a procura e a recuperação são conduzidas, tendo em vista, entre outras coisas, a rememoração. (*ibidem*, p. 18-20).

Alguns dos eventos que produzem o ato de aprendizagem são externos ao estudante e, por isso mesmo, são mais facilmente observáveis: a estimulação que atinge o estudante e os produtos (incluindo a informação escrita e falada) resulta de sua resposta. Já muitos eventos da aprendizagem são internos e são inferidos de observações externas. As atividades internas são chamadas de processos de aprendizagem. Estes processos são subentendidos pelos pontos de transição entre cada conjunto de estruturas do modelo básico de aprendizagem (*ibidem*). Os componentes deste modelo demonstram, ao final, que a aprendizagem ocorre como resultado da interação entre o indivíduo e seu ambiente, reforçando o caráter neobehaviorista reflexões de Gagné.

### 3. Sistema de informação aberto

O papel das bibliotecas desde a Antiguidade até a Idade Contemporânea tem sido universalizar o conhecimento produzido e acumulado pela humanidade, transmitindo-o a todas as gerações. Por isso, pode-se inferir que, ao permitir o acesso à informação, as bibliotecas e arquivos se constituem como os primeiros sistemas de informação (SI) estruturados.

Considerando a teoria de sistemas, de Bertalanfy, pode-se afirmar que o SI é um sistema aberto, tendo em vista estar em constante interação com o meio ambiente e em permanente geração de fluxo informacional, sendo organizado e integrado por pessoas, equipamentos e métodos. Para Bertalanfy (2009), os sistemas abertos realizam interação com o meio ambiente através de entradas (input) e saídas (output), e tem como base a interação dinâmica dos seus componentes, com objetivo de formar um todo para determinados fins, mantendo uma relação de troca.

Laudon e Laudon (1999, p. 4) definem o SI como “[...] componentes inter-relacionado, trabalhando juntos para coletar, recuperar, processar, armazenar e distribuir informação com finalidade de facilitar o planejamento, o controle, a coordenação, a análise e a tomada de decisão nas organizações.”

Os SI são pensados, conforme Araújo (2009), a partir da lógica dos processos de entrada (entrada de dados, com a aquisição de itens informacionais, a seleção destes itens para a composição de determinado acervo), de processamento (os itens informacionais que dão entrada num sistema de informação precisam ser descritos, catalogados, classificados, indexados) e de saída (pelo acesso aos itens informacionais por

parte dos usuários, na forma de disseminação, entrega da informação, empréstimo, etc).

Os sistemas de informação aberto (SIA) gerados por bibliotecas/arquivos e centros de documentação – são sistemas sociais que têm a função de executar processos de gestão/organização da informação e de sua comunicação/difusão para o usuário/sociedade, objetivando ampliar a base de conhecimento do indivíduo e, conseqüentemente, da sociedade. Eles tornaram-se cada vez mais presentes na realização dessas atividades, possibilitando o acesso e uso da informação. A biblioteca se constitui, como um SIA, ao buscar interagir e favorecer o desenvolvimento da autonomia dos usuários, no acesso, utilização e aplicação das informações, com foco no paradigma social da Ciência da Informação, refletindo a respeito das mudanças cognitivas e sociais dos sujeitos.

A **entrada** diz respeito às influências que são recebidas pelo ambiente externo ao sistema, são os insumos, a matéria prima, a informação, e outros elementos. O **processamento** é responsável por transformar os elementos provenientes do meio em produtos ou resultados. O sistema aberto se relaciona com o ambiente mantendo uma relação de troca. Nessa relação entre o sistema e o ambiente há troca de energia: importa insumos e exporta resultados. Nota-se também que o ambiente é um sistema formado por outros sistemas. O **ambiente** é constituído por elementos que não são parte do sistema, porém o influencia todo o tempo provocando modificações e necessidade de novas adaptações. As **saídas** são resultados (produtos ou serviços) do processamento realizado pelo sistema. Tanto as entradas quanto as saídas permitem que o sistema interaja com outros sistemas que fazem parte do ambiente. (Santana; Varela, 2014). A **retroalimentação** é o retorno daquilo que foi liberado pela saída ao ambiente, como opinião e sugestão da clientela ou usuários, lucros, entre outros, e ele pode ser positivo ou negativo, possibilitando correções, ajustes, modificação ou reforçamento do sistema. São saídas que retornam para o sistema para dar suporte na avaliação ou correção de entrada. (Laudon; Laudon, 1999).

A figura 3, anexo C, representa um sistema de informação aberto que pode ser utilizado em uma biblioteca para organizar o conhecimento e possibilitar o aprendizado dos usuários. Este sistema reúne elementos como a mediação da informação, que é conceituada na CI como uma ação de interferência, tendo a comunicação como aliada do profissional da informação para analisar previamente as necessidades e os conhecimentos dos usuários, facilitando a entrada

de insumos que serão processados. É no processamento que ocorre a transformação dos elementos provenientes do meio ambiente, que entraram no sistema. A saída está relacionada à análise do conhecimento obtido pelos usuários, após participarem de experiências pedagógicas, podendo, para isso, estabelecer níveis de codificação e uma trajetória desenvolvida a partir do modelo básico de aprendizagem e memória subjacente nas teorias de aprendizagem, a exemplo da teoria de processamento de informação de Gagné. A retroalimentação teria como base o paradigma social da Ciência da Informação, que pressupõe a atuação para e com os sujeitos, em uma ação compartilhada, participativa, baseada na responsabilidade social do profissional da informação.

De acordo com Varela e Barbosa (2012), no atual contexto de mudanças constantes na sociedade da informação, os sistemas de informação abertos - a partir dos ambientes virtuais de organização e disseminação da informação - criam e oferecem serviços interativos que respondam às questões do usuário, efetivando a disseminação, significação e usabilidade da informação, sobre os princípios da objetivação e subjetivação da condição humana na dinâmica de busca e uso da informação, mediante a explicitação da necessidade informacional do usuário, tendo como subsídios a leitura, as competências informacionais e a mediação dos profissionais da informação.

Ainda segundo as autoras, um sistema de informação aberto deve, por meio de uma ação comunicativa entre o estoque de conhecimentos e o usuário, tornar possível a recuperação e a produção de novos conhecimentos. Para Varela e Barbosa (2012), os procedimentos e técnicas que vão concretizar a ação de transferência da informação para o usuário estão assentados em duas ações: por um lado, é preciso representar o conteúdo do item informacional – é o ponto de acesso; por outro, é preciso criar instrumentos de busca – os índices. Neste ponto de confluência há que se lançar mão das teorias cognitivas e de aprendizagem, porque, para representar o conteúdo informacional é preciso simular o comportamento do usuário na busca da informação.

#### 4. Outros pensadores sobre o processo de aprendizagem

Focaliza-se aqui a compreensão do complexo fenômeno da aprendizagem humana, que é abordada, a partir de uma galeria de teóricos, postos em discussão acerca das diferentes

perspectivas teóricas e metodológicas ante as diferentes concepções do fenômeno.

Piaget (1974) argumenta que a epistemologia genética trata de compreender os processos do conhecimento científico por meio de uma combinação de análises psicogenéticas e de formalização de estruturas, de sorte que seja possível atingir e compreender as condições psicológicas da formação dos conhecimentos elementares e coordenar esses resultados com o estudo das condições de formalização.

A epistemologia genética trata das relações de conhecimento entre o sujeito (aquele que conhece) e os objetos (tudo aquilo a ser conhecido). Além das questões lógicas ou de validade que estão implícitas nessas relações, há duas questões especificamente relacionadas ao sujeito (aquele que conhece): de um lado, a questão do objeto a ser conhecido por meio da experiência e, de outro, o fato de que essa experiência é sempre a experiência de um sujeito. E assim, necessita-se desenvolver uma psicogênese, ou seja, uma análise genética do desenvolvimento do pensamento, a fim de se reconstituir a formação das estruturas pelo estudo das ações do sujeito, passo a passo. Assim, a contribuição fundamental da teoria de Piaget está em considerar o homem um ser ativo na construção do conhecimento.

Segundo Piaget (1971), a estrutura cognitiva é construída em etapas e cada etapa incorpora as anteriores, dando-se a construção do conhecimento pela ação recíproca e interativa do sujeito com os objetos (meio). A organização da realidade dá-se por meio do pensamento estruturado, que se expressa mediante o processo de adaptação. Deste modo, a estrutura mental e o conhecimento são construídos em uma relação dialética entre a maturação biológica e o ambiente.

Para entender o processo de apreensão da realidade por meio da estrutura cognitiva, que organiza os estímulos do mundo, Piaget (1971) desenvolveu e definiu os dois pilares de sua teoria: a assimilação e a acomodação. O conceito de assimilação diz respeito à capacidade de a estrutura cognitiva do indivíduo atuar no ambiente. O organismo assimila a realidade de acordo com a capacidade de sua estrutura cognitiva interna. Em contrapartida, a estrutura cognitiva é mobilizada e modificada em função da realidade, acomodando-se a ela. A adaptação justifica-se como um estado de equilíbrio entre a acomodação e a assimilação.

A aprendizagem é uma modificação duradoura (equilibrada) do comportamento, em função das aquisições decorrentes da experiência. O co-

nhecimento constrói-se em um movimento contínuo de equilíbrio, daí a importância de que a ação pedagógica deve provocar o sujeito aprendiz, por meio de situações desequilibradoras e dar espaço para que ele possa criar e/ou descobrir as soluções a partir do próprio esforço para superar o desequilíbrio. (Piaget, 1987).

Vygotsky (1979) reafirma que a consciência é uma propriedade essencial do homem, mediada pela sua experiência com o mundo e se propõe a estudar o comportamento, seus mecanismos, suas partes componentes e sua estrutura. Seu argumento central é que tudo que se tem nos experimentos de reflexos mostra quão complexa cada resposta é e o quanto essa complexidade depende da estrutura dos mecanismos comportamentais dos quais a resposta é uma parte, e a impossibilidade de estudar a resposta de forma abstrata.

Este autor declarou que os processos mentais elementares constituem a base a partir da qual se dá o desenvolvimento superior mediado semioticamente. A tese de Vygotsky defende o domínio dos signos ou dos “instrumentos psicológicos” como modificador qualitativo do funcionamento mental, o que significa dizer que a ontogênese é de fato, governada pela organização dos sistemas de signos, e a única maneira de se descobrir a natureza desse funcionamento, no indivíduo, consistia em retomar suas origens.

A abordagem sócio-histórica de Vygotsky (1991) baseia-se na ideia central de que o ser humano se desenvolve pela interação social, quando o desenvolvimento cognitivo mantém estreita relação com a aprendizagem. A aprendizagem, quando significativa, estimula e desencadeia o avanço para um nível de maior complexidade que, por sua vez, serve de base para novas aprendizagens.

Vygotsky formulou os conceitos de zona de desenvolvimento real, zona de desenvolvimento proximal e zona de desenvolvimento potencial. A zona de desenvolvimento real é o nível de desenvolvimento das funções mentais do indivíduo que se estabeleceram como resultado de certos ciclos de desenvolvimento já completados, ou de um processo de desenvolvimento específico. A zona de desenvolvimento proximal é a distância entre o nível de desenvolvimento real, que se costuma determinar com a solução de problemas de forma independente, e a zona de desenvolvimento potencial determinado, com a solução de problemas sob a orientação de outro indivíduo ou em colaboração com companheiros mais capazes.

Todo ser humano, inserido em uma realidade sócio-histórica, somente adquire a condição humana se for, em sua relação com o mundo, mediado por instrumento de sua cultura – signo, palavra, símbolo. O conhecimento é, portanto, uma produção cultural, diretamente relacionada com a linguagem e com a interação social. A mediação é a ação que se interpõe entre sujeito e objeto de aprendizagem, sendo a palavra de fundamental importância.

A obra de Paulo Freire (1980) se constitui em uma tentativa de recuperar e proporcionar conceitos e instrumentos metodológicos do estruturalismo genético e da razão dialética, mostrando sua aplicabilidade ao processo pedagógico. No conjunto de sua obra, podem ser destacados alguns elementos ontológico-epistemológicos, entre eles o conceito de “educação dialógica” versus “educação bancária”. Na “educação bancária” duas estruturas se confrontam – a do educador que sabe e comunica e a do educando que não sabe e aprende – e na educação dialógica, dois processos de estruturação gnoseológica interagem, um ensinando/aprendendo com o outro e ambos construindo o conhecimento do mundo.

Freire (1979, 2002) ressalta a importância e a necessidade de se entender a existência humana a partir de sua substancialidade, ou seja, o reconhecimento de todos os homens como verdadeiros sujeitos históricos. Os atributos dos seres humanos não podem, assim, sobrepujar o reconhecimento da existência humana: a sua presença no mundo como sujeito.

Para Feuerstein (1980), a modificabilidade é um fator tanto biológico quanto cultural, o que o remete a Vygotsky. O conceito de modificabilidade implica uma estrutura cognitiva permeável aos estímulos culturais e torna cada indivíduo único, imprevisível, capaz de superar suas condições atuais. Feuerstein parece apropriar-se da ênfase vygotskiana da mediação cultural dos processos psicológicos, enfatizando epistemologicamente a necessidade e importância da ação mediada nos processos de construção intelectual do sujeito.

Em síntese, percebe-se que Feuerstein (1980) e Vygotsky (1979) desenvolveram suas teorias sob forte influência de Piaget. Contudo, eles foram percebendo que certos aspectos da diretriz piagetiana poderiam ser enriquecidos. Para Vygotsky, a insatisfação referia-se ao individualismo epistemológico da teoria de Piaget e à negligência da mediação social. Para Feuerstein, havia a questão dos mecanismos concretos da aprendizagem sobre a mediação de outro ser humano (Gomes, 2002).

Feuerstein (1980) explica a aprendizagem humana, não só pela integridade biológica dos genes e cromossomos ou pela simples exposição direta a objetos, acontecimentos, atitudes e situações, mas emerge de uma relação indivíduo-meio, que é mediatizada por outro indivíduo mais experiente, cujas práticas e crenças culturais são transmitidas, promovendo zonas mais amplas de desenvolvimento crítico e criativo – desenvolvimento de processos psicológicos superiores – rumo à autonomia cognitiva, portanto decorre de uma aprendizagem mediada.

Segundo Morin (2000), um conhecimento pertinente deve enfrentar a complexidade. Complexus significa “o que foi tecido junto”. Há complexidade quando elementos diferentes, constitutivos do todo, são inseparáveis (como o econômico, o político, o sociológico, o psicológico, o afetivo, o mitológico), há um tecido interdependente, interativo e inter-retroativo do objeto de conhecimento com seu contexto – as partes e o todo, o todo e as partes, as partes entre si. A complexidade, portanto, é a união da unidade com a multiplicidade, sabendo-se, de antemão, que sempre haverá a incerteza no pensamento complexo, o que permite escapar do dogmatismo que reina nos pensamentos não complexos.

Para Morin (1986) a capacidade de aprender está ligada ao desenvolvimento das competências inatas do indivíduo em adquirir conhecimentos, associadas às influências e estímulos externos, da cultura e o ato de conhecer está presente nas ações biológicas, cerebrais, espirituais, culturais, linguísticas, sociais políticas e históricas. Desse modo, o ser condiciona o conhecer, que ao mesmo tempo condiciona o ser.

#### 4.1. O caráter técnico, social e mediador da organização do conhecimento

A organização da informação e sua estruturação em sistemas abertos podem ser consideradas verdadeiras estratégias de mediação, desde que utilizem princípios teóricos e metodológicos consoantes com o complexo fenômeno da aprendizagem humana. Deste modo, garantir-se-ia a facilitação da aprendizagem e a consequente promoção do acesso ao conhecimento.

A organização da informação, do ponto de vista da Biblioteconomia e da Ciência da Informação, transcende às questões operacionais e logísticas de processar/tratar a informação, na medida em que os princípios da disciplina estão atrelados ao ideal de tornar acessível a todos o conhecimento produzido.

Com efeito, analisando-se a trajetória histórica da área, evidenciam-se que os vários processos

de organização e disseminação da informação estão centrados nos eixos da criação e padronização das ferramentas de representação da informação, recursos mediadores do acesso ao conhecimento, marcando o caráter social, cognitivo e mediador, em essência, da disciplina.

Atualmente, nas redes informacionais, os conteúdos se entrelaçam, maximizando e dificultando, as possibilidades de busca e recuperação, como a mediação na internet por motores de busca, que revelam preciosidades ou conteúdos respostas sem significado. Estudiosos de renome do campo da indexação e recuperação da informação, como Lancaster e Tefko Saracevic, ressaltam a necessidade de que princípios de relevância e precisão sejam preservados para garantir qualidade nas pesquisas bibliográficas.

A Web Semântica surgiu para a superação das barreiras mencionadas por meio das várias soluções tecnológicas implementadas em sua arquitetura, como metadados, linguagem de marcação XML e a arquitetura *RDF-Resource Description Framework*, que garantem a interoperabilidade dos dados em nível semântico, estrutural e sintático; além das ontologias, que proporcionam a definição semântica; e os agentes inteligentes ou motores de busca, que, espera-se, proporcionarão uma busca e recuperação mais efetiva a partir do estabelecimento de regras e do acesso a coleções de recursos devidamente estruturados, representados e definidos semanticamente (Pickler, 2007, p. 17).

Os processos e ferramentas de organização e representação do conhecimento, no âmbito da biblioteconomia e ciência da informação, têm funcionalidades nitidamente mediadoras, tendo em vista que objetivam a localização da informação na perspectiva do acesso pleno ao conhecimento.

Sobre esta relação intrínseca da cognição com os processos de tratamento e difusão da informação, já na década de 70 do século XX, Shera (1973) chamava a atenção sobre as questões de mediação que envolviam o usuário, os registros gráficos e a cognição. Conforme Castells (2003) os modos como as informações são construídas e apresentadas são diversificados de acordo com os esquemas de quem as disponibiliza e não de quem os busca, o que pressupõe a necessidade de mediação e de conhecimento de teorias cognitivas.

Para Latour (2001) a mediação ocorre no espaço social de interações e produção do conhecimento, incluindo entidades humanas e artificiais num mesmo continuum; já Vygotsky (2012) assevera que as funções mentais superiores do

ser humano devem ser consideradas produtos de uma atividade mediada e que, para acessar a informação a mediação é desempenhada pelos instrumentos materiais e psicológicos, construídos artificialmente, observando que ambos têm caráter social por natureza. Estes pensamentos remetem ao papel das ferramentas naturais e artificiais representativas do conhecimento que atuam como pontes para acesso às fontes de informação.

Ingwersen (2002), na concepção do modelo cognitivo do processo de recuperação da informação, identifica os processos de cognição que ocorrem durante o processamento da informação. Por seu turno, nos anos 90, Hjørland (1995) acrescenta que a indexação é um processo intelectual dependente da cognição e do domínio do contexto físico, psicológico e sócio-cognitivo, supondo que o indexador necessita realizar a identificação e a seleção de conceitos na concepção orientada para o conteúdo.

## 5. Considerações finais

Neste artigo, procurou-se refletir sobre a utilização de sistemas de informação abertos, organização do conhecimento e da aprendizagem como movimento das tecnologias colaborativas e de acesso livre, no sentido de criar facilidades de acesso e uso da informação para o usuário, na perspectiva de incrementar a aprendizagem, por meio de estratégias mais flexíveis, atrativas e dialógicas. No bojo desta proposta, está implícita a necessidade de aprofundamento de estudos sobre a complexidade do processo de aprendizagem, a complexa manobra cognitiva do sujeito, sobre o emaranhado de fontes de informação advindas do avanço tecnológico; sobre o desempenho das políticas institucionais e profissionais que lidam com a informação, ao aplicar as inovações científicas e tecnológicas e sobre o papel do mediador neste contexto, de sofisticados mecanismos e ferramentas de busca em razão das múltiplas sintaxes.

Buscar e usar a informação constituem competências cruciais na sociedade da aprendizagem, envolve a busca ativa ou passiva da informação, planejamento, estratégias e motivação para atingir objetivos, monitoração de estratégias, conhecimento e definição de canais ou fontes de informações potenciais, competências para usar tecnologias da informação e avaliação desse processo.

O uso da informação compõe-se de atividades em que o indivíduo se engaja para apreender a informação e transformá-la em conhecimento, compreendendo habilidades intelectuais como decodificação, interpretação, controle e organi-

zação do conhecimento. A decodificação e a interpretação por sua vez incluem atividades de leitura, estabelecimento de relações e conhecimento prévio e as novas informações, comparação de vários pontos de vista e avaliação. Controle e organização relacionam-se propriamente à organização da informação por meio de uso de instrumentos cognitivos, como recursos, esquemas, mapas conceituais e elaboração de textos.

Nem sempre a compreensão se dá na plenitude, por isso fala-se em níveis de construção do sentido, pois os esquemas cognitivos são particulares a cada indivíduo, construídos pelas experiências vivenciadas. As conexões entre os conhecimentos adquiridos ao longo dos anos vão sendo estruturadas continuamente, de uma forma que é particular a cada um. A compreensão envolve uma rede complexa que se entrelaça com a visão, com o reconhecimento, com a memória, dedução, análise, conexão e avaliação no processo de apreensão da informação.

A essência da temática em questão aponta para o investimento de uma educação direcionada ao desenvolvimento de competências em informação, fazendo com que profissionais e usuários assumam o papel de mediadores da informação, buscando reciprocidade, significado e transcendência desta informação, percorrendo o fluxo de um processo cognitivo nas dimensões de acesso, análise, interpretação, avaliação, produção etc., fechando o ciclo informacional e pedagógico – a informação como ponto de saída e de entrada.

Devido a toda essa complexidade, os profissionais de informação dedicados à organização da informação e do conhecimento, lidando com sistemas de informação abertos, na perspectiva do processo de aprendizagem, devem observar que:

- o caráter processual e progressivo da compreensão garante a possibilidade de passar a estruturas sempre mais profundas na abordagem do conhecimento;
- a representação do conhecimento compreende várias maneiras nas quais nossas mentes criam e modificam as estruturas mentais, que representam o que se conhece sobre o mundo externo, envolvendo tanto a forma declarativa (saber o que), como a forma não-declarativa (saber como) de conhecimento;
- pela imaginação, criam-se estruturas mentais que representam coisas que no momento não estão sendo percebidas pelos órgãos sensoriais;
- a unidade fundamental do conhecimento simbólico é o conceito, e os conceitos podem ser organizados em esquemas, que podem incluir outros esquemas, variar em aplicação e em abstracionismo, e incluir informações sobre as relações entre conceitos, atributos, contextos e conhecimento geral, bem como informações sobre relações causais.

## Referencias

- Araújo, C A A (2009). Correntes teóricas da Ciência da Informação. *Ciência da Informação*. 38:3 (Setembro 2009) 192-204.
- Bertalanffy, L V (2009). *Teoria Geral dos Sistemas: fundamentos, desenvolvimento e aplicações*. Petrópolis, RJ: Vozes.
- Castells, M A (2003). *Galáxia da Internet. Reflexões sobre a internet, os negócios e a sociedade*. Rio de Janeiro: Jorge Zahar Ed.
- Farias, M G G. (2014). *Análise da produção, implementação e avaliação de um modelo de mediação da informação no contexto de uma comunidade urbana*. 283f. Tese. (Doutorado em Ciência da Informação) – Programa de Pós-Graduação em Ciência da Informação, Universidade Federal da Bahia, Salvador, 2014.
- Feuerstein, R (1980). *Instrumental enrichment: an intervention program for cognitive modifiability*. Baltimore: University Park Press.
- Freire, P (2002). *Pedagogia do oprimido*. Rio de Janeiro: Paz & Terra.
- Freire, P (1979). *Educação e mudança*. Rio de Janeiro: Paz & Terra.
- Freire, P (1980). *Conscientização*. São Paulo: Moraes.
- Gagné, R M (1980). *Princípios Essenciais da Aprendizagem para o Ensino*. Porto Alegre: Ed. Globo.
- Gomes, C M A (2002). *Feuerstein e a construção mediada do conhecimento*. Porto Alegre: Artmed.
- Hjorland, B (1995). Toward a new horizon in information science: domain analyses. *Journal of the American Society for Information Science*. 46:6 (1995) 400-425.
- Ingwersen, P (2002). *Information retrieval interaction*. London: Taylor Graham.
- Latour, B (2001). *A esperança de Pandora: ensaios sobre a realidade dos estudos científicos*. Bauru: Edusc.
- Laudon, K C; Laudon, J P (1999). *Sistemas de informação*. Rio de Janeiro: LTC.
- Kuhlthau, C C (1993). A Principle of Uncertainty for Information Seeking. *Journal of Documentation*. 49:4 (1993) 339-355.
- Morin, E (2000). *Ciência com consciência*. Rio de Janeiro: Bertrand.
- Morin, E (1986). *Para sair do século XX*. Rio de Janeiro: Nova Fronteira.
- Piaget, J (1987). *O nascimento da inteligência na criança*. Rio de Janeiro: Guanabara.
- Piaget, J (1971). *A Epistemologia Genética*. Petrópolis: Vozes.
- Piaget, J (1974). *Aprendizagem e conhecimento*. Rio de Janeiro: Freitas Bastos.
- Pickler, M E V (2007). Web Semântica: ontologias como ferramentas de representação do conhecimento. *Perspectivas em Ciência da Informação*. 12:1 (Janeiro/Abril 2007) 65-83.

- Santana, I F; Varela, A V (2014). Sistema de educação a distância na UFBA: acesso, uso e comportamento informacional dos usuários. En: XV Encontro Nacional de Pesquisa em Ciência da Informação – Enancib. Belo Horizonte. Além das nuvens: expandindo as fronteiras da Ciência da Informação. *Anais...* Belo Horizonte: An-cib/PPGCI/UFMG, 2014. 2724-2743.
- Shera, J (1973). Toward a theory of librarianship and information science. *Ciência da Informação*. 2:2 (1973) 87-97.
- Varela, A V (2007). *Informação e Autonomia: a mediação segundo Feuerstein*. São Paulo: Editora SENAC.
- Varela, A V; Barbosa, M L A (2012). The Complexity of Finding Information in Collaborative Information Systems: Cognitive Needs. En: Currás, E; Romero, N L. (eds.). (2012) *Systems Science and Collaborative Information Systems: Theories, Practices and New Research*. Hershey PA - Estados Unidos: IGI Global, 87-120.
- Vygotsky, L S (1991). *A formação social da mente: o desenvolvimento dos processos psicológicos superiores*. São Paulo: Martins Fontes.
- Vygotsky, L S (1979). *Pensamento e Linguagem*. Lisboa: Ed. Antídoto.
- Vygotsky, L S (2012). *Pensamiento y lenguaje*. Barcelona: Paidós.
- Wilson, T D (1999). Models of information behavior research. *Journal of Documentation*. 55 (1999) 249-70.
- Wilson, T D (1981). On user studies and information needs. *Journal of Documentation*. 37:1 (1981) 03-15.
- Wilson, T D (1996). *Information behaviour, an interdisciplinary perspective*. 1996. Recuperado 23-04-2014 de <http://informationr.net/tdw/publ/infbehav/>.

## Anexo A

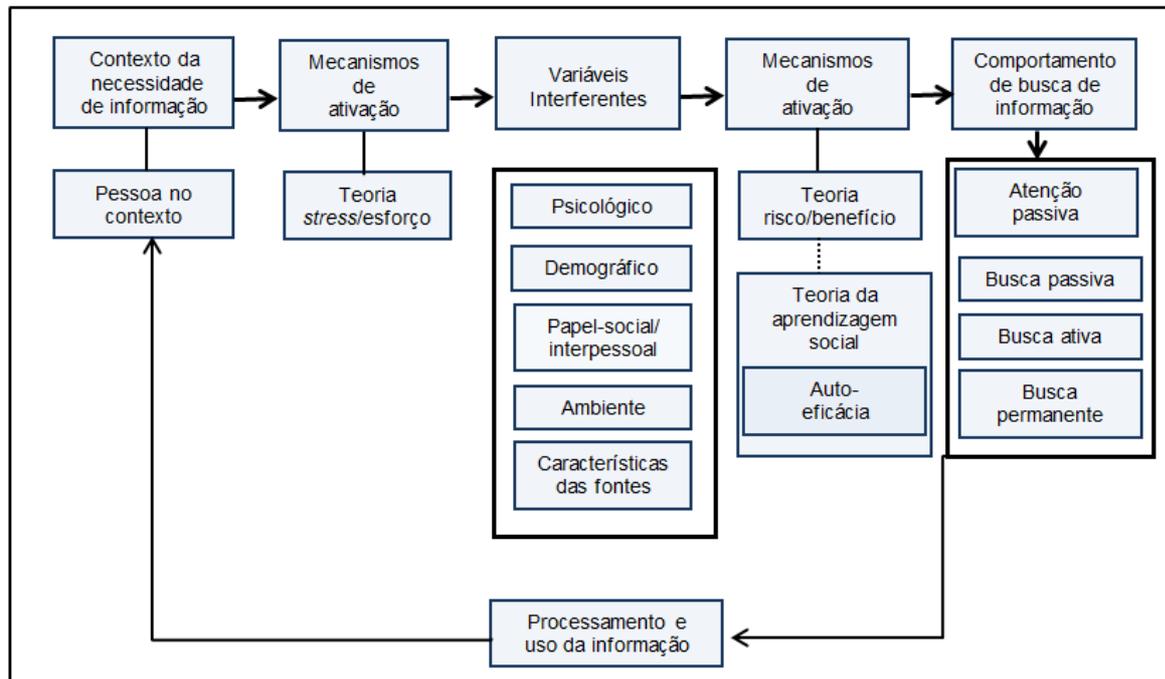


Figura 1. Modelo de comportamento informacional de Wilson de 1996

## Anexo B

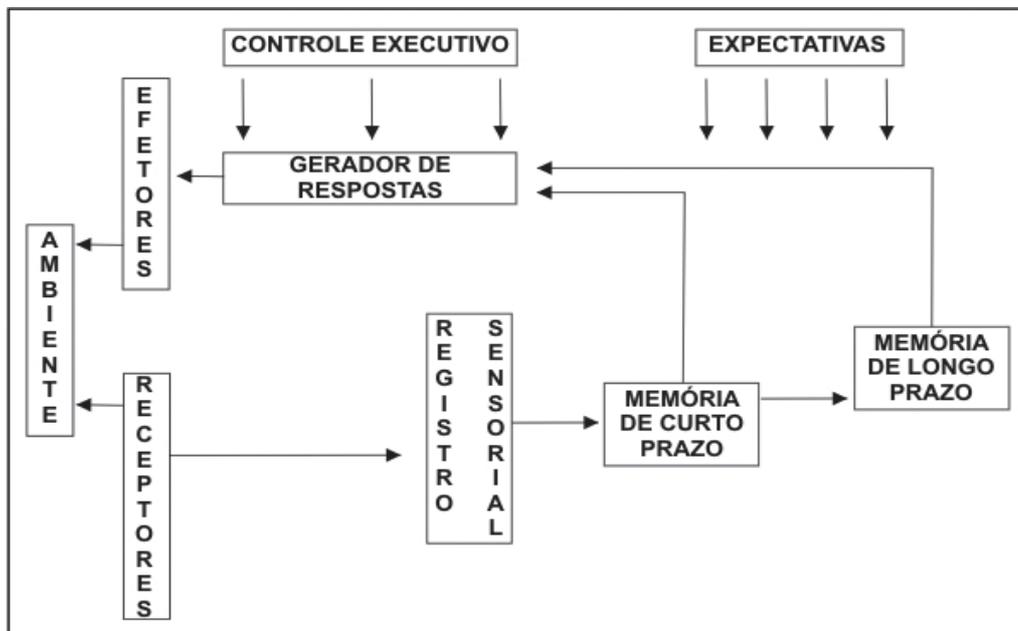
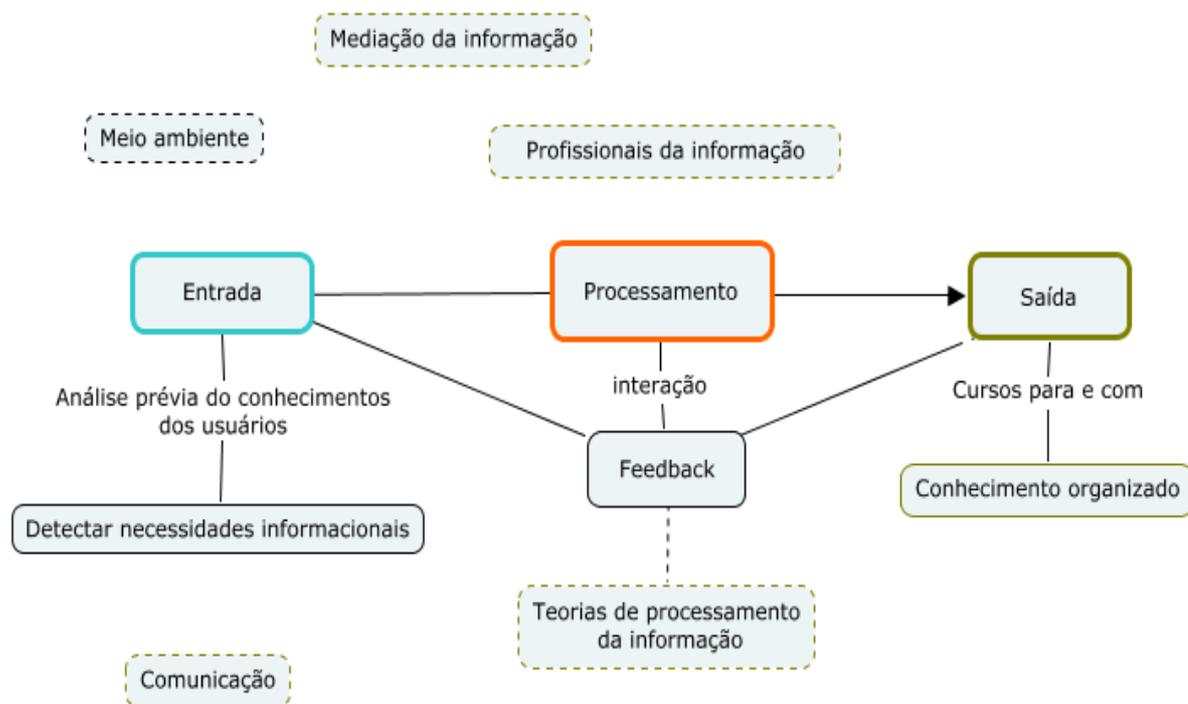


Figura 2. Modelo básico de aprendizagem e memória subjacente nas teorias de "processamento de informação" de Gagné 1980

## Anexo C



**Figura 3.** Sistema de informação aberto em um ambiente de informação. Elaborado pelas autoras, 2015.

**Hito 1: Modelos de organización del conocimiento. Modelaje de sistemas y ciclo de vida de la información**

**Hito 2: Entornos y sistemas para la organización del conocimiento. Diseño, normalización e interoperabilidad de sistemas de información abiertos.**

**Hito 3: Instrumentos para el acceso y transparencia de la información: gestión de documentos, plataformas digitales y archivos**

**Hito 4: Usuarios y usos de la información. Patrones vinculados a la transparencia y el gobierno abierto**

**Hito 5: Gobernanza y Auditoría de información en la organización del conocimiento.**

**Hito 6: Representación del conocimiento: procesos y herramientas en sistemas abiertos y multilingües**

**Hito 7: Epistemología en la organización del conocimiento**

**Hito 8: Igualdad, género y diversidad en la organización del conocimiento**

---

# A justiça na *web*: estudo sobre a gestão da informação em processos judiciais electrónicos no Brasil.

*Justice on the web: study on the management of information in electronic legal proceedings in Brazil*

---

Marieta Marks Löw (1)

(1) Programa Doutoral em “Informação e Comunicação em Plataformas Digitais”, Faculdade de Letras da Universidade do Porto, Portugal, marietalow@gmail.com

## Resumo

Apresenta os resultados preliminares de pesquisa sobre gestão da informação judicial em tribunais brasileiros onde houve a adopção de plataformas digitais para a produção e tramitação de processos. O artigo discute a produção e gestão da informação e documentação judicial em ambiente digital no contexto da justiça brasileira. Apresenta o contexto jurídico que permitiu a mudança de suporte dos autos processuais. Avalia o papel dos profissionais de informação nos tribunais brasileiros no contexto do paradigma pós custodial. Analisar a produção, uso e acesso da documentação processual em ambiente digital. Discute as estratégias de preservação digital no contexto da documentação judicial. Nos últimos anos tribunais brasileiros passaram a usar plataformas digitais de informação em ambiente *web* para produzir e dar andamento aos autos processuais. A partir de pesquisa documental e empírica em tribunais brasileiros, observa-se que a diversidade de soluções tecnológicas desenvolvidas nos diferentes ramos de justiça e instâncias processuais dificulta a gestão da informação judicial. Há uma expectativa de unificação dos sistemas a partir das políticas desenvolvidas pelo Conselho Nacional de Justiça. Entretanto a ausência de profissionais da ciência da informação e de políticas de preservação digital podem comprometer o acesso futuro à documentação produzida.

**Palavras chave:** Gestão da Informação; Gestão Documental; Justiça; Desmaterialização; E-Justiça; Preservação Digital;

## 1. A “Justiça Eletrônica” e a gestão da informação judicial

O chamado “processo eletrônico” tem substituído, nos tribunais brasileiros, as rotinas manuais que envolvem a produção e tramitação da documentação em papel por rotinas automatizadas em sistemas informacionais através da rede mundial de computadores. Essas mudanças impactam tanto a rotina cartorária como a própria relação da sociedade com a justiça. Trata-

## Abstract

The article presents partial results of a research on information management in Brazilian courts with digital proceedings. It discusses the production and management of information and legal documentation on the digital environment in Brazilian justice. It presents the legal framework that allowed the change to the digital context. It assesses the role of information professionals in the Brazilian courts in the context of post custodial paradigm. It's analyzed the production, use and access in the digital environment. It's discussed the digital preservation strategies in the context of judicial documentation. In recent years Brazilian courts began to use digital platforms of information over the web to produce documents in judicial proceedings. From documentary and empirical research in Brazilian courts, it was observed that the diversity of technological solutions developed in the different branches and hierarchical levels of justice hinders the management of judicial information. It's expected that the policies developed by the National Council of Justice may unify the technological solutions already been used. However the absence of information science professionals and of digital preservation policy may compromise future access to the documentation now produced.

**Keywords:** Information Management; Document Management; Justice; Dematerialization; E-Justice; Digital Preservation;

se da chamada “desmaterialização do processo” no âmbito do fazer da justiça.

Neste artigo discutiremos os resultados preliminares de pesquisa de doutoramento que analisa a gestão da informação judicial em plataformas digitais, no contexto da justiça brasileira. A pesquisa busca investigar como a adopção desses sistemas informacionais impacta a produção, uso, acesso e preservação da informação no sistema de justiça. Para tanto, foram selecionados quatro diferentes tribunais brasileiros, nos

principais ramos de justiça do país, para serem como estudos de caso. Com o uso de metodologia qualitativa foram realizadas observações e entrevistas nos seguintes órgãos: Tribunal de Justiça do Rio Grande do Sul, o Tribunal Regional do Trabalho da 12ª Região, em Santa Catarina, o Tribunal Regional Federal da 5ª Região, em Pernambuco, e o Superior Tribunal de Justiça. Foram realizadas observações nos locais de trabalho, bem como entrevistas com servidores e magistrados em cada tribunal. Também houve uma recolha de inquéritos com o público externo para complementar os dados

Atualmente o poder judiciário naquele país é um do três poderes que compõe o Estado Federativo. Existente desde o período colonial, é somente à partir da instalação da República que os tribunais passam a ter uma maior estruturação institucional e o poder judiciário ganha uma crescente autonomia em relação aos governos e ao próprio Estado (G. S. RIBEIRO, 2006). A Justiça chega ao final do século XX com um fazer tradicional, com excesso de trabalho, um processo judicial burocratizado, com formalidades e regras a serem seguidas, em um mundo em rápida transformação, onde a informação e a celeridade são as chaves para se adaptar à nova realidade tecnológica e complexa que se apresenta.

A implantação de estratégias de Tecnologias de Informação e Comunicação (TIC) pode, então, ser considerada como parte de um programa mas amplo de reforma e modernização da justiça brasileira (ANDRADE & JOIA, 2012b). Essas reformas no país, entretanto, acompanham um movimento global, iniciado nos anos 90 do século XX, o qual buscava, na modernização do sistema judicial, um meio de fortalecer a democracia e de resolver conflitos sociais (FABRI, 2007; VELICOGNA, 2008).

O uso das TIC pelo judiciário também relaciona-se com o desenvolvimento tecnológico de uma maneira mais ampla, e pode ser entendido dentro da perspectiva de emergência de uma Sociedade em Rede, ou Sociedade da Informação (CASTELLS, 2001b, 2001c).

Mi punto de partida, y no soy el único que lo asume, es que al final del siglo XX, hemos vivido uno de esos raros intervalos de la historia. Un intervalo caracterizado por la transformación de nuestra "cultura material" por obra de un nuevo paradigma tecnológico organizado en torno a las tecnologías de la información (CASTELLS, 2001a, pp. 59–60).

Para o autor, não se trata apenas de uma mudança tecnológica, mas de uma revolução que modifica a própria estrutura social, um novo paradigma sociotécnico, chamado paradigma

tecnológico. A *Internet* é o tecido das nossas vidas, diz Castells (2001c), onde a dimensão social da tecnologia da informação parece seguir a lei proposta por Melvin Kranzberg, para o qual a tecnologia não é boa nem má, tampouco neutra, é uma força que penetra no núcleo da vida e da mente, (CASTELLS, 2001a, pp. 109–110), e modifica a maneira de se comunicar e se relacionar dos indivíduos e destes com o Estado.

Assim, como consequência do que Castells chama de emergência do paradigma tecnológico, têm crescido, nas últimas décadas, o uso das tecnologias na produção da informação judicial. pelos tribunais Isso acabou melhorando a forma de relacionamento entre os cidadãos e a justiça, baseada em maior transparência e acesso à informação judicial (FABRI, 2011).

Para Velicogna (VELICOGNA, 2008) o uso das Tecnologias de Informação e Comunicação (TIC) é considerado um dos elementos-chave para a melhoria da administração da Justiça, permitindo que as cortes revejam suas funções e atividades:

ICT can be used to enhance efficiency, access, timeliness, transparency and accountability, thus helping judiciaries to provide adequate services. New possibilities are emerging for the integration and automation of court procedures and practices. In addition, the use of the Internet, can offer the chance to open the judiciary to the public providing both general and specific information on its activities, thereby increasing also legitimacy (VELICOGNA, 2008).

Os serviços disponíveis são vários e de diversas naturezas: bases de dados e ferramentas para pesquisa online de decisões e jurisprudência, acesso às informações processuais, consulta do histórico de tramitação, disponibilização das decisões às partes e advogados, petição eletrônica, e, mais recentemente, o próprio processo passou a existir apenas em formato digital (VELICOGNA, 2007).

Marco Fabri, ao analisar a implantação das tecnologias nos tribunais italianos diz:

L'utilizzo delle tecnologie può e deve dare un contributo importante per migliorare significativamente la nostra amministrazione della giustizia. La disponibilità di servizi web, lo scambio di documenti giudiziari per via telematica, la possibilità di avere on-line legge e giurisprudenza, l'elaborazione di dati statistici attendibili e immediati per il controllo di gestione, sono solo alcuni esempi delle opportunità che gli uffici giudiziari possono cogliere, seguendo anche gli esempi migliori disponibili in Europa, per ripensare radicalmente funzioni e attività svolte e riguadagnare

progressivamente una certa efficacia ed efficienza. (FABRI, 2011, p. 7).

Diversas iniciativas para implantar a tecnologia na geração, tramitação e uso da informação judicial foram feitas em diferentes tribunais do mundo nos últimos anos, em especial na última década. Tornou-se possível a disponibilização não só da movimentação processual como da íntegra das decisões, o recebimento de petições e outros documentos em forma eletrônica, tanto para os interessados, como para outros órgãos do sistema de justiça, e finalmente a própria geração e tramitação do processo judicial tornou-se eletrônica. Além das informações processuais, outras informações jurídicas puderam ser disponibilizadas na net, como bases de dados de jurisprudência e de decisões, atos normativos e súmulas de tribunais. (1)

Tais informações, sem o uso da tecnologia, acabam por ter seu uso restrito aos chamados operadores do direito, como juízes, advogados, serventuários de justiça e outros profissionais que conhecem os espaços de disponibilização das informações nos tribunais, como arquivos, bibliotecas e outras unidades de informação. O contexto de produção e disponibilização da informação acaba sendo determinante para os usos e acessos que esta terá. Ao ser disponibilizada na rede, novos usos e usuários poderão surgir. Para tanto a gestão dessa informação digital é tão necessária quanto a gestão da documentação em outros suportes.

O acesso aos acervos documentais pelo cidadão não é apenas um serviço da administração mas um direito dentro de uma sociedade democrática e transparente na perspectiva da sociedade em rede..

Nesse sentido, alerta F. Ribeiro, acerca dos acervos judiciais:

Um dos grandes desafios da sociedade digital é, com efeito, a compatibilização da gestão da informação nos contextos organizacionais, com a preservação da memória orgânica das entidades produtoras dessa mesma informação. Essa problemática, quando transferida para o caso concreto da informação judicial, coloca grandes desafios e implica que se equacionem diversas questões (F. RIBEIRO, 2010, p. 39).

Desse modo, a gestão da informação judicial, a partir da adoção de plataformas digitais, ganha novas perspectivas, tanto na produção, quanto no uso institucional e social (F. RIBEIRO, 2010), e pode ser vista pela abordagem da Ciência da Informação, enquanto fenômeno humano e social. (SILVA & RIBEIRO, 2011; SILVA, 2006);

A informação social está implicada no processo de gestão de qualquer entidade ou organização e, assim sendo, as práticas informacionais decorrem e articulam-se com as concepções e práticas dos gestores e atores e com a estrutura e cultura organizacionais (SILVA & RIBEIRO, 2011, p. 60).

Tal contexto de produção documental precisa, também, ser pensado de modo mais amplo, como espaço de representação social, pois reflete as práticas e discursos de um determinado grupo social. Pode-se pensar no processo judicial como o lugar de materialização de tais práticas e discursos e o Poder judiciário dentro daquilo que Pierre Bourdieu define como "Campo Jurídico". Para o autor, este é um campo de disputa simbólica de poder que pode ser definido como:

o lugar de concorrência pelo monopólio do direito de dizer o direito, quer dizer, a boa distribuição ou a boa ordem, na qual se defrontam agentes investidos de competência ao mesmo tempo social e técnica que consiste essencialmente na capacidade reconhecida de interpretar (de maneira mais ou menos livre ou autorizada) um corpus de textos que consagram a visão legítima, justa, do mundo social (BOURDIEU, 2004, p. 212).

Assim como as práticas estão contextualizadas em um ambiente organizacional próprio, a informação judicial não pode ser analisada apenas como resultado de uma atividade institucional, ele deve ser percebida de forma sistêmica, como um processo infocomunicacional complexo (SILVA, 2006). Articulam-se, assim, práticas, culturas e contextos institucionais, dotando a informação judicial de características inerentes à instituição que a produz. Nessa perspectiva torna-se relevante perceber as alterações e o impacto da adoção de novas plataformas tecnológicas pelos tribunais.

## 2. O contexto jurídico para o processo eletrônico

A mudança de suporte da informação foi possível devido à mudança legislativa que introduziu novas regras para a produção da documentação judicial com o uso de recursos digitais. A nova legislação permitiu que os tribunais de justiça fizessem uso exclusivo de meios eletrônicos para a tramitação dos processos judiciais, bem como para a comunicação de atos e transmissão de peças processuais, o que até então não era explicitamente previsto. A Lei n.º 11.419/06 (Brasil, 2006), além de uma das principais leis a tratar do tema, tinha por objetivo proporcionar maior celeridade no trâmite processual e, assim,

atingir maior efetividade do acesso à justiça, através de uma mudança no suporte do processo (MILBRADT, 2010).

A informatização de atos e procedimentos judiciais foi possível já a partir de 2001, de forma parcial, com a Lei 10.259, a qual permitiu o uso de sistemas eletrônicos para algumas situações específicas nos processos dos Juizados Especiais da Justiça Federal, como a intimação eletrônica de partes e advogados. O § 2º do art. 8º explicita: “Os tribunais poderão organizar serviço de intimação das partes e de recepção de petições por meio eletrônico” (Brasil, 2001)

Com isso, os primeiros sistemas de gestão eletrônica de documentos começaram a ser desenvolvidos ou ampliados, levando ao desenvolvimento de sistemas híbridos, onde os documentos eram gerados e, em alguns casos, assinados eletronicamente, mas impressos e anexados ao processo em papel. Um dos primeiros sistemas com tramitação exclusivamente eletrônica foi o sistema dos Juizados Especiais do Tribunal Regional Federal da 4ª Região, o e-Proc V1 (Justiça Federal do Rio Grande do Sul, 2015).

A partir de 2006, com a Lei 11.419, foi possível a desmaterialização e a tramitação exclusivamente eletrônica dos processos judiciais em todos os ramos de justiça e para todos os tipos de processo judicial no Brasil. Diversos tribunais começaram a desenvolver aplicações ou melhorar seus sistemas de gestão visando a chamada virtualização do processo (2). Com o tempo o Conselho Nacional de Justiça (CNJ) passou a coordenar os esforços de desenvolvimento tecnológico visando a construção de um sistema que pudesse atender as demandas de todos os ramos de justiça do país.

Diversas normas administrativas foram lançadas pelos diferentes tribunais, regulamentando internamente o uso dos recursos eletrônicos, a partir da nova legislação. Normas de uso nacional também foram editadas pelo CNJ. Destacam-se: a Resolução nº 91, de 29 de setembro de 2009, que institui o Modelo de Requisitos para Sistemas Informatizados de Gestão de Processos e Documentos do Poder Judiciário (MoReq-Jus) (Conselho Nacional de Justiça - CNJ, 2009); a Resolução Conjunta nº 3 de 1 de abril de 2013, que institui o Modelo Nacional de Interoperabilidade do Poder Judiciário e do Ministério Público (Conselho Nacional de Justiça - CNJ & Conselho Nacional do Ministério Público - CNMP, 2013); e a Resolução nº 185 de 18 de dezembro de 2013 que institui o Sistema Processo Judicial Eletrônico - PJe como sistema de processamento de informações e prática de

atos processuais e estabelece os parâmetros para sua implementação e funcionamento (Conselho Nacional de Justiça - CNJ, 2013).

### **3. O profissional da informação nos tribunais e o paradigma pós custodial**

O desenvolvimento dos sistemas de informação, especialmente os sistemas de processo eletrônicos, pelos tribunais tem sido considerados essenciais para o funcionamento da Justiça brasileira. Os investimentos feitos em tecnologia da Informação resultaram em melhorias na produtividade e eficiência de alguns tribunais (ANDRADE & JOIA, 2012a). A tecnologia da informação é uma das áreas estabelecidas no mapa estratégico do Poder Judiciário.

Percebe-se, entretanto, que, o desenvolvimento destes sistemas de informação é feito pelas áreas de tecnologia da informação e área fim, sem a participação dos profissionais de gestão da informação e documentos, ou responsáveis pelos acervos documentais. Nos tribunais pesquisados verificou-se que os gestores das áreas de acervo não participavam diretamente dos comitês de gestão dos sistemas.

Essa dissociação entre a área de informação do tribunal e a área de tecnologia pode ser compreendida como reflexo do papel que os arquivos e centros de documentação tiveram na gestão da informação institucional, dentro de um paradigma custodial que dominou as instituições arquivísticas nas décadas anteriores ao desenvolvimento tecnológico.

Responsáveis, essencialmente pela custódia e preservação da documentação, os arquivos e centros de documentação recebem a documentação que já tramitou pelas áreas fim e deve ser mantida para fins de consulta pelos prazos legais. Historicamente a função de gestão da documentação produzida e em tramitação não estava entre suas atribuições principais. Tais áreas historicamente sofreram com escassez de recursos materiais e de pessoal qualificado.

Analisar a legislação que regulava a área de arquivos pode jogar alguma luz sobre esta questão: a Lei nº 6.546, de 4 de julho de 1978 (Brasil, 1978) regulamenta as profissões de arquivista e técnico de arquivo. Essa lei representou, à época, um avanço na definição legal das atribuições dos arquivistas e no reconhecimento dos arquivos, enquanto órgãos responsáveis pela custódia e preservação dos acervos. Ao exigir curso superior em arquivologia (art. 1º, cc art. 5º), a lei confere à profissão uma distinção técnica, somente conseguida através da formação universitária (3).

É possível inferir uma intenção do legislador em profissionalizar os arquivos, garantindo assim, o correto tratamento dos documentos administrativos e históricos, e também adequar os arquivos nacionais às práticas profissionais já consolidadas no exterior, dentro das políticas de modernização do Estado, buscadas pelo regime militar em vigor naquele período.

No entanto, a promulgação de uma lei, por si não garante sua aplicação. Como ressalta Jardim (JARDIM, 2003, p. 38), a atuação dos arquivistas é que vai garantir a aplicação das normas:

(...) um marco legal só provoca impactos arquivísticos quando vai além de uma declaração de princípios conceituais bem estruturados, amparado em redefinições institucionais oportunas e promissoras. O desafio maior para as instituições arquivísticas, seus profissionais e a sociedade é a construção cotidiana da legislação no fazer arquivístico. O contrário disso pode redundar num inferno de boas intenções cujo epicentro é uma legislação tornada 'letra morta'.

As atribuições previstas na lei evidenciam uma forte tendência ao tecnicismo e à guarda dos documentos em arquivo, caracterizando as atribuições profissionais dentro do chamado paradigma custodial (SILVA & RIBEIRO, 2011). Percebe-se aqui como tal paradigma é fortalecido pela actuação dos profissionais de Arquivos, Bibliotecas e centros de documentação científica e técnica, no contexto do Estado Nação (e sua metamorfose, o Estado Cultural). Pode ser reconhecida, aqui, uma sobrevalorização da custódia, conservação e restauro, pela ênfase dada à memória e à cultura enquanto instrumentos de legitimação do Estado Nação através da documentação de arquivo.

Segundo Eliana Mattar (2003, pp. 20–21), a legislação arquivística existente até a promulgação da Constituição Federal de 1988 não era suficiente para a consolidação de uma política nacional que salvaguardasse a documentação pública, e garantisse o correto tratamento arquivístico dos documentos relevantes à memória nacional. Com a nova Carta

dá-se início à política nacional de arquivos públicos e privados, na qual a gestão tem papel fundamental por meio da aplicação racional de técnicas arquivísticas, visando à seleção para a eliminação de documentos sem valor e à guarda dos considerados de valor permanente. Evita-se, dessa forma, a preservação indiscriminada e custosa.

Sua implantação vem a ocorrer ao longo da década de 1990. Primeiros, com a Lei de Arquivos, depois com seus decretos regulamentados. Seguem-se as instruções do Arquivo Nacional (CONARQ), criada pela referida lei e presidido pelo diretor-geral daquele órgão (MATTAR, 2003, p. 22).

A chamada Lei de Arquivos, Lei nº 8.159/1991 (Brasil, 1991) trouxe conceitos e definições claras, evidenciando a função dos arquivos e quais as responsabilidades dos gestores. Utilizando termos e designações arquivísticas consolidadas, na prática profissional da chamada disciplina arquivística. O primeiro artigo já deixa clara a necessidade de tutela do Estado aos documentos de arquivo, para garantir sua proteção e uso (grifo nosso):

Art. 1º - É dever do Poder Público a gestão documental e a proteção especial a documentos de arquivos, como instrumento de apoio à *administração, à cultura, ao desenvolvimento científico* e como elementos de prova e informação.

Importante ressaltar que o acesso à documentação é determinado, tanto pelo artigo 5º, inciso XXXIII da CRFB/88, quanto pelos artigos 4º e 5º e pelo capítulo V da referida lei. Essa é uma mudança importante em relação à lei anteriormente analisada, onde o acesso aos documentos pelo cidadão não estava claramente garantido. No entanto, na prática, esse acesso ficou dificultado pela própria estrutura dos arquivos, onde poucos possuíam infraestrutura, mão-de-obra ou tecnologia para viabilizar o acesso rápido e amplo aos seus documentos. Essa situação pouco se modificou e até os dias atuais é bastante comum a necessidade de uma solicitação prévia e autorização do órgão para se ter acesso à documentação pública.

Os artigos 2º e 3º definem arquivos e gestão de documentos, respectivamente (grifo nosso):

Art. 2º - Consideram-se arquivos, para os fins desta lei, os conjuntos de documentos produzidos e recebidos por órgãos públicos, instituições de caráter público e entidades privadas, em decorrência do exercício de atividades específicas, bem como por pessoa física, qualquer que seja o suporte da informação ou a natureza dos documentos.

Art. 3º - Considera-se gestão de documentos o conjunto de procedimentos e operações técnicas à sua produção, tramitação, uso, avaliação e arquivamento em fase corrente e intermediária, visando a sua eliminação ou recolhimento para guarda permanente.

Pela leitura do artigo 2º, depreende-se algumas questões relevantes: o suporte da informação não justifica um tratamento diferenciado, assim, os documentos gerados ou armazenados em meio digital ou magnético também são passíveis de tratamento arquivístico, nos moldes da lei; a natureza, a tipologia ou a espécie documental também não influenciam no seu tratamento, uma vez que o que determina se o documento é arquivístico ou não é a sua função, ou seja, são produzidos e recebidos em decorrência das atividades de uma instituição ou pessoa, e não a sua forma.

A lei também prevê, com uma clara intenção preservacionista, a responsabilização para o caso de destruição da documentação em arquivo permanente, ou considerada de interesse público e social, para o caso de arquivos privados:

Art. 25 - Ficarà sujeito à responsabilidade penal, civil e administrativa, na forma da legislação em vigor, aquele que desfigurar ou destruir documentos de valor permanente ou considerado como de interesse público e social.

A lei separa, por fim, os documentos dos arquivos públicos em *correntes*, *intermediários* e *permanentes*, e considera os últimos inalienáveis e imprescritíveis. Essas denominações estão associadas à chamada teoria das três idades, desenvolvida ao longo do século XX, como uma forma de atender à gestão dos arquivos administrativos, e que consolidou-se como um dos pilares da disciplina arquivística (ROUSSEAU & COUTURE, 1998).

A separação dos arquivos significa também uma diferenciação de tratamentos. Os documentos gerados na administração são tratados com classificações e tabelas diferenciadas daquelas existentes no arquivo permanente. Durante sua fase corrente e intermediária estão à serviço das necessidades do órgão produtor. Quando são transferidos para o arquivo permanente, muda-se o seu *status*, e o seu tratamento será feito para o uso social, enquanto registo e patrimônio histórico. Na prática, o mesmo documento, com a mesma informação, passa por processos de gestão diferenciados, na tentativa de conciliar o uso institucional e administrativo, com a função social e histórica. Essa prática, adotada no Brasil e em vários outros países, é originária da tradição anglo-americana de separação entre o que ficou chamado de *records management* e *archives*. Assim, de uma mesma entidade produtora, teremos documentos, às vezes da mesma classe, mas com tratamentos diferenciados, dependendo em que arquivo se encontra.

Como ressalta Fernanda Ribeiro, falando do contexto português, essa separação entre arquivos históricos e arquivos correntes, além de dicotômica, é antinatural e faz com que se perca a perspectiva do conjunto da documentação.

O divórcio entre as entidades produtoras de informação nos serviços públicos e a documentação mais antigas dessas mesmas entidades, incorporada em arquivos históricos (Torre do Tombo e arquivos distritais), fez com que se separassem artificialmente partes de um todo e que as mesmas fossem organizadas e disponibilizadas em moldes diferenciados, perdendo-se a unidade sistêmica entre elas (F. RIBEIRO, 2010, p. 35).

O capítulo IV da Lei de Arquivos trata da organização e administração das instituições arquivísticas públicas. Fica vedado ao Arquivo Nacional a guarda e o tratamento de documentos que não aqueles do Executivo Federal. Assim, sem um destino externo para guarda e conservação, os processos judiciais acabaram sendo manidos nos arquivos e áreas de documentação de cada tribunal.

Assim, no caso brasileiro não há um arquivo histórico centralizado para a documentação judicial. Os arquivos do Judiciário Federal são geridos pelo próprio poder, e os arquivos estaduais tem normativo próprio sobre a vinculação de seus arquivos aos arquivos públicos estaduais. Entretanto, a falta de uma regulamentação ao longo do século XX fez com que diversos conjuntos documentais ficassem apartados de suas instituições originárias. Em muitos tribunais brasileiros, a documentação das últimas décadas está no arquivo do Tribunal, no entanto a documentação antiga foi encaminhada aos arquivos públicos e históricos dos estados. O Arquivo Nacional também possui parte da documentação do Supremo Tribunal Federal, e conjuntos de documentos do Superior Tribunal de Justiça e alguns outros conjuntos documentais de tribunais e fóruns extintos.

A tradição do fazer arquivístico nos arquivos judiciais foi, desse modo, orientada mais à guarda e preservação da sua documentação do que voltada à gestão integrada e ao acesso. Por outro lado o processo judicial em si, enquanto resultado das atividades fins dos tribunais, possui características diferenciadas da documentação administrativa, seguindo normas próprias de tramitação.

Assim, com áreas documentais voltadas para a guarda, preservação e acesso dos processos em fase intermediária e permanente, os profissionais da informação pouco estavam envolvi-

dos com a produção documental e gestão da informação em fase corrente.

Essa ausência de arquivistas e profissionais da área da ciência da informação no planejamento e desenvolvimento dos sistemas de informação pode gerar dificuldades futuras na gestão, acesso e preservação da documentação judicial, uma vez que a legislação arquivística não diferencia o suporte da informação. Assim, as normas e procedimentos hoje realizados pelas áreas de acervo, deverão ser também realizadas nos processos mantidos em ambientes digitais.

#### **4. A complexidade da documentação judicial: da produção à preservação**

Pode-se perceber o processo como um documento individual, ou, pensando a partir da perspectiva do fluxo da informação, como um conjunto complexo de documentos. Nesse sentido, os documentos que formam o processo são reunidos em sequência durante seu “processamento”, de acordo com leis e normas próprias de funcionamento, as quais garantem a validade jurídica e a autenticidade do documento como um todo. A forma do processo judicial brasileiro está regrada nos códigos de processo (Código de Processo Civil e Código de Processo Penal). No entanto, diversas outras normas são associadas ao fazer da justiça: desde preceitos constitucionais, passando pelas normas legais complementares, até as normas supralegais, que são decisões e regulamentos das diversas instâncias administrativas dentro do próprio Judiciário (Súmulas, Regimentos, Normativos de corregedorias, etc.).

O processo judicial, ao ser produzido, torna-se registro (documento) e também a própria expressão da Justiça. Nesse sentido pode ser percebido como “o” documento judicial em si. No entanto é múltiplo, e pode apresentar-se com característica as mais diversas, com fazeres diferenciados, e formas distintas, sendo assim uno e diverso ao mesmo tempo. Pode ser classificado por diferentes critérios. Ao ser classificado de acordo com a classe, será separado por matéria (cível, criminal) e tipologia (rito sumário, ordinário, recurso, etc.). A classificação por assunto é complementar à de classe, e extremamente diversificada (Direito Administrativo, de Família, Ambiental, etc.) (4). É importante destacar que só recentemente foram criados padrões nacionais de classificação processual e que estes padrões não estão totalmente implementados.

O assunto, valor da ação, ou tipo de crime podem determinar o órgão julgador e o rito pro-

cessual (especializações, juizados, conciliações). A instância julgadora também pode ser determinada pelas partes envolvidas (foro especial), ou pelo tipo de recurso. O processo pode ter objetivos distintos (ação cautelar, ação de conhecimento, ação executória), ou tramitar em diferentes instâncias (1º. Grau, 2º. Grau, tribunais superiores), pode conter processos apenas ou acessórios (demandas paralelas, ou questões incidentais que correm em autos apartados, ou são julgadas separadamente em instâncias diferentes), e, ainda, mudar de classe, de assunto, ou até de órgão julgador ao longo da sua tramitação.

A análise acima demonstra o grau de complexidade da produção do documento, o qual assume as mais diversas formas e tamanho ao longo de seu processamento, transitando por diferentes instituições e instâncias. A implementação de soluções informatizadas passa por estabelecer diferentes registros e fluxos de informação para dar conta de toda a diversidade do processo. É preciso, também, que sejam preservadas as características de autenticidade e fidedignidade do documento, garantindo o valor probatório do documento, preservando, assim, os direitos ali constituídos.

Nos tribunais pesquisados percebe-se, inicialmente, a complexidade da gestão e da padronização dos procedimentos cartorários que levam à produção documental. Cada unidade ou juízo possui regras e entendimentos da norma legal, o que determina os procedimentos, prazos e formas de tramitação de um processo. A produção ainda é pouco automatizada pelos sistemas existentes, levando a muitos casos de retrabalho. Também há diferentes sistemas de informação em uso, como os sistemas de geração de correspondência, sistemas de controle de documentos e protocolo, etc. Percebe-se uma dificuldade de integração das informações entre o sistemas.

As regras para consulta aos processos judiciais também são bastante variadas. Em alguns tribunais as partes podem aceder ao sistema e consultar apenas partes do processo, enquanto advogados, mesmo que não diretamente envolvidos com o processo podem aceder à integralidade dos documentos. As citações de advogados também podem ocorrer de formas distintas, tanto por meio eletrônico como por correspondência impressa enviada pelos correios.

A integração dos sistemas existentes com outros órgãos do sistema de justiça ainda é incipiente e pouco automatizada. Em alguns dos tribunais pesquisados o sistema utilizado pela primeira instância e pelo tribunal recursal não

atuavam de forma sincronizada, exigindo migração de dados não automatizadas. A comunicação com tribunais de outras esferas (superiores ou de outros ramos de justiça) ou com outros órgãos (como polícias e Ministério Público) ainda é incipiente ou inexistente.

Preocupa, ainda a falta de políticas e ações desenvolvidas para garantir a gestão da documentação em fases de arquivamento intermediário e permanente.

Tais dificuldades são sentidas também pelas instituições e usuários dos sistemas. Como há um constante desenvolvimento destes, muitas das questões aqui levantadas estão aguardando desenvolvimento pelo tribunais, e podem ainda ser contornadas.

## 5. A Informação digital e sua preservação

A preservação digital é tema de crescente preocupação entre arquivistas e profissionais da informação e das tecnologias. A migração para o formato digital exige não só uma política constante de atualização tecnológica, como também uma eficaz gestão da informação, de forma a identificar a o que será preservado, quais os prazos e quais as ações a serem submetidas a cada tipo de informação e suporte.

A discussão sobre a preservação do processo judicial electrónico está inserida no contexto mais amplo da preservação da informação em suporte electrónico. Uma análise mais aprofundada sobre esse tema pode ajudar a compreender as dificuldades enfrentadas atualmente pelos tribunais brasileiros.

Como reflexo das políticas de governo aberto e governo electrónico, o *e-justice* surge para o judiciário não só como um caminho para ampliar o acesso à informação, como também um caminho para solucionar os entraves e a morosidade dos sistemas de justiça, tradicionalmente espaços conservadores e burocráticos. A mudança para o meio digital facilita o acesso e uso da informação, entretanto o documento digital apresenta novas dificuldades envolvendo a guarda e a preservação do documento electrónico.

No Brasil, com a implantação do processo eletrônico em todos os ramos do poder judiciário, o assunto torna-se de extrema importância para diversos órgãos. O Conselho Nacional de Justiça CNJ, órgão criado com a função de controle administrativo do sistema de justiça brasileiro, lançou oficialmente o chamado PJe, ou Processo Judicial Eletrônico em 21 de junho de 2011. O sistema foi desenvolvido pelo próprio Conselho, com colaboração de outros tribunais. O

projeto foi idealizado para servir à todos os ramos de justiça, tornando-se um sistema único para toda a justiça brasileira. O material disponibilizado no *website* (do CNJ - Conselho Nacional de Justiça) deixa clara a preocupação da instituição em reduzir despesas do Judiciário, em geral, com aquisições de *softwares*, com a segurança e com a interoperabilidade. No entanto, uma análise mais detalhada faz notar que não há menção à preservação da informação digital ali produzida. Há algumas tabelas de padronização de metadados como classes, assuntos, agentes, *etc.*, no entanto elas parecem ter sido pensadas em termos de interoperabilidade mais do que para preservação.

Atualmente vários tribunais já adotaram o PJe, e muitos estão trabalhando na integração dos sistemas legados. Os outros sistemas em uso pelos tribunais participantes da pesquisa foram, em geral, desenvolvidos para funcionalidades específicas e acabaram sendo ampliados para atender às necessidades de tramitação electrónica. Alguns permitem a captura, o processamento e a disponibilização de documentos em meio eletrônico. Em muitos casos, o sistema atua em uma fase específica da tramitação processual, ou em apenas uma das instâncias, e não em todo o curso do processo até a decisão final. Sabe-se que os tribunais superiores ainda não finalizaram a implantação do sistema PJe, recebendo os processos recursais em sistemas próprios.

Se, por um lado, os sistemas de gestão foram desenvolvidos para cumprir demandas da fase de tramitação e não propriamente para as fases de arquivamento e para a preservação digital, por outro, algumas ações do CNJ foram pensadas especificamente para garantir o acesso continuado da documentação judicial, como o desenvolvimento do já citado Moreq-Jus e o Programa Nacional de Gestão Documental e Memória do Poder Judiciário – Proname.

O Moreq-Jus (Conselho Nacional de Justiça - CNJ, 2009) teve origem em modelos de requisitos desenvolvidos no Brasil e na Europa, como o modelo do Conselho da Justiça Federal (CJF, 2007), o modelo do Conselho Nacional de Arquivos (CONARQ - Conselho Nacional de Arquivos), o modelo português da Torre do Tombo e as versões da Comissão Europeia que o originaram, o MoReq e o MoReq2, além do modelo de metadados PREMIS, versão 2008.

Assim, além de um conjunto de requisitos para o desenvolvimento dos sistemas informatizados, o Modelo prevê também a utilização de metadados de segurança, de auditoria e de preservação, cujos elementos são apresentados como

anexo ao modelo de requisitos. Os capítulos correspondem, de maneira geral, aos diversos âmbitos que devem compor um sistema de informação: classificação e manutenção, captura, armazenamento, preservação, segurança, tramitação e fluxo de trabalho, avaliação e destinação, pesquisa, localização e apresentação de documentos, funções administrativas, usabilidade, interoperabilidade, disponibilidade, desempenho e escalabilidade.

A organização do texto e os conceitos utilizados no modelo deixam clara a importância da gestão documental e dos conceitos arquivísticos para a elaboração do modelo. Também fica evidente a adoção de conceitos vindos dos estudos sobre preservação digital, uma vez que esta é separada em física, lógica e conceitual, sendo citada, na introdução, a conhecida Carta sobre a Preservação do Patrimônio Digital da UNESCO.

A análise do Modelo de requisitos permite inferir que o documento foi criado como ferramenta essencial aos desenvolvedores de sistemas de informação para o Poder Judiciário. Há ali uma preocupação com a preservação dos documentos digitais. No entanto, o modelo trata o tema ainda de maneira incipiente. Por outro lado, considerando as grandes disparidades tecnológicas existentes entre os diferentes tribunais, e a natureza da documentação judicial para a memória social do país, a efetiva implementação do conjunto de requisitos apresentados no modelo, representa um importante avanço para a preservação destes documentos.

Uma outra ação de preservação da memória do judiciário é o desenvolvimento, no CNJ, do Programa Nacional de Gestão Documental e Memória do Poder Judiciário – PRONAME (Conselho Nacional de Justiça - CNJ, 2015a). Como o nome indica, trata-se de um programa amplo para a gestão da documentação produzida pelos diversos órgãos do Poder Judiciário, com representantes dos diferentes ramos de justiça.

Em 2011 foi lançada pelo órgão a recomendação n.º 37 (posteriormente alterada pela Recomendação 4/2013), com um manual para a gestão dos documentos, incluindo classificação, avaliação e destinação dos documentos, além dos instrumentos de gestão, como tabelas de classificação e temporalidade. Sabe-se que há a preocupação do CNJ com a preservação da documentação digital, e também com a disponibilização em repositórios digitais da documentação em papel de valor histórico, no entanto, ainda não há nada divulgado pelo grupo nesse sentido.

Considerando as características do documento judicial, não basta a preservação apenas da informação. Para cumprir suas funções, o processo judicial digital precisa ser preservado de forma a não pôr em risco sua autenticidade, integridade, organicidade, e em alguns casos a confidencialidade. Muitas das soluções desenvolvidas a partir das preocupações com acervos arquivísticos buscam preservar a autenticidade do documento em meio digital. Entretanto, em muitos casos, parte-se de contextos onde o repositório é o ambiente de custódia de documentos produzidos em outras instituições, dentro da perspectiva arquivística do ciclo vital dos documentos.

Para o caso do Poder Judiciário brasileiro, isso não se verifica, pois, como anteriormente mencionado, não há um arquivo central que recolha a documentação histórica judicial. Se, por um lado, isso permite uma gestão “integrada” dos documentos eletrônicos, desde a produção até o arquivamento, por outro, é preciso uma certa uniformidade e a interoperabilidade dos sistemas dos diversos tribunais de modo a facilitar a consulta e o acesso da população a essa documentação. Nesse sentido, a adoção, pelos diferentes tribunais, de um único sistema garantiria maior uniformidade e a possibilidade de uma política integrada para a gestão e preservação dos documentos judiciais. Entretanto, as bases para uma política integrada de preservação digital ainda precisam ser construídas.

A literatura sobre preservação digital é vasta, apesar de relativamente recente. Pode-se dizer que o tema torna-se mais complexo pela natureza do próprio objeto de estudo, a tecnologia, em constante desenvolvimento. Os diversos objetivos envolvendo a preservação digital produzem diferentes soluções tecnológicas.

De maneira geral pode-se dizer que a preservação digital envolve ações e políticas que permitam o acesso à informação produzida em ambientes digitais, isso é, computadores. Uma definição bastante conhecida, redigida para a UNESCO no documento *Guidelines for the Preservation of the Digital Heritage*, diz que a preservação digital abrange um conjunto de atividades que garantam acesso continuado de longo prazo à informação e ao patrimônio cultural em formatos digitais, e envolve a manutenção de aspectos físicos, lógicos e conceituais do objeto digital (United Nations Educational Scientific and Cultural Organization - UNESCO, 2003).

Entretanto, o termo “informação digital”, por sua vez, tem recebido diversas nomenclaturas como documento digital, objeto digital, arquivo digital,

documento electrónico, registo digital, etc. O *Dicionário Brasileiro de Terminologia Arquivística* apresenta documento digital como o documento codificado em dígitos binários, acessíveis por meio de sistema computacional (Arquivo Nacional do Brasil, 2005) e é considerado sinónimo de arquivo digital. Também apresenta uma separação entre documento digital e documento electrónico, sendo o segundo o gênero do primeiro, onde também se incluem outros documentos acessíveis por equipamentos electrónicos, como cartões perfurados.

Por outro lado, a partir dos estudos sobre meta-informação surge a noção de *objeto digital*, que seria rapidamente adotada pelas comunidades de preservação digital (5). O grupo de trabalho *Preservation Metadata: Implementation Strategies – PREMIS*, o define como “*a discrete unit of information in digital form*”, composto por três diferentes elementos:

O *file* (ficheiro – é uma sequência ordenada de bytes com um nome e reconhecida por um sistema operativo), o *bitstream* (é um conjunto contíguo ou não contíguo de bits dentro de um ficheiro ex. um gráfico dentro de um ficheiro Word) e a *representation* (representação – conjunto de *files*, necessários para a apresentação de uma entidade intelectual) (PINTO, 2009).

K. Thibodeau, em conhecida publicação de 2002, propõe uma definição mais sintética e, ao mesmo tempo, mais completa e abrangente: “*a digital object is an information object, of any type of information or any format, that is expressed in digital form*”. E complementa dizendo que:

All digital objects are entities with multiple inheritance; that is, the properties of any digital object are inherited from three classes. Every digital object is a physical object, a logical object, and a conceptual object, and its properties at each of those levels can be significantly different. A physical object is simply an inscription of signs on some physical medium. A logical object is an object that is recognized and processed by software. The conceptual object is the object as it is recognized and understood by a person, or in some cases recognized and processed by a computer application capable of executing business transactions (THIBODEAU, 2002, p. 6).

Essa definição é então ampliada pela Biblioteca Nacional da Austrália e a UNESCO, que incorpora o que ficou conhecido como dimensão essencial (*essential elements*): *as bundles of essential elements that embody the message, purpose, or features for which the material was chosen for preservation* (United Nations

Educational Scientific and Cultural Organization - UNESCO, 2003). O termo *objeto digital* acaba por se sobrepor às definições anteriores já que permite uma melhor análise das diferentes características da informação no ambiente digital. Considerando a diversidade e complexidade desse objeto, Pinto, em estudo sobre a preservação digital em Portugal, propõe que o uso do termo *documento digital* se restrinja às variantes digitais dos formatos tradicionais de documentos analógicos (PINTO, 2009).

No entanto, é possível perceber que, seja um documento originariamente digital, ou seja ele digitalizado, o conteúdo e o suporte se apresentam como duas componentes distintas desse documento, e não mais uma unidade indissociável como se verifica para o documento em papel. Portanto, é preciso repensar o que se busca preservar.

A realidade digital relativiza efetivamente o conceito de documento como o concebemos na realidade analógica e torna-o ‘um’ entre os diversos ‘tipos’ de informação em ambiente digital, o que faz colocar, definitivamente, a informação no centro das atenções, apesar da necessidade de lhe associar o termo digital reforçando, no nosso ponto de vista, a peculiaridade dos atributos e a pluridimensionalidade das unidades de informação criadas em ambiente digital (PINTO, 2009).

Há que se considerar também que o ambiente digital produz informações que não podem ser expressas nos formatos analógicos tradicionais de documentos, como as bases de dados dinâmicas, por exemplo. Assim, perceber a multidimensionalidade do objeto digital torna mais evidente a complexidade envolvida na sua preservação.

Se pode, dizer, então, que a preservação do processo judicial eletrônico passa por assegurar que a informação produzida em ambiente digital continuará acessível nas suas dimensões física, lógica, conceitual e essencial. Mas como? Para preservar os processos em papel não é preciso muito, basta preservar o suporte e a informação se mantém acessível. Mesmo com políticas de preservação negligentes ou inexistentes a documentação em papel pode levar séculos para deixar de existir. O objeto eletrônico, por sua vez, pode desaparecer em instantes, basta um clique.

As estratégias para se evitar uma “amnésia” digital vão desde a preservação de equipamentos até a recriação de uma “pedra de roseta” tecnológica, com a preservação de diferentes linguagens e tecnologias. Os estudos atuais não permitem dizer que há um método totalmente confiável que resolva definitivamente a questão,

ao contrário, apontam vários caminhos possíveis, com vantagens e desvantagens. Por outro lado, as pesquisas nessa área são cada vez mais interdisciplinares, e reunindo grupos de pesquisas nacionais ou de grande abrangência, já que esta temática desperta interesse de diferentes sectores da sociedade.

Nesse sentido é preciso destacar os trabalhos da Comissão Europeia, da Torre do Tombo e do Arquivo Nacional brasileiro no desenvolvimento de modelos de requisitos para sistemas de informação e gestão de documentos eletrônicos (MOREQ). Estes trabalhos, provenientes de instituições arquivísticas, tinham por objetivo definir os requisitos para os sistemas de gestão de documentos em ambientes digitais, e serviram de modelo para o desenvolvimento do Moreq-Jus, no âmbito do Poder Judiciário brasileiro. Originalmente desenvolvido pelo CJF, esta foi uma das primeiras ações para a preservação da informação judicial gerada em sistemas eletrônicos.

Para além dos requisitos funcionais dos sistemas, foi preciso também desenvolver esquemas de metadados, a fim de garantir a preservação das informações sobre o próprio documento/informação. Modelos como o Moreq 2007, programas para o desenvolvimento de *e-government* e o *PREMIS Data Dictionary for Preservation Metadata* merecem destaque.

Os trabalhos para construção de normas e padrões de preservação, igualmente, permitiram o desenvolvimento de diversos projetos em escala global. Citamos entre outros o modelo OAIS – *Open Archival Information System*, que desenvolveu um padrão de componentes funcionais destinados ao funcionamento de um sistema de preservação da informação digital. Este modelo originou a norma ISO 14721:2003 (FERREIRA, 2006, p. 27).

Outras diversas iniciativas surgiram no âmbito de bibliotecas e arquivos nacionais, como os projetos da British Library, da National Archives of the Netherlands, da National Library of Australia e do National Archives of Australia, do Library and Archives Canada, da New Zealand National Library, e da Library of Congress nos Estados Unidos (6).

Entre os projetos desenvolvidos por universidades, cita-se o Projeto InterPARES, desenvolvido na University of British Columbia, no Canadá, com equipes de diversos países do mundo trabalhando em um modelo teórico-prático para a preservação de longa duração de documentos digitais autênticos. Na Universidade da Virginia, nos Estados Unidos, um outro projeto busca desenvolver um repositório *open source*, voltado

para armazenamento do conhecimento científico e demais objetos digitais.

Para a equipe do projeto InterPARES, o desenvolvimento de novas soluções foi buscado através de uma nova abordagem em conceitos da antiga diplomática, buscando garantir a autenticidade dos documentos transferidos ao arquivo. As dificuldades tecnológicas para a preservação do documento digital, dentro dos pressupostos arquivísticos do ciclo vital dos documentos geraram a polêmica conhecida como custódia e pós-custódia, ou o que ficou conhecido como teoria da arquivística integrada e o conceito de documento contínuo, desenvolvido na Austrália. Para Erlandsson (*apud* RONDINELLI, 2005, p. 72) o modelo de documentos contínuos

rejeita o modelo tradicional de ciclo vital dos documentos, o qual incorpora uma divisão rígida entre documentos arquivísticos correntes e permanente. O objetivo do modelo é promover regimes de gerenciamento integrado de documentos e processos arquivísticos (...) O ciclo vital foi recontextualizado de tal maneira que considerações históricas podem agora ser acomodadas desde o momento em que os documentos foram criados.

Atualmente essa polêmica foi, de algum modo, superada, mas fica a percepção de que o documento eletrônico não pode ser pensado sob os mesmos conceitos que eram válidos para o documento em papel. A preservação do documento digital deve ser pensada desde a sua criação, sob pena de suas características serem adulteradas ao longo do que se chamava ciclo de vida.

Além dos projetos citados acima, muitos outros, de menor escala estão em desenvolvimento. Não é objetivo deste trabalho fazer um levantamento completo, interessa, entretanto, demonstrar a variedade de projetos e objetivos nos estudos para a preservação digital (7).

Percebe-se que os diferentes objetivos com a preservação digital relacionam-se com os diferentes resultados buscados por cada grupo envolvido. Por exemplo, a gestão dos arquivos digitais busca a preservação de longo prazo de documentos correntes com valor permanente, de forma a garantir sua autenticidade, já as preocupações das bibliotecas estão voltadas para o acesso à informação, enquanto fonte de conhecimento. Essas diferenças de perspectiva resultam em estratégias diferentes de preservação.

Não cabe aqui detalhar cada uma dessas estratégias (8), mas é preciso ressaltar que, apesar

das diversas soluções possíveis, os debates têm ficado entre duas: emulação vs. migração.

Pode-se dizer que estas duas alternativas conduzem a dois caminhos distintos e, de certo modo, agregam muitas das demais soluções. Conforme Ferreira (FERREIRA, 2006, pp. 63–64)

tem-se assistido a uma batalha ideológica entre aqueles que defendem estratégias baseadas em migração e os que defendem a utilização de estratégias de emulação. Esta discussão tem na sua origem, questões relacionadas com a autenticidade e integridade dos materiais. Nas estratégias derivadas da migração, assume-se que os objetos digitais irão sofrer modificações ao longo do tempo. Determinadas migrações poderão mesmo originar perdas substanciais de informação. Para os defensores da emulação, assumir de antemão que a informação que se procura preservar será sistematicamente adulterada ao longo do tempo viola os pressupostos mais elementares da preservação.

Por outro lado, o autor acredita que este debate tende a se esgotar na medida em que cresce o sentimento de que as estratégias de preservação devem ser implementadas de acordo com o contexto específico de cada organização e do objeto que se quer preservar:

A seleção de estratégias de preservação deve ter em conta diversos factores como as características intrínsecas dos objetos, o custo de implementação e manutenção, os interesses do arquivo e da sua comunidade de interesse, entre outros. Para diversos autores este último ponto é de extrema importância. A informação terá pouca utilidade se não for preservada e disseminada de acordo com as necessidades da sua comunidade de interesse (FERREIRA, 2006, p. 65).

Assim, ao invés de buscar uma única solução, trata-se, antes de tudo, de avaliar as características do material e os objetivos da preservação, para selecionar políticas possíveis para o caso concreto, considerando as diferentes camadas de preservação do documento: física, lógica, e conceitual.

As diferentes estratégias deverão ser implantadas tendo como base um sistema de arquivamento digital compatível, que garanta os objetivos da preservação. Atualmente existem diversas plataformas capazes de abrigar a informação digital, e outras tantas estão em desenvolvimento para atender os requisitos técnicos da preservação digital (DSpace, Fedora, ICA-AtOM, por exemplo), mas não se trata apenas de defi-

nir qual o *software* que será utilizado. É preciso pensar uma política ampla que atenda as peculiaridades da documentação judicial, desde a produção até a fase de arquivamento e disponibilização dessa informação à sociedade.

## 6. Considerações finais

O desenvolvimento da *Internet* e do que vem sendo chamado de *Sociedade em Rede* modificou as relações da sociedade com as instituições públicas, isso acabou por chegar também aos órgãos de justiça. Acompanhando um movimento internacional de modernização, a justiça brasileira ampliou seu investimento em tecnologia como parte de uma ampla reforma buscando maior produtividade, transparência, e, principalmente, celeridade na resolução dos conflitos por ela mediados.

A adoção de sistemas informatizados permitia não só a melhoria na administração judiciária, como uma melhor gestão da informação. Assim foram desenvolvidos sistemas para gestão e controle dos documentos administrativos e judiciais, com sistemas de metadados e até de produção electrónica dos documentos. Com o surgimento da *web* e das tecnologias associadas, tornou-se possível a tramitação de um processo totalmente virtual, sem a impressão dos documentos em papel.

A total desmaterialização dos processos judiciais foi possível tanto pelo desenvolvimento de sistemas com assinatura electrónica e ambientes seguros na rede, como pela flexibilização da legislação e publicação de um conjunto de atos normativos que permitiam essa mudança de suporte.

Diversos tribunais passaram a desenvolver internamente ou compraram soluções para a produção e tramitação electrónica de documentos. Surge assim o chamado processo electrónico, que se distingue do processo em papel apenas pelo suporte onde a informação é mantida.

O processo judicial, é entretanto, composto por uma série de documentos que representam as ações legais realizadas pelos diferentes atores do processo legal. Essa documentação torna-se um conjunto complexo de dados e metadados em um sistema. Além de diferentes atores internos e externos ao tribunal de origem, esta documentação percorre diferentes órgãos do sistema de justiça. Essas passagens para diferentes instâncias processuais, quando realizadas no ambiente digital exigem a migração dos dados para diferentes sistemas, já que não há uma efetiva interoperabilidade entre os sistemas dos diversos órgãos envolvidos.

Verificou-se, ainda, uma dissociação entre os procedimentos de gestão documental nos processos em papel para os processos produzidos em ambiente digital. A tradição custodial das áreas de documentação e acervo fez com que as políticas de gestão da informação e dos documentos em fase corrente fossem desenvolvidas por usuários do sistema e profissionais da tecnologia da informação, e não por profissionais da ciência da informação.

Com acervos documentais gigantescos em papel ainda por serem adequadamente organizados e disponibilizados, os tribunais acumulam uma nova massa documental a ser tratada, agora em suporte eletrônico, com todos os riscos que a tecnologia apresenta. O desenvolvimento de políticas de gestão e preservação dessa documentação ainda são incipientes.

O desenvolvimento de algumas políticas estratégicas por parte do Conselho Nacional de Justiça iluminam esse quadro e permitem acreditar em uma melhoria dessa situação para os próximos anos. O desenvolvimento de um sistema único, ou de uma política de integração e interoperabilidade entre os sistemas legados poderá trazer mais segurança na tramitação do processo entre os diferentes órgãos e instâncias. A melhoria das políticas de gestão e preservação da documentação também são necessárias, assim como o desenvolvimento tecnológico, para garantir o acesso e a autenticidade dos documentos.

Sugere-se, ainda, o desenvolvimento de um modelo unificado para a gestão da informação, permitindo um acesso amplo e facilitado à documentação, de forma a garantir a transparência e o controle das instituições pela sociedade.

## Notas

- (1) Para um histórico do processo de implantação da tecnologia nos tribunais brasileiros ver LÖW (2012).
- (2) Conforme depoimento do Juiz do Trabalho e Conselheiro do CNJ Rubens Curado, para esta pesquisa, em 04/12/2014.
- (3) Ressalvados os casos mencionados no art. 1º, III, para aqueles que possuíam tempo de atividade profissional na época da vigência da Lei.
- (4) Para uma visão mais completa das classes e assuntos, ver o Sistema de Gestão de Tabelas Processuais Unificadas do Conselho Nacional de Justiça (2015b).
- (5) Para uma revisão mais abrangente sobre conceitos e definições ver PINTO (2009, pp. 119–133).
- (6) Para um amplo levantamento das pesquisas e iniciativas sobre preservação de documentos e informação digital, em âmbito nacional e internacional, ver publicações de Manuela PINTO (2009).
- (7) Um levantamento mais abrangente pode ser encontrado em MCDONALD (2006).
- (8) Para uma revisão sobre as diferenças técnicas de cada uma das estratégias de preservação digital ver FERREIRA (2006).

## Referências

- ANDRADE, A., & JOIA, L. A. (2012a). Investimento em Tecnologia da Informação e Eficiência da Justiça Brasileira. Em AMCIS 2012 Proceedings. Obtido de <http://aisel.aisnet.org/amcis2012/proceedings/SpanishPortuguese/2/>
- ANDRADE, A., & JOIA, L. A. (2012b). Organizational structure and ICT strategies in the Brazilian Judiciary System. *Government Information Quarterly*, 29(SUPPL. 1), S32–S42. doi:10.1016/j.giq.2011.08.003
- Arquivo Nacional do Brasil. (2005). Dicionário brasileiro de terminologia arquivística (No. 51). Rio de Janeiro.
- BOURDIEU, P. (2004). *O Poder Simbólico*. Lisboa: DIFEL.
- Brasil. Lei N. 6.546 de 4 de julho de 1978. Dispõe sobre a regulamentação das profissões de Arquivista e de Técnico de Arquivo, e dá outras providências. (1978).
- Brasil. Lei n. 8.159, de 8 de janeiro de 1991. Dispõe sobre a política nacional de arquivos públicos e privados e dá outras providências. (1991). Obtido de [http://www.planalto.gov.br/ccivil\\_03/leis/L8159.htm](http://www.planalto.gov.br/ccivil_03/leis/L8159.htm)
- Brasil. Lei N. 10.259, de 12 de julho de 2001. Dispõe sobre a instituição dos Juizados Especiais Cíveis e Criminais no âmbito da Justiça Federal. (2001). Obtido de [http://www.planalto.gov.br/ccivil\\_03/leis/LEIS\\_2001/L10259.htm](http://www.planalto.gov.br/ccivil_03/leis/LEIS_2001/L10259.htm)
- Brasil. Lei Federal 11.419 de 19 de dezembro de 2006. Dispõe sobre a informatização do processo judicial; altera a Lei no 5.869, de 11 de janeiro de 1973 – Código de Processo Civil; e dá outras providências. (2006). Obtido de [http://www.planalto.gov.br/ccivil\\_03/\\_ato2004-2006/2006/lei/l11419.htm](http://www.planalto.gov.br/ccivil_03/_ato2004-2006/2006/lei/l11419.htm)
- CASTELLS, M. (2001a). *La era de la Información: fin del milenio (Vol.3) (Vol. 3)*. Madrid: Alianza Editorial.
- CASTELLS, M. (2001b). *La era de la Información: la sociedad red (Vol.1) (Vol. 1)*. Madrid: Alianza Editorial.
- CASTELLS, M. (2001c). *The Internet Galaxy: Reflections on the Internet, Business, and Society*. Oxford: Oxford University Press.
- Conselho Nacional de Justiça - CNJ. Resolução N. 91 de 29 de setembro de 2009. Institui o Modelo de Requisitos para Sistemas Informatizados de Gestão de Processos e Documentos do Poder Judiciário e disciplina a obrigatoriedade da sua utilização no desenvolvimento e manutenção de sistemas in (2009). Obtido de <http://www.cnj.jus.br/atos-normativos?documento=78>
- Conselho Nacional de Justiça - CNJ. Resolução N. 185 de 18 de dezembro de 2013. Institui o Sistema Processo Judicial Eletrônico - PJe como sistema de processamento de informações e prática de atos processuais e estabelece os parâmetros para sua implementação e funcionamento. (2013). Obtido de <http://www.cnj.jus.br/busca-atos-adm?documento=2492>
- Conselho Nacional de Justiça - CNJ. (2015a). Programa Nacional de Gestão Documental e Memória do Poder Judiciário (Proname). Obtido 21 de Abril de 2015, de <http://www.cnj.jus.br/programas-e-acoas/pj-proname>
- Conselho Nacional de Justiça - CNJ. (2015b). Sistema de Gestão de Tabelas Processuais Unificadas. Obtido 21

- de Maio de 2015, de <http://www.cnj.jus.br/sistemas/eficiencia-modernizacao-e-transparencia/sistema-de-gestao-de-tabelas-processuais-unificadas-resolucao-n-46>
- Conselho Nacional de Justiça - CNJ, & Conselho Nacional do Ministério Público- CNMP. Resolução Conjunta N. 3 de 16 de abril de 2013. Institui Modelo Nacional de Interoperabilidade do Poder Judiciário e do Ministério Público e dá outras providências. (2013). Obtido de <http://www.cnj.jus.br/atos-normativos?documento=1721>
- FABRI, M. (2007). Information and Communication Technology for the Public Prosecutor's Office. Bologna: CLUEB.
- FABRI, M. (2011). L'altra E-Justice: Analisi delle sentenze e durata dei procedimenti. Bologna: CLUEB.
- FERREIRA, M. (2006). Introdução à Preservação Digital: conceitos, estratégias e actuais consensos. Escola de Engenharia da Universidade do Minho.
- JARDIM, J. M. (2003). O inferno das Boas intenções: legislação e políticas arquivísticas. Em E. MATTAR (Ed.), Acesso à informação e política de arquivos (pp. 37–45). Rio de Janeiro: Arquivo Nacional.
- Justiça Federal do Rio Grande do Sul. (2015). História. Obtido 20 de Maio de 2015, de [https://www2.jfrs.jus.br/?page\\_id=2969](https://www2.jfrs.jus.br/?page_id=2969)
- LÖW, M. M. (2012). Da automatização à virtualização: a criação do processo eletrônico no Brasil. *Scire*, 18(2), 143–146.
- MATTAR, E. (2003). Acesso à Informação e Política de Arquivos. Rio de Janeiro: Arquivo Nacional.
- MCDONALD, J. (2006). A Review of Relevant International Initiatives. Obtido 6 de Fevereiro de 2012, de <http://www.collectionscanada.gc.ca/digital-initiatives/012018-3320-e.html>
- MILBRADT, P. A. (2010). A Efetividade do Acesso à Justiça por Meio do Processo Eletrônico. Faculdade de Direito/UFRGS., Dissertação de Mestrado.
- PINTO, M. M. (2009). PRESERVMAP: Um roteiro da preservação na era digital. Porto: Edições Afrontamento.
- RIBEIRO, F. (2010). Da gestão da informação judicial à memória orgânica do sistema judiciário: o contexto dos arquivos portugueses. *Páginas a&b - arquivos & bibliotecas*, 6(2), pp. 27–44.
- RIBEIRO, G. S. (2006). O Povo na Rua e na Justiça, a Construção da Cidadania e Luta por Direitos: 1889-1930. Em M. F. SAMPAIO, M. C. BRANCO, & P. LONGHI (Eds.), *Autos da Memória: a história brasileira no Arquivo da Justiça Federal* (pp. 155–223). Rio de Janeiro: Tribunal Regional Federal da 2a. Região.
- RONDINELLI, R. C. (2005). Gerenciamento arquivístico de documentos eletrônicos: uma abordagem teórica da diplomática arquivística contemporânea (IV.). Rio de Janeiro: Editora FGV.
- ROUSSEAU, J.-Y., & COUTURE, C. (1998). Os fundamentos da disciplina arquivística. Lisboa: Publicações Dom Quixote.
- SILVA, A. M. da. (2006). A Informação - Da compreensão do fenómeno e construção do objecto científico. Porto: Edições Afrontamento.
- SILVA, A. M. da, & RIBEIRO, F. (2011). Paradigmas, serviços e mediações em Ciência da Informação. Recife: Néctar.
- THIBODEAU, K. (2002). Overview of technological Approaches to digital preservation and challenges in coming years. *The State of Digital Preservation: An International Perspective*. Washington, D.C.
- United Nations Educational Scientific and Cultural Organization - UNESCO. (2003). Guidelines for the preservation of Digital Heritage. Sidney.
- VELICOGNA, M. (2007). Justice Systems and ICT, What can be learned from Europe? *Utrecht Law Review*, 3(1), 129–147.
- VELICOGNA, M. (2008). Use of Information and Communication Technologies (ICT) in European Judicial Systems. Strasbourg: European Commission for the Efficiency of Justice. Obtido de [http://www.coe.int/t/dghl/cooperation/cepej/series/Etudes/7TIC\\_en.pdf](http://www.coe.int/t/dghl/cooperation/cepej/series/Etudes/7TIC_en.pdf)

---

# Abordagem sobre vocabulários controlados para arquivos: conceitos, aplicações e metodologias

*Observations on controlled vocabularies for archives: concepts, applications and methodologies*

---

**Walter Moreira (1), Luciana Davanzo (2), Isabela Santana de Moraes (3)**

(1) Universidade Estadual Paulista (UNESP), Av. Hygino Muzzi Filho, nº 737, CEP: 17525-000, Marília, SP, Brasil, walter.moreira@marilia.unesp.br

(2) luciana.davanzo@gmail.com

(3) isabela@marilia.unesp.br

## Resumo

A utilização de vocabulários controlados em unidades arquivísticas é tomada como objeto de estudo dessa pesquisa, cujo objetivo é identificar e descrever essa temática no interior da literatura periódica científica da arquivologia. Foram utilizadas as seguintes bases de dados para o estabelecimento do corpus documental: SCOPUS, BRAPCI e LIST&A. O período de cobertura considerado foi de dez anos e a estratégia de busca utilizada contemplou os termos “tesauro”, “vocabulário controlado” e “arquivologia”. A leitura analítica do corpus foi realizada com subsídios da técnica de análise de conteúdo. Os resultados apontam a relativa incipiência do tema na arquivologia. Observou-se também que ainda é necessário uma aproximação maior da arquivologia com os vocabulários controlados, uma vez que, majoritariamente a literatura remete-se à biblioteconomia. A ausência de padronização na descrição de assunto e a baixa oferta de soluções pela literatura científica especializada tem se convertidos nos principais obstáculos para as ações de interoperabilidade requeridas aos sistemas de informação modernos.

**Palavras-chave:** Arquivologia. Organização da informação arquivística. Representação da informação arquivística. Vocabulários controlados.

## 1. Introdução

A arquivologia, enquanto campo de investigação e de práticas relacionadas à produção, organização, guarda, preservação e utilização das informações arquivísticas, tem sido cada vez mais desafiada a oferecer soluções eficazes e sintonizadas com os processos contemporâneos de tratamento e disseminação da informação, marcados pela presença das tecnologias de informação e comunicação.

Nesse cenário e tendo em vista o documento arquivístico digital (Levy, 1994; Rondinelli,

## Abstract

The use of controlled vocabularies in archives units is taken as an object of study of this research, which aims to identify and describe this subject within the scientific periodical literature of archives science. The following databases were used to establish the documentary corpus: SCOPUS, BRAPCI, and LIST&A. The coverage period considered was ten years and the search strategy included the terms "thesaurus", "controlled vocabulary" and "archives science". The analytical reading of the corpus was carried out with grants from the content analysis technique. The results show the relative incipient theme in archives science. It was also noted that it is still necessary to promote further approximation of archives science with controlled vocabularies, since, mostly literature refers to the library science. The lack of standardization in the subject description and the low supply of solutions for specialized scientific literature has been converted into the main obstacles to the interoperability of cases brought to modern information systems.

**Keywords:** Archives science, Information organization, Information retrieval, Controlled vocabularies.

2011), as questões teóricas e metodológicas impostas como problemas científicos para a consolidação epistemológica da Arquivologia, ganham novos contornos.

O conceito de documento arquivístico, como é próprio dos conceitos, *lato sensu*, ainda não é absolutamente pacífico. Há uma longa tradição nessa discussão que envolve pesquisadores como Casanova (1928), Schelenberg (1956), Duranti, Eastwood e Macneil (2002) e Rondinelli (2011), entre outros.

Rondinelli (2011), por exemplo, aponta a divergência que existe entre os termos “documento

arquivístico” e “informação arquivística”, sendo o termo “documento arquivístico”, mais comum em países de língua inglesa e o termo “informação arquivística” mais comumente utilizado em países de língua latina.

Independentemente da denominação, entretanto, para que sejam considerados como documentos de arquivo, esses documentos precisam ser produzidos e recebidos por uma determinada instituição. O documento arquivístico (ou a informação arquivística), portanto, já possui em sua gênese um objetivo. Num primeiro momento serve à própria instituição que o produziu e, posteriormente, pode assumir caráter histórico e, desse modo, servir aos pesquisadores ou aos usuários da sociedade civil, de modo geral.

Os documentos arquivísticos são produzidos, reunidos e organizados a partir de um contexto específico de produção, seja o produtor desse contexto, uma pessoa, uma família ou uma instituição (pública ou privada). Os documentos estabelecem um conjunto de relações entre si que precisam, portanto, ser mantidas. Para que sejam localizados e recuperados é preciso que sejam definidos os pontos de acesso correspondentes, isso é, é preciso descrevê-los em função dos interesses do sistema onde se localiza e também de outros sistemas com os quais espera-se que dialogue.

O conceito de informação e, naturalmente, o conceito de informação arquivística, tradicionalmente assentados em concepções epistemológicas desenhadas para modelos analógicos também precisam ser repensados em função do cenário hipertextual contemporâneo (Lamarca Lapuente, 2006). Em comum tem-se que ambos os cenários são marcados pela presença e pela intermediação da linguagem.

Nesse ambiente hipertextual, em que se confundem produtores e usuários da informação e no qual a própria configuração conceitual do documento é objeto de constantes revisões, indicadores de credibilidade, precisão, exatidão e racionalidade da informação são ainda mais requeridos.

Nesse cenário, as formas de produzir, acessar e recuperar a informação mudaram de maneira significativa. Não é mais absolutamente necessário, por exemplo, que o pesquisador ou usuário esteja no mesmo local físico em que os documentos estão custodiados, isso pode ocorrer por meio do ambiente *web*, por exemplo (Buckland, 1999). Os documentos, inclusive, podem não estar fisicamente armazenados no local de acesso, pois nesse cenário a duplicação dos documentos, para fins de acesso, é um processo com custo extremamente baixo. Há, eviden-

temente, outros motivos, alheios a essa pesquisa, pelos quais se pode questionar e fazer objeção à duplicação.

Naturalmente, propostas modernas de organização da informação devem considerar os recursos tecnológicos disponíveis. A base tecnológica influencia a produção, a circulação e aumenta a diversidade de suportes documentais. Essa mesma base, contudo, favorece o desenho de instrumentos mais precisos de representação e de recuperação da informação.

Há, como se pode depreender, questões de fundo que subjazem às questões tecnológicas propriamente ditas. O problema básico que configura os campos da arquivologia, da biblioteconomia, da documentação e da ciência da informação, nesse prisma, ainda vigora: para que uma informação possa cumprir seu papel no ciclo de produção, organização, disseminação e geração de novo conhecimento, isso é, para que seja socializada, é necessário que seja localizável e recuperável pelo usuário, o que é possível por meio de tratamento linguístico informacional adequado.

A produção de informação arquivística interpõe-se como intermediária entre a produção e o uso do conhecimento. Essa intermediação ocorre, basicamente, por meio da linguagem do sistema, o que requer consistência terminológica como condição necessária ao diálogo eficaz entre a representação da informação pela unidade arquivística (entrada) e representação da informação pelo usuário (saída). As designações “entrada” e “saída” nesse caso consideram obviamente o ponto de vista do sistema. Considerando-se o ponto de vista do usuário essas designações serão, naturalmente, invertidas.

Não é possível falar em consistência terminológica em ações de representação nas quais são envolvidos vários sujeitos, com diferentes formações discursivas, sem algum recurso da padronização que lhes garanta o acordo. Deve-se atentar, contudo, para o caráter mediador, não coercitivo, da padronização que se pode alcançar com o uso de vocabulários controlados consensualmente compartilhados.

Observando-se esse caráter mediador dos vocabulários controlados, ele é requerido, de modo geral, em todas as ações que envolvem tratamento, organização e recuperação da informação. Desse modo, também desempenham um papel vital em instrumentos mais *modernos* como ontologias, as folksonomias e os recursos de *linked data*.

Em suma, se há um processo em que se requer a descrição de elementos não contemplados

pelas normas arquivísticas e mais suscetíveis às variações culturais, institucionais e terminológicas, como é o caso da descrição do assunto do documento, é necessário considerar a função de centralidade que o vocabulário controlado assume (Buckland, 1999; Lancaster, 2002).

Desse modo, a utilização de vocabulários controlados em unidades arquivísticas é tomada como objeto de estudo, considerando-se que esses instrumentos de organização do conhecimento oferecem subsídios não somente para a organização como também para a recuperação da informação, independentemente de suporte analógico ou digital.

Um vocabulário controlado, como resultado do processo de controle vocabular utilizado num determinado sistema de informação, é uma linguagem construída para fins documentários, constituída de termos organizados em estrutura relacional. A utilização de um vocabulário controlado visa a padronizar e facilitar a entrada e a saída de dados em um sistema de informação documentária, promovendo eficácia na comunicação entre os usuários e o sistema. A estrutura relacional revela as relações hierárquicas, associativas e de equivalência que os conceitos guardam entre si.

Um vocabulário controlado permite representar a informação e o conhecimento por meio de um conjunto padronizado e finito de termos para facilitar a recuperação (Kobashi, 2008; Buckland, 1999).

Observando-se a relevância teórico-prática do tema e tendo-se em vista sua relativamente baixa abordagem na literatura mais específica da arquivologia, objetiva-se com essa pesquisa identificar e analisar a abordagem (conceitos, aplicações e metodologias) sobre vocabulários controlados e sua aplicação em unidades arquivísticas, conforme sua expressão na literatura científica da arquivologia e da ciência da informação.

Dentre os objetivos da arquivologia, é sempre importante lembrar, está a criação de condições para proporcionar acessibilidade aos documentos. Desenvolver e utilizar instrumentos que contribuam para esse fim torna-se, portanto, fundamental. Nessa conjuntura, acredita-se que os vocabulários controlados podem contribuir de maneira bastante satisfatória, pois se não houver acordo terminológico não será possível a precisão na comunicação.

Os vocabulários controlados são elaborados de maneira a facilitar e tornar mais preciso o acesso à informação, portanto, sua elaboração leva em consideração alguns fatores, tais como, os

autores, o catalogador, a estrutura sindética, o pesquisador e a consulta formulada" (Buckland, 1999).

Os vocabulários controlados são elaborados a partir de necessidades específicas de cada instituição, assim, cada vocabulário controlado será único de acordo com cada instituição e em função de suas particularidades. Entretanto, independentemente das especificidades de cada instituição, o vocabulário controlado será elaborado a partir de algumas vertentes fundamentais, conforme explicita Buckland (1999): os vocabulários do autor, do catalogador, da estrutura relacional, do pesquisador e da expressão de busca.

Para que um vocabulário controlado se torne efetivamente útil é preciso que seja organizado de modo a refletir tanto os objetivos do sistema de informação para o qual foi elaborado, isso é a linguagem do sistema, quanto a linguagem dos usuários (Kobashi, 2008). Desta forma, sua construção requer trabalho integrado, envolvendo tanto os gerenciadores do sistema quanto os usuários da informação. O vocabulário controlado deve servir ao sistema, o que implica dizer que sua aplicação ocorre tanto no processo de representação para fins de organização da informação, realizados pelo indexador ou arquivista, quanto em processos de representação para fins de recuperação da informação, realizados pelo usuário.

Os termos, num vocabulário controlado, são dispostos em ordem hierárquica e alfabética (macroestrutura e microestrutura). A macroestrutura é o eixo temático global, em que se apresentam as categorias mais amplas dos campos de conhecimentos abrangidos de forma hierárquica. A microestrutura dispõe, em forma alfabética os termos do vocabulário (Kobashi, 2008).

Considerando-se que o vocabulário controlado é formado a partir das linguagens de especialidade, isso é das terminologias específicas dos campos que pretende organizar e representar, e considerando-se que essas terminologias formam subsistemas dentro do conjunto geral da língua, é natural que os vocabulários controlados sejam instrumentos dinâmicos, que requerem revisões constantes.

O vocabulário controlado é utilizado para indicar o significado de uma determinada expressão, mas, sendo a língua é instável, as palavras podem assumir novos significados, independentemente daqueles que um lexicógrafo possa ter "fixado" (Buckland, 1999). No discurso, ou seja, na utilização efetiva da língua é que surgem

diferenças, ambiguidades e incertezas de significado.

O vocabulário controlado também é importante para questões de identidade, que, por sua vez, são indispensáveis para as instituições e para a sociedade, pois o uso da linguagem é inevitavelmente baseado em cultura (Buckland, 1999).

### 1.1. Procedimentos metodológicos

A pesquisa tem caráter bibliográfico, exploratório e descritivo. A escolha desse delineamento para a pesquisa justifica-se em função de se ter verificado, em estudo anterior (Davanzo; Moreira, 2014), a insuficiência do tratamento do tema relativo aos vocabulários controlados na literatura brasileira de ciência da informação.

Desse modo, para o estabelecimento do *corpus documental 1*, foram utilizadas as seguintes bases de dados: SCOPUS (Elsevier); Base Referencial de Artigos de Periódicos de Ciência da Informação (BRAPCI) e *Library, Information Science & Technology Abstracts with Full Text (EBSCO)*. A escolha das bases considerou sua representatividade, confiabilidade e cobertura, atestadas em trabalhos como Jacso (2005) e Bufrem, Costa, Gabriel Júnior, Pinto (2010) e Grácio e Oliveira (2012).

O período de cobertura considerado foi de dez anos (2005-2014) e a estratégia de busca utilizada contemplou os termos “tesauro”, “vocabulário controlado” e “arquivologia”, considerando-se também suas variações de equivalência sinonímica, de número e de idioma (inglês, português e espanhol).

Observou-se, como critério de seleção para a composição do *corpus documental 2*, a disponibilidade do artigo em texto completo e, preenchida essa condição, a presença (simultânea ou não) dos termos de busca nos campos de título, palavras-chave e resumos.

Foram identificados inicialmente 23 artigos na SCOPUS, 21 artigos na LISTA e apenas um artigo na BRAPCI. Após a eliminação dos itens duplicados e dos artigos que traziam as expressões apenas nos campos de resumo, obteve-se um total geral de 28 artigos.

Para fins de identificação dos artigos relevantes para a composição do *corpus de análise*, os 28 artigos foram pontuados e organizados num *ranking*. Desse modo, conforme a ocorrência da palavra nos diversos campos de busca, o artigo obteve as pontuações indicadas no Quadro I.

| Campo(s) de ocorrência          | Valor |
|---------------------------------|-------|
| Palavra-chave                   | 3     |
| Título                          | 2     |
| Resumo                          | 1     |
| Palavra-chave + título + resumo | 6     |
| Palavra-chave + Título          | 5     |
| Palavra-chave + resumo          | 4     |
| Título + resumo                 | 3     |

Quadro I. Critérios de pontuação dos artigos para composição do corpus de análise

Considerando-se que a pontuação mínima obtida foi 3 e a máxima foi 12, estabeleceu-se como nota de corte a média de 6 pontos. Desse modo, o *corpus* de análise foi composto de 14 artigos.

A leitura analítica dos textos completos selecionados foi realizada com subsídios da técnica de análise de conteúdo, conforme a proposta de Bardin (2009).

## 2. Resultados e discussão

Conforme apontado nos objetivos, interessaram à análise desenvolvida nessa pesquisa os artigos que apresentam a conjunção de tratamento dos temas “vocabulários controlados” e “arquivos”. Interessou saber como essas temáticas tem se aproximado.

A ordenação dos trabalhos conforme o critério de pontuação apresentado na subseção 1.1 (procedimentos metodológicos) está apresentada no Quadro II.

| Referência                                   | Pontuação |
|--|-----------|
| Schenkolewski-Kroll e Tractinsky (2006)      | 12        |
| Neal (2008)                                  | 10        |
| Nilbe & Tarkpea (2014)                       | 10        |
| Siqueira (2011)                              | 10        |
| Eito Brun (2011)                             | 9         |
| Balkan e Bell (2014)                         | 7         |
| Baum e Kempf (2012)                          | 7         |
| Jenkins, Oppenheim, Probets & Hubbard (2008) | 7         |
| Lin, Murray, Dorr, Hajic & Pecina (2009)     | 7         |
| Reid & Simmons (2007)                        | 7         |
| Young & Mandelstam (2013)                    | 7         |
| Clarke (2008)                                | 6         |
| Gimenez Chornet (2011)                       | 6         |
| Jenkins, Probets, Oppenheim & Hubbard (2007) | 6         |

Quadro II. *Artigos ordenados conforme a pontuação recebida*

É importante observar a quantidade relativamente baixa de trabalhos recuperados, considerando-se que foram consultadas três bases de dados de periódicos, sendo duas internacionais, com cobertura de dez anos.

Apenas um trabalho (Schenkolewski-Kroll & Tractinsky, 2006) apresentou a ocorrência dos termos (“tesauro” ou “vocabulário controlado” e “arquivo”) em todos os campos (título, resumo e palavra-chave) e obteve pontuação máxima conforme os critérios estabelecidos.

Dentre os quatorze artigos que compuseram o *corpus* de análise, havia três que, após leitura analítica e categorização, verificou-se que abordavam temática não exatamente relacionada aos objetivos desta pesquisa. Desse modo, foram incluídos na categoria “não se aplica”: Neal (2008) e Young e Mandelstam (2013), que abordam a aplicação de tesouros em bibliotecas, e Nilbe e Tarkpea (2014), que trata da aplicação de tesauro em um repositório.

Seguindo-se a técnica de análise de conteúdo (Bardin, 2009), foram identificados os temas e estes foram organizados em categorias semânticas definidas *a posteriori*, ou seja, após a análise do material.

Recomenda-se, na aplicação da análise de conteúdo (Bardin, 2009), que as categorias sejam autoexcludentes. Essa recomendação não foi observada com rigor por dois motivos: a) o pequeno número de trabalhos identificados e b) a identificação de mais de um tema (categoria) nos trabalhos. A definição do foco temático do artigo ocorreu, portanto, em conformidade com os interesses pragmáticos específicos dessa investigação. Além do mais, a categorização, ensina a mesma Bardin (2009), não deve ser encarada de modo descontextualizado e se reveste sempre de uma hipótese de organização.

Sendo assim, foram analisados 11 artigos, distribuídos nas seguintes categorias de análise: a) conceitos: trabalhos mais conceituais, com foco de abordagem mais teórica nas relações entre vocabulários controlados e arquivos; b) aplicações: trabalhos que priorizam a descrição de aplicações de soluções, isso é, que apresentam casos de sucesso da aplicação de vocabulários controlados em arquivos; c) metodologias: trabalhos que oferecem mais detalhes a respeito das metodologias utilizadas para o desenvolvimento e aplicação de vocabulários controlados em arquivos.

Os trabalhos foram agrupados conforme apresentado no Quadro III e a discussão é apresentada nas subseções subsequentes.

| <b>Categoria</b>      | <b>Autor(es)</b>   |
|-----------------------|--|
| Trabalhos conceituais | Clarke (2008), Lin, Murray, Dorr, Hajic e Pecina (2009), Siqueira (2011), Baum e Kempf (2012), Balkan e Bell (2014)  |
| Foco na aplicação     | Reid e Simmons (2007), Eito Brun (2011), Baum e Kampf (2012), Balkan e Bell (2014)   |
| Foco nas metodologias | Schenkolewski-Kroll e Tractinsky (2006), Jenkins, Oppenheim, Proberts & Hubbard (2007), Jenkins, Oppenheim, Proberts & Hubbard (2008), Giménez-Chornet e Escrig-Giménez (2011) |
| Não se aplica         | Neal (2008), Young (2013), Nilbe (2014)  |

Quadro III. *Categorização do corpus de análise*

### 2.1. Trabalhos conceituais

Nessa categoria foram incluídos os seguintes trabalhos: Clarke (2008), Lin, Murray, Dorr, Hajic e Pecina (2009), Siqueira (2011), Baum e Kempf (2012), Balkan e Bell (2014).

Os vocabulários controlados facilitam o acesso a coleções digitais e outras pela representação explícita dos princípios subjacentes de organização (Lin, Murray, Dorr, Hajic & Pecina, 2009)

Um conceito associado ao conceito de vocabulário controlado que parece próximo (se não o estiver fato) do consensual é o conceito de interoperabilidade. Isso está previsto, inclusive, em normas como a recentemente revisada ISO 25964, que orienta a elaboração de tesouros. Dividida em duas partes, a primeira refere-se às características gerais dos tesouros e dos elementos que o compõem, a segunda parte trata especificamente da interoperabilidade entre vocabulários controlados (International Organization for Standardization [ISO], 2011).

A ISO 25964 privilegia o conceito em detrimento das palavras ou termos. A mesma norma, em versão anterior, considerava o tesauro como um conjunto de termos interconectados por uma rede de relações semânticas. A proposta atual define o tesauro como um conjunto de conceitos relacionados semanticamente. Ainda que a tipologia de relações seja a mesma de um tesauro baseado nas relações entre termos, a terminologia sintetiza uma diferença substancial de concepção, mais alinhada aos fundamentos da teoria comunicativa da terminologia.

Num tesouro baseado na relação entre termos, as relações são definidas entre os termos. Nesse tipo de relação podem-se relacionar termos que denotam um mesmo conceito, como ocorre, por exemplo, com as abreviaturas. Num tesouro baseado em conceito, algumas relações ocorrem entre os conceitos e outras entre os termos. (Pastor-Sanchez, Martínez Mendez & Rodríguez-Muñoz, 2009; Clarke & Zeng, 2012).

Entre os autores que destacam a necessidade de aspectos de interoperabilidade nos vocabulários controlados estão Balkan e Bell (2014).

A preocupação com os aspectos de interoperabilidade em vocabulários controlados decorre dos crescentes esforços de internacionalização em ambientes científicos. Baum e Kempf (2012) consideram, contudo, que a normalização necessária à indexação internacional coerente, a qual é vital para a recuperação da informação, ainda não está suficiente desenvolvida. Os autores não fazem referência a nenhuma norma de modo específico, não se sabe se tinham conhecimento da norma ISO citada anteriormente ou se ela está incluída na crítica.

De qualquer modo, os processos de internacionalização, pelo que carregam de complexidade cultural, andam a passos lentos. Esforços no sentido de promover a padronização e internacionalização existem desde, pelo menos, as propostas da documentação, realizadas por Paul Otlet. Baum e Kempf (2012) situam os trabalhos sobre essas propostas nos anos 1970 e atribuem a falta de diretrizes coerentes à ausência de pesquisas sobre as diferenças nos modelos de indexação de resultados de pesquisa realizados pelos diversos países.

Clarke (2008) ratifica o importante avanço verificado por meio do conjunto de metadados *Dublin Core* e chama a atenção para o papel que os vocabulários controlados podem assumir no preenchimento do elemento “assunto”.

O único trabalho brasileiro localizado, conforme a estratégia de busca estabelecida para essa investigação e apresentada nos procedimentos metodológicos, é de Siqueira (2011). Este trabalho, analisado apressadamente, poderia ser colocado na categoria “aplicações”. Sua leitura atenta, contudo, revela que, apesar de analisar um arquivo específico, o Sistema de Arquivos da USP (SAUSP), isso é realizado com o objetivo de demonstrar a eficácia do vocabulário controlado como instrumento que agrega precisão na descrição e na localização da informação arquivística.

## 2.2. O foco nas aplicações

Nessa categoria foram incluídos os seguintes trabalhos: Reid e Simmons (2007), Eito Brun (2011), Baum e Kampf (2012) e Balkan e Bell (2014).

O tesouro multilingue *European Language Social Science Thesaurus* (ELSST) foi objeto de dois artigos: Baum e Kampf (2012) e Balkan e Bell (2014). O ELSST é administrado pelo UK *Data Archive*, que também administra o tesouro monolíngue *Humanities and Social Science Electronic Thesaurus* (HASSET), do qual o ELSST é derivado.

Eito Brun (2011) descreve a experiência de aplicação de padrões desenvolvidos em XML para a gestão e publicação das descrições e documentos do fundo Engenheiro Carlos Fernández Casado. O trabalho foi desenvolvido para o Centro de Estudos Históricos de Obras Públicas e Urbanismo do Ministério do Desenvolvimento, na Espanha.

A solução envolveu o desenvolvimento de um ambiente para edição e criação de instrumentos de descrição *Encoded Archival Description* (EAD) compatíveis com as normas internacionais ISAD(G) e ISAAR(CPF). Este ambiente permite acesso a vocabulários controlados locais ou remotos para a realização da descrição.

Reid e Simmons (2007) apresentam a experiência de aplicação de arquivos de autoridade e tesouros na definição dos pontos de acesso do *Archival Research Catalog* (ARC). O ARC é um projeto desenvolvido pela *National Archives and Records Administration* (NARA), no âmbito do governo federal dos Estados Unidos. O uso de vocabulário controlado, relatam os autores, assegura ao ARC consistência e precisão na recuperação da informação.

## 2.3. Foco nas metodologias

Nessa categoria foram incluídos os seguintes trabalhos: Schenkolewski-Kroll e Tractinsky (2006), Jenkins, Oppenheim, Proberts & Hubbard (2007), Jenkins, Proberts, Oppenheim & Hubbard (2008), Giménez-Chornet e Escrig-Giménez (2011).

Giménez-Chornet e Escrig-Giménez (2011) descrevem uma aparentemente complexa experiência de integrar quatro unidades arquivísticas espanholas. Para isso, treze especialistas em arquivo utilizaram o mesmo servidor para criar descrições padronizadas com recurso das normas ISAD(G) e ISAAR(CPF) e de três tesouros para fins de indexação, definindo instituições, lugares geográficos e assuntos.

Este trabalho descreve os aspectos de construção de um protótipo de um tesouro incorporado num sistema de busca situado na intranet que atende a quatro arquivos históricos geograficamente separados.

Jenkins, Oppenheim, Probeta e Hubbard (2008), descrevem o processo de criação de um vocabulário controlado para analisar processos de autoarquivamento de acordos de transferência de direitos autorais realizados por editores de revistas. O vocabulário controlado consiste de três categorias que possibilitam descrever: o que pode ser autoarquivado, as condições e as restrições do autoarquivamento. O vocabulário controlado, na avaliação dos autores do trabalho, mostrou-se eficaz, mas requer revisão e atualização contínuas em função da dinamicidade inerente ao assunto que cobre.

Dentre as dificuldades encontradas durante o desenvolvimento do vocabulário controlado, relatam a necessidade de reduzir o número de descritores em função da grande variedade de termos utilizados pelos editores, sem consistência e sem precisão.

O primeiro estágio na criação do vocabulário, que não seguiu, segundo informam os autores, um método formal, foi familiarizar-se com as políticas de autoarquivamento e com os acordos de transferência de direitos autorais. A partir dessa análise, os termos foram colhidos e organizados nas seguintes categorias: “que” (*pre-prints, pos-prints etc.*), “onde” (repositório público, página pessoal do autor etc.), “qual versão” (versão do autor, versão formatada do editor etc.) e “requisitos” (o que o autor deve fazer para que o autoarquivamento seja permitido pelo editor).

Além do artigo descrito anteriormente, no qual Jenkins, Oppenheim, Probeta & Hubbard (2008) falam sobre a criação de vocabulário controlado para a análise de acordos de transferência de direitos autorais, denominado *RoMEO studies 7*, os mesmos autores, em estudo publicado um ano antes, discutem aspectos lógicos do autoarquivamento, denominado *RoMEO studies 8* (Jenkins, Oppenheim, Probeta & Hubbard (2007)).

O artigo de Schenkolewski-Kroll e Tractinsky (2006) foi o que melhor respondeu à estratégia de busca utilizada para localizar material para essa pesquisa, conforme já apontado anteriormente. É também o que apresenta mais claramente a relação entre a descrição de material arquivístico, indexação e recuperação da informação, bem como aspectos sobre a construção de tesouros para fins arquivísticos.

Esse trabalho relata a experiência desenvolvida nos arquivos israelenses, bem como a situação geral dos arquivos nesse país em função das normas internacionais e sua adaptação às necessidades arquivísticas nacionais.

Os autores descrevem dois projetos específicos que representam diferentes visões: o tesouro do *National Insurance Institute of Israel* e os arquivos *Religious Zionism at Barllan University*. Em conclusão, o artigo destaca uma proposta de criação de um sistema de descrição e recuperação da informação arquivística baseado em ISAD(G), ISAAR(CPF) e um tesouro geral adequado a todos os arquivos israelenses.

### 3. Considerações finais

Os resultados apontam para a relativa incipiência do tema na arquivologia e sinalizam uma possível confusão conceitual em relação ao termo “vocabulário controlado”, tomado, em alguns casos, como sinônimo de “tesouro”.

Observou-se, ainda, a ausência de um referencial teórico forte e específico sobre a aplicação de vocabulários controlados em arquivos, já que, normalmente, os textos são construídos com o referencial teórico e o modelo de abordagem da biblioteconomia em relação a esses instrumentos. A esse respeito é preciso desenvolver novas e mais profundas investigações.

Embora pareça predominar na literatura arquivística a denominação “vocabulários controlados”, a literatura recente da *International Society of Knowledge Organization* (ISKO) tem adotado de formal preferencial o termo “knowledge organization system” (KOS), como mais generalizante.

Essa investigação é limitada pelo recorte que estabelece, como, de fato ocorre, aliás, com qualquer outra investigação. Investigações mais amplas devem incluir termos como “sistemas de organização do conhecimento”, “instrumentos de organização do conhecimento”, “linguagens documentárias”, “ontologias”, “taxonomias”, “folksonomias” e “classificação facetada” e suas relações com a arquivística. É preciso verificar também a confusão terminológica que se estabelecem entre essas expressões.

A pesquisa que motivou esse relatório foi desenvolvida com a função precípua de organizar e analisar o material bibliográfico (artigos de periódicos) disponível sobre as relações entre vocabulários controlados e arquivos. Desse modo, pretende-se que sirva como guia para a condução de futuras investigações sobre o tema.

As normas ISAD(G) e ISAAR(CPF) orientam a descrição de material arquivístico com base na proveniência e auxiliam na descrição dos fundos em vários níveis. Falta, ainda, uma norma ou um conjunto de diretrizes que sejam aplicadas à descrição e indexação de conteúdo dos registros. A ausência de padronização nesse sentido, bem como a baixa oferta de soluções pela literatura científica especializada tem se tornado um dos principais obstáculos para as ações de interoperabilidade requeridas aos sistemas de informação modernos. Dificultam, outrossim, como atestado pelos diversos estudos citados, a precisão na representação e na recuperação da informação arquivística.

## Referências

- Balkan, Lorna; Bell, Lucy. (2014). Linking thesauri - ELSST as hub for social science data terms. *IASSIST Quarterly*, 16-21.
- Bardin, L. (2009). Análise de conteúdo. Lisboa: Edições 70.
- Baum, K., Kempf, A. O. (2012). Thesaurus-based indexing of research data in the social sciences: opportunities and difficulties of internationalization efforts. *IASSIST Quarterly*, 12-15.
- Buckland, M. (1999, May). Vocabulary as a central concept in library and information science. In T. Arpanac et al. (Eds.). *Digital libraries: interdisciplinary concepts, challenges, and opportunities: proceedings of the Third International Conference on Conceptions of Library and Information Science*. Dubrovnik, Croatia. 3-12.
- Bufrem, L. S., Costa, F. D. O.; Gabriel Júnior, R. F. & Pinto, J. S. P. (2010). Modelizando práticas para a socialização de informações: a construção de saberes no ensino superior. *Perspectivas em Ciência da Informação*, 15(2), 22-41.
- Casanova, E. (1928). *Archivistica*. Arti Grafiche Lazzeri.
- Clarke, S. G. D. (2008). The last 50 years of knowledge organization: a journey through my personal archives. *Journal of Information Science*, 34(4), 427-437.
- Davanzo, L. & Moreira, W. (2014). O vocabulário controlado como ferramenta do processo de organização e recuperação da informação. *Anais do XIV Encontro Nacional de Pesquisa em Ciência da Informação*. Belo Horizonte: ECI/UFGM. 1142-1148.
- Clarke, S. G. D. & Zeng, M. L. (2012). From ISO 2788 to ISO 25964: the evolution of thesaurus standards towards interoperability and data modeling. *Information Standards Quarterly*, 24(1), 20-26.
- Duranti, L; Eastwood, T; Macneil, H. (2002). *Preservation of the integrity of electronic records*. Dordrecht: Kluwer Academic.
- Eito Brun, R. (2011). Aplicabilidad de EAD en archivos históricos de ingeniería: el fondo Carlos Fernández Casado. *Scire*, 17(1), 35-39.
- Giménez-Chornet, V. & Escrig-Giménez, M. (2011). Designing a thesaurus to give visibility to the historical archives in the Archivo del Reino in Valencia. *Knowledge Organization*, 38(2), 154-166.
- Grácio, M. C. C. & Oliveira, E. F. T. (2012). Visibilidade dos pesquisadores no periódico *Scientometrics* a partir da perspectiva brasileira: um estudo de cocitação. Em *Questão*, 18(edição especial), 99-113.
- International Organization for Standardization (2011). ISO 25964: information and documentation- Thesauri and interoperability with other vocabularies. Geneva: ISO.
- Jacso, P. (2005). As we may search: comparison of major features of the Web of Science, Scopus and Google Scholar citation-based and citation-enhanced databases. *Current Science*, 89(9), 1537-1547.
- Jenkins, C., Oppenheim, C., Proberts, S. & Hubbard, B. (2008). RoMEO studies 7: creation of a controlled vocabulary to analyse copyright transfer agreements. *Journal of Information Science*, 34(3), 209-307.
- Jenkins, C., Proberts, S., Oppenheim, C. & Hubbard, B. (2007). RoMEO studies 8: self archiving: the logic behind the colour-coding used in the copyright knowledge bank. *Program*, 41(2)124-133.
- Kobashi, N. Y. (2008). *Vocabulário controlado: estrutura e utilização*. [Texto preparado para subsidiar o debate sobre vocabulário controlado da pesquisa Mapeamento para a reunião da Rede de Escolas de Governo em 15 e 16 dez. 2008].
- Lamarca Lapuente, M. J. (2006). *Hipertexto: el nuevo concepto de documento en la cultura de la imagen*. Tesis de doctorado en Fundamentos, Metodología y Aplicaciones de las Tecnologías Documentales y Procesamiento de la Información, Universidad Complutense de Madrid. Recuperado em 30 jan. 2015, de <http://www.hipertexto.info>.
- Lancaster, F. W. (2002). *El control del vocabulário en la recuperación de información*. Valencia: Universitat de Valencia.
- Levy, D. M. (1994). Fixed or fluid? Document stability and new media. In: *Proceedings of the European Conference on Hypermedia Technology (ECHT 94)*. Edinburgh, UK. *Proceedings...* Edinburgh, p. 24-31.
- Lin, J., Murray, G. C., Dorr, B. J., Hajic, J. & Pecina, P. (2009). A cost-effective lexical acquisition process for large scale thesaurus translation. *Lang Resources & Evaluation*, n. 43, p. 27-40, 2009.
- Neal, D. (2008). News photographers, librarians, tags, and controlled vocabularies: balancing the forces. *Journal of Library Metadata*, 8(3), 199-219.
- Nilbe, S. & Tarkpea, T. (2014). Using the estonian subject thesaurus in the digital environment. *Cataloging & Classification Quarterly*, 52(1), 32-41.
- Pastor-Sanchez, J. A., Martínez Mendez, F. J., Rodríguez-Muñoz, J. V. (2009). Advantages of thesaurus representation using the Simple Knowledge Organization System (SKOS) compared with proposed alternatives. *Information Research*, 14(4).
- Reid, L. J. E. & Simmons, C. J. (2007). Authority control at the National Archives and Records Administration. *Journal of Archival Organization*, 5(1/2), 95-120.
- Rondinelli, R. C. (2011). *O conceito de documento arquivístico frente à realidade digital: uma revisão necessária*. Tese de doutorado em Ciência da Informação, Universidade Federal Fluminense, Brasil, RJ, Niterói.
- Schenkolewski-Kroll, S. & Tractinsky, A. (2006). Archival description, information retrieval, and the construction of thesauri in israeli archives. *Archival Science*, 6, 69-107.
- Schellenberg, T. R. (1956). *Modern archives: principles and techniques*. Chicago: University of Chicago Press.
- Siqueira, J. C. (2011). Recursos linguísticos para análise de vocabulário controlado: o caso do SAUSP. *Biblionline*, 7(2), 52-62.
- Young, J. L.; Mandelstam, Y. (2013). It takes a village: developing Library of Congress genre/form terms. *Staff Publications*, Paper 2.

---

# Clasificación y indización de la información en Registros y Archivos de Gestión de las universidades

*Classification and information indexing in University Registries and Current Records*

---

**Raquel Luise Pret (1), Rosa Inês de Novais Cordeiro (2)**

(1) Departamento de Ciência da Informação, Universidade Federal Fluminense, Rua Lara Vilela, 126, São Domingos, Niterói, Rio de Janeiro, Brasil, pretraquel@vm.uff.br

(2) rosacordeiro@vm.uff.br

## Resumen

Investigación sobre la clasificación e indización de documentos en los registros y archivos de gestión de las universidades, teniendo en cuenta las especificidades de las actividades en estas instituciones. El proceso de clasificación y descripción de la representación documental se analiza en base a los usos de los documentos que se realizan por las personas que producen y requieren a estos. Parece que los usos de los documentos son múltiples y ubicuos, no limitados a la discusión sobre archivos de gestión y permanentes. Sin embargo, también se identifican tres tipos de usos de los documentos recurrentes en los archivos: presunción de prueba, acciones mediadoras e el informativo. El estudio trata de distanciarse de las teorías llamadas representacionistas de la información y se basa en secciones de Wittgenstein sobre la idea del lenguaje a partir de las experiencias y las prácticas cotidianas. Este artículo es el resultado de un estudio teórico, cuyo campo empírico son los registros y archivos de gestión de la Universidade Federal Fluminense, en Brasil.

**Palabras clave:** Clasificación. Indización. Registros. Archivos de gestión. Archivos universitários. Usos de los documentos.

## 1. Introdução

Os arquivos são instituições historicamente marcadas pelas relações jurídico-administrativas que seus documentos possuem. Neste sentido, eles diferenciam-se de outras unidades de informação, pois, para além dos valores informativos, seus documentos servem também para viabilizar ações administrativas, muitas vezes reguladas e legitimadas pelo ordenamento jurídico em que a instituição produtora faz parte.

Diversas funções e atividades são realizadas pela mediação destes dispositivos. Portanto, os documentos de arquivo não são apenas artefa-

## Abstract

Research on the classification and indexing of records in the registries and current records of universities, considering the specificities of the activities in these institutions. The process of classification and description for the documentary representation analysis is based on the uses of records made by people who produce and require these records. It appears that the uses of the records are multiple and ubiquitous, not limited to the boundaries of current and permanent archives. However, three types of recurrent form of covered uses in the archival field are also identified: proof presumption, mediator actions and informative. The study seeks to distance itself from the so called representationalist of information theories and is based on Wittgenstein's lectures on the thought of language from experiences and daily practice. This article is the result of a theoretical study, whose empirical field are registries and current records of the Universidade Federal Fluminense, in Brazil.

**Keywords:** Classification. Indexing. Registry records. Current records. University archives. Record uses.

tos produzidos por suas instituições de origem para atestar uma ação, mas eles são produzidos porque há imposições, coerções, consensos, múltiplas relações que exigem a sua produção. Pensar em documentos de arquivo é pensar no sistema jurídico legal da sociedade em que tais arquivos estão inseridos. Os documentos de arquivo são dispositivos produzidos, regulados, distribuídos e escalonados por convenções socialmente instituídas.

Esta condição, dependente dos atos jurídicos, administrativos, comerciais, familiares que os homens vão desenvolvendo através dos anos, dos séculos, faz com que os documentos apresentem outra característica muito particular, a de constituir séries de unidades relacionadas entre si por algum ele-

mento, que as convertem em conjuntos orgânicos facilmente identificáveis e, por tanto, formando unidades complexas de milhares de documentos simples (Cortés Alonso, 2008, p. 409).

O documento surgiu em nossa sociedade *ocidental* como "instrumento de normalização do discurso da verdade, um meio da validação da verdade que possuía características que permitiam tal validação" (Pret, 2013, p. 53). O sistema jurídico, como sistema de regulação, desenvolveu-se criando diversas normatizações e ordenamentos, na tentativa de controlar as relações sociais, e o documento foi uma delas. Percebemos assim que o documento de arquivo somente é possível de ser produzido e de tramitar inserido no ordenamento jurídico do qual faz parte, enfatizando-se a afirmativa de Bellotto (2006) de que uma das principais características do documento de arquivo é de ser este presunção de prova.

Heredia Herrera (1991, p. 30), em consonância com Bellotto, afirma que o documento de arquivo está marcado por duas coordenadas naturais: o produtor que o gera em virtude das funções que tem atribuídas ou reconhecidas e a função/atividade que o determina.

Esta apropriação é necessária, pois para pensarmos a classificação e a indexação dos documentos de arquivo, alvo deste trabalho, é fundamental analisarmos os usos que os sujeitos que os produzem e os manipulam fazem deles. Assim, identificamos preliminarmente alguns usos do documento historicamente instituídos nos arquivos: o de presunção de prova, o mediador de ações e o informativo. Alguns autores que se transformaram em referências da literatura arquivística dedicaram tempo e esforços para tratá-los, tais como: Schellenberg (2006); Cruz Mundet (1994); Cortés Alonso (2011); Duranti (1997); Rosseau e Couture (1998), além de Bellotto (2004), todos eles, de diferentes formas, analisaram a questão do uso dos documentos, uns com mais ênfase no uso do documento para viabilizar ações, atividades e funções da instituição produtora; outros como registro de ação, presunção de prova, além dos estudos do uso do documento como testemunho, como referência informativa para a tomada de decisão, mas todos os autores citados identificaram tais usos dos documentos de arquivo. Usos provenientes da configuração assumida pelos arquivos ao longo do tempo e das práticas dos sujeitos que produzem e manuseiam tais documentos. Não buscamos a universalidade dos usos dos documentos de arquivo, pois eles podem ser múltiplos, e qualquer tentativa neste sentido já nasceria fracassada, no entanto, a multiplicidade de obras que

trataram de tais usos e em diferentes "universos arquivísticos", instiga-nos a reflexão sobre eles na análise sobre classificação e indexação nos arquivos universitários.

Ao trabalhar com a análise dos usos informacionais dos documentos de arquivo, nossa pesquisa procura distanciar-se das teorias representacionistas. Fogl (1979), por exemplo, em sua concepção sobre informação, apresenta os conceitos como representações diretas do real que determinam um domínio e que se encontram encapsulados nas palavras. A informação pode de ser representada de forma genérica porque há uma linguagem da mente que os indivíduos compartilham. Percebemos que esta forma de tratar a informação é influenciada pelo representacionismo aristotélico, uma vez que seus esforços em descrever conteúdos informacionais, mapear leis e propriedades de objetos informativos, partem do princípio que há um sistema de significados universal baseado na essência das coisas e no repertório cognitivo semelhante dos seres humanos.

Essa teoria representacionista sobre a informação, realizada por diversos autores como Shera e Egan (1961), Soergel (1999), Vickery (1973), entre outros, trata os enunciados informativos como aspectos objetivos do significado, produtos do conhecimento enquanto entidade mental. A informação é identificada como um aspecto material do processo de cognição, ou seja, expressão definitiva do conhecimento. Identificamos a influência da tradição aristotélica neste pensamento, uma vez que parte do pressuposto que os sistemas de representação *traduzem* para a realidade as construções, as configurações e as deduções da mente humana.

Uma definição é uma frase que significa a essência de uma coisa. Apresenta-se ou sob a forma de uma frase em lugar de um termo, ou de uma frase em lugar de outra frase; pois às vezes também é possível definir o significado de uma frase. Aqueles cuja explicação consiste apenas num termo, por mais que façam, não conseguem dar definição da coisa em apreço, porque uma definição é sempre um certo tipo de frase (Aristóteles, *Metafísica*, 983a).

Dessa forma, o conhecimento se daria a partir de regras lógicas, de padrões preestabelecidos combinados entre si a partir de cada situação apresentada pela realidade. Nesta concepção instrumentalista, as palavras possuem funções, elas designam as coisas do mundo. Portanto, caberia ao profissional da informação compreender, organizar e descrever tais combinações.

A crítica de Wittgenstein, em *Investigações Filosóficas*, feita ao representacionismo aristotélico na linguagem nos é útil para tratarmos a infor-

mação, uma vez que a base da informação é a linguagem. De acordo com o filósofo, o problema de associarmos cada palavra a um conceito é o engano que tal crença nos leva: a ilusão que os significados são entidades autônomas e correspondências diretas do real. Essa ânsia por generalidades cria o mito que a essência das coisas estão nos conceitos, sublimando assim a linguagem do pulsar da vida, apagando os contextos de uso das palavras. De forma semelhante, os estudos representacionistas da informação baseiam suas operações em sistemas de regras definidos pelos sentidos apreendidos dos textos registrados nos documentos.

Wittgenstein propõe então uma subversão da concepção da linguagem: em vez de voltarmos para o apagamento forçado dos contextos de uso das palavras em prol de conceitos universalizantes, mas precários, pensemos na linguagem a partir das experiências e da práxis cotidianas. Dessa forma, os significados vão forjando-se na dinâmica da vida. As atividades humanas são inúmeras e a linguagem atravessa todas elas, são relações mutuamente constitutivas. "[...] Seguir uma regra é uma práxis. E acreditar seguir uma regra não é seguir uma regra. E daí não podemos seguir uma regra privadamente, porque senão acreditar seguir a regra seria o mesmo que seguir a regra." (Wittgenstein, *Investigações Filosóficas*, p. 93, 1999).

Destarte, para abordarmos as classificações dos documentos de arquivo, precisamos admitir que estas sempre são arbitrariedades, "*igualações do que não é igual*", como afirma Nietzsche (*Sobre verdade e mentira no sentido extra moral*, § 2), enquadramentos impositivos. Assim, assumimos, da mesma maneira que Walter Benjamin (1987) o fez, que "*toda ordem é precisamente uma situação oscilante à beira de um precipício*". Essa instabilidade se dá por nenhuma forma de organização conseguir captar a dinâmica da vida, das relações sociais, da realidade das coisas, as complexidades dos sujeitos. Ressalvados por essas características dos sistemas classificatórios é que propomos pensar a classificação nos protocolos e arquivos correntes universitários a partir dos usos que a entidade produtora faz de seus documentos, em seu micro-universo, em suas práticas cotidianas.

## 2. A classificação e os vínculos arquivísticos

Refletindo sobre os documentos de arquivo corrente, eles e as ações da entidade produtora são mutuamente constitutivos. Os documentos

são gerados pelas ações, como registros, e viabilizam sua execução. Não há documento sem ação e há certas ações, de acordo com o ordenamento jurídico e administrativo em que a entidade produtora faça parte, que somente podem ser desenvolvidas por meio dos documentos.

O contexto de produção do documento justifica-se por essa relação orgânica entre o documento de arquivo e a ação que o originou, um não existe sem o outro. Na maioria das vezes, a ação está inserida em uma atividade mais ampla, havendo inúmeros processos para o seu desenvolvimento. Os documentos, assim como as atividades e as funções fazem parte da vida da instituição produtora, estão na vida cotidiana de seus produtores. Por estes usos dos documentos de arquivo, por esta relação mutuamente constitutiva entre o documento e as ações é que a classificação funcional se justifica.

De acordo com Foscarini (2009), a prática da classificação de documentos vem da necessidade de tornar explícito o vínculo arquivístico entre os documentos. Este vínculo existe entre todos aqueles registros que participam da mesma atividade, desde o momento da criação destes. Através do ato de classificação, a rede de relações inerentes à natureza de qualquer documento não só é trazida à luz, mas também é estabelecida e perpetuada. Deste modo, o significado de cada documento em relação a todos os outros, bem como a estrutura do conjunto de documento (isto é, o fundo arquivístico) pode ser compreendida e transmitida ao longo do tempo (Foscarini, 2009, p. 2).

A autora mencionada defende ainda um cuidadoso estudo sobre as atividades e as funções da entidade produtora dos documentos para a construção do sistema de classificação no arquivo. Entendemos que esta defesa trata-se de uma reflexão acerca dos usos que a entidade produtora tem de seus documentos, uma vez que eles estão o tempo inteiro viabilizando suas ações, atividades e funções.

Para melhor compreendermos as diferenças entre *ação*, *função* e *atividade*, utilizamos a definição clássica de Schellenberg:

Os documentos de arquivo são produto de uma ação. Uma ação pode ser tratada em termos de funções, atividades e atos (transactions). O termo função é usado com referência a todas as responsabilidades atribuídas a um órgão a fim de atingir os amplos objetivos para os quais foi criado. Geralmente, essas funções são definidas nas leis ou regulamentos que criam o órgão. [...] Cada função pode ser subdividida em diversas atividades, consideradas como uma série de ações, levadas a

efeito de uma função específica (Schellenberg, 2006, p. 84)

Portanto, entendemos como função o conjunto de atividades destinadas a um propósito, uma abstração e, como tal, precisa de uma estrutura feita por regras convencionadas e recursos para se materializar. Concordamos com Sabourin (2001), que define função como qualquer finalidade de alto nível, responsabilidade, tarefa ou atividade que, dentro da hierarquia da instituição produtora, definida na agenda de sua instituição com regulações legais, políticas ou mandatárias. As funções de uma instituição são o próprio motivo de sua existência, uma vez que podem ser compreendidas como um conjunto ou série de atividades, ou como define Foscarini, *processos de trabalho*, que, quando realizados de acordo com uma sequência prescrita, irão resultar em uma produção institucional ou individual de resultados esperados, isto é, a execução de serviços delegados.

A classificação nos arquivos serve, sobretudo, para a instituição produtora compreender seu universo documental e as ações, atividades e funções amalgamadas por eles. O compromisso da classificação dos documentos de arquivo não é com o conteúdo destes, em sua unidade, é com os vínculos arquivísticos, ou seja, com seu registro e fixação em um plano de classificação. Classificar, nesta lógica, é um exercício de compreensão da entidade produtora de si mesma, de seu cotidiano, de suas práticas que justificam o seu *ser* e o seu *estar* no mundo. Desse modo, a classificação funcional para os documentos de arquivo justifica-se por estar baseada nos seus usos pela entidade produtora. A compreensão das funções, das atividades, das ações e dos documentos, ou seja, da *práxis* da entidade produtora que motiva a operação da classificação funcional.

O que distingue um arquivo (no sentido de uma pluralidade de documentos) como um todo orgânico, ou *universitas rerum*, ou seja, uma entidade estruturada, são as circunstâncias contingentes da sua criação, a partir da formação de um mero conjunto ou soma de itens únicos, unidos artificialmente para a realização de quaisquer fins. Este link incremental é original e necessário (conhecido como "vínculo arquivístico"), ele existe entre todos os documentos que pertencem juntos ao mesmo processo, porque eles são originados durante, e em virtude da mesma atividade ou processo de negócios (Foscarini, 2009, p.2, *tradução nossa*)

A classificação evidencia a rede de relações mediadas por qualquer documento e a perpetua. Ao classificar, a instituição produtora não aproxima documentos semelhantes e separa os diferentes, como a interpretação representacionista sugere, a classificação depende do con-

texto de produção e uso dos documentos porque é baseada neles. Os esquemas classificatórios baseados em assuntos atribuídos às unidades documentais quebram com os vínculos entre as ações e os documentos a ponto de não serem mais perceptíveis, causando o apagamento dos contextos, tornando incompreensível, muitas vezes, o porquê da existência dos próprios documentos.

A Arquivística para atender às demandas de uso informacional dos documentos de arquivo, ou seja, os documentos usados como referências, desenvolveu ao longo dos séculos a descrição como operação arquivística capaz de possibilitar o acesso ao conteúdo informacional presente nos documentos de arquivo.

De acordo com Oliveira (2010), a descrição arquivística é um processo intelectual de representação dos documentos de arquivo que tem como finalidade o controle e o acesso aos mesmos.

A descrição exige um pleno conhecimento do contexto de produção dos documentos, dos vínculos arquivísticos, do contexto histórico no qual o arquivo foi produzido que extrapola a mera utilização dos instrumentos de pesquisa (Oliveira, 2010, p. 45)

A descrição trata de uma complexidade maior, de um todo, no que se refere a recuperação da informação no documento de arquivo. Oliveira (2010, p. 46) afirma que a descrição é uma representação formulada pelo arquivista e a sua formulação envolve processos e métodos de pesquisa complexos e particulares que convergem para a compreensão de um arquivo. Um elemento fundamental para essa compreensão é a reconstrução do contexto arquivístico. Podemos destacar algumas atividades arquivísticas inseridas no processo de descrição como a elaboração de mecanismos de acesso; a explicitação dos relacionamentos entre os documentos de arquivo e as funções e atividades que lhes deram origem; e, também, o controle do acervo.

Todavia, a operação arquivística da descrição, de acordo com os pressupostos apresentados, a formulação histórica de seu conceito e sua padronização indicada pelas normas - a *Manual for Archival Description*, elaborado pela Universidade de Liverpool, sob a direção de Michael Cook; *Rules for Archival Description*, produzidas pelo Bureau of Canadian Archivists e pelo Conselho Canadense de Arquivos; a *International Standard for Archival Description*, desenvolvida pelo Conselho Internacional de Arquivos; e a *Norma Brasileira de Descrição Arquivística* - apontam para representações de um tipo de informação característico dos arquivos perma-

nentes. A preocupação com o contexto do documento de arquivo, característica da descrição como operação de representação da informação, atende a um dos dois níveis de informação presentes nos arquivos, relatados por Rosseau e Couture (1998, p. 251), a informação do conjunto documental que revela as ações, atividades e funções das atividades produtoras. De acordo com os autores canadenses, há outro nível de informação presente nos arquivos, que é a informação contida no documento em si, isoladamente, relacionada mais ao conteúdo do que ao contexto.

Bellotto (2006, p. 33) ao abordar a descrição afirma que é uma atividade característica dos arquivos permanentes, sendo a "chave de utilização dos arquivos como que fontes primárias da história". Nesta ótica, os usos informativos dos documentos nos arquivos permanentes podem ser considerados distintos dos usos informativos dos documentos junto a entidade produtora, ainda nos setores de trabalho, ou seja, nos protocolos e arquivos correntes. Não negamos o *continuum* do documento de arquivo, sendo este o mesmo em todas as idades (corrente, intermediária e permanente) e em todo orgânico. No entanto, seus usos apresentam particularidades que precisam ser analisadas.

Nos arquivos correntes, os usos informativos estão relacionados às próprias atividades da entidade produtora, podem ajudar na tomada de decisão, permitir o controle fiscal, aprimorar funções, ter ciência das decisões tomadas, identificar em que momento estão os processos de trabalho, entre outros usos. Já os documentos dos arquivos permanentes podem ter diferentes usos uma vez que o seu público é ampliado. O acesso aos documentos dos arquivos correntes está restrito a sua entidade produtora. Já os documentos nos arquivos permanentes têm o seu acesso franqueado a qualquer interessado.

Diversos autores do campo da arquivística como Cruz Mundet (1994), Mendo Carmona (1995), Heredia Herrera (1991), Schellenberg (2006), Rosseau e Couture (1998), Bellotto (2006) estabeleceram diferenças acerca dos valores de uso dos documentos de arquivo junto à entidade produtora e dos documentos junto aos arquivos históricos. Há consenso entre os autores de que os documentos nas unidades de trabalho possuem valor primário, isto é, estão relacionados às utilizações imediatas e administrativas que lhe deram os seus criadores (Rosseau e Couture, 1998, p. 117), seu valor está vinculado às razões para as quais os documentos foram criados (Schellenberg, 2006). Já os

documentos dos arquivos históricos possuem valor secundário, ou seja, eles não estão vinculados diretamente às intenções e razões que motivaram sua criação. As dinâmicas de uso para viabilizar as atividades e funções cessaram. Nesta perspectiva, podemos pensar os documentos como objetos *semióforos*, utilizando uma categoria de análise de Pomian<sup>(1)</sup>. Os documentos são desligados dos seus valores imediatos para ocuparem funções sociais mais próximas aos usos referenciais, culturais, históricos, testemunhais.

O valor primário define-se como sendo a qualidade de um documento baseado nas utilizações imediatas e administrativas que lhe deram os seus criadores, por outras palavras, nas razões para as quais o documento foi criado. A noção de valor primário está diretamente ligada à razão de ser de documentos e recobre exatamente a utilização dos documentos para fins administrativos. [...] O valor secundário define-se como sendo a qualidade do documento baseada nas utilizações não imediatas ou científicas. Esta qualidade radica essencialmente no testemunho privilegiado e objetivo que o documento fornece (Schellenberg, 2006, p. 117)

Reconhecemos os esforços da área em estabelecer as necessárias distinções entre os usos dos documentos nos arquivos correntes e permanentes, no entanto, ressaltamos que nossa proposta não é limitá-los, nem estabelecer barreiras e demarcações. Os usos dos documentos podem ser múltiplos e ubíquos. No entanto, reconhecemos que as demandas informacionais nos arquivos correntes são diferentes daquelas dos arquivos permanentes pelos documentos apresentarem dinâmicas diferentes. Os documentos nos arquivos históricos, não fazem parte da dinâmica dos processos de trabalho, não estão presentes nas relações entre os sujeitos e unidades que compõem a entidade produtora, não justificam sua presença pela necessidade de viabilizar atividades e funções da entidade produtora.

O foco de nossa análise encontra-se nos usos informacionais dos documentos de arquivos correntes e protocolos. Entendemos que os usos informacionais são afetados pela dinâmica dos processos de trabalho das unidades que compõem a entidade produtora. Diante desta perspectiva, consideramos necessário pensar formas de recuperação da informação que não estejam limitadas ao processo de classificação, tampouco à descrição pelos motivos já mencionados.

### 3. Protocolos, arquivos correntes e os processos de trabalho nas universidades

Os protocolos, unidades de registro, procuram registrar não somente a entrada e/ou produção de um documento, mas o trâmite dos processos de trabalho a fim de facilitar a própria organização dos documentos (Foscarini, 2009, p. 24). Os protocolos agrupam em um mesmo processo todos os documentos relacionados a mesma atividade. O registro dos trâmites pelo protocolo também facilita a percepção do contexto de produção e de uso dos documentos dentro da entidade produtora.

Cada ação realizada no processo deve ser registrada pelos sistemas de protocolo. De acordo com Foscarini (2009, p. 26), os protocolos são unidades essenciais em um sistema de arquivo confiável, seu uso é obrigatório a todos os órgãos públicos. O registro de um documento no protocolo significa certificar o exato momento em que ele foi produzido ou recebido, identificando-o de modo único, inserindo-o na classificação e realizando sua indexação de forma precisa, colocando a unidade documental dentro de seu contexto processual e arquivístico. Por essas características cada unidade de trabalho deve possuir seu próprio protocolo, pois a produção e recebimento dos documentos se dão no desenvolvimento dos processos de trabalho, e os protocolos devem ser integrados pelo sistema de arquivo da entidade.

No Brasil, em especial na Administração Pública, os arquivos correntes atuam em conjunto com os protocolos nas unidades de trabalho, geralmente ocupando o mesmo espaço físico inclusive. Na Universidade Federal Fluminense, por exemplo, na qual os protocolos e arquivos correntes são o campo desta pesquisa, os processos após tramitarem ou enquanto aguardam a próxima ação prevista são arquivados nos arquivos correntes localizados nos próprios protocolos setoriais. As características que distinguem protocolos de arquivos correntes, portanto, não são tão claramente delimitadas. Aquelas que podemos ressaltar é que os protocolos funcionam para registrar, controlar o fluxo documental, informar sobre o trâmite dos documentos em um sistema de arquivo; já os arquivos correntes ordenam e armazenam os documentos que estão em plena vigência, isto é, fazem parte da dinâmica dos processos de trabalho. Ainda sobre as diferenças entre protocolos e arquivos correntes, Schellenberg ressaltou:

A única diferença aparente é que os serviços alemães usam o registro ou o protocolo para controlar

a entrada e o movimento de documentos, mas ainda aqui a diferença pode ser considerada mais aparente do que real, pois os serviços de arquivos correntes americanos podem se valer dos arquivos de prosseguimento ou fichas-lembretes ou de qualquer outro dispositivo semelhante, para conseguir o mesmo controle (Schellenberg, 2006, p. 113)

Outra característica que distingue protocolos de arquivos correntes é a própria existência das unidades. Enquanto, o surgimento do protocolo está relacionado às necessidades de controle da produção e do trâmite dos documentos, método surgido entre os séculos XVII e XVIII, o surgimento dos arquivos correntes está ligado à teoria das três idades e à gestão de documentos, fundamentos arquivísticos mais contemporâneos.

Segundo Rosseau e Couture (1998), o crescimento exponencial dos documentos a partir dos anos de 1940, devido o aumento da complexidade das administrações e da realização de suas atividades através de documentos, fez com que os arquivos voltassem seus esforços para elaborar maneiras de melhor organizar sua documentação. Como demonstraram Rosseau e Couture (1998), Mendo Carmona (1995), Heredia Herrera (1991), Schellenberg (2006) e outros autores da Arquivística, razões de espaço e de ordem econômica levaram aos arquivos a elaborar operações que evitassem o acúmulo de documentos durante demasiado tempo nos locais de administração de origem. No entanto, o descarte ou a transferência precisavam passar por rigorosos controles para evitar a migração precoce de documentos importantes ainda no desenvolvimento das atividades, evitando também a eliminação de documentos de valor permanente, que gerassem efeitos jurídicos, legais, fiscais ou administrativos ainda. Assim, surgiu a *teoria das três idades* para melhor organizar os documentos, dividindo-os em três períodos: denominam de *arquivos correntes*, *documentos ativos*, *arquivos de gestão*, *current archives*, *current records*, os documentos que têm alta taxa de utilização. Por outro lado, empregam-se as expressões *arquivos intermediários*, *archivos intermedios*, *documentos semi-ativos*, *semi-current records*, *semi-active records*, para documentos que por razões legais, financeiras ou administrativas precisam ser preservados, mas que são utilizados com menos frequência que os documentos nas unidades de trabalho. Enfim, utiliza-se expressões como *arquivo definitivo*, *arquivo de custódia*, *arquivo permanente*, *documentos inativos*, para designar os documentos que deixam de ter valor previsível no plano administrativo ou financeiro, mas que possuem valores históricos, de teste-

munho, de referência e geram direitos ou obrigações legais de modo permanente. A teoria das três idades surgiu para facilitar a organização dos documentos, no entanto, como ressalta Cortés Alonso (2011), essa divisão serve para facilitar a manipulação e compreensão da massa documental, não deve se perder de vista a integralidade dos documentos. Outra importante teoria arquivística está associada às três idades, a *gestão de documentos*.

De acordo com Llansó Sanjuan (2006, p. 42), para entendermos o conceito de gestão de documentos é necessário compreendermos primeiramente o ciclo de vida dos documentos, originalmente chamada teoria das três idades. O objetivo de sua formulação é garantir a presença do arquivista e dos métodos que aplica em cada fase de organização dos documentos, garantindo desse modo que eles recebam o tratamento adequado enquanto existirem.

Os estudos acerca da gestão de documentos, desde o aparecimento do conceito após a II Guerra Mundial, influenciados pela teoria da Administração Científica baseada nos pressupostos tayloristas, dedicaram seus esforços na organização e classificação dos documentos de arquivo, sobretudo nos arquivos correntes e intermediários, visando a eficiência e a eficácia na execução das atividades da entidade produtora.

Para Cruz Mundet (1994, p. 143), a gestão de documentos é o conjunto de tarefas e procedimentos orientados a conseguir uma maior eficácia e economia na utilização dos documentos por parte das administrações. Este processo se estende ao ciclo de vida completo dos documentos, desde sua produção até a sua eliminação final ou seu envio ao arquivo para sua conservação permanente. O objeto está dirigido a assegurar a preservação de uma documentação adequada, evitar o não essencial, simplificar os sistemas de criação e uso da papelada, melhorar a forma em como se organizam e recuperam os documentos, proporcionar o cuidado adequado e o armazenamento a baixo custo dos documentos nos arquivos intermediários e assegurar a eliminação adequada dos documentos.

Llansó Sanjuan (2006, p. 46) defende que um bom programa de gestão de documentos deve incluir: a) plano de classificação, b) programa de conservação ou tabela de temporalidade, c) gestão dos arquivos intermediários e dos documentos semi-ativos, d) transferência e seleção de documentos inativos destinados a ser conservados, e) eliminação de documentos inativos sem valor histórico.

A literatura arquivística tem se preocupado com a organização dos arquivos para viabilizar as atividades e funções das unidades de trabalho a partir da gestão de documentos; outra preocupação tem sido com a avaliação dos documentos para evitar o extravio daqueles importantes às atividades da entidade produtora e diminuir os custos com armazenamentos desnecessários. No entanto, os esforços para se pensar em métodos e operações nos arquivos correntes que facilitem a recuperação da informação nas unidades documentais ainda não atingiram a mesma proporção no campo arquivístico do que as operações anteriormente mencionadas.

Abordando as universidades como entidades produtoras, precisamos considerar suas especificidades em relação às suas dinâmicas organizacionais, estruturas hierárquicas, práticas decisórias que diferem de outros tipos de organização. Os processos de trabalho são mais autônomos, respondendo muitas vezes apenas ao seu departamento, instituto ou faculdade. Os trâmites da documentação nem sempre seguem os caminhos usuais pela própria cultura de liberdade de ação e independência na tomada de decisão que a universidade possui. Assim, docentes, administradores, coordenadores, colegiados, reitores, entre outros funcionários da instituição não necessariamente seguem os padrões e ordenamentos para a produção e registro de seus documentos, previstos pela gestão de documentos da universidade. Cohen e March (1986) nomeiam as universidades por essas características como *anarquias organizadas*.

No entanto, se as atividades dos departamentos, colegiados e institutos não estão sendo registradas pelos protocolos e muitas decisões não estão sendo incorporadas aos sistemas de informação, percebe-se que o problema está na própria elaboração e execução dos programas de gestão de documentos e das unidades de informação que não estão baseadas nas práticas administrativas e nos processos de trabalho, mas em representações abstratas que não atendem as especificidades da entidade produtora, no caso, a universidade.

Yakel e Bost (1994) apontam que a maioria das abordagens sobre arquivos universitários está voltada para a investigação do comportamento e das necessidades dos usuários nos arquivos permanentes, no entanto, os administradores, em muitos casos, são a primeira população usuária dos arquivos.

As autoras afirmam que os arquivistas têm sido lentos em estudar os usos administrativos dos arquivos universitários. Elas qualificam como uso administrativo qualquer demanda direciona-

da aos arquivos para dar assistência aos administradores universitários, aos funcionários da administração, aos docentes, ou qualquer atividade ou função relacionada à missão da universidade.

Observamos, a partir de nosso trabalho na Universidade Federal Fluminense, que as pesquisas realizadas nos arquivos pelos administradores e funcionários, geralmente, são para resolver questões específicas, com começo e fim imediatos. Os documentos de arquivo em unidades de trabalho apresentam características peculiares: podem ser soluções potenciais para a resolução de problemas administrativos, assim esperam por essas demandas, ao mesmo tempo, estão em circulação, pois precisam viabilizar e mediar outras ações dos processos de trabalho. Os arquivistas necessitam elaborar meios que disponibilizem as informações de tais documentos e localizem-nos no exato momento da demanda gerada pelo administrador.

Yakel e Bost (1994) afirmam que muitos problemas administrativos nas universidades não são resolvidos pela falta de comunicação entre os arquivos e as unidades de trabalho. Segundo as autoras, isto se dá por causa da falta de gestão de documentos e da má administração da base de dados dos arquivos que não analisam os usos de sua documentação. Concluímos então que os estudos sobre o comportamento de busca dos usuários na recuperação das informações também precisam ser considerados para melhorarmos os sistemas de arquivo.

#### **4. Indexação: processo baseado nos usos dos documentos**

O arquivo como um sistema de recuperação da informação precisa basear seus procedimentos e mecanismos a partir da análise da maneira em que os pesquisadores-administradores definem as suas ações, atividades e funções em termos de assunto. Isto se dá porque os administradores têm usos particulares da informação, uma vez que, muitas vezes, eles são os próprios produtores das informações que demandam.

Yakel e Bost (1994, p. 602) evidenciam em seu estudo que os usuários administrativos possuem demandas informacionais diferentes dos usuários dos arquivos permanentes, estas estão relacionadas na execução das suas atividades, as respostas precisam ser imediatas devido as restrições administrativas de tempo no desenvolvimento dos processos de trabalho.

Por esses usos informativos dos documentos de arquivo pelos administradores é que a gestão

de documentos deve priorizar a indexação para facilitar a recuperação da informação, da mesma forma que faz com a avaliação dos documentos, visando organização dos arquivos. Os processos de trabalho precisam dos documentos que viabilizem seu trâmite e muitas vezes o conteúdo informacional contido neles. Os usos administrativos e os usos informativos dos documentos são simultâneos e intrínsecos, não devem tratados de formas isoladas nos arquivos.

A teoria arquivística dedicou-se durante um longo período aos estudos sobre organização dos arquivos e que os procedimentos adotados respeitassem os princípios da proveniência e da ordem original, ademais tais procedimentos também deveriam garantir a organicidade dos documentos de arquivo, sendo possível a recuperação do contexto. No entanto, urge pensar em processos que melhorem a recuperação dos documentos de arquivo e das informações contidas neles. Estes processos devem estar pautados nas análises dos usos dos documentos dos arquivos e nos comportamentos de busca da informação por seus usuários.

A indexação apresenta-se como processo capaz de viabilizar a recuperação dos documentos e de seus conteúdos informacionais nos protocolos e arquivos correntes universitários. De acordo com Wellisch (1995, p. 15), o documento é “um meio sobre ou em que a mensagem é registrada”. Em consequência, é na operação da indexação, mediante as etapas de análise e tradução dos documentos com base no seu uso, que as mensagens (conceitos-chave) dos documentos são atribuídas e nomeadas de forma condensada, podendo-se para isto utilizar uma linguagem documentária verbal (lista de cabeçalhos de assunto, tesouros, etc). Entretanto, é conveniente lembrar que estas padronizações são acordos, regras instáveis, cuja eficácia está vinculada aos usos informativos que cada comunidade discursiva tem em seu micro-universo.

O processo de indexação não é uma atividade nova adotada nos arquivos. Os instrumentos de pesquisa historicamente utilizam procedimentos da indexação em seus catálogos, índices e repertórios. Em 1955, Martin já indicava a necessidade de aprimorar a padronização das entradas dos assuntos nos índices dos documentos de arquivo. A *Norma Brasileira de Descrição Arquivística* (Nobrade) estabelece como área de descrição, além das sete previstas na *ISAD(G)* (identificação, contextualização, estrutura, condições de acesso e uso, fontes relacionadas, notas e controle da descrição), a área oito, nomeada de *pontos de acesso e indexação de*

*assuntos*, prevendo a criação de políticas de indexação, a produção de índices e vocabulários controlados. Essas abordagens sobre indexação apresentam a característica em comum de realizar representações descritivas da informação, com a padronização de catálogos, inventários e índices.

A compreensão da indexação da *Nobrade*, tal como em Borko e Bernier (1978, p.8): "processo de analisar o conteúdo informacional dos registros do conhecimento e sua expressão na linguagem do sistema de indexação", ou em Vanslype (1991, p. 137) "operação que consiste em enumerar os conceitos sobre os quais trata um documento e representá-los por meio de uma linguagem combinatória: lista de descritores livres, lista de autoridades e o thesaurus de descritores", é importante na recuperação dos documentos uma vez que está baseada em um uso documento que é o querer dizer do produtor do documento. No entanto, considerar apenas tal uso no processo de indexação é limitador e se mostra insuficiente para a recuperação da informação.

Na literatura de indexação encontramos algumas abordagens quanto ao procedimento de realização da indexação nos serviços de informação e podem ser resumidas nas seguintes perspectivas: indexação centrada no documento, indexação centrada no usuário e indexação centrada no domínio do conhecimento. Neste trabalho, gostaríamos de acentuar esta última perspectiva que foi proposta por Mai (2005). O autor sugere que uma série de questões sejam observadas no âmbito institucional e quatro dimensões consideradas no processo de indexação: análise do domínio, necessidades dos usuários, desempenho dos indexadores e análise do documento.

Segundo Blair (2003, p. 4), a recuperação da informação é, primeiramente, um processo linguístico. Precisamos descrever o que queremos e como queremos recuperar. No entanto, a descrição intelectual do conteúdo pode ser realizada erroneamente de diferentes maneiras. Por isso, é aconselhável basearmos a representação nos usos da informação, pois a probabilidade de descrevermos de forma efetiva os conteúdos dos documentos é maior. Os usos que as pessoas fazem da informação, a filosofia da linguagem e a comunicação científica tiveram um interesse secundário na recuperação da informação e tal situação se acentua no contexto arquivístico.

Neste sentido, Mai (2005), leitor de Blair, propõe uma mudança de abordagem da indexação, em vez deste processo centrar-se no documento e

em seu texto, voltar-se à análise dos usos da informação e ao comportamento de busca dos sujeitos na recuperação da informação, além da análise do domínio. De acordo com o autor, a indexação trabalha necessariamente com a função e o significado do documento alvo de análise que só podem ser definidos a partir dos usos que as pessoas fazem dele.

Os estudos *pragmáticos* relacionados à semântica nos dão importantes contribuições para trabalhar o processo de indexação. Este processo poderá ser potencializado se passarmos a baseá-lo nas práticas cotidianas, naquilo que as pessoas fazem com a informação e na análise de domínio.

Portanto, devemos reavaliar a crença na homogeneidade dos signos para elaborar índices, vocabulários controlados, linguagens documentárias. Em vez disso, devemos voltar nossas atenções para os significantes e os destinatários, mais do que pensar em significados universais, sublimados do cotidiano das unidades de informação, neste caso, dos protocolos e arquivos correntes universitários. A informação, assim como seu sentido, não possui uma essência apriorística, ela se dá em seu uso. "A informação é aquilo que pode ser considerado informativo para alguém" (Frohmann, 2001, p. 14). O que pode ser informativo para uma comunidade discursiva, em um contexto sócio-histórico, em uma instituição, não necessariamente servirá a outras comunidades em contextos distintos.

Desse modo, para se realizar o processo de indexação, que segundo Beghtol (1986), trata-se de estabelecer os assuntos dos documentos e sistematizá-los em linguagem controlada, é preciso também considerar o contexto de produção e uso da informação como apontam Silva e Fujita:

Durante a atividade de indexação e no momento de atribuir o termo que melhor represente uma palavra, o indexador deve considerar o contexto em que a palavra será usada, tomando como base, por exemplo, como os usuários interessados no documento recuperariam por "aquele" determinado termo (Silva & Fujita, 2004, p. 152)

No caso dos protocolos e arquivos correntes universitários, há particularidades nos usos da informação que carecem ser consideradas. Nos arquivos correntes, os documentos estão em função da administração, de seus próprios produtores, para servir ao trâmite dos assuntos e das ações do cotidiano. A indexação deve servir para este fim, isto é, localizar os documentos e as informações que são solicitados com o fim de

utilizar seus dados ou certificar os atos que contém (Cortés Alonso, 2011).

Em face do exposto, entendemos que o perfil de usuário no arquivo corrente é diferente do usuário do arquivo permanente. O usuário do arquivo corrente é ao mesmo tempo produtor da informação. A sua demanda informacional requer uma recuperação rápida e eficiente para a tomada de decisão, para o andamento e execução de uma atividade, para a conclusão de um ato. O não atendimento de determinada demanda informacional pode gerar algum dano na execução da atividade administrativa ou mesmo o seu impedimento.

Outra questão que deve ser analisada pelo protocolo/arquivo corrente no processo de indexação da informação de seus documentos são as estratégias de busca de seus usuários, por ser também produtor da informação, seu comportamento de pesquisa estará relacionado às configurações do documento, pois enquanto produtor, o usuário conhece o próprio alvo de sua demanda informacional.

De acordo com Smit (2011), no arquivo corrente, os documentos visam a viabilização de transações institucionais ao registrá-las. O usuário está envolvido no processo, participa de alguma forma (informacional, decisória, avaliadora etc) da instituição.

Diante desta perspectiva, a indexação nos protocolos e arquivos correntes universitários não deve estar pautada apenas na análise e síntese dos textos dos documentos. Como afirma Barité (2007), a indexação deve estar pautada na opinião especializada, ou seja, na opinião dos utilizadores do sistema de informação, além do contexto cultural, social, histórico, organizacional que as unidades de informação estão inseridas. O autor aponta que a indexação deve considerar a garantia organizacional que está relacionada aos usos e às rotinas das pessoas que trabalham e realizam atividades nas organizações de grandes dimensões. Este comportamento gera uma sublinguagem particular que inclui termos próprios que precisam ser considerados na elaboração das linguagens documentárias.

## 5. Considerações Finais

Destarte, procuramos registrar as primeiras impressões acerca dos processos de classificação e indexação nos protocolos e arquivos correntes universitários. No entanto, admitindo os aspectos arbitrários, restritivos e instáveis de tais processos. Devido a tais características propomos o afastamento de teorias representa-

cionistas da informação. A sublimação dos contextos da informação em prol de generalizações e padronizações causa enganos, pois naturaliza os processos de classificação e indexação como se fossem entidades superlativas e abstratas. Estas operações como formas de enquadramentos, regras e convenções somente são eficazes se estiverem em consonância com os usos dos documentos e das informações que os indivíduos fazem nas unidades de informação.

A classificação nos protocolos e arquivos correntes universitários serve para melhor compreender a própria entidade produtora, assim como as suas atividades e funções. Os documentos encontram-se nessas unidades para viabilizar ações, os processos de trabalho de seu produtor. O processo de classificação é uma tentativa da instituição entender o seu universo e suas práticas. Ele não surge para recuperar a informação, embora possa fazê-lo, no entanto, não é sua prioridade.

A indexação apresenta-se como um processo capaz de atender as necessidades de recuperação da informação nos protocolos e arquivos correntes universitários. Como processo de atribuição de assuntos dos documentos, ela deve ser baseada nos usos da informação e no comportamento de busca que os sujeitos possuem ao acessar a unidade de informação.

As universidades possuem especificidades em suas atividades que precisam ser consideradas no desenvolvimento de seus esquemas classificatórios e nas políticas de indexação como a autonomia de decisão dos colegiados, os processos de trabalho isolados nos institutos, a diversidade de missões como ensino, pesquisa e extensão, entre outras questões.

Esta investigação procurou discutir a organização e a recuperação da informação no âmbito da Arquivística e na interlocução com a Ciência da Informação. A compreensão de tais processos como a classificação e a indexação se deu a partir do mapeamento dos usos do documento e da informação nos protocolos e arquivos correntes universitários. Por fim, é oportuno destacar que esta análise parte de uma pesquisa que está sendo realizada junto ao Programa de Pós Graduação em Ciência da Informação da Universidade Federal Fluminense, em nível de doutorado.

## Notas

- (1) Pomian cria a categoria *objetos semióforos* para estabelecer a diferença entre os objetos do uso cotidiano e aqueles que são isolados, separados das funções para os quais foram criados a fim de serem cultuados ou servirem de testemunhos his-

tóricos. *Vf. Pomian, K. (1984). Coleção. In: Enciclopédia Einaudi. Porto: Casa da Moeda, 51-86.*

## Referencias

- Aristóteles. (1990). *Metafísica*. (V. García Yebra, Trad.) Madrid: Editorial Gredos.
- Barité, M. (2007). La garantía literaria: vigencia y proyección teórico-metodológica. *Anais do VIII Enancib*.
- Beghtol, C. (1986). Semantic validity: concepts of warrant in bibliographic classification systems. *Library Resources & Technical Services*, 109-123.
- Bellotto, H. (2004). *Arquivos permanentes: tratamento documental*. Rio de Janeiro: FGV.
- Benjamin, W. (1987). Desempacotando minha biblioteca. In: W. Benjamin, *Rua de mão única* (R. R. Torres Filho, & J. C. Barbosa, Trads., pp. 227-234). São Paulo: Brasiliense.
- Blair, D. (2003). Information retrieval and philosophy of language. *Annual Review of Information and Technology*, 37(1), 3-50.
- Borko, H., & Bernier, C. (1978). *Indexing concepts and methods*. New York: Academic Press.
- Bureau of Canadian Archivists. (1993). *Rules for archival description- RAD*. Ottawa.
- Cohen, M. C., & March, J. G. (1986). *Leadership and ambiguity: the American College President*. Boston: Harvard Business School Press.
- Conselho Internacional de Arquivos. (2000). *ISAD(G): norma internacional de descrição arquivística*. Rio de Janeiro: Arquivo Nacional.
- Conselho Nacional de Arquivos (Brasil). (2005). *Norma brasileira de descrição arquivística*. Rio de Janeiro: Arquivo Nacional.
- Cook, M., & Procter, M. (1989). *A manual of archival description*. London: Gower.
- Cortés Alonso, V. (2008). *Archivos de España y América: materiales para un manual (vol. 2)*. Madrid: Fondo de Cultura Económica de España.
- Cortés Alonso, V. (2011). *Documentación y documentos*. Madrid: Anabad.
- Cruz Mundet, J. R. (1994). *Manual de arquivística*. Madrid: Fundación Germán Sánchez Ruipérez.
- Duranti, L. (1997). *I Documenti Archivistici: la gestione dell'archivio da parte dell'ente produttore*. Roma: Ministero per i beni culturali e ambientali, Ufficio centrale per i beni archivistici.
- Fogl, J. (1979). Relations of the concepts 'information' and 'knowledge'. *The Hague*, 4(2), 21-24.
- Foscarini, F. (2009). *Function-based records classification systems: an exploratory study of records management practices in central banks*. Vancouver: University British Columbia.
- Frohmann, B. (2001). Discourse and documentation: some implications for pedagogy and. *Journal of Education for Library and Information Science*, 42(1), 12-26.
- Heredia Herrera, A. (1991). *Arquivística general: teoría y práctica*. Sevilla: Disputación Provincial de Sevilla.
- Llansó Sanjuan, J. (2006). Sistemas arquivísticos y modelos de gestión de documentos en el ámbito internacional. *Revista Códice*, 2, 53-88.
- Mai, E. (2005). Analysis in indexing: document and domain centered approaches. *Information Processing and Management*, 41, 599-611.
- Martin, D. (1955). Use of cataloging techniques in work with records and manuscripts. *The American Archivist*, 18(4), 317-336.
- Mendo Carmona, C. (1995). El largo camino de la Archivística: de práctica a ciencia. *Revista de Historia de la Cultura Escrita*, 113-132.
- Nietzsche, F. W. (2007). *Sobre verdade e mentira no sentido extra moral*. (F. d. Barros, Trad.) São Paulo: Hedra.
- Oliveira, L. V. (2010). *Modelagem e status científico na descrição arquivística no campo dos arquivos pessoais*. São Paulo: USP.
- Pret, R. L. (2013). Documento e prova: ordenamento jurídico, validações diplomáticas e a Archivística. *Informação Arquivística*, 2(2), 55-73.
- Rousseau, J., & Couture, C. (1998). *Os fundamentos da disciplina arquivística*. Lisboa: Dom Quixote.
- Sabourin, P. (2001). Constructing a function-based classification system: business activity structure classification system. *Archivaria*, 51, 137-154.
- Schellenberg, T. R. (2006). *Arquivos modernos: princípios e técnicas*. Rio de Janeiro: FGV.
- Shera, J., & Egan, M. E. (1961). Exame atual da Biblioteconomia e da Documentação. In: S. C. Bradford, *Documentação* (pp. 15-60). Rio de Janeiro: Fundo de Cultura.
- Silva, M. R., & Fujita, M. S. (2004). A prática de indexação: análise da evolução. *Transinformação*, 16(2), 133-161.
- Smit, J. W. (2011). O processo comunicacional entre o arquivo e seus usuários: algumas questões de linguagem. *Anais do IV Encontro de base de dados sobre informações arquivísticas*, 40-47.
- Soergel, D. (1999). The rise of ontologies or the reinvention of classification. *Journal of the American Society of Information Science*, 50(12), 1119-1120.
- Vanslype, G. (1991). *Los lenguajes de indización: concepción, construcción y utilización en los sistemas documentales*. Madrid: Fundación Germán Sánchez Ruipérez; Pirámide.
- Vickery, B. (1973). *Classification and indexing in science*. London: Butterworths.
- Wellisch, H. (1995). *Indexing from A to Z*. 2nd ed. New York: H.W. Wilson.
- Wittgenstein, L. (1999). *Investigações filosóficas*. (J. C. Bruni, Trad.) São Paulo: Nova Cultural.
- Yakel, E., & Bost, L. (1994). Understanding administrative use and users in university archives. *The American Archivist*, 56, 593-615.

---

# La búsqueda de información en herramientas de descubrimiento y OPAC's: fortalezas y debilidades

*Information Searching Using Discovery Tools and OPACs: Strengths and Weaknesses*

---

**Blanca Rodríguez Bravo (1), Maria da Graça Simões (2), M. Cristina V. de Freitas (3), José Antonio Frías (4)**

(1) Universidad de León, Facultad de Filosofía y Letras, Campus de Vegazana s/n, 24071, blanca.rodriguez@unileon.es

(2) Universidade de Coimbra, Largo da Porta Férrea 3004 530, gsimoes@fl.uc.pt

(3) Universidade de Coimbra, Largo da Porta Férrea 3004 530, cristina.freitas@fl.uc.pt

(4) Universidad de Salamanca, Calle Francisco de Vitoria 6-16, E-37008, frias@usal.es

## Resumen

Con la finalidad de profundizar en las fortalezas y debilidades de Discovery Tools (DTs) y OPACs se han seleccionado cinco universidades españolas y portuguesas que cumplen con los siguientes requisitos: 1) impartir estudios de Información y Documentación, dado que nuestra pertenencia a esta área científica facilita el juicio sobre la pertinencia de los resultados hallados y, 2) implementar diferentes herramientas de descubrimiento. Partiendo de la realización de varias búsquedas temáticas y por autores, el objetivo general de este trabajo se cifra en indagar sobre la eficacia del proceso de búsqueda y la pertinencia de los resultados utilizando ambas herramientas. Los objetivos específicos planteados son: determinar la exhaustividad de los resultados pertinentes ofrecidos en la consulta paralela de OPACs y DTs; analizar dichos resultados; identificar las funcionalidades de filtro y su ordenación en los DTs y en los OPACs; determinar la facilidad para refinar los resultados y/o ampliar la búsqueda haciendo uso de facetas, índices o búsqueda avanzada; y valorar la presencia y utilidad de las recomendaciones de documentos relacionados y de las contribuciones/valoraciones de los usuarios en forma de etiquetas o tags, comentarios y puntuaciones.

**Palabras clave:** Búsqueda de información; Evaluación; Herramientas de descubrimiento; OPACs; universidades españolas; Universidades portuguesas.

## 1. Introducción

Han sido numerosos los estudios que han puesto de relieve las limitaciones de los OPACs como instrumentos intermediarios en la recuperación de información. Es notorio que tradicionalmente los catálogos han proporcionado información detallada y de alta calidad sobre libros impresos, si bien a menudo con un formato que no facilita la agregación de datos ni el trabajo

## Abstract

In order to learn more about the strengths and weaknesses of Discovery Tools (DTs) and OPACs, we selected five universities from Portugal and Spain which met the following requirements: 1) they taught Library and Information Studies, since our expertise in this scientific field would enable us to determine the relevance of the results retrieved, and 2) they had implemented different tools. We conducted several subject and author searches in both tools, with the overall objective of exploring the efficiency of the search process and the relevance of the results retrieved when using each tool. The specific objectives were: to determine the comprehensiveness of the relevant results retrieved in a parallel search of OPACs and discovery tools; to analyse these results; to identify the filter functions for selection and how they were organized in DTs and OPACs; to determine the ease with which the results could be refined and/or the search could be widened by making use of facets, indexes, or advanced search features; and to assess the implementation and usefulness of recommendations of related documents and user opinions/reviews in the form of tags, comments and ratings.

**Keywords:** Evaluation; Discovery tools; Information searching; OPACs; Spanish Universities; Portuguese Universities.

colaborativo; han incluido información sobre revistas de calidad variable y de escasa utilidad para los usuarios; no solían proporcionar información sobre artículos de revista y contenían irregularmente descripciones de recursos electrónicos.

El camino hacia la consecución de la nueva generación de catálogos es reciente. Para ello se han utilizado varias vías: la integración de las

descripciones de revistas y de recursos electrónicos; la incorporación de las aplicaciones web 2.0 al catálogo, convirtiéndolo en un OPAC social, participativo; y la implementación de herramientas de búsqueda federada de los materiales de la biblioteca que han evolucionado hacia Discovery Tools (DTs).

Las bibliotecas pueden adoptar un envoltorio web que permite camuflar su catálogo debajo de otra interfaz de usuario. Esta situación parece ser la predominante en la actualidad. El abandono del OPAC en favor de una nueva generación de herramientas (DTs) es una tendencia posible pero que hoy por hoy no parece hallarse consolidada.

Estudios varios (OCLC, 2009) han constatado las diferencias existentes entre las prioridades de los usuarios finales, en relación a la calidad de los datos de los catálogos, y aquellas de los bibliotecarios. Así, como observan Fagan et al. (2012) la implementación de interfaces de descubrimiento, que enmascaran el catálogo sin sustituirlo, puede considerarse una etapa de compromiso entre las preferencias de ambos.

Desde un punto de vista arquitectónico y visual, los DTs son reconocibles principalmente por su utilización de la búsqueda facetada de un ítem catalogado que permite al usuario añadir o eliminar puntos de acceso y expandir o limitar su búsqueda. No obstante, más allá de las facetas, la utilidad de los DTs radica en la habilidad para combinar múltiples fuentes de metadatos en una única caja de búsqueda sencilla. La combinación de registros MARC, repositorios OAI, bases de datos, imágenes, materiales multimedia y otros recursos online conduce a la obtención exhaustiva de resultados hasta ahora nunca alcanzados de una manera tan sencilla.

El acceso a volúmenes de información ocultos anteriormente en silos de catálogos o bases de datos particulares, sin posibilidad de ser recuperados simultáneamente y desde un único punto de acceso, es una revolución tanto para los investigadores como para los usuarios inexpertos. Así, Breeding (2011) señala que los DTs están pensados para encontrar, no solo para buscar.

Los DTs difieren de los metabuscadores en que estos últimos buscan en múltiples bases de datos y, posteriormente, agregan los resultados. Con la utilización de software de búsqueda federada los resultados ofrecidos dependen tanto de los algoritmos de búsqueda y de los rankings de relevancia del metabuscador como de los correspondientes a cada herramienta utilizada.

Contrariamente, los DTs importan los metadatos a un único índice y aplican un único conjunto de algoritmos de búsqueda y una única fórmula de ordenación de resultados. Los catálogos se extienden para incluir búsquedas a artículos de revistas (WorldCat) y grandes índices de artículos de revistas engloban a los catálogos de biblioteca y a los repositorios institucionales (Summon, Primo Central, EDS, etc.). Cualquiera de las dos rutas conduce al resultado deseado de ofrecer un acceso unificado a todos los recursos que la biblioteca posee o contrata.

En trabajos anteriores se indagó en el estadio de evolución de los catálogos de las bibliotecas universitarias españolas y portuguesas (Rodríguez y Travieso, 2013; Rodríguez et al., 2014). El propósito iba encaminado a conocer qué bibliotecas habían implementado interfaces de descubrimiento, qué herramientas habían sido las elegidas y qué funcionalidades ofrecían. Los resultados mostraron que en 2014 casi el 60% de las bibliotecas universitarias españolas habían instalado interfaces de descubrimiento. En el caso portugués el porcentaje se limitaba al 17%. Todas las herramientas evaluadas permitirían implementar todas o la mayoría de las funcionalidades sometidas a evaluación. En el caso de las bibliotecas españolas las herramientas predominantes en la actualidad, según un estudio de reciente publicación, son: Summon (28,4%) EDS (12,2%), Primo Central (9,5%) y WorldCat Local (8,1%) (Ávila, Ortiz y Rodríguez, 2015).

## 2. Objetivos e metodología

El objetivo general de este trabajo se cifra en indagar sobre la eficacia del proceso de búsqueda y la relevancia y características de los resultados obtenidos en ambas herramientas, OPACs y DTs. Los objetivos específicos planteados son: i) determinar la exhaustividad de los resultados pertinentes ofrecidos en la consulta paralela de OPACs y DTs implementados en universidades españolas y portuguesas; ii) determinar la facilidad para refinar los resultados y/o ampliar la búsqueda haciendo uso de facetas, índices o búsqueda avanzada; iii) indagar en la tipología documental, actualidad y lengua de los 25 primeros documentos relevantes recuperados en las búsquedas; iv) identificar las funcionalidades de filtro (selección) y su ordenación en los DTs y en los OPACs; v) indagar en la existencia y uso de recomendaciones de documentos relacionados y de contribuciones/valoraciones de los usuarios en forma de etiquetas o tags, comentarios y puntuaciones.

Estos objetivos son acordes con los criterios a valorar en la elección de un DT puestos de relieve por Freund, Poehlmann y Seale (2012): actualidad de los resultados, facilidad para filtrar y obtener resultados así como ausencia de duplicados, tiempo de respuesta, contenido obtenido y relevancia adecuada. De todos ellos no hemos atendido al tiempo de respuesta habida cuenta de que esta es inmediata en todas las herramientas analizadas.

Con la finalidad de indagar en las fortalezas y debilidades de DTs y OPACs, se seleccionan las herramientas de universidades que cumplan con el requisito de implementar diferentes DTs y de impartir estudios de Información y Documentación, dado que nuestra proximidad a esta área científica facilita juzgar la pertinencia de los resultados hallados en las búsquedas llevadas a cabo.

Así, las universidades seleccionadas han sido: Universidad de Coimbra, Facultad de Letras (UC), Universidad de León (ULE), Universidad de Murcia (UMU), Universidad de Oporto (UP) y Universidad Politécnica de Valencia (UPV). Las herramientas que tienen implementadas son Drupal, WorldCat Local, Summon, Ebsco Discovery Service (EDS) y Primo Central respectivamente.

Se ha partido de la realización de tres búsquedas de materia y dos búsquedas de autor en la caja de búsqueda sencilla de los DTs y de los OPACs para valorar las diferencias en cuanto al número de documentos recuperados en ambas herramientas. Los resultados obtenidos se han filtrado utilizando la búsqueda por campos de la búsqueda avanzada. Las funcionalidades de filtro o selección de los DTs y de los OPACs se han explorado, así mismo, en el proceso de llevar a cabo las búsquedas mencionadas.

Como señalan Villén-Rueda, Senso y Moya-Anegón (2007) los usuarios realizan búsquedas poco sofisticadas que hemos tratado de emular. Destacan estos autores la escasa frecuencia del uso de operadores en las búsquedas por materias donde se recurre a la combinación con la [y], que el sistema realiza por defecto, en el 66% de las ocasiones. En el caso de la búsqueda por materias se trató de combinar temas específicos y generales. Los términos de la búsqueda fueron: indización automática, evaluación de interfaces y bibliotecas digitales. Por lo que se refiere a las búsquedas por autores se eligieron dos autores anglosajones reputados y cuyas obras son bien conocidas en la Península Ibérica: Frederick W. Lancaster y Jennifer Rowley. Las búsquedas se han efectuado entre el 16 de abril y el 10 de mayo de 2015.

Para valorar la pertinencia de los resultados revisamos los 25 primeros, siguiendo el trabajo de Noe (2012) realizado sobre resultados obtenidos de Summon y Google Scholar.

Los datos recogidos se han presentado en tablas, dando preferencia en la presentación, para su mejor comprensión, a las cifras relativas por delante de las absolutas.

A continuación se realiza el análisis y la discusión de los resultados más relevantes obtenidos.

### 3. Resultados

Tras el análisis de los DTs realizado en 2013 y 2014 poníamos de relieve que el mantenimiento en paralelo de dos interfaces de usuario suponía una indefinición que no servía sino para crear confusión al usuario que no sabe dónde dirigir su búsqueda. En este sentido, el análisis efectuado en esta ocasión ha servido para observar ciertas mejoras. Las cinco universidades exploradas presentan en la página principal de la biblioteca una única caja de búsqueda sencilla que permite efectuar la búsqueda en todos los recursos accesibles.

La Universidad de Murcia es quien facilita más la tarea al usuario. En la primera página muestra que se puede “buscar en el catálogo” o “buscar en todo”. La Universidad de León desde la página principal de la biblioteca facilita la búsqueda diferenciada en el “Catálogo ULE”, en el “Portal” o en el DT “WorldCat ULE”.

La Universidad Politécnica de Valencia da acceso al DT desde dos páginas diferentes. En una de ellas la caja de búsqueda sencilla lanza la búsqueda en el DT denominado “Polibuscador”. Desde otra página, o una vez realizada la búsqueda en la primera, se puede restringir la búsqueda al “catálogo”, a la “biblioteca digital”, a la “bibliografía de las asignaturas” o realizar una “búsqueda global”.

La Universidad de Coimbra en la primera página muestra que se puede buscar en “Todos” o en el “OPAC”. Lo mismo se aprecia en la primera página de la Biblioteca de la Universidad de Oporto. Sin embargo, en este caso, la información no es tan explícita, resaltando más el buscador del DT que el acceso al OPAC.

Sin duda se constata el intento de acercamiento al usuario aunque creemos que solo está bien resuelto en la Universidad de Murcia.

Se ha comprobado en este estudio la existencia de búsqueda simple y avanzada en todas las páginas en cuatro de las universidades analizadas. La única excepción es la Universidad de

Coimbra que dispone tan solo de la caja de búsqueda simple. Se valora positivamente la existencia de una caja de búsqueda simple en cada página que se complementa con un enlace a la búsqueda avanzada.

### 3.1. Exhaustividad y pertinencia de los resultados

Se ha constatado que los DTs son un avance en la consecución del fin último de las bibliotecas que no es otro que el de ayudar a los usuarios a descubrir los contenidos accesibles desde la biblioteca en cualquier formato. No obstante, ello debe lograrse sin menoscabo de la pertinencia de la respuesta (Freund, Poehlmann y Seale, 2012).

Los resultados se ordenan por relevancia por defecto aunque puede modificarse la ordenación bajo otros criterios como la fecha de publicación. En el caso de la ULE se activa por defecto la ordenación por "Biblioteca y relevancia". Los algoritmos de ordenación distan mucho de ser claros al igual que no son transparentes los depósitos documentales donde los DTs realizan sus búsquedas.

Se incluyen en la Tabla I (Anexo 1) los resultados obtenidos utilizando la caja de búsqueda global de los DTs y de los OPACs.

Los resultados muestran que la recuperación de registros es muy superior en los DTs que en los OPACs. Como sabemos, la búsqueda en los OPACs se limita generalmente a la colección en propiedad de la biblioteca y, tradicionalmente, a libros aunque se ha enriquecido en los últimos años con tesis doctorales, capítulos de monografías, ponencias en actas de congresos, artículos de revistas y recursos electrónicos además de documentos en otros formatos como documentos de audio y vídeo. Por el contrario, los DTs ofrecen documentos extraídos de muy variados depósitos, índices y bases de datos además de aquellos documentos propiedad o suscritos por la biblioteca.

Como sabemos no existe sistema que ofrezca un 100% de respuesta y un 100% de pertinencia en la recuperación. Por tanto, cuanto más resultados se recuperen mayor nivel de ruido se hallará. Gracias a la ordenación por relevancia el usuario puede, no obstante, limitar ese ruido comenzando su indagación por los resultados presentados en las primeras páginas.

Con respecto a los DTs se han observado diferencias sustanciales entre el número de documentos recuperados en las universidades de Oporto y Murcia respecto a las universidades restantes. Podría ser que esta situación obedeciera a que EDS y Summon proporcionen acceso a más depósitos de documentos que Primo Central y WorldCat Local. Las empresas propietarias de los DTs no proporcionan información suficiente a este respecto, como ya se ha señalado. En el caso de la Universidad de Oporto la razón puede residir también en su tamaño, superior al de las restantes instituciones analizadas.

En lo que se refiere a las búsquedas por conceptos se puede concluir que cuanto más amplio sea el concepto buscado mayor número de documentos recuperaremos. Así, la búsqueda por "bibliotecas digitales" devuelve más resultados que aquella por "indización automática" en todas las universidades y tanto en los DTs como en los OPACs.

En el caso de temas que contienen una combinación de conceptos donde uno de ellos es más genérico ("evaluación" de interfaces) la recuperación es poco homogénea. Se da la circunstancia de que se pueden encontrar más documentos que en la búsqueda de "bibliotecas digitales" algo que se observa en las universidades de Coimbra, León y Oporto. Generalmente este elevado número de resultados en el DT trae aparejado un elevado número de resultados no relevantes como se observa en la Tabla IV, Anexo 1. En el caso de Murcia y Oporto, el OPAC, por el contrario, no recupera nada.

Respecto a las búsquedas por autor se ha constatado que el número de resultados obtenidos guarda una estrecha relación con el número de partes que compongan el nombre del autor y de lo frecuentes que sean estos componentes. Se han constatado mayores dificultades en la búsqueda de las obras de Lancaster por las distintas formas en que las bibliotecas han registrado esta autoridad: F. W. Lancaster, Frederick W. Lancaster y Frederick Wilfrid Lancaster.

En las búsquedas por autor, los catálogos, en algunos casos, siguen requiriendo que se introduzca el apellido delante. Es el caso del catálogo de las universidades de Murcia, Coimbra y Oporto. En la búsqueda avanzada en el campo de materias de la ULE también se requiere la inversión. En las universidades de Coimbra y Oporto, por el contrario, no se requiere dicha inversión.

Para filtrar o seleccionar los resultados relevantes de las búsquedas cuando la respuesta es tan elevada que hace difícil el cotejo de los registros existen dos posibilidades básicas. Hacer uso de la búsqueda avanzada o utilizar las facetas, en concreto, filtrar por temas, materias o disciplinas, autores, fechas, lenguas, tipos documentales, etc.

Todos los DTs utilizados facilitan diferentes filtros que comentaremos posteriormente. Hemos ensayado la limitación por facetas temáticas y hemos observado que solo la Universidad de Murcia facilita la búsqueda disciplinar. Realizamos un ensayo limitando la búsqueda a Biblioteconomía e Informática para las búsquedas temáticas y los resultados obtenidos eran relevantes en un porcentaje que superaba el 75%.

Las restantes herramientas de descubrimiento no facilitan la búsqueda por disciplinas sino por palabras clave que mezclan términos en español e inglés de manera bastante inconsistente y que dificultan en mayor medida la obtención de resultados pertinentes. Esta situación es puesta de relieve por Breitbach (2012) quien afirma que en los DTs no queda clara la forma en que las facetas temáticas ayudan a descubrir contenido disciplinar específico.

En estos casos la utilización de la búsqueda por campos de la búsqueda avanzada parece más eficaz y, por ello, ha sido la opción elegida. Comentamos a continuación (Tabla II, Anexo 1) los resultados de las búsquedas por campos de materia y autor efectuadas.

Como se puede observar, los resultados obtenidos limitando las búsquedas a un campo específico - materia o autor - son sustancialmente inferiores a los obtenidos de la búsqueda global. En algún caso coinciden porque en la búsqueda por materias o autor no se encuentra la entrada buscada y se ejecuta la búsqueda por palabras, como es el caso de los OPACs de la UC, ULE, UP y UMU.

Los resultados de la búsqueda por campos devuelve, así mismo, más resultados de los DTs que de los OPACs. La mayor parte de ellos provienen de la búsqueda conceptual de "bibliotecas digitales". En esta ocasión es el DT de la ULE la herramienta que más documentos recupera, seguida de las de las universidades de Oporto y Murcia. En el caso de la Universidad de Coimbra la herramienta instalada, Drupal, no facilita la búsqueda por campos.

Describimos a continuación (Tabla III, Anexo 1) el porcentaje que representan los resultados de la búsqueda por campos con respecto a la búsqueda global para valorar la capacidad que como filtro tiene la utilización de la búsqueda avanzada utilizando los campos de materias y de autores.

Observamos que en los DTs la búsqueda por campos filtra los resultados eliminando entre el 52% de los resultados en la ULE y el 99% en la UP. A mayor número de resultados obtenidos

en la búsqueda global más alta es porcentualmente la reducción.

En los OPACs, donde los resultados son menos abultados, la disminución es muy inferior situándose en torno al 50%, alcanzando el 60% en la Universidad de León.

Consideramos, por tanto, que la búsqueda por campos funciona adecuadamente como mecanismo de filtro o selección.

### 3.2. Relevancia y actualidad de los resultados

Como se puede apreciar en la Tabla IV (Anexo 1), destaca la pertinencia en los resultados ofrecidos por los DTs de las universidades españolas. Con respecto a los OPACs, destacan los resultados de la ULE y UMU.

Se aprecian diferencias significativas entre unas búsquedas y otras. La búsqueda en la que se han obtenido resultados más pertinentes es en la de "bibliotecas digitales". En la búsqueda de "evaluación de interfaces", por el contrario, es en la que se ha encontrado más ruido debido a que se utilizan dos conceptos que se han de combinar y uno de ellos es muy genérico, es el caso de "evaluación", como ya se ha puesto de relieve.

En todas las universidades españolas hemos detectado duplicados en los resultados obtenidos en las búsquedas. Principalmente en las acometidas en los DTs y tanto en la búsqueda global como en la búsqueda por campos. Tampoco los OPACs están totalmente exentos de duplicados. En las universidades portuguesas se detectan también duplicados en los dos tipos de búsquedas acometidas en los DTs. En los OPACs ello es especialmente verificable en el caso de las búsquedas por materias. La Tabla V (Anexo 1) muestra la pertinencia de los resultados obtenidos en la búsqueda por campos. Sorprendentemente no se muestran diferencias apreciables. Los resultados más pertinentes siguen siendo los de la búsqueda de "bibliotecas digitales" y pese a la dificultad ya mencionada de las distintas formas en que las bibliotecas han utilizado para las obras de Lancaster, los resultados de la búsqueda de este autor son más pertinentes que los de Rowley en la búsqueda por campos y en la global.

Tampoco resultan más pertinentes los resultados de los OPACs que los de los DTs como habría cabido esperar.

En las Tablas VI y VII (Tipos de los documentos relevantes, Anexo 1) se aprecia el dominio de los libros entre los documentos relevantes recuperados. Este predominio es significativamente

superior en las búsquedas de los OPACs que en las de los DTs donde predominan los artículos de revista. La nota discordante la proporcionan las universidades de León y Oporto, cuyos DTs recuperan más libros que artículos.

Ninguno de los OPACs recupera más que marginalmente artículos de revista. Sin embargo, el de la ULE recupera un elevado número de capítulos de libros resultado de la tarea habitual que la biblioteca lleva a cabo de vaciado de actas de congreso principalmente. Es evidente que no realiza, sin embargo, registros analíticos de revistas.

En el apartado de "otros" se recogen principalmente tesis, proyectos de fin de carrera, trabajos de fin de máster o trabajos de fin de grado con una presencia significativa en el DT de la Universidad de Oporto.

En las mismas Tablas VI y VII (Fechas de los documentos relevantes recuperados, Anexo 1) se aprecia que más de la mitad de los documentos recuperados han sido publicados durante la primera década del siglo XXI y, complementariamente, en la última del siglo XX. Los documentos anteriores a 1990 son escasos o inexistentes.

Por lo que se refiere a los documentos de los últimos años, son significativos principalmente en los DTs de la UP, UPV y UMU. En el caso del OPAC de Murcia la mayor actualidad de los documentos recuperados con respecto a otros OPACs obedece a que los resultados se ordenan por fecha descendente. También en el caso del DT de la ULE, cuando se realiza la búsqueda en el campo de materias, los resultados se ordenan por fecha aunque se solicite su ordenación por relevancia lo que supone un incremento artificial en la actualidad de los documentos recuperados.

Como puede apreciarse en las mismas Tablas VI y VII (Lenguas de los documentos relevantes) la mayoría de los documentos recuperados son documentos en español, portugués y/o inglés. Cabe señalar que las herramientas portuguesas han recuperado documentos en español mientras que las españolas no han recuperado documentos en portugués.

Las restantes lenguas son meramente testimoniales, limitándose, cuando aparecen, al catalán y francés por este orden en el caso de España.

### 3.3. Funcionalidades de filtro y ordenación

La abundancia o la escasez en los resultados obtenidos pueden requerir limitar o extender la búsqueda. La caja de búsqueda única y global

de los DTs privilegia la exhaustividad sobre la pertinencia por lo que suele ser poco frecuente la obtención de resultados demasiado limitados y mucho más común la obtención de resultados abrumadores. Breitbach (2012) apunta la sobrecarga informativa y la creación de falsas expectativas en el usuario como debilidades de los DTs.

Las facetas permiten limitar o extender la búsqueda mostrando conceptos relacionados, formatos, rangos de fechas, idiomas, etc. Como indica Noe (2012) es imprescindible acudir a las facetas para conseguir precisión dado que implican el uso transparente para el usuario de la lógica booleana. No obstante, el usuario suele desconocer su utilidad, que se debe explicar con más claridad, y su disposición no es intuitiva (Holman et al., 2012).

En el caso de los OPACS una recuperación más precisa pasa por la combinación booleana de conceptos en la búsqueda avanzada y por la utilización de índices (Villén-Rueda, Senso y Moya-Anegón, 2007).

Creemos necesario subrayar la utilidad de los índices de autores, títulos y materias que no se hallan presentes con carácter general. Hofmann y Yang (2012) consideraban que el reemplazo definitivo del OPAC por el DT pasa porque los DTs integren de manera generalizada la búsqueda avanzada y las funcionalidades de navegación por índices alfabéticos de autor, título y materia.

En este análisis hemos constatado que la búsqueda avanzada ya se ha generalizado y está disponible tanto para la búsqueda en los OPACS como en los DTs. No así la navegación por índices alfabéticos, inexistentes en todas las herramientas analizadas a excepción del OPAC de la Universidad de Murcia que permite navegar por índices de autor, materias, lugar de publicación, editor, signatura y CDU.

Con respecto a las posibilidades de ampliación de resultados, en Murcia al igual que en León se puede ampliar la búsqueda a resultados ajenos a la biblioteca. La UPV permite extender la búsqueda a referencias sin texto completo e informa sobre el número de ejemplares de cada recurso. Se ha observado que se implementa FRBR pues se muestran conjuntamente las manifestaciones de la misma obra. El mismo documento pero de diferentes ediciones, por ejemplo. Al igual que Metz-Wiseman et al. (2012) consideramos positivamente la organización jerárquica de los resultados siguiendo este modelo, algo que facilitan algunas de herramientas como Primo y WorldCat.

En lo que se refiere a la presentación de los resultados todas las herramientas ofrecen por defecto la ordenación por relevancia salvo ULE donde el DT los presenta por “biblioteca y relevancia”, como se ha apuntado previamente, lo que obliga a seleccionar la ordenación por relevancia específicamente. La UPV facilita además la ordenación por popularidad.

Presentamos en la Tabla VIII (Anexo 1) las 15 funcionalidades de selección que hemos observado en alguna de las herramientas estudiadas.

En general podemos señalar que las funcionalidades de selección son numerosas y variadas y más abundantes en los DTs que en los OPACs, salvo en las universidades de Coimbra y Murcia. No obstante, ninguna de las herramientas facilita el uso de todas las posibilidades de filtro seleccionadas. El abanico mayor de posibilidades corresponde a los DTs de UP, UPV y ULE y al OPAC de UC.

Todas las herramientas permiten seleccionar por temas, materias o disciplinas, autores, idiomas y fechas. Muy frecuente es la posibilidad de seleccionar por localización y formato. Con respecto a la primera, la mayoría de las herramientas permite limitar la búsqueda a una biblioteca o colección particular o bien a una particular ubicación. En relación a la segunda, es preciso apuntar que en algunas herramientas se diferencia claramente entre tipos documentales y formatos y otras presentan conjuntamente ambas posibilidades. También es relativamente común la facilidad de refinar por tipo de publicación o contenido. Algunos DTs permiten limitar la búsqueda a artículos peer-reviewed, algo que apuntan Chapman et al. (2012) es apreciado por los usuarios.

En el extremo opuesto algunas funcionalidades solo se hallan implementadas en una herramienta. En estos casos siempre se corresponden con un particular DT. La Universidad de Oporto facilita la selección por área geográfica y tipo de préstamo.

Consideramos de extrema utilidad la posibilidad de realizar búsquedas específicas de los materiales recomendados en las bibliografías de las asignaturas, posibilidad ofrecida por el DT de la UPV. Es imprescindible la existencia de una transparente y efectiva conexión entre la bibliografía y el DT implementado que permita acceder con facilidad a los recursos recomendados por los profesores en sus guías docentes.

Por lo que respecta a las funcionalidades de ordenación y visualización (Tabla IX, Anexo I), en las universidades de Murcia, Oporto y León los resultados globales son muy superiores en

los DTs en relación a los OPACs. En las demás universidades (UPV y UC) son idénticos. En cuanto a los resultados parciales, los DTs de las universidades de Oporto y León son los que implementan un mayor número de funcionalidades. Por el contrario, en la Universidad de Coimbra solo se halla implementada la funcionalidad relacionada con la relevancia. Asimismo, en lo referente a las funcionalidades, las implementadas con mayor frecuencia ya sea en los OPACs ya sea en los DTs son las más comunes (“fecha de publicación” y “relevancia”). No obstante, sería conveniente que tanto los OPACs como los DTs implementasen el mayor número posible de funcionalidades de ordenación y visualización, dados los beneficios evidentes que tienen para los usuarios en el ámbito de la identificación, selección y obtención de registros/obras de su interés.

### 3.4. Recomendaciones de documentos relacionados, contribuciones y valoraciones de los usuarios

En un trabajo anterior (Rodríguez y Travieso, 2013) se constataba que Summon mostraba carencias en la implementación de funcionalidades relacionadas con los aspectos sociales o participativos del OPAC pues no realizaba recomendaciones a los usuarios ni admitía sus contribuciones. En relación a Primo Central en el apartado de recomendaciones y/o materiales relacionados permitían no sólo sugerencias de materias relacionadas sino también de obras relacionadas o títulos similares.

En lo que se refiere a la aportación de contenido enriquecido, las nubes de etiquetas eran poco comunes. En cuanto a las contribuciones de los usuarios, lo más habitual era la posibilidad de añadir etiquetas de descripción, pero algunos sistemas permitían aportar comentarios, críticas o reseñas, como era el caso de WorldCat Local.

En el presente análisis (Tabla X, Anexo 1) se ha constatado un avance significativo en el uso de los aspectos sociales o colaborativos de las herramientas estudiadas. Por lo general, observamos su implementación en ambas herramientas. No obstante, cabe señalar que son los DTs de las universidades de Coimbra y León y los OPACs de las universidades de Oporto y Murcia los que mejor cumplen los requisitos colaborativos. De las cuatro funcionalidades colaborativas observadas, las más frecuentes son los “comentarios” y las “etiquetas o tags”. De la observación individual de las herramientas, se pone de relieve que, salvo los DTs de las universidades de

UMU y UP y el OPAC de la UC, todos los demás implementan los comentarios.

Las etiquetas o tags, por su parte, no se implementan en los DTs de las universidades de UP y UMU ni en los OPACs de ULE y UC. Aparte de los DTs de la Universidad de León y Coimbra y el OPAC de la Universidad de Oporto no hemos visto puntuaciones, pero sí hemos comprobado que se facilita la ordenación de resultados por popularidad en el caso de la UPV, como se ha apuntado anteriormente. A su vez, las “recomendaciones de documentos relacionados” se encuentran tanto en los DTs de UP, ULE y UC, aunque no de modo uniforme, como en los OPACs de UMU y UP.

#### 4. Conclusiones y discusión

La finalidad de los bibliotecarios debe ser ayudar a los usuarios a descubrir los contenidos accesibles desde la biblioteca en cualquier formato, sin descuidar la pertinencia que tradicionalmente ha perseguido la tarea de representación y organización de la información llevada a cabo por ellos mismos. Son muchos y variados los estudios que constatan las diferencias existentes entre las prioridades de los usuarios finales, en relación a la calidad de los datos de los catálogos, y aquellas de los bibliotecarios (OCLC, 2009). Observan Fagan et al. (2012) que la implementación de DTs, que enmascaran el catálogo sin sustituirlo, pueden considerarse una etapa de compromiso en el camino de conseguir satisfacer las preferencias del usuario y el bibliotecario.

Del trabajo realizado podemos concluir que se sigue apreciando cierta indefinición a la hora de dirigir al usuario a la nueva herramienta y/o al catálogo tradicional (Rodríguez y Travieso, 2013). En el caso español solo la UMU presenta al usuario las posibilidades existentes de manera clara y sencilla. Por lo que respecta a las universidades portuguesas, se puede apreciar que la Universidad de Coimbra tiene esta información igualmente explícita.

Los DTs suponen un avance considerable. La combinación del catálogo de la biblioteca, los índices de artículos y otras fuentes de información en una interfaz unificada es una mejora indudable. Pese a esta consideración favorable existen carencias señaladas por Fagan et al. (2012) como es la dificultad para llevar a cabo estrategias de búsqueda disciplinares específicas y la ausencia de control de autoridades que inciden negativamente en la precisión de la recuperación de información.

En el mismo sentido opina Breitbach (2012) quien afirma que en los DTs los metadatos han perdido valor habida cuenta de que la misma terminología tiene distintos significados porque las fronteras entre las disciplinas desaparecen y los mismos términos se utilizan con distintos significados. La información está descontextualizada.

En este estudio hemos constatado que las búsquedas disciplinares y la posibilidad de precisar la búsqueda utilizando índices escasean. Solo la universidad de Murcia lo facilita. De ahí que para filtrar los resultados hayamos acudido a la búsqueda avanzada.

Las búsquedas en la caja única de búsqueda global o en la búsqueda por campos devuelven un número significativamente más alto de resultados en los DTs que en los OPACs. El número de resultados obtenido varía en función de que se busquen conceptos más o menos específicos. Hoseth (2012) apunta que los DTs funcionan bien para búsquedas generales o interdisciplinares.

La relevancia de los resultados obtenidos no parece superior en los OPACs como se esperaba y no se aprecia tampoco una diferencia significativa entre la búsqueda global y la búsqueda por campos.

Los documentos relevantes recuperados son fundamentalmente libros y artículos de revista. La recuperación de libros es más frecuente en los OPACs y la de artículos en los DTs. La presencia de capítulos de libros, recursos electrónicos y otros tipos documentales es marginal. Solo en ULE y UP hemos visto un predominio de libros en las búsquedas en el DT. En el caso de la Universidad de León se ha constatado una considerable abundancia de capítulos en el OPAC.

Los documentos recuperados se adscriben fundamentalmente a la primera década del siglo XXI y a la última década del siglo XX. Son poco frecuentes los documentos anteriores a 1991 y sí hay una cierta presencia de publicaciones de los últimos cinco años afectados, en cierta medida, por la disminución de compras y suscripciones en las bibliotecas universitarias españolas y portuguesas. En relación a las lenguas predominantes de los documentos relevantes analizados cabe destacar que están escritos principalmente en las lenguas de las bibliotecas analizadas y en inglés. Esta última lengua es más frecuente en el caso de los libros. La presencia de otras lenguas es marginal.

Con respecto a las funcionalidades de selección o filtro, en general los DTs ofrecen un mayor

abanico de posibilidades que los OPACS. Todas las herramientas permiten seleccionar por temas, materias o disciplinas, autores, idiomas y fechas. También es frecuente la posibilidad de seleccionar por localización, tipo de publicación, contenido y formato.

Por lo general, las funcionalidades de ordenación y visualización se encuentran implementadas en la mayoría de las universidades analizadas, independientemente de la herramienta, subrayándose que “fecha de publicación” y “relevancia” son las que más se observan tanto en los OPACs como en los DTs. Sin embargo, desde el punto de vista del usuario, dichos resultados no son suficientemente satisfactorios, por lo que se recomienda el incremento en la implementación de dichas funcionalidades.

En relación a los aspectos colaborativos de las herramientas analizadas consideramos positivamente la opción de incluir la contribución de los propios usuarios, fomentando que éstos conciban estas herramientas como algo más cercano y participativo, y la opción de sugerir búsquedas relacionadas o documentos similares. No obstante se observa la limitada utilización de estas funcionalidades sociales. Se requiere una labor de difusión por parte de las bibliotecas para fomentar la participación de los usuarios.

Por último, consideramos conveniente ampliar el contenido enriquecido que se ofrece con cada registro, que se limita mayoritariamente a la inclusión de las portadas, con la aportación de sumarios, reseñas, críticas, etc. que permitan una selección más fundamentada y, por tanto, eficaz.

Es preciso señalar que el número de sedes exploradas ha sido muy limitado. Por los motivos expuestos consideramos necesario seguir abundando en esta línea de trabajo y realizar una exploración y evaluación más exhaustiva que permita conocer el impacto de las diferentes herramientas implementadas y/o sus desarrollos y personalizaciones en la cantidad y relevancia de los resultados hallados.

Así mismo conviene seguir indagando en la influencia del trabajo propiamente bibliotecario en estos mismos resultados. Hemos comprobado que la recuperación de documentos se ve enriquecida por la integración de registros analíticos o de tesis doctorales y otros trabajos de los estudiantes como son los TFM o TFG.

## Referencias

Ávila-García, L., Ortiz-Repiso, V., & Rodríguez-Mateos, D. (2015). Herramientas de descubrimiento: ¿una ventani-

lla única? *Revista Española de Documentación Científica*, 38 (1), enero-marzo, doi: <http://dx.doi.org/10.3989/redc.2015.1.1178>

Breeding, M. (2011). Discovering Harry potterly barn. *Computers in Libraries*, 31 (2), 21-23.

Breitbart, W. (2012). Web-Scale discovery: a library of Babel. En: M. P. Popp, & D. Dallis (eds.). *Planning and implementing resource discovery tools in Academic Libraries*. Hershey, PA: IGI Global, 637-645.

Chapman, S., Dennis, S., Folger, K., & Varnum, K. (2012). Developing a user-centered article discovery environment. En: M. P. Popp, & D. Dallis. (eds.). *Planning and implementing resource discovery tools in Academic Libraries*. Hershey, PA: IGI Global, 194-224.

Fagan, J. C., Mandernach, M. A., Nelson, C. S., Paulo, J. R., & Saunders, G. (2012). Usability test results for a discovery tool in an academic library. *Information Technology and Libraries*, march, 83-112. Disponible en: <http://ejournals.bc.edu/ojs/index.php/ital/article/view/1855/1745>

Freund, L., Poehlmann, C., & Seale, C. (2012). From metasearching to discovery: the University of Florida experience. En: M. P. Popp, & D. Dallis (eds.). *Planning and implementing resource discovery tools in Academic Libraries*. Hershey, PA: IGI Global, 22-43.

Hofmann, M. A., & Yang, S. Q. (2012). "Discovering" what's changed: a revisit of the OPACs of 260 academic libraries. *Library Hi Tech*, 30 (2), 253-274.

Holman, L. Darraj, E., Glaser, J., Hom, A., Mathieson, H., Nettles, D., & Waller, A. (2012). How users approach discovery tools. En: M. P. Popp, & D. Dallis (eds.). *Planning and implementing resource discovery tools in Academic Libraries*. Hershey, PA: IGI Global, 252-267.

Hoseth, A. (2012). Criteria to consider when evaluating eeb-based discovery tools. En: M. P. Popp, & D. Dallis (eds.). *Planning and implementing resource discovery tools in Academic Libraries*. Hershey, PA: IGI Global, 90-103

Metz-Wiseman, M. Griffin, M., Borchert, C. A., & Henry, D. (2012). Best practices for selecting the best fit. En: M. P. Popp, & D. Dallis (eds.). *Planning and implementing resource discovery tools in Academic Libraries*. Hershey, PA: IGI Global, 77-89.

Noe, D. E. (2012). Replicating top user's searches in Summon and Google Scholar. En: M. P. Popp, & D. Dallis (eds.). *Planning and implementing resource discovery tools in Academic Libraries*. Hershey, PA: IGI Global, 225-239.

OCLC (2009). *Online Catalogs: what users and librarians want*. Disponible en: <http://www.oclc.org/content/dam/oclc/reports/onlinecatalogs/fullreport.pdf>

Rodríguez-Bravo, B. & Travieso-Rodríguez, C. (2013). Catálogos de nueva generación en las bibliotecas universitarias españolas: una primera exploración. En: *Informação e/ou conhecimento: as duas faces de Jano. I Congresso ISKO Espanha e Portugal-XI Congreso ISKO España: 7 a 9 de novembro de 2013*. Porto: Faculdade de Letras de Universidade do Porto: Cetac-Media, 864 - 878.

Rodríguez-Bravo, B., Travieso Rodríguez, C., Simões, M. G. de M. & Freitas, M. C. V. de. (2014). Evaluating discovery tools in Portuguese and Spanish Academic Libraries. En: *Knowledge Organization in the 21st Century: Between Historical Patterns and Future Prospects*. Wieslaw Babik. Würzburg: Ergon-Verlag, 2014, 516-523.

Villén-Rueda, L., Senso, J. A., & Moya-Anegón, F. (2007). The use of OPAC in a large academic library: a transactional log analysis study of subject searching. *Journal of Academic Librarianship*, 33 (3), 327-337.

## Anexo 1

| Términos de búsqueda        | Universidad/herramienta |      |        |      |      |      |      |      |     |      |
|-----------------------------|-------------------------|------|--------|------|------|------|------|------|-----|------|
|                             | UMU                     |      | UP     |      | UPV  |      | ULE  |      | UC  |      |
|                             | DT                      | OPAC | DT     | OPAC | DT   | OPAC | DT   | OPAC | DT  | OPAC |
| Indización automática       | 105                     | 6    | 212    | 0    | 42   | 2    | 87   | 3    | 0   | 0    |
| Evaluación de interfaces    | 2624                    | 0    | 6722   | 0    | 403  | 7    | 237  | 5    | 290 | 2    |
| Bibliotecas digitales       | 10557                   | 67   | 1464   | 7    | 841  | 162  | 121  | 135  | 68  | 39   |
| Frederick Wilfrid Lancaster | 176                     | 15   | 358449 | 1    | 180  | 2    | 389  | 10   | 210 | 2    |
| Jennifer Rowley             | 11384                   | 5    | 11485  | 1    | 857  | 6    | 1592 | 9    | 2   | 3    |
| n                           | 24846                   | 93   | 378332 | 9    | 2323 | 179  | 2426 | 162  | 570 | 46   |

Tabla I. Resultados de la búsqueda simple.  
Fuente: Autores

| Términos de búsqueda        | Universidad/herramienta |      |      |      |     |      |      |      |    |      |
|-----------------------------|-------------------------|------|------|------|-----|------|------|------|----|------|
|                             | UMU                     |      | UP   |      | UPV |      | ULE  |      | UC |      |
|                             | DT                      | OPAC | DT   | OPAC | DT  | OPAC | DT   | OPAC | DT | OPAC |
| Indización automática       | 14                      | 6    | 56   | 0    | 8   | 0    | 60   | 3    | -  | 0    |
| Evaluación de interfaces    | 10                      | 3    | 258  | 0    | 13  | 4    | 23   | 5    | -  | 0    |
| Bibliotecas digitales       | 308                     | 61   | 240  | 7    | 276 | 156  | 896  | 81   | -  | 42   |
| Frederick Wilfrid Lancaster | 529                     | 15   | 835  | 1    | 32  | 2    | 237  | 10   | -  | 2    |
| Jennifer Rowley             | 307                     | 5    | 608  | 1    | 217 | 6    | 974  | 6    | -  | 2    |
| n                           | 1168                    | 90   | 1997 | 9    | 546 | 168  | 2190 | 105  | 0  | 46   |

Tabla II. Resultados de la búsqueda avanzada.  
Fuente: Autores

| Tipos de búsquedas/% | Universidad/herramienta |       |        |       |       |       |       |       |      |      |
|----------------------|-------------------------|-------|--------|-------|-------|-------|-------|-------|------|------|
|                      | UMU                     |       | UP     |       | UPV   |       | ULE   |       | UC   |      |
|                      | DT                      | OPAC  | DT     | OPAC  | DT    | OPAC  | DT    | OPAC  | DT   | OPAC |
| Simple               | 24846                   | 93    | 378332 | 9     | 2323  | 179   | 2426  | 162   | 570  | 46   |
| Avanzada             | 1168                    | 90    | 1997   | 9     | 546   | 168   | 2190  | 105   | 0    | 46   |
| % simple             | 95,51                   | 50,81 | 99,47  | 50,00 | 80,96 | 51,58 | 52,55 | 60,67 | 100  | 50   |
| % avanzada           | 4,48                    | 49,18 | 0,52   | 50,00 | 19,03 | 48,41 | 47,44 | 39,32 | 0,00 | 50   |

Tabla III. Porcentajes de los resultados de la búsqueda simple y de la búsqueda avanzada.  
Fuente: Autores

| Términos de búsqueda        | Universidad/herramienta |      |     |      |     |      |     |      |    |      |
|-----------------------------|-------------------------|------|-----|------|-----|------|-----|------|----|------|
|                             | UMU                     |      | UP  |      | UPV |      | ULE |      | UC |      |
|                             | DT                      | OPAC | DT  | OPAC | DT  | OPAC | DT  | OPAC | DT | OPAC |
|                             | %                       |      | %   |      | %   |      | %   |      | %  |      |
| Indización automática       | 68                      | 24   | 100 | 0    | 100 | 8    | 32  | 12   | 0  | 0    |
| Evaluación de interfaces    | 32                      | 0    | 0   | 4    | 36  | 0    | 36  | 20   | 8  | 0    |
| Bibliotecas digitales       | 100                     | 100  | 100 | 32   | 100 | 100  | 100 | 100  | 84 | 100  |
| Frederick Wilfrid Lancaster | 100                     | 60   | 16  | 4    | 100 | 8    | 100 | 40   | 8  | 8    |
| Jennifer Rowley             | 100                     | 0    | 8   | 4    | 36  | 20   | 40  | 36   | 8  | 12   |

Tabla IV. Resultados relevantes de la búsqueda en todos los campos (n=25).  
Fuente: Autores

| Términos de búsqueda        | Universidad/herramienta |      |     |      |     |      |     |      |    |      |
|-----------------------------|-------------------------|------|-----|------|-----|------|-----|------|----|------|
|                             | UMU                     |      | UP  |      | UPV |      | ULE |      | UC |      |
|                             | DT                      | OPAC | DT  | OPAC | DT  | OPAC | DT  | OPAC | DT | OPAC |
|                             | %                       |      | %   |      | %   |      | %   |      | %  |      |
| Indización automática       | 20                      | 12   | 100 | 0    | 28  | 0    | 20  | 0    | 0  | 0    |
| Evaluación de interfaces    | 12                      | 0    | 4   | 0    | 16  | 0    | 16  | 0    | 0  | 0    |
| Bibliotecas digitales       | 100                     | 100  | 100 | 24   | 100 | 100  | 100 | 100  | 0  | 100  |
| Frederick Wilfrid Lancaster | 92                      | 60   | 0   | 4    | 100 | 8    | 100 | 40   | 0  | 8    |
| Jennifer Rowley             | 48                      | 0    | 28  | 4    | 40  | 20   | 36  | 24   | 0  | 8    |

Tabla V. Resultados relevantes de la búsqueda por campos materia y autor (n=25).  
Fuente: Autores

| Criterios           | Universidad/herramienta   |      |    |      |     |      |     |      |    |      |
|---------------------|---------------------------|------|----|------|-----|------|-----|------|----|------|
|                     | UMU                       |      | UP |      | UPV |      | ULE |      | UC |      |
|                     | DT                        | OPAC | DT | OPAC | DT  | OPAC | DT  | OPAC | DT | OPAC |
|                     | Tipos de documentos       |      |    |      |     |      |     |      |    |      |
| Libros              | 24                        | 40   | 8  | 11   | 10  | 30   | 45  | 24   | 11 | 25   |
| Artículos           | 66                        | -    | 32 | -    | 80  | -    | 27  | 3    | 4  | 1    |
| Capítulos de libros | 1                         | -    | -  | -    | -   | -    | -   | 24   | -  | -    |
| Otros               | 9                         | 6    | 16 | -    | 3   | 4    | 5   | 1    | 12 | 4    |
|                     | Fechas de los documentos  |      |    |      |     |      |     |      |    |      |
| 2011-2015           | 22                        | 6    | 21 | -    | 19  | 1    | 7   | 3    | 2  | 5    |
| 2000-2010           | 47                        | 24   | 29 | 9    | 37  | 24   | 42  | 35   | 22 | 21   |
| 1990-2000           | 23                        | 14   | 5  | 2    | 17  | 8    | 20  | 13   | 3  | 3    |
| (...)-1990          | 8                         | 2    | 1  | -    | 20  | 1    | 8   | 1    | -  | 1    |
|                     | Lenguas de los documentos |      |    |      |     |      |     |      |    |      |
| Español             | 53                        | 26   | -  | 1    | 59  | 13   | 38  | 37   | 7  | 1    |
| Portugués           | 1                         | -    | 49 | 2    | -   | -    | -   | -    | 10 | 3    |
| Inglés              | 45                        | 19   | 6  | 8    | 34  | 19   | 39  | 15   | 10 | 26   |
| Otras               | 1                         | 1    | 1  | -    | -   | -    | -   | -    | -  | -    |

Tabla VI. Búsqueda simple. Documentos relevantes.  
Fuente: Autores

| Criterios           | Universidad/herramienta   |      |    |      |     |      |     |      |    |      |
|---------------------|---------------------------|------|----|------|-----|------|-----|------|----|------|
|                     | UMU                       |      | UP |      | UPV |      | ULE |      | UC |      |
|                     | DT                        | OPAC | DT | OPAC | DT  | OPAC | DT  | OPAC | DT | OPAC |
|                     | Tipos de documentos       |      |    |      |     |      |     |      |    |      |
| Libros              | 31                        | 39   | 28 | 8    | 10  | 30   | 44  | 40   | -  | 26   |
| Artículos           | 33                        | -    | 8  | -    | 60  | -    | 23  | -    | -  | 1    |
| Capítulos de libros | -                         | -    | 1  | -    | -   | -    | -   | 1    | -  | -    |
| Otros               | 4                         | 4    | 21 | -    | 1   | 2    | 1   | -    | -  | 2    |
|                     | Fechas de los documentos  |      |    |      |     |      |     |      |    |      |
| 2011-2015           | 5                         | 4    | 11 | -    | 15  | -    | 4   | -    | -  | 1    |
| 2000-2010           | 36                        | 26   | 44 | 7    | 25  | 25   | 35  | 30   | -  | 24   |
| 1990-2000           | 18                        | 11   | 2  | 1    | 11  | 5    | 23  | 8    | -  | 4    |
| (...)-1990          | 9                         | 2    | 1  | -    | 20  | 2    | 6   | 3    | -  | -    |
|                     | Lenguas de los documentos |      |    |      |     |      |     |      |    |      |
| Español             | 23                        | 23   | -  | 1    | 27  | 8    | 50  | 11   | -  | 3    |
| Portugués           | -                         | -    | 26 | 1    | -   | -    | -   | -    | -  | 5    |
| Inglés              | 43                        | 19   | 32 | 6    | 39  | 24   | 18  | 30   | -  | 21   |
| Otras               | 2                         | 1    | -  | -    | 5   | -    | -   | -    | -  | -    |

Tabla VII. Búsqueda avanzada. Documentos relevantes.  
Fuente: Autores

| Funcionalidades de selección                                   | Universidad/herramienta |      |        |      |     |      |     |      |    |      | n  |
|--|-------------------------|------|--------|------|-----|------|-----|------|----|------|----|
|  | UMU                     |      | UPORTO |      | UPV |      | ULE |      | UC |      |    |
|  | DT                      | OPAC | DT     | OPAC | DT  | OPAC | DT  | OPAC | DT | OPAC |    |
| Recurso electrónico: directorio o base de datos que lo incluye | -                       | -    | -      | -    | x   | -    | x   | -    | x  | x    | 4  |
| Tipo de publicación/contenido                                  | x                       | x    | x      | -    | x   | x    | x   | -    | x  | -    | 7  |
| Formato  | -                       | x    | x      | x    | x   | x    | x   | x    | x  | x    | 9  |
| Materias/Disciplinas   | x                       | x    | x      | x    | x   | x    | x   | x    | x  | x    | 10 |
| Autores  | x                       | x    | x      | x    | x   | x    | x   | x    | x  | x    | 10 |
| Título de revista  | -                       | x    | -      | -    | x   | -    | -   | -    | -  | x    | 3  |
| Idiomas  | x                       | x    | x      | x    | x   | x    | x   | x    | x  | x    | 10 |
| Fechas   | x                       | x    | x      | x    | x   | x    | x   | x    | x  | x    | 10 |
| Área geográfica  | -                       | -    | x      | -    | -   | -    | -   | -    | -  | -    | 1  |
| Biblioteca/colección/Localización                              | x                       | -    | x      | x    | x   | x    | -   | -    | x  | x    | 9  |
| Tipo de préstamo   | -                       | -    | x      | -    | -   | -    | x   | -    | -  | -    | 1  |
| Ejemplares disponibles   | -                       | -    | -      | -    | -   | -    | x   | -    | -  | x    | 2  |
| Texto completo   | -                       | -    | x      | -    | -   | -    | x   | -    | -  | -    | 2  |
| Proveedor/Editor   | -                       | -    | x      | -    | -   | x    | -   | x    | -  | x    | 4  |
| Asignaturas  | -                       | -    | -      | -    | x   | -    | -   | -    | -  | -    | 1  |
| n  | 6                       | 7    | 11     | 6    | 10  | 8    | 10  | 7    | 8  | 10   |    |

Tabla VIII. Funcionalidades de selección. Fuente: Autores

| Funcionalidades de ordenación y visualización  | Universidad/herramienta |      |    |      |     |      |     |      |    |      | n |
|--|-------------------------|------|----|------|-----|------|-----|------|----|------|---|
|  | UMU                     |      | UP |      | UPV |      | ULE |      | UC |      |   |
|  | DT                      | OPAC | DT | OPAC | DT  | OPAC | DT  | OPAC | DT | OPAC |   |
| Alterar nº de resultados mostrados por página  | -                       | -    | x  | -    | -   | -    | -   | -    | -  | -    | 1 |
| Tipo de visualización de la página (columnas)  | -                       | -    | x  | -    | -   | -    | -   | -    | -  | -    | 1 |
| Relevancia ( <i>ranking</i> )  | x                       | -    | x  | x    | x   | x    | x   | x    | x  | -    | 8 |
| Visualización de los resultados (extendida o abreviada, considerándose solo el conjunto de los resultados) | -                       | -    | x  | -    | -   | -    | -   | -    | -  | -    | 1 |
| Fecha de publicación (del más actual al más antiguo)   | x                       | x    | x  | x    | x   | x    | x   | x    | -  | x    | 9 |
| Muestra/oculta ventanas automáticamente (con datos del registro, al cargar el mismo)                       | x                       | -    | -  | -    | x   | -    | -   | -    | -  | -    | 2 |
| Título (en orden alfabético)   | -                       | -    | -  | -    | -   | x    | x   | x    | -  | -    | 3 |
| Autor (en orden alfabético)  | -                       | -    | -  | -    | -   | -    | x   | -    | -  | -    | 1 |
| n  | 3                       | 1    | 5  | 2    | 3   | 3    | 4   | 3    | 1  | 1    | - |

Tabla IX. Funcionalidades de ordenación y visualización de los resultados. Fuente: Autores

| Funcionalidades colaborativas              | Universidad/herramienta |      |    |      |     |      |     |      |    |      | N |
|--|-------------------------|------|----|------|-----|------|-----|------|----|------|---|
|  | UMU                     |      | UP |      | UPV |      | ULE |      | UC |      |   |
|  | DT                      | OPAC | DT | OPAC | DT  | OPAC | DT  | OPAC | DT | OPAC |   |
| Recomendaciones de documentos relacionados | -                       | x    | x  | x    | -   | -    | x   | -    | x  | -    | 5 |
| Etiquetas o tags                           | -                       | x    | -  | x    | X   | x    | x   | -    | x  | -    | 6 |
| Comentarios                                | -                       | x    | -  | x    | X   | x    | x   | x    | x  | -    | 7 |
| Puntuaciones                               | -                       | -    | -  | x    | -   | -    | x   | -    | x  | -    | 3 |
| n  | 0                       | 3    | 1  | 4    | 2   | 2    | 4   | 1    | 4  | 0    | - |

*Tabla X. Funcionalidades colaborativas.  
Fuente: Autores.*

---

# Gestão do Conhecimento nas Organizações de Ensino: recorrência de registros na *Web of Science*

*Knowledge Management in Education organizations: recurring records in Web of Science*

---

**Santos, Cintia Almeida da Silva (1), Hoffmann, Wanda Aparecida Machado (2)**

(1) Doutoranda do Programa de Pós-Graduação em Ciência, Tecnologia e Sociedade da UFSCar – *Campus* São Carlos, São Carlos – Brasil. Bibliotecária-Documentalista do Instituto Federal de Educação, Ciência e Tecnologia de São Paulo – *Campus* Araraquara, São Paulo – Brasil – [cintia@ifsp.edu.br](mailto:cintia@ifsp.edu.br), Ramal de Acesso Eng. Heitor de Souza Pinheiro, s/n, Jardim dos Manacás, Araraquara – São Paulo, Brasil, CEP. 14801-600, telefone 55 16 3303-2341

(2) Pós-Doutorado em Prospecção de Informação Tecnológica. Docente do Programa de Pós-Graduação em Ciência, Tecnologia e Sociedade e do Departamento de Ciência da Informação da UFSCar – *Campus* São Carlos, São Carlos – Brasil. [wanda@nit.ufscar.br](mailto:wanda@nit.ufscar.br), Rodovia Washington Luiz Km 235, São Carlos – São Paulo, Brasil, CEP. 13.565-905, telefone 55 16 3351-8351

## Resumo

Objetivos: A) realizar revisão de literatura sobre Gestão do Conhecimento no contexto educacional; B) realizar pesquisa na base de dados *Web of Science* para verificar a recorrência de registros relacionados à temática no período de 2010 a 2014.

Metodologia: abordagem exploratória, de caráter inventariante. A pesquisa foi realizada na *Web of Science*. Resultados: a) a recorrência de registros sobre Gestão do Conhecimento está diminuindo com o passar dos anos; b) os periódicos com maior recorrência de registros são das áreas: Negócios/Economia, Ciência da Computação, Engenharias, Ciência da Informação e Educação; c) a dificuldade na escolha dos termos para a busca na *Web of Science* é fator a ser ponderado; d) A Universidade de São Paulo aparece como a única instituição brasileira no estudo. Considerações finais: com relação a recorrências de registros sobre Gestão do Conhecimento na Educação Profissional e Tecnológica ainda há um *gap*, pois a recorrência de registros é inócua. São necessários estudos, publicações científicas e proposições de modelos de Gestão do Conhecimento voltados para as organizações de ensino, especificamente para as organizações de ensino profissional e tecnológico com vistas a fortalecer a criação e a disseminação do conhecimento científico e tecnológico.

**PALAVRAS-CHAVE:** Gestão do Conhecimento. Sociedade do Conhecimento. Sociedade da Informação. *Web of Science*,

## 1. Introdução

A Sociedade da Informação (SI), com vistas à Sociedade do Conhecimento (SC), de acordo com Dziekaniak e Rover (2011), centra-se nos

## Abstract

Objectives: A) perform a literature review on knowledge management in the educational context; B) conduct research on Web of Science database in order to verify the recurrence of theme-related records in the 2010-2014 period. Methodology: exploratory and inventorial approach. The survey was conducted on Web of Science. Results: a) the recurrence of records on Knowledge Management is decreasing over the years; b) the journals with higher recurrence of the records are from the following areas: Business / Economics, Computer Science, Engineering, Information Science and Education; c) the difficulty in choosing the terms for searching on Web of Science is a factor to be considered; d) The University of São Paulo appears as the only Brazilian institution in the study. Final considerations: regarding the recurrences of records on Knowledge Management in Vocational and Technological Education, there is still a gap, since the recurrence of records is innocuous. It is necessary the development of studies, scientific publications and the proposition of knowledge management models addressed to educational organizations, specifically to the vocational and technological ones, in order to strengthen the creation and dissemination of scientific and technological knowledge.

**KEYWORDS:** Knowledge Management. Knowledge Society. Information Society. Web of Science.

denominados ativos intangíveis. Para Kayo, os ativos intangíveis são compreendidos como todos os bens e direitos que uma organização econômica possui, valorizados em termos monetários. Os ativos intangíveis são entendidos

como bens que não podem ser vistos ou tocados, ou seja, são bens incorpóreos, mas com valor econômico. Exemplos de ativos intangíveis: conhecimento, talento, processos, direitos autorais, entre outros (KAYO, *et al*, 2006).

Nesse cenário, o indivíduo possui papel fundamental nos processos estratégicos organizacionais. No contexto da SI, o caráter é mercadológico, ao passo que na SC o caráter deverá ser atrelado ao compartilhamento de informações que busquem uma sociedade justa e igualitária.

Conhecimento individual e coletivo, se administrados favoravelmente, propiciam condições organizacionais de estabilidade e crescimento estratégico para as organizações. Nesse sentido, as atenções organizacionais voltam-se ao indivíduo, como pode comprovar a teoria do capital humano que elucida a importância do conhecimento retido nos indivíduos. De acordo com Lucci, a Teoria do Capital Humano desenvolveu-se a partir da década de 60 por dois economistas que posteriormente receberam o prêmio Nobel (Theodore Schultz e Gary Becker). De acordo com essa teoria, pode-se dizer que o progresso de um país é alavancado pelo investimento que se faz nas pessoas (LUCCI, 2006).

Como o artigo em sua essência apresenta uma metodologia de caráter inventariante e descritivo, o mesmo constituiu-se em duas vertentes. A primeira foi a realização de revisão literária sobre a Gestão do Conhecimento (GC), afinando para o contexto educacional, com vistas a obter informações referenciais sobre a GC no contexto especificamente de Educação Profissional e Tecnológica (EPT). Já a segunda vertente caracterizou-se pela pesquisa empírica na base de dados *Web of Science*, propondo identificar a recorrência de registros, e por consequência, ressaltar alguns pontos, tais como: principais autores e áreas de publicação; quantitativo de publicações ao ano, no período de 2010 a 2014; principais periódicos, instituições e idiomas de publicação.

Prospectou-se como resultados que a GC é temática nova no âmbito da EPT. Nesse sentido, foi apresentada a hipótese de que a GC ainda é temática pouco recorrente no campo da educação, especificamente nas organizações de EPT. Entendeu-se que as abordagens em GC estão em sua maior parte vinculadas aos aspectos econômicos e administrativos das organizações, em detrimento dos aspectos pedagógicos das organizações de ensino, especificamente com relação à adoção da terminologia GC nessas organizações.

Como explana os autores Carvalho (2012) e Pires (2013), toda organização, de alguma forma, possui alguma prática relacionada à GC, basta voltar o olhar para a GC. Como explana Pires (2013), “a organização deve trocar a lente para enxergar a importância e a aplicação da GC”.

Carvalho (2012, p.4) relata ainda que a compreensão sobre a GC possibilita a identificação e o aperfeiçoamento das práticas já existentes na organização.

O artigo foi estruturado com as seções: Introdução; Referencial Teórico; Percurso Metodológico; Resultados e Discussões e Considerações Finais.

## 2. Referencial Teórico

### 2.1. Sociedade do Conhecimento ou Sociedade da Informação? A Gestão do Conhecimento no Contexto da Educação: uma visão preambular.

Para Castells (2007), Dziekaniak e Rover (2011), a SI precisa compreender respectivamente informação e conhecimento como fenômenos amplos de cunho: social, cultural, econômico, tecnológico, científico e político. Nesse sentido, o uso adequado e coerente das Tecnologias da Informação e da Comunicação (TICs) se faz necessário para que a sociedade possa, com maior facilidade, garantir o direito de acesso e uso da informação para, assim, produzir conhecimento.

Almeida (2006) discorre que o avanço das TICs causou grandes impactos na economia mundial e na forma como as organizações operam seus negócios. As TICs têm permitido, às organizações e aos indivíduos, perceberem a importância e a aplicabilidade dos processos de criação, acesso e disseminação da informação e do conhecimento.

Não obstante, informação e conhecimento, ambos indispensáveis para a construção e avanço de uma sociedade, ainda constituem domínio de poucos, pois se encontram arquivados, armazenados e manipulados pela minoria da sociedade. Ainda que o fenômeno da Internet tenha contribuído com as diversas possibilidades de acesso e disseminação, o uso da informação e do conhecimento como mola propulsora de mudanças estratégicas se encontra nas mãos de poucos, proposição sustentada por Lucci (2006).

Nantes (2005) disserta que o conhecimento sempre fora necessário para a elaboração de produtos e o desenvolvimento de diversas atividades, mas sua importância aumentou apres-

sadamente com o desenvolvimento da Ciência e da Tecnologia, especificamente nas últimas décadas do século XX.

Dziekaniak e Rover (2011) discorrem sobre a evolução dos conceitos que cercam tanto a SI quanto a SC. Para elucidar, apresenta-se a Tabela 1, que aponta algumas características da Sociedade da Informação *versus* a da Sociedade do Conhecimento, conforme Dziekaniak e Rover:

| SOCIEDADE DA INFORMAÇÃO                       | SOCIEDADE DO CONHECIMENTO                              |
|---|--|
| Cunho mercadológico e econômico               | Cunho social e acadêmico                               |
| Destaque para Informação e TICs               | Destaque para a Sociedade                              |
| Detentores do Conhecimento – classe dominante | Detentores do Conhecimento - Sociedade                 |
| TI como forma de manter os lucros             | TI a serviço da Sociedade                              |
| Foco na tecnologia – transmissão de dados     | Foco na ideologia de uma Sociedade justa e igualitária |

TABELA 1. Sociedade da Informação versus Sociedade do Conhecimento. Fonte: Adaptado de Dziekaniak e Rover (2011).

As primeiras manifestações que culminaram na ambiência e caracterização da SI se deram no final da década de 70. O precursor da SI (ainda que não usuário dessa terminologia) foi o sociólogo Daniel Bell (1978), há pouco mais de três décadas.

Bell (1978) instituiu a tese sobre a importância do conhecimento como mola propulsora de concorrência entre as organizações, isto é, dominaria a sociedade quem detivesse o conhecimento. Nesse sentido, pesquisas e desenvolvimento de novos produtos e serviços pelas organizações seriam ações recomendadas. O discurso de Bell alinha-se à Teoria do Capital Humano, explanada por Lucci (2006), e foi muito bem incorporado pelas grandes organizações econômicas.

Alguns dos autores contemporâneos que se debruçam por investigações nessa temática são Castells (2007) e Lévy (2010). Em linhas gerais, os autores enfatizam que o conhecimento converteu-se no principal fator de produção na SI ou Sociedade Informacional, terminologia adotada e preferida por Castells (2007).

Dziekaniak e Rover (2011) não consideram informação e conhecimento sinônimos, tendo em vista que informação pode ser considerada insumo básico para a produção do conhecimento e não o conhecimento em si.

Lévy (2010) retrata que vivemos em uma economia baseada em conhecimento. O termo “Economia do Conhecimento” foi cunhado por Drucker em 1969 para se referir à aplicação do conhecimento existente em qualquer campo ou fonte, novo ou velho, como estímulo ao desenvolvimento econômico. Porém, o teórico que mais chamou a atenção para o impacto do conhecimento nas economias das sociedades industriais avançadas foi Daniel Bell (GUILLE, 2008)

Assim, quanto mais os indivíduos conseguem interagir e compartilhar conhecimento, mais a sociedade se solidifica e esses indivíduos se estabelecem no cenário competitivo. Percebe-se, com essa afirmativa, a importância do trabalho em equipe e do compartilhamento de informações, que trabalhadas e interpretadas, auxiliam na produção de novos conhecimentos e na tomada de decisão. Contudo, ainda vivenciamos a SI e não a SC, pois o caráter da SI ainda é muito mais mercadológico do que ideológico (DZIEKANIAK; ROVER, 2011).

Assim, os autores supracitados (DZIEKANIAK; ROVER, 2011) defendem a tese de que a sociedade ainda se encontra na SI, e o caminho para se chegar efetivamente até a SC é longo e depende de várias prerrogativas.

Para que a SC se consolide, ela deverá estar alicerçada na ideologia de uma sociedade igualitária. Nesse sentido, a inclusão digital e informacional deverá ser dirimida.

Enquanto a SI está focada na Informação como matéria-prima, a sociedade ainda volta os olhos para a solidificação de mercado e para o acúmulo de riquezas. Nesse sentido, a SC deverá pautar-se na erradicação da exclusão digital e informacional, para que o saber seja compartilhado e construído por todos, de maneira igualitária.

Nesse contexto, o Estado tem papel fundamental, pois deverá intervir através da Educação, oferecendo oportunidades igualitárias para todos. Compreende-se, com essa afirmativa, que as organizações de ensino devem exercer papel fundamental no processo de transformação da SI para a SC.

Abdul Khan (*apud* Burch, 2005) estabelece diferenças entre a SI e a SC. Enquanto a primeira está relacionada estritamente à tecnologia, a segunda apresenta possibilidades de transformação social.

O posicionamento que se apresenta nesse estudo é reflexivo, visto que, para a SC existir, a sociedade precisará passar por reformas invasivas em diferentes áreas, principalmente na

Educação. Primeiramente, o Estado deverá estruturar políticas e projetos públicos que deverão ser construídos e manipulados por toda a sociedade. Os investimentos em Ciência, Tecnologia e Inovação deverão ser feitos nas universidades, agências de fomento, institutos, entre outros órgãos; mas esses investimentos deverão ser partilhados pela sociedade, como apontam Dziekaniak e Rover (2011).

As TICs deverão ser enxergadas apenas como ferramentas de apoio à sociedade e não como personagens centrais, como acontece nos dias atuais. A tecnologia é apenas um meio, uma ferramenta de auxílio nesse processo.

Dziekaniak e Rover (2011) destacam que, quanto mais periférico, em termos econômicos, é o país, maiores são as dificuldades de vivenciar plenamente a SC.

O conceito de GC tem fundamentação multidisciplinar. Como aponta Azevedo (2004), essa fundamentação ancora-se nas Ciências Cognitivas, na Administração e na Ciência da Informação. A GC envolve variados processos que objetivam fomentar a cultura do conhecimento coletivo dentro das organizações. Nesse sentido, compreende-se a GC inserida no âmbito da SI como um modelo de gestão que pode ser adotado nas diferentes organizações.

O foco deste estudo volta-se para as organizações de ensino, pois compreende-se que é por meio da Educação que transformações e inserções poderão ser feitas, para que resultados surjam a médio e longo prazo a fim de impactar diretamente na sociedade, tais como: a formação de cidadãos críticos, questionadores, dotados de capacidade de empreender e de inovar.

Isto posto, a necessidade de dialogar sobre em qual sociedade estamos inseridos é premente. Assim, propõe-se, com esta breve contextualização, abrir discussões e reflexões sobre os caminhos de nossa sociedade, sobre a sociedade na qual nos encontramos e sobre a sociedade para qual estamos caminhando.

## 2.2. Mas e o conhecimento? É fácil defini-lo?

Para Platão, filósofo grego, discípulo de Sócrates, o conhecimento é memória, é lembrança. Na filosofia platônica, é somente através da reflexão, do pensamento e das ideias que o indivíduo pode conhecer as coisas (CEZAR, 2009).

O conhecimento inicialmente é implícito e individualizado, parece simples. Contudo, definir conhecimento não é uma tarefa fácil, aponta Carvalho (2012). Para o autor, temos em mente o

significado de conhecimento, mas explicar uma definição não é simples.

De acordo com Carvalho (2008), os dados são registros de um evento sem significado. Quando estes passam a ter significado em determinado contexto, passam a ser informação. Já o conhecimento é a informação dotado de significado e que serve para a tomada de decisão, ou seja, possui uma finalidade, um significado para o indivíduo, e é passível de mutação e compartilhamento; portanto, conhecimento.

É oportuno observar, de acordo com Xavier e Costa, que:

A informação e o conhecimento são simultaneamente causa e efeito um de si mesmos (grifo nosso), numa interação dinâmica em que a sucessão pode ser plenamente invertida mas não gera nenhuma contradição, pois se é causa e efeito com relação a coisas diferentes em momentos distintos, quer dizer que se é causa só quando o outro é efeito e se é efeito apenas quando o outro for causa, gera assim expansão benéfica a ambos. Disponibilizar informação é promover a geração de conhecimento, que por sua vez produzirá mais informação e assim sucessivamente (grifo nosso); somente com uma política nacional de desenvolvimento em ciência e tecnologia, mas não só a partir dela, viabilizar-se-á manutenção desse fluxo (XAVIER; COSTA, 2010, p. 80).

Segundo Nonaka e Takeuchi (1997) “o conhecimento, ao contrário da informação, diz respeito a crenças e compromissos. O conhecimento é uma função de uma atitude, perspectiva ou intenção específica”. Nesse sentido, o conhecimento é representado por processos, modelos, informações úteis, ou seja, pode ser considerado um conjunto de determinadas informações que levarão um indivíduo a tomar decisões, resolver problemas ou criar novos conhecimentos.

No âmbito organizacional, o conhecimento classifica-se em dois tipos: conhecimento tácito, aquele compreendido como o conhecimento particular e interno que o indivíduo possui, e que está ligado às experiências vividas; e conhecimento explícito, aquele que pode ser transmitido entre os indivíduos e que está exteriorizado e registrado.

Existe uma complementação entre os dois tipos de conhecimento. Conforme discorrem Nonaka e Takeuchi (1997), a imbricação dos conhecimentos tácito e explícito é fundamental para o desenvolvimento do indivíduo e das organizações. O cenário não é diferente para as organizações de ensino (NONAKA; TAKEUCHI; 1997, p. 57).

Nesse sentido, destacamos a Espiral do Conhecimento, compreendida pela interação das quatro formas de conversão do conhecimento: socialização; exteriorização; combinação e internalização. O conhecimento tácito e explícito na organização torna-se maior na escala na medida em que aumentam seus níveis ontológicos. O conjunto dessas quatro formas é que compõe a espiral do conhecimento, elemento indispensável para a implementação da Gestão do Conhecimento.

Como anteriormente explanado, para que a SI transforme-se em SC, reformas profundas são necessárias, a começar pela Educação (DZIEKANIAK; ROVER, 2011). Nesse sentido, as organizações de ensino possuem muito a contribuir para tal transformação, e a GC, enquanto modelo de gestão, pode ser potencialmente explorada.

Seguindo a linha de raciocínio de Hoffmann (2012), o conhecimento pode até ser confundido com uma mercadoria, mas conhecimento não é isso; antes, ele se traduz em um processo complexo e dinâmico, pois encontra-se em constante mutação.

Assim, o conhecimento, com diversas definições, sumariamente expostas, trata-se de uma abstração interior (SETZER, 1999), elemento individual, que quanto mais compartilhado, mais valor possui; e, desta maneira, novos conhecimentos são criados.

Isto posto, não há uma definição única para conhecimento. Deve-se levar em consideração as experiências, valores, cultura e interação entre os indivíduos e entre as organizações.

### 2.3. Importância da Gestão do Conhecimento para as Organizações de Ensino

Tendo em vista a importância do conhecimento para a sociedade, o cenário é igualitário para as organizações de ensino. Nesse contexto, a GC pode ser amplamente adotada, já que seus objetivos se traduzem em estimular, desenvolver e disseminar o capital intelectual das organizações.

A GC pode ser considerada um processo associativo que tem como foco as estratégias organizacionais. Dessa maneira, envolve as competências dos indivíduos, pois, como afirmam Nonaka e Takeuchi (1995), todo conhecimento se inicia no indivíduo.

As organizações de ensino são ambientes copiosos para implementação da GC. De acordo com Pires (2013), toda e qualquer organização

pode detectar e mapear o conhecimento ao menos nas seguintes esferas:

- No ambiente de negócios;
- No relacionamento com o público de interesse;
- Nos documentos, padrões e processos;
- Nas pessoas e cultura.

Nas organizações empresariais, a GC volta-se à competitividade e à tomada de decisão (TERRA, 2005). Mas e nas organizações de ensino? Como a GC é compreendida e qual sua importância?

A literatura sobre GC no contexto educacional com relação ao contexto corporativo empresarial é incipiente, especificamente com relação às organizações de EPT. Nesse sentido, no ambiente educacional, os olhares sobre as inúmeras possibilidades de exploração da GC devem se voltar não somente aos aspectos administrativos, mas também aos aspectos pedagógicos, ou seja, os processos de ensino-aprendizagem.

A GC, para Azevedo (2004), pode ser compreendida como fortalecedora dos processos de ensino-aprendizagem, pois ela pode auxiliar na construção de ambientes propícios com vistas à estimulação das práticas pedagógicas e do levantamento das competências dos recursos humanos da organização.

Complementando, os autores Assumpção e Dimitrov (2010, p. 5) visualizam a GC como um “*processo educacional complexo de contínua redução do tempo entre a identificação de problemas [...] e a geração de novos saberes*”.

A organização de ensino deve posicionar-se favoravelmente à GC à medida que o cenário educativo vem sofrendo mudanças, especialmente com relação à inserção e utilização das TICs nos processos de ensino-aprendizagem. Dessa maneira, adotar um modelo de gestão que favoreça a aquisição, armazenamento e compartilhamento da informação e do conhecimento, potencializando o uso das TICs, é fator agregador para qualquer organização de ensino.

Quando se faz referência à GC como modelo de gestão, esse modelo pode ser utilizado pela organização e pode apresentar diferentes abordagens, tais como: Aprendizagem Organizacional; Gestão do Capital Intelectual; Gestão por Competências; Inteligência Organizacional ou Competitiva, como mostra o *Quadro 1* que apresenta a Gestão do Conhecimento e as suas diferentes abordagens, de acordo com Hoffmann (2012).

|   |  |
|---|--|
| <p><b>Aprendizagem organizacional</b></p> <p>Objetiva aproveitar os recursos já existentes na organização para que os indivíduos se identifiquem e utilizem as melhores práticas já existentes. O processo de aprendizagem se dá como um desenvolvimento natural e harmônico entre os indivíduos. Processo contínuo para detecção e correção de erros em busca de atingir os objetivos.</p> | <p><b>Gestão do Capital Intelectual</b></p> <p>Somatório do conhecimento de todos da organização. A Gestão do capital intelectual pressupõe alinhar esforços corporativos da junção do – capital humano (indivíduos) + capital estrutural (patentes, marcas, processos, manuais, banco de dados, entre outros produtos) + capital de clientes e fornecedores (valor dos relacionamentos e compartilhamentos da organização).</p>                                 |
| <p><b>Gestão por competências</b></p> <p>As competências devem ser identificadas na organização e são abordadas em três dimensões: competência organizacional, competência essencial, competência individual. Nessa abordagem, são compreendidos os processos organizacionais tanto dos equipamentos quanto dos indivíduos.</p>   | <p><b>Inteligência Organizacional ou Competitiva</b></p> <p>Compreendida como a integração de um processo de aprendizagem motivado pela competição em consonância com um processo analítico, transformando informações desagregadas em conhecimento estratégico, ou seja, um programa organizacional sistêmico de coleta, tratamento e análise de informações sobre as atividades de outras organizações (consideradas concorrentes) a vigilância ambiental.</p> |

QUADRO 1. *Gestão do Conhecimento – diferentes abordagens. Fonte: adaptado de HOFFMANN (2012).*

Deve-se salientar que a informação é insumo básico para gerar conhecimento. Resgata-se Le Coadic quando o autor explana que a informação pode ser considerada o sangue da ciência. O autor aponta que, sem a informação, a ciência sequer pode se desenvolver e viver. Sem informação, a pesquisa seria inútil e o conhecimento não existiria (LE COADIC, 1996). Destaca-se, portanto, a importância do conhecimento documentado e da organização de ensino nesse processo.

As organizações, especialmente as de ensino, devem oferecer condições capacitadoras necessárias para que o indivíduo e o grupo no qual está inserido possam criar e fomentar o conhecimento organizacional. Para isso, são necessárias cinco fases, de acordo com Escrivão, Nagano e Escrivão Filho (2011):

intenção organizacional, que pode ser definida como a aspiração de uma organização às suas metas; autonomia aos membros da organização, o que introduz oportunidades inesperadas e possibilita a automotivação dos indivíduos para a criação de conhecimento; flutuação e caos criativo, que estimulam a interação entre a organização e o meio ambiente externo, possibilitando a exploração da

ambigüidade, da redundância e dos ruídos dos sinais ambientais, aprimorando seu sistema de conhecimento; redundância, informações que transcendem as exigências operacionais imediatas dos membros da organização, aumentando o volume de informações a serem processadas; e variedade de requisitos dos membros da organização, o que facilita que se enfrentem diferentes situações (ESCRIVÃO; NAGANO; ESCRIVÃO FILHO, 2011, p. 100).

Desse modo, é notória a importância da GC para as organizações de ensino, pois é na esfera da Educação que deverão advir grandes mudanças para a sociedade. Como apontado por Dziekaniak e Rover (2011), é fundamental que informação e conhecimento sejam tópicos centrais nessas organizações. É necessário que os profissionais envolvidos nessas organizações sejam capacitados e incentivados para que realizem trabalhos em equipe, compartilhando experiências e transformando conhecimento tácito em explícito.

### 3. Percurso Metodológico

Como o artigo em sua essência apresenta uma metodologia de caráter inventariante e descritivo, o mesmo consistiu em duas vertentes. A primeira se deu através da realização de revisão literária sobre a GC e sua importância para as organizações de ensino; a segunda caracterizou-se pela pesquisa empírica sobre a recorrência de registros na base de dados *Web of Science*.

A GC tem sido alvo de inúmeras pesquisas em diferentes áreas do conhecimento, tais como: Negócios/Economia, Ciências da Computação, Ciências da Informação, Engenharia, Educação, entre outras.

As subseções detalham as duas vertentes, de maneira a ressaltar a importância da revisão da literatura aliada à pesquisa empírica.

#### 3.1. Revisão de Literatura – Primeira Vertente

A revisão de literatura se deu por meio de pesquisas bibliográficas realizadas em livros, bases de dados, vídeos e periódicos específicos sobre GC em seu caráter amplo e no contexto educacional.

A revisão de literatura, etapa fundamental em toda pesquisa, se dá por meio da pesquisa bibliográfica, conforme aponta Boccato:

a pesquisa bibliográfica busca a resolução de um problema (hipótese) por meio de referenciais teóricos publicados, analisando e discutindo as várias contribuições científicas. Esse tipo de pesquisa trará subsídios para o conhecimento sobre o que foi pesquisado, como e sob que enfoque e/ou pers-

pectivas foi tratado o assunto apresentado na literatura científica (BOCCATO, 2006, p. 266).

Pizanni et al (2012, p. 65) enfatizam, com relação às pesquisas bibliográficas em bases de dados, que os pesquisadores devem se atentar à importância de usar vocabulário controlado, que seja reconhecido pelas bases de dados, assim como fazer uso de estratégias de busca para combinar os termos da pesquisa e realizar as pesquisas em bases de dados confiáveis.

### 3.2. Pesquisa Empírica: Coleta de Dados – Segunda Vertente

As pesquisas na *Web of Science* foram realizadas por meio do Portal de Periódicos Capes, criado oficialmente no ano de 2000. Com a criação do Portal, a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes), fundação do Ministério da Educação (MEC), desempenha papel fundamental na expansão e consolidação da pós-graduação *stricto sensu* em todos os estados brasileiros (CAPES, 2008).

A Capes passou a centralizar e otimizar a aquisição desse tipo de conteúdo por meio da negociação direta com editores internacionais. Pode-se definir o Portal de Periódicos Capes como:

uma biblioteca virtual que reúne e disponibiliza a instituições de ensino e pesquisa no Brasil o melhor da produção científica internacional. Ele conta com um acervo de mais de 37 mil títulos com texto completo, 126 bases referenciais, 11 bases dedicadas exclusivamente a patentes, além de livros, enciclopédias e obras de referência, normas técnicas, estatísticas e conteúdo audiovisual, o Portal de Periódicos Capes é atualmente o maior portal de produção científica-acadêmica brasileira (CAPES, 2008).

O *Quadro 1* – Descrição da *Web of Science* apresenta uma definição da base, instrumento multidisciplinar de grande relevância no cenário acadêmico mundial:

| Base de Dados                           | Descrição Sumária  |
|---|--|
| <b>Web of Science – acesso restrito</b> | O Portal de Periódicos, por meio de assinatura junto à <i>Thomson Reuters Scientific</i> , oferece acesso à coleção principal da <i>Web of Science</i> , permitindo acesso a referências e resumos em todas as áreas do conhecimento. Por meio da <i>Web of Science</i> , estão disponíveis ferramentas para análise de citações, referências, índice, permitindo análises bibliométricas. Cobre aproximadamente 12.000 periódicos. A assinatura desse conteúdo oferece a possibilidade de consulta a 5 coleções: <i>Science Citation Index Expanded</i> (SCI-EXPANDED) - com disponibilidade de acesso desde 1945 até o presente; <i>Social Sciences Citation Index</i> (SSCI) - com disponibilidade de acesso desde 1956 até o presente; <i>Arts &amp; Humanities Citation Index</i> (A&HCI) - com disponibilidade de acesso desde |

1975 até o presente; *Conference Proceedings Citation Index- Science* (CPCI-S) - com disponibilidade de acesso desde 1991 até o presente; e *Conference Proceedings Citation Index - Social Science & Humanities* (CPCI-SSH) - com disponibilidade de acesso desde 1991 até o presente. Esse conteúdo pode ser localizado pela opção *Buscar Base*.

*QUADRO 1. Descrição sumária da Web of Science. Fonte: Portal de Periódicos Capes (2010).*

Escolheu-se a *Web of Science* por ser a base de dados referencial de maior cobertura no Portal de Periódicos Capes.

Primeiramente, foi feita pesquisa sobre GC sem definição de temática. Posteriormente, foram feitas pesquisas sobre GC no âmbito da Educação. O *Quadro 2* (Anexo 1) apresenta informações referentes à Pesquisa sobre GC e o *Quadro 3* (Anexo 2) refere-se aos termos utilizados nas pesquisas sobre GC no contexto educacional.

Foram feitas pesquisas com diferentes termos de consulta, porém respeitando o refinamento da busca. Os termos utilizados para busca no campo título das bases de dados apresentam-se no Anexo 2 que refere-se ao *Quadro 3*.

Escolheu-se o campo título tendo em vista que tal campo é de grande importância nas publicações científicas, pois pode ser considerado o “cartão de visita” das publicações. Nesse sentido, quando o termo é de grande relevância, geralmente é mencionado no título.

O idioma não foi selecionado; porém, reconhece-se que o idioma Inglês é o mais recorrente no meio acadêmico. Cabe lembrar que a não restrição do idioma pode apresentar outros idiomas e por consequência países com produções sobre a temática.

Salienta-se que foi realizada a recuperação somente de registros publicados no período entre 2010 e 2014, com recuperação apenas no campo título dos mesmos, sendo realizadas duas frentes de pesquisa, conforme mencionados nos Anexos 1 e 2.

De posse dos dados extraídos nas pesquisas realizadas no período de 06 a 8 de março de 2015, foram utilizadas as análises da própria *Web of Science* para elucidar alguns apontamentos e resultados obtidos.

Isto posto, as informações foram analisadas com base nas pesquisas:

- Pesquisa sobre GC em Âmbito Geral;
- Pesquisa 1 (*Quadro 3*).

As discussões com relação às pesquisas realizadas são apresentadas na seção de Resultados e Discussões.

#### 4. Resultados e Discussões

Nesta sessão, são apresentados os dados obtidos com a Pesquisa Geral e Pesquisa 1, realizadas na *Web of Science*.

Primeiramente, destacou-se que a quantidade de registros recuperados na base *Web of Science* para o termo *Knowledge AND Management* apresentou uma recorrência de 3.200 registros, e a busca para o termo *Knowledge AND Management AND Education* apresentou uma recorrência de 103 registros. Foram esses os números analisados no presente estudo.

A Pesquisa Geral apresentou uma recorrência de 3.200 registros, ao passo que a Pesquisa 1 apresentou uma recorrência de 103 registros, conforme ilustram as *Figuras 1 e 2*, respectivamente, ambas com declínio na recorrência de registros ao ano, conforme apresentam os *Quadros 2 e 4*.

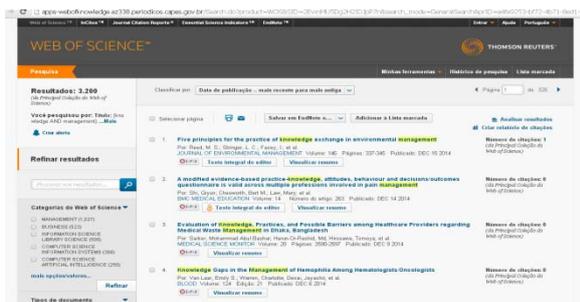


Figura 1. Pesquisa Geral realizada sobre GC, no período de 2010-2014.

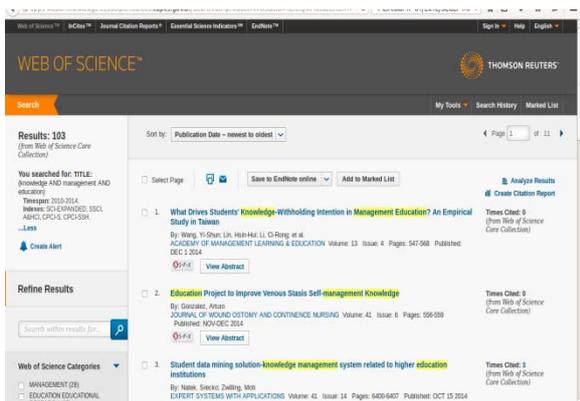


Figura 2. Pesquisa Realizada sobre GC e Educação na Web of Science (2010-2014).

| Termo de Busca   | Web of Science   |
|--|--|
| <b>Knowledge AND Management AND Education</b><br>– período: de 2010 a 2014<br>- recuperação no campo <u>título</u> | <b>103</b><br><br><b>2010</b> – 28 registros<br><b>2011</b> – 26 registros<br><b>2012</b> – 24 registros<br><b>2013</b> – 13 registros<br><b>2014</b> – 12 registros |
| <b>DIMINUIÇÃO DE RECORRÊNCIA DE REGISTROS COM O PASSAR DOS ANOS</b>  |  |

Quadro 4. Pesquisa sobre GC na Educação – Pesquisa 1.

Como as pesquisas foram realizadas em março, presume-se que o ano de 2014 poderá ter aumento no quantitativo de registros indexados. Outro fato que deve ser ressaltado é que a pesquisa recuperou registros que continuam os termos pesquisados apenas no campo título e não nos campos palavras-chave e/ou resumo.

Verificou-se que existem poucos artigos indexados na *Web of Science* referentes à GC no contexto educacional. Restringindo a pesquisa especificamente ao contexto da EPT, os números recorrentes foram inócuos. Nesse sentido, admite-se a hipótese de que publicações sobre GC no campo da EPT ainda são incipientes.

Assim, a pesquisa que seria inicialmente realizada sobre GC no contexto da EPT teve seu curso ampliado para a GC no contexto educacional.

Através dos registros obtidos, foram feitas análises, por meio de mecanismos oferecidos pela própria base de dados, em que são destacados: países de publicação; autores; área da publicação; periódicos; filiações institucionais; idioma de publicação; e tipo de documento, como segue:

- **Países de Publicação:** na Pesquisa Geral, tem-se, em primeiro lugar, China, seguido dos Estados Unidos da América (EUA); posteriormente, Inglaterra, Malásia e Espanha. O Brasil aparece em 10º lugar com 93 registros indexados (Figura 3).

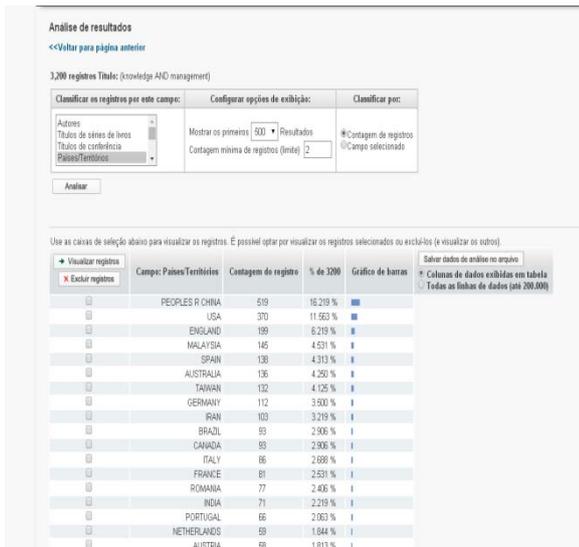


Figura 3. Pesquisa Geral – Países de Publicação.

Já na Pesquisa 1, os cinco primeiros países são: China, EUA, Austrália, Iran e Malásia. O Brasil aparece em 6º lugar, com recorrência de 4 registros (Figura 4). Nesse sentido, a expressividade do Brasil é quantitativamente maior se comparada a produções em GC em esfera ampla.

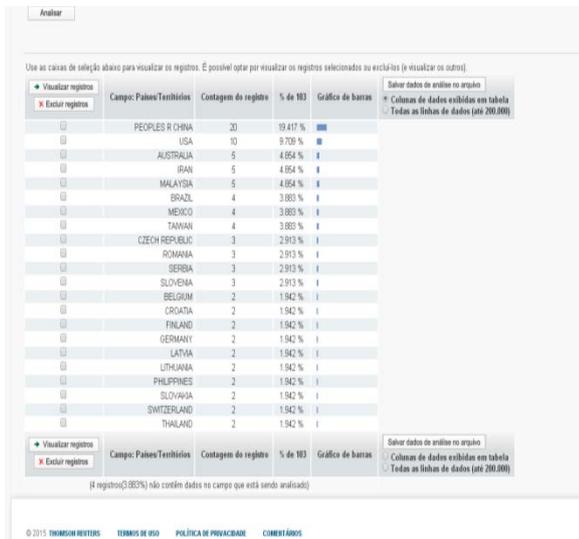


Figura 4. Pesquisa 1 – Países de Publicação.

- **Autores:** na Pesquisa Geral, os cinco primeiros autores são: Wong K.Y. (12 registros); Cardoso L. (11 registros); Salleh K. (10 registros); Nazem (8 registros); Bures V. (7 registros), conforme aponta a Figura 5:



Figura 5. Pesquisa Geral – Autores.

Na Pesquisa 1, os cinco primeiros autores são: Aguilar J. M. P.; Lopez M. S. H; Perez L. R. V. (recorrência de 3 registros para cada autor); Arsenijevic J.; e Escrivão G. (recorrência de 2 registros para cada autor), conforme Figura 6.

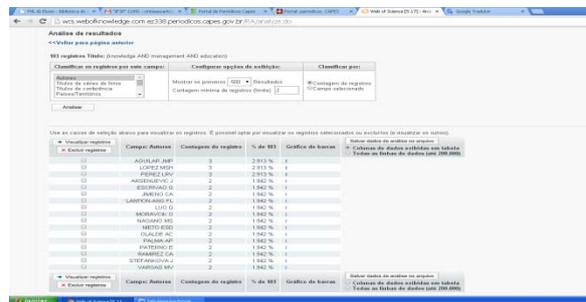


Figura 6. Pesquisa 1 – Autores.

- **Áreas de publicação:** As Figuras 7 e 8 apresentam as principais áreas que possuem publicações em GC.

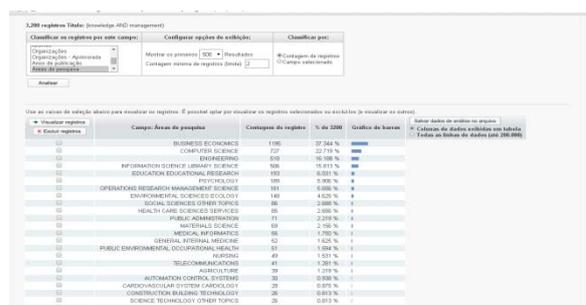


Figura 7. Pesquisa Geral – Áreas de Publicação.

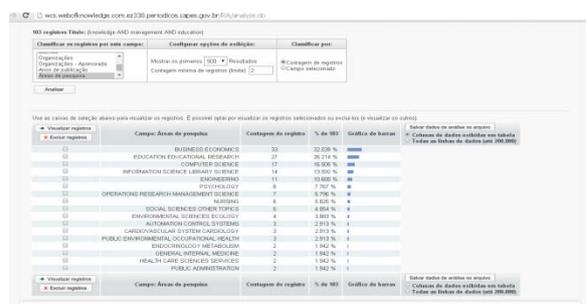


Figura 8. Pesquisa 1 – Área de Publicação.

Santos, Cintia Almeida da Silva Santos; Hoffmann, Wanda Aparecida Machado. *Gestão do conhecimento nas organizações de ensino: recorrência de registros na Web of Science.* En XII Congreso ISKO España y II Congreso ISKO España-Portugal, 19-20 de noviembre, 2015, Organización del conocimiento para sistemas de información abiertos. Murcia: Universidad de Murcia.

Tanto na Pesquisa Geral quanto na Pesquisa 1, a principal área de publicação é a de Economia/Negócios. Foram destacadas as cinco primeiras áreas em ambas as pesquisas: Economia/Negócios; Ciência da Computação, Educação, Ciência da Informação e Engenharias.

A Educação aparece em 5º na Pesquisa Geral (123 registros) e em 2º na Pesquisa Temática (27 registros).

- **Periódicos:** As Figuras 9 e 10 apresentam os principais periódicos com publicações sobre GC.

Figura 9. Pesquisa Geral – Periódicos.

Figura 10. Pesquisa 1 – Periódicos.

Com relação aos periódicos, na Pesquisa Geral, o periódico *Journal of Knowledge Management* (recorrência de 85 registros) apresenta maior recorrência, correlacionando-se com a área de maior publicação da Pesquisa Geral (Negócios/Economia), enquanto que, na Pesquisa 1, o periódico com maior recorrência de registros é o *Procedia Social and Behavioral Sciences* (recorrência de 4 registros), fator coerente, analisando as áreas principais de publicação de ambas as pesquisas.

- **Instituições:** as Figuras 11 e 12 apresentam as instituições de filiações dos autores:

Figura 11. Filiações institucionais – Pesquisa Geral.

As instituições, em sua maioria, são universidades ou institutos de pesquisa. Tanto na Pesquisa Geral quanto na Pesquisa 1, a *Islamic Azad University* aparece em primeiro lugar, com uma recorrência de 40 registros na Pesquisa Geral e 3 na Pesquisa 1. Em segundo lugar, aparece a instituição brasileira Universidade de São Paulo (USP), com recorrência de 17 registros na Pesquisa Geral e 2 registros na Pesquisa 1.

Destaca-se que a USP é considerada a principal universidade brasileira (BBC Brasil, 2015).

Figura 12. Filiações institucionais – Pesquisa 1.

- **Idiomas:** O Inglês domina as publicações científicas. As Figuras 13 e 14 apresentam os principais idiomas: em primeiro lugar, o Inglês, seguido do Chinês, são os principais idiomas tanto na Pesquisa Geral, quanto na Pesquisa 1. O Português aparece em 5º lugar na Pesquisa Geral e em 3º lugar na Pesquisa 1.

Figura 13. Idioma de publicação – Pesquisa Geral.

Figura 14. Idioma de publicação – Pesquisa 1.

- **Tipo de Documento:** as Figuras 15 e 16 apresentam os principais tipos de documentos indexados na *Web of Science*. No contexto da pesquisa, é relevante verificar a recorrência dos artigos publicados durante o período investigado.

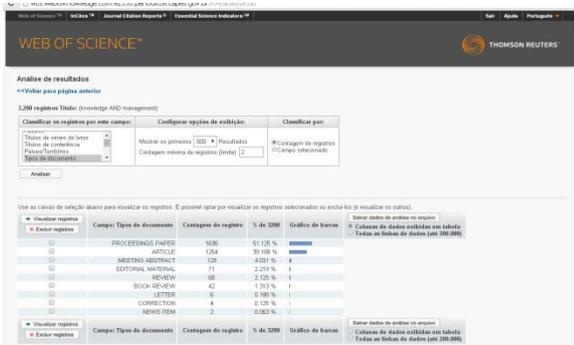


Figura 15. Tipo de Publicação – Pesquisa Geral.

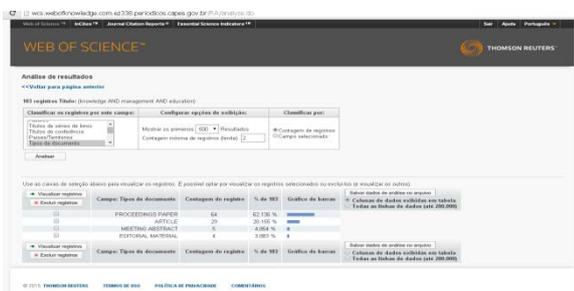


Figura 16. Tipo de Publicação – Pesquisa 1.

Os artigos científicos, um dos principais produtos científicos, aparecem com expressividade tanto na Pesquisa Geral quanto na Pesquisa 1. Representam, na Pesquisa Geral, 1.254 registros, e, na Pesquisa 1, 29 registros.

Esta amostra apresentou, de maneira breve, como diferentes análises podem ser feitas a partir da utilização das ferramentas disponíveis nas bases de dados. São ressaltados alguns pontos:

- a) a recorrência da literatura sobre GC na *Web of Science* está diminuindo com o passar dos anos;
- b) com relação à pesquisa empírica, os periódicos com maior recorrência de registros são das áreas: Negócios/Economia, Ciência da Computação, Engenharias, Ciência da Informação e Educação;
- c) a dificuldade na escolha dos termos para a busca na *Web of Science* é fator a ser ponderado;
- d) a USP aparece como a única instituição brasileira na amostra da pesquisa.

- e) as universidades ainda se mantêm na dianteira com relação à produção científica.

## 5. Considerações Finais

Ainda não há consenso sobre uma definição e uma metodologia clara para a disciplina GC, haja vista sua natureza multifacetada (PEREZ-MONTORO GUTIÉRREZ, 2006, p. 118). Não obstante, devido à importância do conhecimento para a sociedade, compreende-se que a GC pode ser implementada nas organizações de ensino, já que seus objetivos se traduzem em estimular, desenvolver e disseminar o capital intelectual das organizações.

De acordo com a recorrência de registros, a GC na esfera educativa, principalmente na EPT, apresenta-se como incipiente, ao menos no período pesquisado e utilizando como ferramenta uma das maiores bases de dados do universo acadêmico-científico.

A GC está muito mais recorrente em temáticas de áreas como: Negócios/Economia, Ciências da Computação e Engenharias em detrimento das áreas da Educação.

Com relação à *Web of Science*, destaca-se que a verificação da recorrência de registros na base possibilita, ainda que de maneira qualitativa, realizar análises proeminentes para desdobramentos investigativos com relação às demais abordagens quantitativas e qualitativas.

Ressalta-se a importância de se explorar com maior aprofundamento as possibilidades e vantagens para as organizações educacionais adotarem modelos de gestão baseados em conhecimento.

Um fator de relevância com relação ao estudo diz respeito à diminuição da recorrência de registros sobre GC tanto na esfera geral quanto no contexto da Educação. Ao menos no campo título, o declínio é comprobatório.

Ainda é válido mencionar a importância de estudos, proposições, construções de modelos de GC voltados para as organizações de EPT, que carecem de contribuições que visem ao fortalecimento, à criação e à disseminação do conhecimento científico e tecnológico, assim como é importante a publicação dos estudos realizados.

Para encerrar, enfatiza-se que os eventos relacionados à Organização do Conhecimento são catalizadores de experiências e arenas férteis que propiciam atualizações dos aportes teórico-metodológicos entre os pesquisadores participantes.

## Referências

- ALMEIDA, F. C. (2006). Do saber intelectual ao conhecimento ação. Disponível em: <<http://www.hottopos.com/convenit4/fernalm.htm>>. *Convenit Internacional*. v. 4. Acesso em: 28 nov. 2014.
- ASSUMPCÃO, R. P. S.; DIMITROV, P. (2010). *Gestão do conhecimento: uma estratégia para o desenvolvimento do Sistema Único de Saúde – SUS*. São Paulo: FUNDAP. 12 p.
- AZEVEDO, A. M. A. M. P. (2004). *Gestão do conhecimento em organizações educacionais de nível médio*. Biblioteca SEBRAE. Disponível em: <[http://201.2.114.147/bds/bds.nsf/B72246EF3F957E4703256F01006A5D66/\\$File/NT000948E6.pdf](http://201.2.114.147/bds/bds.nsf/B72246EF3F957E4703256F01006A5D66/$File/NT000948E6.pdf)>. Acesso em: 04 jan. 2015.
- BELL, D. (1978). *O advento da sociedade industrial: uma tentativa de previsão social*. São Paulo: Cultrix.
- BBC Brasil. (2015). Ranking traz 5 universidades brasileiras entre 10 melhores da América Latina. Disponível em: <[http://www.bbc.com/portuguese/noticias/2015/06/150603\\_ranking\\_universidades\\_mdb](http://www.bbc.com/portuguese/noticias/2015/06/150603_ranking_universidades_mdb)>. Acesso em: 10 jun. 2015>.
- BOCCATO, V. R. C. (2006). Metodologia da pesquisa bibliográfica na área odontológica e o artigo científico como forma de comunicação. *Rev. Odontol. Univ. Cidade São Paulo*, São Paulo, v. 18, n. 3, p. 265-274. Disponível em: <<http://www.scielo.br/pdf/pci/v11n2/v11n2a10.pdf>>. Acesso em: 15 maio 2015.
- CAPES. (2008). História e missão. Disponível em: <<http://www.capes.gov.br/historia-e-missao>>. Acesso em 20 maio 2015.
- BURCH, S. (2005). Sociedade da informação/ sociedade do conhecimento. In: AMBROSI, A.; PEUGEOT, V.; PIMENTA, D. *Desafios das palavras*. Editora: VECAM. Disponível em: <<http://vecam.org/article519.html>>. Acesso em: 30 nov. 2014.
- CARVALHO, F. C. A. (2012). *Gestão do conhecimento*. São Paulo: Pearson. 298 p.
- CARVALHO, F. C. A. (2008). *Gestão do conhecimento: curso de férias*. São Paulo: ESPM.
- CASTELLS, M. (2007). A era da intercomunicação. In: CASTELLS, M, et al. *Caminhos para uma comunicação democrática*. São Paulo: Editora e Livraria Instituto Paulo Freire.
- CEZAR, A. (2009). Sócrates e Platão. *Fragments de filosofia*. 2009. Disponível em: <<http://giulianofilosofo.blogspot.com.br/2009/06/socrates-e-platao.html>>. Acesso em: 04 jan. 2015.
- DZIEKANIAK, G.; ROVER, A. (2011). Sociedade do conhecimento: características, demandas e requisitos. *DataGramaZero - Revista de Ciência da Informação*, v. 12, n. 5, out. Disponível em: <[http://www.dgz.org.br/out11/Art\\_01.html](http://www.dgz.org.br/out11/Art_01.html)>. Acesso em: 04 nov. 2014.
- ESCRIVÃO, G.; NAGANO, M. S.; ESCRIVÃO FILHO, E. (2014). A gestão do conhecimento na educação ambiental. *Perspectivas em Ciência da Informação*, v.16, n.1, jan./mar. 2. Disponível em: <<http://www.scielo.br/pdf/pci/v16n1/a06v16n1>>. Acesso em 10 jul. 2015.
- GUILLE, D. (2008). O que distingue a economia do conhecimento? Implicações para a educação. *Cad. Pesqui.* São Paulo. v. 38, n. 35, set.-dez. Disponível em: <[http://www.scielo.br/scielo.php?pid=S0100-15742008000300004&script=sci\\_arttext](http://www.scielo.br/scielo.php?pid=S0100-15742008000300004&script=sci_arttext)>. Acesso em: 04 jan. 2015.
- HOFFMANN, W. A. M. (2012). *Gestão do conhecimento: aprender e compartilhar*. São Carlos: EdUFSCar. 159p.
- KAYO, E. K. et al. (2006). Ativos intangíveis, ciclo de vida e criação de valor. *Revista de Administração Contemporânea*. v. 10, n. 3, p. 73-90. Disponível em: <<http://www.scielo.br/pdf/rac/v10n3/a05v10n3.pdf>>. Acesso em: 16 jul. 2014.
- LE COADIC, Y. F. (1996). *A Ciência da Informação*. Brasília: Briquet de Lemos.
- LÉVY, P. (2010). *Collaborative learning in the digital social medium*. Anotações da palestra proferida pelo autor, realizada na II Semana da Ciência da Informação da FURG. Rio Grande: FURG. Disponível em: <<http://ow.ly/1nhGp>>. Acesso em: 30 nov. 2014.
- LUCCHI, E. A. (2006). *A sociedade do conhecimento e a educação para o pensar* (notas de conferência para alunos e professores de ensino médio em diversos estados do Brasil). São Paulo: Saraiva. Disponível em: <<http://www.hottopos.com/vidlib7/e2.htm>>. Acesso em: 02 jan. 2015.
- NANTES, C. (2005). A importância do conhecimento. *Interativa: artigos e negócios*. abr. Disponível: <[http://internativa.com.br/artigo\\_conhecimento\\_04.html](http://internativa.com.br/artigo_conhecimento_04.html)>. Acesso em: 30 nov. 2014.
- NONAKA, I.; TAKEUCHI, H. (1997). *Criação de conhecimento na empresa: como as empresas japonesas geram a dinâmica da inovação*. Rio de Janeiro, Campus.
- NONAKA, I; TAKEUCHI, H. *The knowledge-creating company: how japanese companies create the dynamics of innovation*. New York: Oxford University Press, 1995. 65-80 p.
- PEREZ-MONTORO GUTIÉRREZ, M. (2006). O conhecimento e sua gestão em organizações. In: TARAPANOFF, K. *Inteligência, informação e conhecimento em corporações*. Brasília, DF: IBICT, UNESCO. p. 117-138.
- PIRES, C. (2013). Gestão do conhecimento: uma mudança de olhar. Disponível em: <<https://www.youtube.com/watch?v=y8mFqLUBRmQ>>. Acesso em: 02 jan. 2015. (vídeo).
- PIZZANI, L. et al. (2012). A arte da pesquisa bibliográfica na busca do conhecimento. *Rev. Dig. Bibl. Ci. Inf.*, Campinas, v.10, n.1, p.53-66, jul./dez. ISSN 1678-765X. Disponível em: <<http://www.sbu.unicamp.br/seer/ojs/index.php/rbci/article/view/522>>. Acesso em: 15 maio 2015.
- PORTAL DE PERIÓDICOS DA CAPES. (2010). Web of Science: coleção principal. Disponível em: <[http://www.periodicos.capes.gov.br/?option=com\\_pollction&mn=70&smn=79&cid=81](http://www.periodicos.capes.gov.br/?option=com_pollction&mn=70&smn=79&cid=81)>. Acesso em: 11 jan. 2015.
- SETZER, V. W. D. (1999). ado, informação, conhecimento e competência. *Datagramazero*, Rio de Janeiro, n. 0, dez. Disponível em: <[http://www.dgz.org.br/dez99/Art\\_01.htm](http://www.dgz.org.br/dez99/Art_01.htm)>. Acesso em: 06 jan. 2015.
- TERRA, J. C. (2005). Gestão do conhecimento: o grande desafio empresarial. *Terra Fórum Consultores*. Disponível em: <[http://www.terraforum.com.br/biblioteca/Documents/libdoc0000011v002Gestao%20do%20Conhecimento\\_%20O%20grande%20desafio%20e.pdf](http://www.terraforum.com.br/biblioteca/Documents/libdoc0000011v002Gestao%20do%20Conhecimento_%20O%20grande%20desafio%20e.pdf)>. Acesso em: 02 jan. 2015.
- XAVIER, R. C. M.; COSTA, R. O. (2010). Relações mútuas entre informação e conhecimento: o mesmo conceito? *Ci. Inf.*, Brasília, DF, v. 39 n. 2, p.75-83, maio-ago.

## Anexo 1

| Termo de Busca  | Web of Science   |
|---|--|
| <b>Knowledge AND Management</b><br>– período: de 2010 a 2014<br>– recuperação no campo título | <b>3200</b><br><br><b>2010</b> – 697 registros<br><b>2011</b> – 667 registros<br><b>2012</b> – 666 registros<br><b>2013</b> – 641 registros<br><b>2014</b> – 529 registros |
| <b>DIMINUIÇÃO DE RECORRÊNCIA DE REGISTROS COM O PASSAR DOS ANOS</b>                           |  |

*Quadro 2. Pesquisa sobre GC em Âmbito Geral*

## Anexo 2

| Pesquisa / Termos de Busca<br>- período: de 2010 a 2014<br>- recuperação no campo <u>título</u> | Recorrência de Registros na <i>Web of Science</i>                    |
|---|--|
| 1 - Knowledge AND Management AND Education  | 103 (pesquisa utilizada para análise, denominada <i>Pesquisa 1</i> ) |
| 2- Knowledge AND Management AND Educational   | 27   |
| 3 - Knowledge AND Management AND High School  | 3  |
| 4 - Knowledge AND Management AND Professional AND Educational                                   | 1  |
| 5 - Knowledge AND Management AND Vocational AND Educational                                     | 3  |
| 6 - Knowledge AND Management AND Technological AND Education                                    | -  |

*Termos utilizados nas pesquisas sobre GC no contexto educacional*

**Hito 1: Modelos de organización del conocimiento. Modelaje de sistemas y ciclo de vida de la información**

**Hito 2: Entornos y sistemas para la organización del conocimiento. Diseño, normalización e interoperabilidad de sistemas de información abiertos.**

**Hito 3: Instrumentos para el acceso y transparencia de la información: gestión de documentos, plataformas digitales y archivos**

**Hito 4: Usuarios y usos de la información. Patrones vinculados a la transparencia y el gobierno abierto**

**Hito 5: Gobernanza y Auditoría de información en la organización del conocimiento.**

**Hito 6: Representación del conocimiento: procesos y herramientas en sistemas abiertos y multilingües**

**Hito 7: Epistemología en la organización del conocimiento**

**Hito 8: Igualdad, género y diversidad en la organización del conocimiento**

---

# Competências infocomunicacionais para acesso e uso da informação

*InfoCommunication literacy o access and use of information*

---

**Jussara Borges (1), Miguel Ángel Marzal García-Quismondo (2)**

(1) Instituto de Ciência da Informação – Universidade Federal da Bahia, Rua Basílio da Gama, s/n, Canela, Salvador, Bahia, Brasil, jussarab@ufba.br

(2) Departamento de Biblioteconomía y Documentación – Universidad Carlos III de Madrid, Calle Madrid, 126, Getafe, Madrid, España, mmarzal@bib.uc3m.es

## Resumo

As competências infocomunicacionais referem-se à capacidade dos usuários para o acesso e uso da informação, bem como a sua comunicação com outros atores sociais. Cada vez mais o domínio dessas competências representa a possibilidade de se integrar socialmente e participar dos processos decisórios da sociedade. Assim, este trabalho tem como objetivos: 1) Aprofundar as abordagens teóricas das competências; 2) Apontar as competências infocomunicacionais desenvolvidas por organizações da sociedade civil de Salvador (Bahia-Brasil) no âmbito da participação cívica. Quanto à metodologia, o primeiro objetivo é basicamente embasado em levantamento e sistematização bibliográfica, enquanto para o segundo, os dados foram coletados por meio de entrevistas com os gestores de organizações em Salvador e observação sistemática de seus sítios, blogs e redes sociais. Os resultados empíricos apontam as competências que organizações têm desenvolvido, destacando-se a competência operacional, se comparada às demais. Conclui-se que a qualificação de organizações para a participação política que se fixe apenas nas competências operacionais é insuficiente, pois não garante capacidade analítica e crítica, estas indispensáveis para a participação ativa e esclarecida nos processos sociais.

**Palabras-chave:** Usuários de informação. Competências infocomunicacionais. Competência em informação. Competência em comunicação. Organizações da sociedade civil. Participação cívica.

## 1. Introdução

A necessidade de democratização no acesso e uso das informações como premissa para o desenvolvimento econômico-social e exercício da cidadania tem fomentado as discussões e iniciativas em torno de sistemas abertos de informação. No entanto, além de sistemas interoperáveis e que permitam uma eficaz gestão da informação, coloca-se a questão do usuário da informação e suas competências para aproveitar as potencialidades desses sistemas.

## Abstract

The infocommunication literacy refers to the ability of users to access and use of information and its communication with other social actors. Increasingly, the mastery of these skills is the ability to integrate socially and participate in decision-making processes of society. This work aims to: 1) To deepen the theoretical approaches of skills; 2) Point the infocommunication literacy developed by civil society organizations in Salvador (Bahia, Brazil) under the civic participation. As for methodology, the first goal is grounded in systematic survey and literature, while for the second, data were collected through interviews with managers of organizations in Salvador and systematic observation of their sites, blogs and social networks. The empirical results show the skills that organizations have developed, operational expertise highlighting, compared to the others. It concludes that the qualification of organizations for political participation that set only in operational skills is insufficient because it does not ensure analytical and critical thought, these indispensable for the active and informed participation in social processes.

**Keywords:** Users of information. Infocommunication literacy. Information literacy. Communication competence. Civil society organizations. Civic participation.

As competências infocomunicacionais referem-se à capacidade dos usuários para o acesso e uso da informação, bem como a sua comunicação com outros atores sociais. Isso porque a cidadania, aqui compreendida como a potencialidade do indivíduo para compreender sua realidade, participar das decisões coletivas com capacidade crítica e receber seus potenciais benefícios, pressupõe um indivíduo capaz de acessar, compreender, assimilar, usar e comunicar informações em seu benefício e de sua comunidade. Portanto, cada vez mais o domínio

dessas competências representa a possibilidade de se integrar socialmente e participar dos processos decisórios da sociedade.

Embora os recursos disponíveis na internet tenham potencializado as possibilidades de acesso à informação e interação, colocam-se demandas mais exigentes ao indivíduo, como selecionar a informação adequada e pertinente às suas necessidades, adequar a linguagem e o meio a uma audiência cada vez mais plural, considerar as regras de convivência do ciberespaço e ainda atentar para aspectos éticos e legais do uso da informação e da comunicação em rede. Assim, se de um lado os recursos eletrônicos de comunicação potencializam a aproximação entre pessoas de diferentes culturas, por outro lado, expõem as fragilidades da sociabilidade intercultural, porque a interação exige respeitar o outro e suas características - ideológicas, linguísticas e emocionais - e ao mesmo tempo interpretar criticamente as mensagens e expressar-se claramente.

Este trabalho (1) faz parte de uma pesquisa de pós-doutoramento, na qual propõe-se “Formular abordagens teóricas e conceituais e procedimentos metodológicos orientados à aferição das competências infocomunicacionais”. Para este *paper* em específico, os objetivos são: 1) Aprofundamento e sistematização das abordagens teóricas das competências infocomunicacionais nos campos da Ciência da Informação e das Ciências Sociais; 2) Apontar as competências infocomunicacionais desenvolvidas por organizações da sociedade civil de Salvador (Bahia-Brasil) no âmbito da participação cívica. As organizações da sociedade civil são um ator social de relevo para a observação empírica porque muitas vezes representam segmentos populacionais em situação de exclusão social no Brasil e estão, cada vez mais, buscando manter-se informadas e alargar sua comunicação aproveitando os recursos disponíveis no ciberespaço.

No intuito de contextualizar os resultados obtidos, na seção seguinte o trabalho discute os conceitos que constituem a base teórica das competências infocomunicacionais.

## 2. Competências infocomunicacionais

O argumento de partida para tratar das competências infocomunicacionais é que a participação nos processos sociais contemporâneos está cada vez mais amparada na capacidade de valer-se dos recursos de informação e comunicação disponíveis. Assim, não há de se perder de vista que a argumentação em torno do desenvolvimento dessas competências tem como objetivo a formação de pessoas capazes a fazer

intervenções informadas e críticas naqueles processos que lhe interessam. Como destacam Sturges e Gastinger (2012), o indivíduo alfabetizado informacionalmente é o que tem competência para escolher recursos, formatos e suportes que tornem realidade seu direito à informação.

Cada vez mais os recursos de informação estão disponíveis on-line, o que permite acesso em tempo real e satisfação imediata das necessidades informacionais. Por outro lado, as fontes estão dispersas; quando possuem sistemas de busca, estes são específicos, demandando constante aprendizagem; não há garantia de veracidade e validade da maior parte do que é divulgado, gerando insegurança; e, como se não bastasse, a avalanche informacional cotidiana gera o sentimento de estar-se tomando decisões sem ter todas as informações necessárias.

El control de calidad y actualidad de los recursos de información digitales se vuelve cada vez más importante y necesario, como asimismo el control de la veracidad que acredita que la información brindada es cierta, tiene un autor responsable y se basa en fuentes fidedignas. Los recursos digitales son fuentes dinámicas, que están en continuo cambio y la accesibilidad que ofrecen las tecnologías para que la información que se ha puesto a disposición de todo el mundo sin pasar por filtros o controles, hacen que la evaluación de las mismas adquiere una relevancia fundamental. (Ceretta & Marzal, 2013, p. 372)

Tem-se então uma mudança de enfoque, que passa da valorização da quantidade de informação estocada para a auto formação utilitária (Montalvo, 2012). Assim, mais importante que memorizar informações é saber quando e onde buscá-las, como analisá-las e processá-las para produzir conhecimento, criar e inovar.

Em termos muito gerais, pessoas e organizações usam o ambiente digital para buscar informação, mas também, e cada vez mais, para se comunicar com outras. Da mesma forma, os últimos relatórios brasileiros TIC Domicílios e Empresas (Comitê Gestor da Internet no Brasil, 2014) vêm apontando o crescimento de atividades relacionadas à *comunicação* (participação em redes sociais, mensagens instantâneas e correio eletrônico) entre os que acessam a internet no Brasil; seguidos dos que utilizam a internet para buscar *informação* (sobre produtos e serviços, para ouvir música). Daí surge a proposta de síntese das competências *infocomunicacionais*. Vale frisar que as competências não são estanques: sua aquisição e desenvolvimento seguem um processo permanente de desenvolvimento, transformação e aperfeiçoamento.

A *competência em informação* é amplamente discutida no âmbito da Ciência da Informação e, em geral, relaciona-se à capacidade de perceber uma necessidade de informação, saber onde buscá-la, identificar o que é relevante, analisar aspectos como validade e veracidade e aplicá-la para resolver questões individuais e coletivas.

Mais recentemente o conceito vem sendo desenvolvido no sentido de se compreender que é um processo social no qual as pessoas estão sendo desafiadas a engajar conhecimentos e habilidades não só para localizar, avaliar e processar informação, mas também para a produção e comunicação da informação e de forma colaborativa:

Las competencias en información, por tanto, deberían ser el compendio de destrezas en el uso de herramientas informáticas y recursos de información, las habilidades en la organización y planificación de los instrumentos de gestión de contenidos, así como la evaluación esos contenidos conforme al plan para su conversión en conocimiento, para llegar a las competencias en la producción y difusión de estos contenidos mediante una edición colaborativa, continua y ética, de modo que este conocimiento pueda llegar a ser saber. (Ceretta & Marzal, 2013, p. 372)

Deve-se ressaltar, portanto, que a competência em informação não se limita ao acesso à informação, porque a apropriação só se realiza na construção de sentido pelo sujeito. Como ressaltam Area e Pessoa (2012, p. 15), a capacidade de interpretar as informações e atribuir-lhes significado útil a um propósito - ou saber empregar a informação para a resolução de um problema - é uma das metas mais relevantes da alfabetização na cultura digital.

Neste ponto reside um dos grandes desafios em termos de igualdade informacional no Brasil. Se por um lado, os números quanto ao acesso às TIC permanecem em crescente, quando se compara a aplicação que as pessoas são capazes de dar a esse acesso, percebe-se um uso limitado. Grandes dificuldades permanecem para atribuir significado à informação e, assim, convertê-la em conhecimento. Essa deficiência advém em parte da qualidade da formação escolar no Brasil: 40% dos estudantes que concluem o ensino fundamental (9 anos de escolaridade) não conseguem identificar o assunto de um texto. Com o ambiente não-linear característico da internet, esse problema tende a agravar-se, pois o hipertexto demanda do leitor a capacidade de articular informações dispersas:

[...] na leitura hipertextual a dinâmica é atribuída pelo leitor, podendo ser mais ou menos articulada, a depender da familiaridade e competência leitora

e informacional daquele que lê/navega através do hipertexto. Isto quer dizer que os sujeitos menos capacitados informacionalmente podem se perder no universo informacional, com quantidade cada vez maiores de informação. (Cuevas-Cervero, Marques, & Paixão, 2014, p. 37)

Outro ponto chave é a capacidade crítica perante a informação – compreensão crítica, negociação de sentidos –, pois com a liberação do polo de emissão (2) (Lemos, 2009), essa capacidade crítica precisa ser direcionada a toda e qualquer fonte de informação e comunicação, e não só para os meios de comunicação de massa.

Por outro lado, a capacidade de interação entre as pessoas tem sido pouco enfatizada. Este é o cerne das competências em comunicação sob o ponto de vista deste trabalho: entender as habilidades, conhecimentos e atitudes que as pessoas estão empregando ou desenvolvendo para relacionar-se umas com as outras através das tecnologias de comunicação. Como diz Cuevas-Cervero, Marques e Paixão (2014, p. 45): “Necessitamos neste momento, incluir os aspectos mais relacionados com os aspectos comunicativos, sociais e participativos da alfabetização.”

A demanda por competências em comunicação ficou mais evidente a partir da emergência das ferramentas 2.0, pois os processos sociais mediados pelo ciberespaço envolvem estabelecer relações com pessoas com quem muitas vezes não haverá a oportunidade de conhecer pessoalmente, manter comunicação durante o período necessário para uma ação, interagir com pessoas de culturas diversas e trabalhar em colaboração com equipes multidisciplinares. Então, propõe-se alargar a problematização e pensar também nessas competências. Como explicam Perrotti e Pieruccini (2007), tanto informar-se, quanto informar estão na base da participação nos processos sociais na atualidade, o que sugere o desenvolvimento de saberes e fazeres apropriados.

De alguna manera, a lo largo de nuestra vida desplegamos la capacidad comunicativa, com mayor o menor dominio em algunos de sus atributos. Adquirimos un repertorio lingüístico que se va ampliando hasta convertirse en un depósito enorme de términos y expresiones, desarrollamos habilidades de interacción intersubjetiva y también, cómo no, formas intermediarias de comunicación; como la escritura y el lenguaje digital suficiente para manejar los dispositivos electrónicos. En suma, alcanzamos, com mayor o menor nivel de dominio, la competencia comunicativa. (Montalvo, 2012, p. 52)

Além da capacidade de expressar-se adequadamente, a comunicação pressupõe considerar o outro. A comunicação é, por natureza, participativa, dialógica, retroalimentada por esse outro

que a aceita ou rechaça; seleciona e interpreta as mensagens dentro de um conjunto de significados e valores que podem ser muito distintos daquele do emissor. Para Wolton (2006), nisso reside o desafio da comunicação: aceitar o outro e defender a sua identidade própria.

A competência em comunicação refere-se, então, à capacidade de estabelecer interação com outras pessoas ou grupos, trocar, criticar e apresentar as informações e ideias de forma a atingir uma audiência e com ela manter uma relação bilateral. Isso pressupõe articular exposições claramente, fundamentando-as com argumentos estruturados, amparados no domínio do vocabulário e de um pensamento abstrato e bem organizado. Como salientam Cuevas-Cervero, Marques e Paixão (2014, p. 38): “[...] o ciberespaço é potencialmente um espaço de interação, onde emissores e receptores podem inverter os polos de interação, numa dinâmica relacional coautores/criadores.”

Subjacentes às competências em informação e comunicação, estão as *competências operacionais*. Dizem respeito à manipulação de computadores e artefatos eletrônicos incluindo um conhecimento básico de hardware, software e redes. O uso dessas ferramentas demanda certa capacidade operativa do usuário para compreendê-las e usá-las, bem como adaptá-las às suas próprias necessidades:

Las TIC han provocado, o al menos han acelerado, una revolución de amplio alcance en nuestra civilización que gira en torno a la transformación de los mecanismos de producción, almacenamiento, difusión y acceso a la información; en las formas y los flujos comunicativos entre las personas; así como en los lenguajes expresivos y de representación de la cultura y el conocimiento. (Area-Moreira & Ribeiro-Pessoa, 2012, p. 14)

Diversos projetos de alcance variado têm direcionado recursos e esforços na promoção das competências. Em seguida surgem métodos e instrumentos de análise ou verificação das competências. Esses padrões são indispensáveis para a conformação de políticas públicas e programas de formação que avancem da conexão para a formação e a inclusão, porque fornecem parâmetros de análise para ajustamento e desenvolvimento dos mesmos. De acordo com Wachholz (2014), precisamos formar as pessoas mediante o uso de novas tecnologias, mas também desenvolver indicadores que avaliem o progresso e as carências remanescentes.

Para verticalizar a compreensão das competências, o Gepicc (3) vem conduzindo pesquisas empíricas com organizações e grupos a partir da construção e aplicação de indicadores que

procuram identificar os componentes de cada competência (Quadro 1).

| Competências   | Componentes<br>O usuário sabe ou é capaz de... |
|----------------|--|
| Operacionais   | Operar computadores e aplicativos              |
|                | Operar um navegador na internet                |
|                | Operar motores de busca de informação          |
|                | Operar mecanismos de comunicação               |
|                | Operar recursos para produção de conteúdo      |
| em Informação  | Perceber uma necessidade de informação         |
|                | Acessar informações                            |
|                | Avaliar a informação                           |
|                | Organizar a informação                         |
|                | Criar conteúdo                                 |
|                | Auto-avaliar a competência em informação       |
| em Comunicação | Estabelecer e manter comunicação               |
|                | Criar laços sociais                            |
|                | Construir conhecimento em colaboração          |
|                | Avaliar a comunicação                          |

Quadro 1. Síntese das competências infocomunicacionais

### 3. Opções metodológicas

Este trabalho origina-se do projeto de pesquisa “Competências infocomunicacionais: desenvolvimento do conceito”. O projeto, como o título indica, pesquisa em âmbito teórico e empírico o conceito das competências infocomunicacionais, bem como sua relevância e aplicabilidade em contextos diversos, principalmente aqueles mediados pela internet. Para este trabalho, contexto empírico de observação é a participação política de organizações da sociedade civil (OSC).

A partir da literatura foram construídos indicadores para observação das competências infocomunicacionais. Os indicadores das competências operacionais e das competências em informação foram observados em 29 organizações sediadas em Salvador, a partir de seu discurso, dos produtos de informação que disponibilizam (sítios eletrônicos, *blogs* e redes sociais *on-line*) e da aplicação de um instrumento de observação específico para as competências em informação. Essa observação partiu da proposta de uma atividade prática: “Por favor, suponha que você tenha de disponibilizar um texto de dois ou três parágrafos no sítio eletrônico ou no *blog* da

organização. Você tem de levantar informações na internet e depois redigir esse texto. Por favor, descreva em voz alta quais os procedimentos que emprega na operação.” À medida que o observado realizava as atividades, foi-lhe solicitado que explicasse os caminhos que estava tomando e os motivos das escolhas. A partir das ações dos observados, foram assinalados os indicadores satisfeitos para avaliar as competências infocomunicacionais. Além disso, muitas vezes, o domínio de uma determinada competência fez-se perceber pelo discurso do entrevistado.

Dentre as 29 OSC originais, para observar as competências em comunicação foram selecionadas as 17 que utilizam a internet de alguma forma para atuação política e mantêm sítios eletrônicos ou blogs ou participam de redes sociais *online*. Algumas informações foram confrontadas com a entrevista previamente realizada para aferição das competências em informação. Assim, sítios eletrônicos, blogs e redes sociais *online* foram analisados no sentido de identificar ou refutar os indicadores.

Os dados foram tratados primariamente por métodos estatísticos simples, levando à representação por meio de gráficos que facilitassem a visualização dos resultados. Em seguida, o conteúdo foi analisado descritivamente, no qual os indicadores foram associados e comparados. Por fim, buscou-se interpretar os resultados, procurando descobrir as relações de causalidade e interpretar as dinâmicas sociais em estudo.

#### 4. Resultados

As organizações da sociedade civil parecem ter visto no ciberespaço uma oportunidade de potencializar sua atuação política e social perante esta sociedade mais conectada à Internet. Portanto, as OSC apresentam a motivação para desenvolver as competências. Esse é um fator fundamental, conforme explica Miranda (Miranda, 2006, p. 113):

Essas competências podem ser desenvolvidas a partir do reconhecimento das necessidades existentes em lidar com o mundo da informação, formando um ciclo criativo e evolutivo de reconhecimento de necessidades e desenvolvimento de habilidades e atitudes apropriadas que vai oferecer possibilidades de solucionar problemas.

De fato, ao adentrar neste ambiente digital mais dinâmico e interativo, as OSC tendem a desenvolver novas competências com vistas a se informar e se comunicar de forma que possam atingir os objetivos pretendidos. Em contrapartida, por se tratar de um processo de aprendizado e adaptação, muitas dificuldades e barreiras

são encontradas. Assim, buscou-se analisar as competências infocomunicacionais que estão sendo desenvolvidas e empregadas por essas organizações em ambientes digitais. O Gráfico 1 (Anexo 1) mostra, sob uma perspectiva global, a média das competências que foram identificadas ao observar o comportamento das organizações diante do uso das ferramentas e recursos disponíveis no ciberespaço. Nota-se a sobreposição das competências operacionais em relação às demais, o que será mais detalhado ao longo do trabalho.

As organizações já estão cientes de que para entrar e se manter no ciberespaço precisam saber lidar com os aparatos digitais e se atualizar constantemente, por isso têm investido neste aprendizado, seja através de cursos e atividades de formação de seus quadros, seja no momento de seleção de novos colaboradores. Com isso, 85,5% dos indicadores observados quanto às *competências operacionais* foram satisfeitos.

Assim, operar computadores e artefatos eletrônicos (95,4%), navegadores de internet (96,5%) e motores de busca por informações (98,8%) não se constituem como problema para a grande maioria das organizações. Apesar de muitas estarem em zona de extrema pobreza, todas elas possuem algum tipo de equipamento eletrônico como computador, *tablets* e *smartphones* e vinte e cinco demonstram ter facilidade em operá-los. Além disso, afirmam não ter dificuldades em lidar com arquivos de imagens, sons e vídeo.

As maiores dificuldades percebidas quanto às competências operacionais se concentram em operar mecanismos comunicação (72,4%) e operar mecanismos de produção de conteúdo (64,6%) como, por exemplo, sítios eletrônicos, blogs e redes sociais. Das 29 organizações que participaram da pesquisa 12 se restringem ao uso do correio eletrônico e dos meios tradicionais como telefone, contato presencial e carta; apenas 17 utilizam meios alternativos e mais econômicos de se comunicar e disponibilizar as informações produzidas, como aplicativos de celular e redes sociais *online*.

Enquanto na pesquisa desenvolvida em 2010 com uma amostra de 44 organizações, apenas cinco mantinham blogs e quatro participavam de redes sociais online (Borges, 2013), atualmente esses recursos apresentam-se bastante disseminados: das 29 OSC que constituem a amostra atual, 9 mantêm blogs e 15 têm perfil em rede social online. Entretanto, a assimilação desses recursos não repercute automaticamente na melhoria das competências em informação e em

comunicação, haja vista os resultados obtidos quanto a essas competências.

O Gráfico 2 (Anexo 2) demonstra como as competências operacionais, informacionais e comunicacionais, respectivamente, distribuem-se irregularmente entre si. Enquanto alguns componentes das competências alcançam níveis elevados, outros alcançaram índices médios.

No processo de apropriação da informação, as competências agem conjuntamente e por isso não há fronteira rígida que as separe. Uma vez que as competências operacionais, informacionais e comunicacionais se envolvem e interagem mutuamente, quando há dificuldades ou deficiências, elas tendem a se ampliar. Por exemplo, as dificuldades operacionais de utilizar mecanismos de comunicação são barreiras para que as OSC possam estabelecer uma comunicação mais próxima de seus interlocutores; assim como a deficiência em operar recursos para a produção de conteúdo pode repercutir na capacidade de criar conteúdo.

Embora o uso da Internet tendo como objetivo final a informação, seja intenso e consideravelmente frequente – 18 organizações apontaram que se mantêm conectadas ininterruptamente, seguidas de outras nove que dizem acessar pelo menos uma vez ao dia – ao observar o Gráfico 2 (Anexo 2) vê-se que em quase 30% dos casos, as OSC enfrentaram problemas para perceber a necessidade de informação diante da situação/problema que foi proposta no momento da realização da pesquisa. As organizações mostraram-se “perdidas”, sem saber como resolver, apesar da questão ser passível de resposta com uma busca simples.

Por outro lado, vencida a barreira da percepção da necessidade de informação, a busca e acesso ganham destaque entre os componentes das competências em informação: 70,6% das organizações atendem aos indicadores. Ou seja, as OSC não apresentam dificuldades para escolher o sistema de busca adequado ao tipo de informação necessária, embora deva-se destacar a proeminência do Google entre os motores de busca.

Assim como apresentado na pesquisa de Lanzi e colegas (2013), o Google é a fonte de informação mais frequente, muitas vezes sem se considerar a adequação à necessidade de informação e a qualidade dos resultados apresentados. Outra semelhança é quanto à inabilidade para criar estratégias de busca mais efetivas. Embora os sujeitos da pesquisa de Lanzi tenham sido estudantes adolescentes, os dados aqui apontam que independente de faixa etária ou formação, há pouca atenção ao uso de ter-

mos e estratégias de busca precisas.

Seis organizações utilizaram termos demasiado abrangentes, levando à alta revocação, mas baixa pertinência. Tampouco é comum a adoção de expressões booleanas ou delimitadores, como as aspas, nos campos de busca. Atitudes assim levam ao comprometimento do processo, já que sem o uso de filtros adequados, o usuário recupera mais itens do que é capaz de avaliar e acaba fazendo escolhas aleatórias, ou seja, usualmente leem os resultados da primeira página, enquanto se despreza todo o restante.

Já o componente avaliar a informação é atendido pelo percentual de 76,5%. Procurou-se então verificar como as organizações avaliam a informação através de questões que denunciasses o seu comportamento. Isso inclui averiguar se consideram aspectos como pertinência, confiabilidade, correção e veracidade, se verificam as fontes quanto à fidedignidade, se diferenciam informação factual de opinião e se selecionam informação pertinente. A maioria (21) disse comparar a informação recuperada com o próprio conhecimento; 10 avaliam a autoridade da fonte e sete confrontam várias fontes.

Quanto à interligação de peças de informação, a média de sucesso das organizações situa-se em 78,4%. A hipertextualidade, característica da internet, permite um alto grau de liberdade ao usuário, que pode navegar através das ligações indefinidamente. Por outro lado, isso exige a construção de mapas mentais para conectar os conceitos e um bom senso de orientação para não se perder no mar de informações. Observa-se que as organizações utilizam estratégias como: comparar informações entre si e com o conhecimento prévio para interpretá-las e conectá-las a um mapa conceitual, manter o senso de orientação entre as várias fontes consultadas, resumir a informação e as organizar visando sua recuperação para um uso atual e futuro.

Apesar do bom empenho em interligar as peças de informação, no momento da produção de conteúdo, apenas 65,5% dos indicadores são atendidos com o mesmo desempenho. Esta é uma questão delicada do ponto de vista da apropriação da informação, pois pode indicar que as organizações não se apropriam da informação com a mesma facilidade com que acessam: “[...] a apropriação implicaria atuação e afirmação dos sujeitos nas dinâmicas de negociação de significados, representaria, no caso específico que nos interessa, transação de significados que diferencia e constitui os negociadores como sujeitos da cultura, protagonistas, cidadãos.” (Perrotti & Pieruccini, 2007, p. 74)

O que se observa entre algumas OSC, portanto, é que ainda se posicionam como consumidoras de informação, mas não protagonizam ações produtivas, a despeito de terem potencialidade para produção de conteúdo, haja vista sua posição de liderança e presença nas comunidades em que atuam.

Como pode ser percebido no Gráfico 2 (Anexo 2), o componente criar conteúdo pode estar interferindo no componente seguinte - estabelecer comunicação - que atinge o percentual de 63,2%. Isso acontece porque a dificuldade de selecionar e reaproveitar o conteúdo acessado para produzir novas informações dificulta a presença da organização no ciberespaço expressando suas ideias. Além disso, percebe-se em algumas OSC, a dificuldade de compreensão das mensagens recebidas e a falta de customização da mensagem e do meio, considerando as características do receptor. Nessa perspectiva, entende-se que além das competências em informação para manterem-se informadas ou mesmo para informar, as OSC são demandadas também em *competências em comunicação*.

Com base nos dados levantados, 70,5% das OSC observadas participam de redes e comunidades virtuais com o objetivo de criar laços sociais. Nesse contexto, ganham destaque os *blogs* e as redes sociais. Destas, as preferidas são o Facebook e o Youtube. Além dessa participação nas redes, para criar laços sociais, 86,2% das organizações divulgam suas ações e possibilitam que o seu público-alvo tenha a oportunidade de interação.

Quanto ao uso de redes sociais *online* para criar laços com diferentes públicos, portanto, observa-se um bom índice de aproveitamento. Assim, é comum encontrar nesses espaços informações sobre acontecimentos importantes ou mesmo experiências compartilhadas pelas organizações com o intuito de se aproximar do público.

Os resultados da pesquisa também indicam que a maioria das OSC está engajada em mobilizar as redes sociais para conseguir ajuda quando precisa: 76,4%. Entre os motivos mais recorrentes estão: pedido de doações, recrutamento de voluntariado, convites para visitas ou, ainda, pedido de apoio às causas que defendem e programas e/ou projetos lançados.

Ademais, também é recorrente se deparar com OSC preocupadas em contribuir com seu próprio conhecimento e ainda argumentar e defender opiniões, como no trecho copiado de uma rede social de uma OSC: “A Família Tradicional, na cultura brasileira, apresenta as figuras paterna, materna e seus filhos. Atualmente há gran-

des debates envolvendo esta estrutura familiar, que em muitos casos, não preenche as características da família do alunado brasileiro, na Casa Pia não é diferente”. Nesse contexto, destaca-se uma atitude importante advinda das OSC: contribuir com seu próprio conhecimento (64,7%). Mesmo sem a intenção explícita, muitas organizações são formadoras de opinião e acabam contribuindo com o seu conhecimento ao disseminar informações sobre o trabalho que realizam. Este componente também está direta e claramente relacionado à capacidade de se apropriar da informação, pois na medida em que a interação promove o exercício da crítica frente aos conteúdos, a capacidade de trocar e se relacionar contribuem para a competência em informação.

As organizações foram questionadas, ainda, no que tange à avaliação da sua comunicação, 70,5% das organizações disseram estar atentas às mensagens indesejáveis, como por exemplo, spams e/ou vírus; 82,3% afirmam julgar questões relacionadas à privacidade e segurança das informações disponibilizadas; 47% garantem compreender as consequências de uma publicação online; e todas as organizações disseram considerar aspectos legais e éticos da comunicação, bem como avaliarem a própria comunicação. Há de se considerar, contudo, algumas incoerências: apesar de todas as OSC afirmarem respeitar os aspectos legais e éticos, poucas são as que, de fato, referenciaram as fontes citadas ou utilizaram recursos simples como aspas para indicar uma citação de terceiros.

## 5. Conclusão

Como Carpentier (2012) salienta, no Brasil observa-se o crescimento da diversidade de práticas políticas orquestradas por atores que frequentemente se situam fora do terreno da política institucionalizada, como grupos de interesse, movimentos sociais, ativistas e organizações da sociedade civil. Essas organizações por já terem uma motivação política anterior, assimilam os recursos eletrônicos que percebem como aliados para suas atividades. Há também no domínio desses recursos uma dimensão de inclusão sociocultural, na medida em que a presença no ciberespaço propicia que organizações – por vezes isoladas ou mesmo a margem da sociedade – possam estabelecer e estreitar suas relações com o mundo.

A partir dos resultados obtidos, conclui-se que as OSC desenvolveram competências operacionais satisfatórias para lidar com os recursos digitais, porém ainda esbarram em dificuldades

de natureza infocomunicacional diretamente ligadas à apropriação da informação. As principais limitações identificadas estão em perceber uma necessidade de informação e criar conteúdo – entre os componentes da competência em informação; e estabelecer comunicação e construir conhecimento em colaboração via Rede – entre os componentes da competência em comunicação.

Por outro lado, entre os indicadores das competências em comunicação com melhores resultados, estão aqueles relacionados com a competência de desenvolver laços sociais. Isso é explicado, de um lado por uma conjuntura nacional - o Brasil já é o terceiro país em usuários do Facebook - e de outro por uma necessidade premente para OSC: elas têm de estreitar laços com parceiros, colaboradores e sociedade civil para manterem-se competitivas. Assim, percebe-se a convergência de um contexto cultural que valoriza a participação em redes sociais online - o que impulsiona as organizações para esse ambiente que, por sua vez, facilita enormemente a divulgação de ações e o compartilhamento de experiências e vivências – com os objetivos institucionais que reclamam visibilidade e aceitação social. Uma vez que as competências infocomunicacionais não podem ser interpretadas como um fim em si mesmo, ter o discernimento de sua aplicação e utilidade é imprescindível: as organizações as têm aplicado para usar os recursos eletrônicos e de rede como um meio para atingir objetivos.

A proeminência da competência operacional frente às demais, no entanto, evidencia que se um lado há uma crescente assimilação de recursos digitais, isso não reverbera diretamente na apropriação analítica e crítica da informação. Como a capacidade de informar-se e comunicar adequadamente estão na base da participação ativa e esclarecida nos processos sociais, a atuação dessas organizações pode ser limitada.

Para se beneficiar de uma cultura cada vez mais mediada pelos recursos digitais, todos nós, diariamente, estamos sendo demandados a lançar mão de competências que nos permitam usar esses recursos de maneira plena para nos consolidarmos como participantes nos processos decisórios da sociedade.

## Notas

- (1) O presente trabalho foi realizado com apoio do CNPq, Conselho Nacional de Desenvolvimento Científico e Tecnológico – Brasil.
- (2) Para Lemos, a liberação do polo de emissão está na base da cultura digital porque o antigo receptor passa a produzir e disseminar sua própria infor-

mação, livremente, em diversos formatos midiáticos e com alcance planetário.

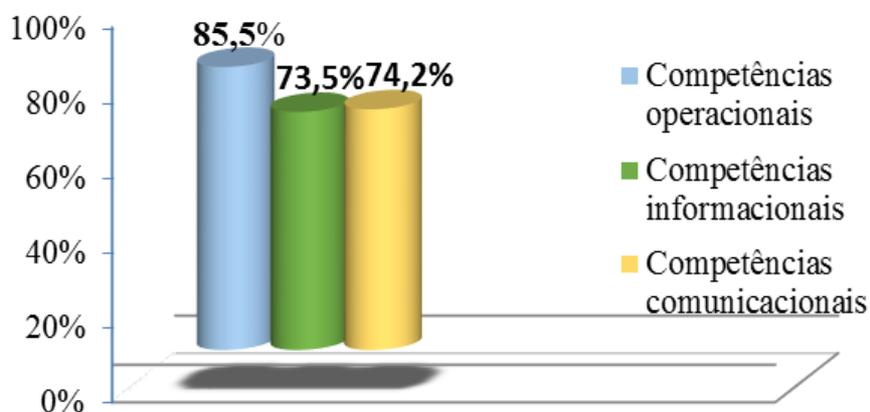
- (3) Grupo de Estudos de Políticas de Informação, Comunicações e Conhecimento. <http://www.gepicc.ufba.br/>

## Referencias

- Area-Moreira, M., & Ribeiro-Pessoa, M. T. (2012). De lo sólido a lo líquido: Las nuevas alfabetizaciones ante los cambios culturales de la web 2.0. *Comunicar: Revista Científica De Comunicación y Educación*, 19(38), 13-20.
- Carpentier, N. (2012). The concept of participation. if they have access and interact, do they really participate? *Revista Fronteiras*, 14(2)
- Ceretta, M. G., & Marzal, M. Á. (2013). Desarrollo de competencias en información: Otra modalidad para fortalecer las competencias lectoras. *Ciência Da Informação*, 40(3)
- Comitê Gestor da Internet no Brasil. (2014). *TIC domicílios e empresas 2013: Pesquisa sobre o uso das tecnologias de informação e comunicação no brasil*. São Paulo.
- Cuevas-Cervero, A., Marques, M., & Paixão, P. B. S. (2014). A alfabetização que necessitamos: Informação e comunicação para a cidadania. [A alfabetização que necessitamos: informação e comunicação para a cidadania] *Informacao e Sociedade*, 24(2), 35-48.
- Lanzi, L. A. C., Vechiato, F. L., Ferreira, Ana Maria Jensen Ferreira da, Vidotti, Silvana Aparecida Borsetti Gregorio, & Casarin, Helen de Castro Silva. (2013). Tecnologias de informação e comunicação no cotidiano dos adolescentes: Enfoque no comportamento e na competências informacionais da geração google. *Informação & Informação*, , 49-75.
- Lemos, A. (2009). Cibercultura como território recombinante. *A Cibercultura e Seu Espelho: Campo De Conhecimento Emergente e Nova Vivência Humana Na Era Da Imersão Interativa*. São Paulo: ABCiber, , 38-46.
- Miranda, S. (2006). Como as necessidades de informação podem se relacionar com as competências informacionais. *Ciência Da Informação*, 35(3), 99-114.
- Montalvo, J. A. C. (2012). La competencia comunicativa en la formación universitaria. Paper presented at the *Competencias Interdisciplinares Para La Comunicación y La Información En La Sociedad Digital*, 47-72.
- Perrotti, E., & Pieruccini, I. (2007). Infoeducação: Saberes e fazeres da contemporaneidade. In M. L. G. Lara, A. Fugino & D. P. Noronha (Eds.), *Informação e contemporaneidade: Perspectivas* (pp. 46). Recife: Néctar.
- Sturges, P., Gastinger, A., & Palacios, A. L. (2012). La alfabetización informacional como derecho humano. Paper presented at the *Anales De Documentación: Revista De Biblioteconomía y Documentación*, , 15(1) 10-14.
- Wachholz, C. (2014). Rumo às sociedades do conhecimento inclusivas: Onde nos encontramos hoje? A medição dos avanços concretizados desde a cúpula mundial sobre a sociedade da informação . In Comitê Gestor da Internet no Brasil (Ed.), *TIC domicílios e empresas: Pesquisa sobre o uso das tecnologias de informação e comunicação no brasil* (pp. 47-56). São Paulo
- WOLTON, D. (2006). É preciso salvar a comunicação. Casal de Cambra: Caleidoscópio.

## Anexo 1

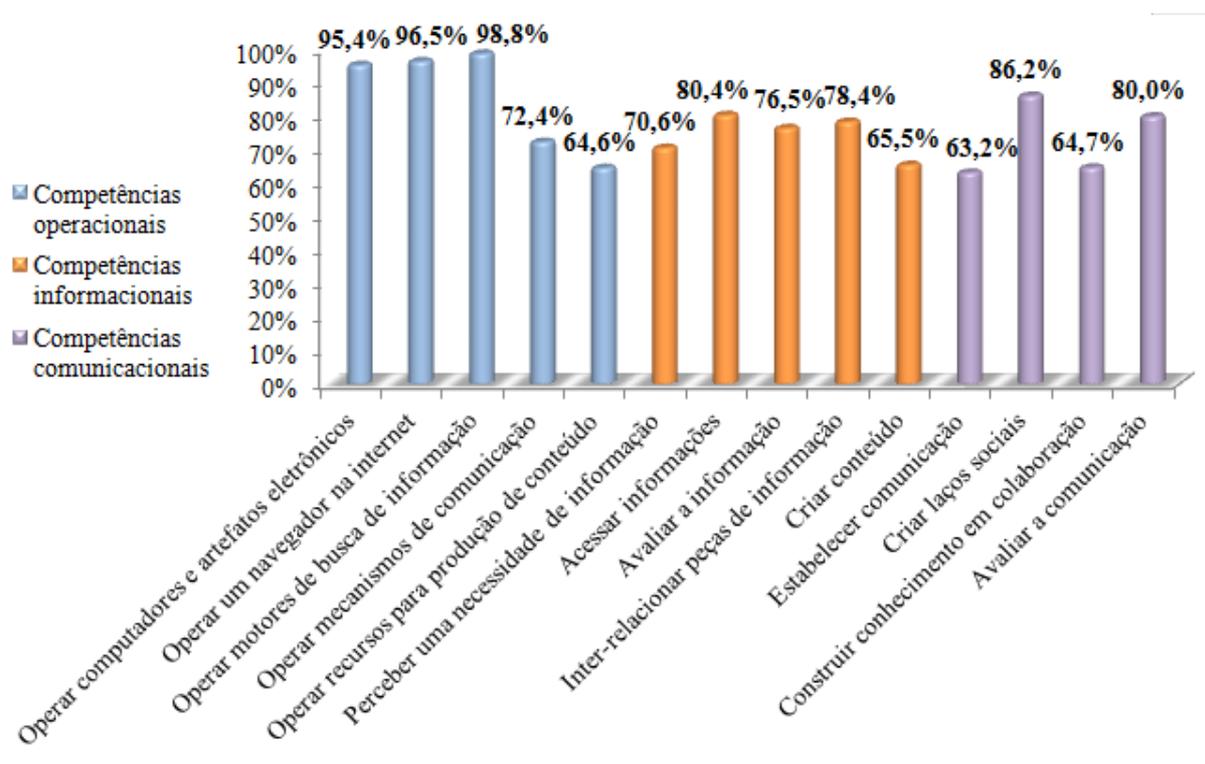
Gráfico 1 - Média das competências infocomunicacionais empregadas pelas OSC



Fonte: Dados da pesquisa empírica, 2013

## Anexo 2

Gráfico 2 - Emprego das competências infocomunicacionais pelas OSC



Fonte: Dados da pesquisa empírica, 2013

**Hito 1: Modelos de organización del conocimiento. Modelaje de sistemas y ciclo de vida de la información**

**Hito 2: Entornos y sistemas para la organización del conocimiento. Diseño, normalización e interoperabilidad de sistemas de información abiertos.**

**Hito 3: Instrumentos para el acceso y transparencia de la información: gestión de documentos, plataformas digitales y archivos**

**Hito 4: Usuarios y usos de la información. Patrones vinculados a la transparencia y el gobierno abierto**

**Hito 5: Gobernanza y Auditoría de información en la organización del conocimiento.**

**Hito 6: Representación del conocimiento: procesos y herramientas en sistemas abiertos y multilingües**

**Hito 7: Epistemología en la organización del conocimiento**

**Hito 8: Igualdad, género y diversidad en la organización del conocimiento**

---

# Uso de taxonomia sobre privacidade para identificação de atividades encontradas em termos de uso de redes sociais

*Using Privacy's Taxonomy to identify activities found on Social Network's terms of use*

---

**Fernando de Assis Rodrigues, M.S. (1), Ricardo César Gonçalves Sant'Ana, Dr. (2)**

(1) Universidade Estadual Paulista, Programa de Pós-Graduação em Ciência da Informação, avenida Hygino Muzzi Filho, 737 – Marília, Brasil – 17.525-000, fernando@elleth.org

(2) Universidade Estadual Paulista, Departamento de Ciência da Informação, avenida Hygino Muzzi Filho, 737 – Marília, Brasil – 17.525-000, ricardosantana@marilia.unesp.br

## Resumo

O objetivo deste trabalho é explicitar, a partir dos termos de uso de redes sociais, elementos que permitam identificar possíveis atividades que apresentem efeitos potencialmente prejudiciais a privacidade, seja por agentes internos ou externos. Para tanto, utilizou-se a Taxonomia da Privacidade de Solove para estabelecer uma categorização destas ações nos termos de uso disponibilizados pelas redes sociais, com enfoque nas diretrizes estabelecidas nestes termos sobre questões inerentes as fases de coleta e armazenamento de dados pessoais. O universo de pesquisa foi delimitado ao estudo dos termos de uso de três das principais redes sociais no momento da pesquisa. Os resultados estão divididos em duas partes: a primeira, identifica indícios de possíveis atividades prejudiciais à privacidade, através da vinculação das características de trechos nos termos de uso com os conceitos da taxonomia, e; a segunda, apresenta comparações entre as classificações possibilitadas pela taxonomia e seu nível de ocorrência nos termos de uso estudados. Concluiu-se que a aplicação de uma taxonomia apropriada pode auxiliar no estudo de termos de uso, possibilitando a explicitação de atividades previstas nestes termos, o que permite inclusive, propostas de futuras aplicações desta metodologia em outros contextos.

**Palavras-chave:** Privacidade. Tecnologias da Informação e Comunicação. Dados pessoais. Redes Sociais. Taxonomia da Privacidade.

## 1. Introdução

A penetração de Tecnologias da Informação e Comunicação (TIC) nas ações e atividades humanas – centradas em uma sociedade em rede, em que o principal ativo é a informação (Castells, 2003) – possibilitou o estabelecimento de um sistema que suportasse a formação de redes com milhares de usuários conectados, com a finalidade de troca de informações sobre inúmeros temas, além de representar um novo

## Abstract

The aim of this study is to explicit elements on social networks who allow the identification of possible activities that shows harmful effects on privacy, either when executed by internal or external agents. To achieve that, this work uses the Privacy Taxonomy by Solove to categorize those potentially harmful actions found on social network's terms of use, specially to emerge aspects inherent in data gathering and data storing phases using personal data. The universe of study was delimited to study terms of use from three prominent social networks at the time of the research. The results are divided into two parts: the first tries to identify evidences of possible activities harmful to privacy, through a linkage of characteristics in terms of use with concepts from Privacy Taxonomy. The second shows a comparison between classifications (enabled when the taxonomy was applied), and occurrences levels of this kind of activities on terms of use. It was concluded that applying an appropriated taxonomy can help on terms of use studies, enabling to explicit predicted activities in those terms of use, which allow, indeed, proposals of new applications of this methodology in others digital contexts.

**Keywords:** Privacy. Information and Communication Technologies. Personal data. Social Networks. Privacy Taxonomy.

local de organização social e cultural (Castells, 2002).

A internet oferece uma infraestrutura que permite o fluxo de inúmeros tipos de dados e documentos e com a criação em 1989 do protocolo *HyperText Markup Language (HTML)*, surgiram plataformas com o intuito de fornecer suporte às redes de informações e inter-relacionamento entre pessoas, denominadas redes sociais on-

line ou redes sociais (1) (Adamic e Adar, 2003; CERN, 2015).

As redes sociais estão presentes desde o início da internet (Adamic e Adar, 2003), e com o amadurecimento das tecnologias envolvidas neste cenário, estas redes sociais oferecem serviços específicos para o inter-relacionamento de seus usuários e para a troca de informações em formato multimídia, como imagens, vídeos, áudios, hiperlinks, e textos (Mislove, et al., 2007).

Algumas das redes sociais em atividade na década de 2010 atingiram números expressivos quando relacionados a população mundial (2). No mês de dezembro de 2014, a rede social Facebook (3) possuía uma média de 1,19 bilhão de usuários mensais e ativos. Portanto seus usuários representavam, aproximadamente, 17% da população mundial. Outras redes sociais apresentavam números totais acima de 250 milhões de usuários ativos, tais como o Instagram (4) e o Twitter (5) (BBC, 2014; Facebook, 2015a; WORLDOMETERS, 2015).

Para Mislove et al. (2007), os sítios destas redes sociais são parte integrante do conjunto dos serviços mais acessados e utilizados da internet.

As redes que são elaboradas e mantidas por instituições e empresas privadas para a troca de informações e experiências pessoais suscitam preocupações já existentes em outros contextos como: exposição de dados sensíveis de usuários para empresas, governos e, inclusive, para outros usuários; a perseguição de pessoas; ações resultantes de intolerância; entre outros. Em todos estes cenários, ocorrem exposições que perpassam questões inerentes a privacidade (Fogel e Nehmad, 2008; Krasnova, et al., 2009; Young e Quan-Haase, 2009; Chen e Zhao, 2012).

Destaca-se ainda que problemas relacionados a privacidade de usuários em redes sociais não são causadas somente pelo uso de TIC (que pode fazer papel de agente catalisador na coleta automatizada de grandes quantidades de dados sobre usuários), mas também pela atividade de usuários, de agentes externos e de controladores que possuam conhecimento suficiente para coletar e processar estes dados com outras fontes, estabelecendo novos dados potencialmente prejudiciais à privacidade (Solove, 2006; Fogel e Nehmad, 2008).

Nesse contexto destacam-se duas dimensões da privacidade a serem estudadas: os resultados de interações entre redes e usuários na fase de coleta de dados, e; a delimitação sobre

o que é realizado com os dados sobre usuários, na fase de armazenamento. Ambas dimensões são efeitos do uso de TIC (Sant'Ana, 2013).

Todavia, generalizar atividades que são prejudiciais à privacidade dos usuários pelos detentores da informação contida nas redes sociais é tema complexo, pois usuários são informados e devem conceder aceitação aos termos de uso da rede social na etapa que realizam o cadastro de uma conta para acessar o serviço. Os termos de uso destas redes têm duplo papel neste processo: pacificador, enquanto elemento de percepção de segurança aos usuários, ao estabelecer limites e garantias legais sobre o que é realizado com dados pessoais, e; como elemento de opacidade entre usuários e os serviços sobre o uso de seus dados, diluído em uma alta complexidade da rede e no volume e variedade de ações e atividades passíveis de realização (Castells, 2002). Desta forma, os ambientes não oferecem condições mínimas de uma plena compreensão ou acompanhamento de potenciais ações (e seus efeitos ou impactos) prejudiciais sobre sua privacidade.

A partir desta complexidade sobre questões de privacidade, Solove (2006) propõe uma taxonomia voltada à dificuldade de segregar tipos de atividades existentes que incidem sobre a privacidade, denominada Taxonomia da Privacidade. Esta taxonomia tem como objetivo simplificar o entendimento das atividades que de alguma forma possam prejudicar ou violar a privacidade dos sujeitos; e está dividida em grupos. Cada grupo possui uma definição de sua atividade prejudicial, descreve o *modus operandi*, as características da violação e prejuízos causados.

## 2. Objetivos e procedimentos

Este trabalho tem como objetivo explicitar, a partir dos termos de uso de redes sociais, elementos que permitam identificar possíveis atividades que apresentem efeitos potencialmente prejudiciais a privacidade, seja por agentes internos ou externos. Para tanto, utilizou-se a Taxonomia da Privacidade, proposta por Solove (2006), para estabelecer uma categorização destas ações nos termos de uso (6) disponibilizados pelas redes sociais, com enfoque nas diretrizes estabelecidas nestes termos sobre questões inerentes as fases de coleta e armazenamento de dados pessoais.

O universo de pesquisa foi delimitado ao estudo das coleções de documentos que compõem os termos de uso das redes sociais: Facebook, Instagram e Twitter. Esta escolha levou em consideração a quantidade de usuários ativos des-

tas redes, apresentada na introdução deste texto.

Não é escopo deste trabalho a análise se estes termos de uso garantem ou não a privacidade de dados pessoais para uma determinada atividade; mas sim, restringe-se a identificar, por meio de uma taxonomia apropriada, possíveis atividades potencialmente prejudiciais à privacidade de seus usuários. Também não é parte integrante as delimitações de localização (externa ou interna ao serviço) do potencial agente de execução destas atividades.

Como metodologia foi traçada uma análise sistematizada, a partir da leitura das coleções de documentos que compõem cada termo de uso das redes sociais, divididas em 3 etapas.

A primeira etapa (segundo capítulo) consiste na explicitação dos grupos de atividades prejudiciais à privacidade, a partir da proposta de categorização apresentada na Taxonomia da Privacidade de Solove (2006).

A segunda, identifica características nos termos de uso sobre questões de dados pessoais de seus usuários, a partir do isolamento de seções e de trechos identificados nos documentos disponíveis, apresentados no terceiro capítulo.

A terceira etapa – quarto capítulo – apresenta os resultados da correlação entre os subgrupos da Taxonomia da Privacidade e as características identificadas nos termos de uso, e estão divididos em duas partes: a primeira, identifica indícios de possíveis atividades prejudiciais à privacidade, através da vinculação das características de trechos nos termos de uso com os conceitos da taxonomia, e; a segunda, apresenta comparações entre as classificações possibilitadas pela taxonomia e seu nível de ocorrência nos termos de uso estudados.

### 3. Estrutura da Taxonomia da Privacidade

A Taxonomia da Privacidade (7) é composta por quatro grupos: Coleta de Informação, Processamento de Informação, Disseminação de Informação e Invasão, que neste trabalho passamos a identificar como Grupo I, II, III e IV, respectivamente.

Os grupos estão divididos em subgrupos, conforme o Quadro I (Anexo 1), que a síntese da taxonomia, agrupando conceitos das atividades prejudiciais à privacidade, relacionando-as com seus respectivos subgrupos.

#### Grupo I: Coleta de Informação

O primeiro grupo, denominado Coleta de Informação, envolve atividades de violação de privacidade no momento da coleta de dados sobre um indivíduo ou um coletivo de indivíduos. Este grupo está dividido em dois subgrupos: Vigilância e Interrogatório.

O subgrupo Vigilância concentra atividades encadeadas com o propósito de vigiar um indivíduo no seu espaço privado ou em espaço público.

Por exemplo, um serviço disponível na internet pode processar dados coletados em diferentes momentos sobre um usuário, sob a justificativa de uso destes na melhoria da experiência do usuário em sua plataforma; e executar ações de vigilância como o direcionamento de conteúdo a partir de dados coletados sobre trajetos do usuário (incluindo as coordenadas geográficas, a umidade, a pressão atmosférica e a altitude); informações do dispositivo de acesso e da rede de dados; histórico de comandos por voz; gostos e experiências sobre locais visitados; tempo de permanência em um local público ou privado; informações sobre a conexão de rede; metadados de imagens, áudios e vídeos; entre outros.

O subgrupo Interrogatório concentra atividades de processos de coleta de dados, baseados em interrogatórios e entrevistas.

Um exemplo são sítios que possuem, no processo de ingresso, formulários com campos de preenchimento obrigatório contendo informações que podem ser sensíveis para certa audiência – e caso o usuário não possua o interesse em compartilhar tais informações, não terá acesso ao serviço.

#### Grupo II: Processamento de Informação

O segundo grupo é denominado Processamento de Informação. Este grupo envolve atividades prejudiciais à privacidade a partir do processo de armazenamento (de persistência), de manipulação e de uso de dados sobre indivíduos. Este grupo está dividido em cinco subgrupos: Agregação, Identificação, Insegurança, Uso Secundário e Exclusão.

O subgrupo Agregação está ligado a atividades vinculadas ao processo de combinação de dados de múltiplas fontes sobre indivíduos, com o propósito de revelar fatos ocultos, quando eram analisados de forma separada.

Por exemplo, um usuário de uma rede social pode fornecer dados sobre sua relação matrimonial (ao preencher informações do seu perfil) e, em outro momento, navegar na internet por páginas sobre locais para o turismo voltado à celebração de lua de mel. Um agente externo,

que possuir o acesso para coletar estes dados, pode inferir, através de algoritmos especializados em agregação de dados, se este indivíduo tem propensão ou não para futuramente comprar produtos e serviços para recém-casados. Ao combinar dados destas fontes, o agente externo realiza uma atividade do subgrupo Agregação.

O subgrupo Identificação consiste em atividades a partir do processo de vinculação de dados que permitam a (re)identificação de usuários (e de seus dados pessoais) com suas respectivas pessoas.

Serviços na internet que oferecem acesso a dados sobre seus usuários podem ser passíveis de coleta destes conjuntos de dados por agentes externos – e estes podem ter a *expertise* necessária para recombinar estes conjuntos de dados coletados com outros dados (de outras plataformas), aumentando o repertório de informações sobre um determinado usuário e (re)identificando-o em diversos domínios (até vinculando estes dados com documentos pessoais, como números de seguro social ou cartão de crédito). Este tipo de atividade é parte integrante do subgrupo Identificação.

O subgrupo Insegurança compreende atividades que não garantem segurança sobre questões de acesso a dados pessoais aos envolvidos.

Por exemplo, quando uma rede é alvo de uma coleta de dados externa não autorizada, através de técnicas como o *exploit*, o resultado é um vazamento de dados pessoais que não há possibilidade de retorno ao estágio anterior (quando não havia o vazamento). Ou seja, como a política de acesso a estes dados foi comprometida e não há mais garantias que não existam cópias destes dados sob tutela de terceiros – uma atividade ligada ao subgrupo Insegurança.

No subgrupo Uso Secundário estão contidas atividades que envolvem o uso de dados coletados para um determinado propósito e utilizados *a posteriori* para outras finalidades.

Quando dados são coletados sobre indivíduos com um propósito e são utilizados para outra finalidade, sem consentimento, tais como o uso e compartilhamento de dados pessoais a terceiros para personalização de propagandas, são atividades que integram o subgrupo Uso Secundário.

O subgrupo Exclusão consiste nas atividades que apresentam opacidade ao indivíduo nos processos: de armazenamento de dados pessoais; no compartilhamento destes dados a terceiros, e; na ausência ou na inabilidade de

participação nas decisões sobre questões envolvendo a coleta, o armazenamento, o uso e o compartilhamento destes dados.

Por exemplo, uma rede social em que o indivíduo desconhece, não tem acesso ou não participa de decisões sobre o uso de seus dados, é uma atividade integrante do subgrupo Exclusão.

### Grupo III: Disseminação de Informação

O terceiro grupo é denominado Disseminação de Informação e envolve atividades de publicação, exposição e disseminação de informações sobre indivíduos. Está dividido em sete subgrupos: Quebra de Sigilo, Divulgação, Exposição, Aumento do Acesso, Chantagem, Apropriação e Distorção.

O subgrupo Quebra de Sigilo consiste em atividades em que ocorrem a quebra de confiança entre as partes em manter a confidencialidade das informações sobre indivíduos.

Um serviço que estabelece em seus termos de uso o não-compartilhamento de dados pessoais a terceiros e, *a posteriori*, estes dados estão disponíveis para um público externo pré-selecionado ou são acessíveis de forma direta ou indireta publicamente, desenvolve-se uma atividade integrante do subgrupo Quebra de Sigilo.

O subgrupo Divulgação consiste em atividades de divulgação e de disseminação de informações sobre um indivíduo, que acarretam mudanças na maneira que outros indivíduos julgam seu caráter.

Quando não está transparente a usuários qual o repertório de informações estará disponível para seus pares e para as conexões de seus pares, este tipo de atividade ofusca a real delimitação da possível audiência, o que pode acarretar um julgamento sobre seu caráter sobre questões de cunho privado a partir de dados pessoais. Este tipo de atividade é parte do subgrupo Divulgação.

O subgrupo Exposição consiste nas atividades vinculadas a exposição para terceiros de atributos emocionais ou físicos de intimidade do indivíduo, tais como a nudez, funções corporais e informações de cunho privado.

Sítios de compartilhamento de conteúdos multimídia podem armazenar e preservar (persistir) dados sobre fotografias e vídeos pessoais sem o consentimento ou pleno entendimento dos envolvidos, revelando sua intimidade a terceiros. Mesmo que o conteúdo possa ser removido a intimidade dos envolvidos já foi revelada, parte de uma atividade do subgrupo Exposição.

O subgrupo Aumento do Acesso agrupa atividades que visam amplificar o acesso a dados pessoais além do previsto ou do combinado entre as partes.

Quando um sítio compartilha dados de usuários com outros serviços (próprios ou de terceiros) amplia o acesso a estes dados além do previamente consentido, mesmo quando este processo está explicitado no termo de uso, pois dados pessoais enviados a estes serviços estão mais sujeitos aos seus respectivos termos de uso, mas sim aos termos de uso dos outros serviços – e estes outros serviços podem ter, em seus termos de uso, delimitações distintas sobre quais dados pessoais serão (re)compartilhados com seus parceiros. Esta atividade é parte do subgrupo Aumento do Acesso;

O subgrupo Chantagem agrupa atividades de controle, de dominação, de intimidação ou de ameaças a pessoas ou grupos, por terceiros.

A ocorrência de chantagens, intimidações e ameaças de grupos ou de indivíduos, com o uso da extorsão para angariar recursos financeiros a partir da revelação de dados pessoais (como fotografias íntimas), integra o subgrupo Chantagem.

No subgrupo Apropriação estão concentradas atividades que utilizam dados pessoais de um determinado sujeito em benefício de um terceiro ou para cancelar um serviço ou um produto, sem o pleno consentimento do sujeito.

Uma rede que utiliza de dados pessoais e fotos de seus integrantes para cancelar produtos de seus colaboradores executa uma atividade integrante do subgrupo Apropriação.

O subgrupo Distorção compreende atividades de disseminação de informações falsas ou interpretadas de maneira dúbia sobre um indivíduo.

Sítios que divulgam parte dos dados pessoais em um catálogo público, a partir da coleta de dados sobre usuários de outras redes sociais, podem descontextualizar estas informações, criando novas interpretações. É uma atividade do subgrupo Distorção.

#### **Grupo IV: Invasão**

O quarto grupo é denominado Invasão e engloba atividades de invasão à privacidade de indivíduos. O grupo está dividido em dois subgrupos: Intromissão e Interferência Decisional.

O subgrupo Intromissão consiste em atividades com o propósito de realizar incursões em assuntos ou em informações de caráter privado.

Por exemplo, a obrigatoriedade de instalar ferramentas ou utilizar serviços com propósito de gravar dados sobre ações em um determinado ambiente, sem consentimento das partes, é uma atividade do subgrupo Intromissão.

O subgrupo Interferência Decisional consiste em atividades de envolvimento do Estado em assuntos de caráter privado, executando decisões em nome do indivíduo.

Quando o Estado decide sobre ações de cunho privado, como partes do corpo humano ou sobre a privacidade dos indivíduos, estas atividades são parte integrante do grupo Interferência Decisional.

### **4. Características dos Termos de Uso**

As redes sociais Facebook, Instagram e Twitter possuem, nos rodapés de seus respectivos sítios, uma área específica para o acesso aos termos de uso dos serviços oferecidos, com o rótulo intitulado *Terms* ou seu equivalente em língua estrangeira (8), e são explicitadas suas características, individualmente, nas seções a seguir.

#### **4.1. Facebook**

O termo de uso é composto por um documento principal denominado *Statement of Rights and Responsibilities* e onze documentos adicionais sobre serviços ou temáticas específicas que a rede social tem envolvimento direto ou indireto: *Data Policy*, com informações sobre a política sobre dados pessoais; *Payment Terms*, com termos adicionais sobre pagamentos pela ferramenta; *Platform Page*, com informações sobre intercâmbio de dados com aplicativos externos; *Facebook Platform Policies*, com as diretrizes para aplicações; *Advertising Guidelines*, com as diretrizes sobre publicidade na rede social; *Self-Serve Ad Terms*, com as diretrizes sobre publicidade para ferramentas conectadas a rede social; *Promotions Guidelines*, com as diretrizes para a elaboração de concursos e premiações; *Facebook Brand Resources*, com as diretrizes para utilização de material de propriedade do Facebook ( *copyright*); *How to Report Claims of Intellectual Property Infringement*, com as diretrizes de direitos autorais; *Pages Terms*, com as diretrizes de uso para criação de páginas, e; *Community Standards*, contendo informações normas de boas condutas e ações aprovadas ou não sobre convivência entre seus usuários. Todos estes documentos estão disponíveis mesmo quando usuário não está identificado na rede social (Facebook, 2015b).

Segundo os documentos disponíveis, o usuário do Facebook, ao postar fotos e vídeos na plata-

forma, garante o direito compartilhado ao Facebook de utilização e transferência de seus dados pessoais a parceiros, em qualquer localidade e a qualquer momento, cessando quando o usuário exclui a sua conta. No documento *Data Policy*, está descrito que a plataforma pode compartilhar dados de usuários com outras empresas parceiras, tais como equipes de vendas externas, empresas de publicidade, escritórios regionais, entre outros (Facebook, 2015c).

O termo de uso não contém informações detalhadas sobre a forma que é realizado o processo de exclusão de dados pessoais a pedido do usuário. Todavia, é presente no texto que dados pessoais que forem relacionados à conteúdos licenciados com *copyright* ou compartilhados a terceiros podem continuar disponíveis, mesmo quando solicitado a sua exclusão (Facebook, 2015b).

Na quarta seção denominada *Registration and Account Security*, do documento *Statement of Rights and Responsibilities*, estão estabelecidas quais condutas são permitidas aos usuários, destacando-se que usuários da rede social devem concordar com a obrigação de informar todos os seus dados pessoais de forma “precisa e atualizada” (Facebook, 2015b, p. 1). A rede social reserva-se no direito de bloquear o acesso do usuário, caso considere que estes dados pessoais ou atividades exercidas na plataforma não estejam de acordo com as condutas preestabelecidas (Facebook, 2015b).

A décima sétima seção (denominada *Definitions*) do documento *Statement of Rights and Responsibilities*, contém a definição com o significado de alguns termos, como a definição para o termo plataforma, que denomina o conjunto de todos os serviços e Interfaces para Programação de Aplicativos externos (API, acrônimo proveniente do termo na língua inglesa) oferecidas pela empresa Facebook (Facebook, 2015b). Estas plataformas habilitam o compartilhamento de dados pessoais do Facebook por “[...] outros, incluindo desenvolvedores de aplicativos e administradores de sítios [...]” (Facebook, 2015b, p. 1).

No documento *Data Policy* é delimitado que metadados identificados nos conteúdos multimídias enviados por usuários são armazenados, como: a localização geoespacial, a data de criação do conteúdo, informações do dispositivo de acesso e do provedor de internet, o endereço IP, o idioma, protocolos, o número do telefone, entre outros (Facebook, 2015c).

O documento *Data Policy* explicita que informações sobre o usuário podem ser utilizadas na personalização de propagandas e em “[...] sis-

temas de medições para [...] exibir anúncios relevantes ou não dos nossos serviços e medir a efetividade e o alcance de nossos anúncios e serviços” (Facebook, 2015c, p. 1).

#### 4.2. Instagram

O termo de uso está dividido em cinco documentos: *Terms of Use*, com informações gerais sobre o termo de uso; *Privacy Policy* com informações vinculadas a política de privacidade adotada pela rede social; *API Terms of Use*, com informações sobre os direitos e deveres de aplicações externas, que utilizam dados de seus usuários coletados através de API; *Community Guidelines*, contendo informações sobre atitudes aprovadas ou não de convivência entre seus usuários, e; *Intellectual Property*, com informações sobre *copyright* e licenciamentos (Instagram, 2015a). Todos estes documentos estão disponíveis mesmo para quando o usuário não está identificado formalmente na rede social.

Segundo os documentos disponíveis, o usuário do Instagram, ao postar fotos e vídeos na plataforma, garante o direito compartilhado ao Instagram de utilização e transferência de seus dados pessoais aos seus parceiros, em qualquer localidade e momento, cessando este direito quando o usuário exclui sua conta. No documento *Terms of Use*, para o usuário ter acesso aos aplicativos desenvolvidos por terceiros em sua conta, consentirá, *a priori*, a publicação de dados pessoais com a empresa terceirizada, como: o nome da conta, o nome de usuário e a foto do perfil (Instagram, 2015a).

O termo de uso não contém informações detalhadas sobre a forma que é realizado o processo de exclusão de dados pessoais a pedido do usuário. Todavia, é presente no texto que os dados pessoais que são compartilhados aos seus parceiros podem continuar disponíveis, sem informar um processo que possibilite a exclusão destes (Instagram, 2015a).

No documento *Terms of Use*, os itens 4, 5 e 6 da seção *General Conditions* definem que o sítio pode: bloquear o acesso, excluir ou forçar a modificação conteúdos ou nome de usuário, e; monitorar atividades de seus usuários a qualquer momento, sem aviso prévio (Instagram, 2015a).

No documento *API Terms of Use* é delimitado que o desenvolvedor de um aplicativo externo deve informar aos usuários como dados pessoais são coletados, armazenados, processados e disseminados (Instagram, 2015b).

O documento *Privacy Policy* informa ao usuário que metadados identificados nos conteúdos multimídias enviados por usuários são armazenados, como: a localização geoespacial, a data de criação, informações do dispositivo de acesso e do provedor de internet, o endereço IP, o idioma, protocolos, o número de telefone, entre outros (Instagram, 2015c).

#### 4.3. Twitter

O termo de uso é composto por um documento principal denominado *Twitter Terms of Service* e cinco documentos adicionais sobre serviços ou temáticas específicas que a rede social tem envolvimento direto ou indireto: *The Twitter Rules*, com os princípios de convivência e uso da ferramenta; *Twitter Privacy Policy*, com informações sobre coleta, uso e compartilhamento de dados pessoais; *Developer Agreement & Policy*, com os direitos e deveres de parceiros desenvolvedores de aplicações externas, que utilizam dados de seus usuários; *Twitter Commerce Terms*, com termos adicionais sobre pagamentos e promoções pela ferramenta, e; *Inactive Account Policy*, com informações sobre exclusão ou a queda por desuso (9) de contas de usuários (Twitter, 2015a). Todos estes documentos estão disponíveis mesmo para quando o usuário não está identificado formalmente na rede social.

O documento do termo de uso principal da rede social estabelece, na seção *Privacy*, que todas as questões relativas ao uso de informações pessoais pelo sítio ou por terceiros são balizadas pelo documento auxiliar *Twitter Privacy Policy* (Twitter, 2015a, 2015b).

O usuário ao postar fotos e vídeos na plataforma garante o direito ao Twitter de utilizar e transferir a terceiros estas informações, em qualquer localidade e momento, salvo quando o usuário exclui o conteúdo. Também reserva-se no direito de compartilhar dados pessoais como: a localização geoespacial, a data de criação de conteúdos, textos, imagens, hiperlinks e o idioma com empresas parceiras e agências de publicidade, através de sua API (Twitter, 2015b).

Por padrão, as postagens de conteúdo do usuário é acessível de forma pública (Twitter, 2015b) e estes conteúdos associados aos dados pessoais sobre o usuário podem ser utilizados para personalizar propagandas e para o entendimento do público e o desenvolvimento do próprio serviço (Twitter, 2015a, 2015b).

## 5. Resultados

A primeira parte dos resultados apresenta os vínculos de indícios de possíveis atividades prejudiciais à privacidade com as características de trechos nos termos de uso com os conceitos da Taxonomia da Privacidade (Solove, 2006), agrupados por rede social. A segunda parte apresenta comparações entre as classificações possibilitadas pela taxonomia e seu nível de ocorrência nos termos de uso estudados.

### 5.1. Vinculação entre a Taxonomia da Privacidade e características dos termos de uso

Nesta seção, cada parágrafo apresenta características identificadas em trechos nos termos de uso e a vinculação com atividades dos subgrupos da Taxonomia da Privacidade são apresentados no final das sentenças. Um trecho pode ser vinculado com um ou mais subgrupos e subgrupos podem ser vinculados com ou mais trechos.

#### Facebook

Na rede social Facebook, o usuário não encontra um local ou documento que estabeleça quais são as empresas parceiras que acessam seus dados pessoais, sendo possível vincular este tipo de atividade ao subgrupo Aumento do Acesso.

Não é possível delimitar de forma transparente quais informações são geradas por empresas parceiras a partir de dados de usuários ou a forma que manipulam e processam estes dados pessoais com outras fontes de dados, o que pode acarretar atividades ligadas aos subgrupos Uso Secundário e Agregação.

Apesar do documento principal do termo de uso estabelecer em sua segunda seção, denominada *Sharing Your Content and Information*, que todos os dados sobre o usuário é de posse do mesmo (Facebook, 2015b), há certa opacidade no *modus operandi* por empresas parceiras após ter o acesso concedido aos dados pessoais, podendo ser um ambiente catalisador de possíveis atividades de quebra de sigilo, como estabelecido no subgrupo Quebra de Sigilo.

Segundo o *Data Policy* (Facebook, 2015c), conteúdos e dados pessoais podem ser compartilhados com empresas de publicidade. Estas empresas de publicidade estão aptas a utilizar dados pessoais (como gostos ou interesses do usuário por páginas, produtos ou serviços) para cancelar ofertas de produtos ou serviços perante o público da rede social, sem a devida transparência ao usuário sobre como este processo é realizado (como, por exemplo, determi-

nar qual foi o público ou o tempo de vigência desta ação), concomitante com atividades ligadas ao subgrupo Apropriação.

Apesar do termo de uso delimitar quais são as atividades permitidas aos usuários na rede social, há opacidade nos procedimentos sobre a arbitragem de supostas violações (como no caso do bloqueio de contas de usuário), diretamente ligado a questões inerentes aos subgrupos Exclusão, Vigilância e Insegurança.

O processo de exclusão de dados pessoais não é transparente ao usuário, principalmente no compartilhamento de conteúdo com seus pares e no vínculo de dados e gostos aos conteúdos *copyright*. Isto cria certa insegurança ao usuário sobre seu modo de agir na rede – já que não há garantias se o conteúdo publicado será excluído por completo, caso o usuário manifeste interesse – vinculando esta atividade ao subgrupo Insegurança.

É possível realizar a coleta de dados através da API, permitindo que possam existir ocorrências de ações de coleta automatizada, por agentes externos ao sítio, sem o consentimento prévio do usuário (pois são permissivos novos acordos com empresas parceiras, ao longo do tempo), podendo desencadear atividades vinculadas ao subgrupo Quebra de Sigilo.

A rede social Facebook pode compartilhar a foto, o nome e outras informações variáveis (não informadas) sobre usuários com empresas parceiras – e estas empresas podem agregar estes dados com suas bases de dados – permitindo que possam existir ocorrências de atividades vinculadas aos grupos: Agregação e Apropriação.

Não é transparente ao usuário se a rede social compartilha metadados dos conteúdos multimídia de usuários com empresas parceiras. Caso estes metadados sejam compartilhados, é possível agregá-los com dados pessoais, garantindo a estas empresas parceiras informações adicionais, como: a localização geoespacial e temporal do conteúdo, atividade que é parte do subgrupo Agregação.

Empresas de publicidade, parceiras da rede social, podem processar dados de usuários com outras fontes, com objetivo de personalizar propagandas e conteúdos a serem apresentados ao usuário. Também é permitindo estender o acesso a estas informações para terceiros, através da compra de espaços publicitários. Estas atividades estão ligadas aos subgrupos Uso Secundário e Agregação.

Como anúncios publicitários personalizados precisam, *a priori*, consumir dados pessoais

para aferir sua precisão, a empresa terceira necessita de acesso identificável sobre dados de cada usuário, permitindo a sua identificação sobre o grupo. Esta atividade é vinculada ao subgrupo Identificação.

### Instagram

Na rede social Instagram, o usuário não encontra um local ou documento que estabeleça quais são as empresas parceiras que acessam seus dados pessoais, sendo possível vincular este tipo de atividade ao subgrupo Aumento do Acesso.

Não é possível delimitar de forma transparente quais informações são geradas por empresas parceiras a partir de dados de usuários ou a forma que manipulam e processam estes dados pessoais com outras fontes de dados, o que pode acarretar atividades ligadas aos subgrupos Uso Secundário e Agregação.

Não há transparência no processo de exclusão de dados pessoais, pois, mesmo com local apropriado para o pedido de exclusão de todos os dados pessoais da ferramenta pelo usuário, a exclusão de dados pessoais não é realizada de forma efetiva nos bancos de dados. A visualização dos dados pessoais da conta excluída é bloqueada para outros usuários, porém parte destes dados continua armazenada nos bancos de dados da empresa. Segundo o Instagram (2015, p. 1), apesar da conta do usuário ser excluída “[...] dados podem persistir e aparecer dentro do serviço”, permitindo que possam existir ocorrências de atividades do subgrupo Insegurança.

Apesar do termo de uso delimitar quais são as atividades permitidas aos usuários na rede social, há opacidade dos procedimentos sobre a arbitragem de supostas violações (como a modificação e a exclusão arbitrária de conteúdos gerados por seus usuários), diretamente ligado a questões inerentes aos subgrupos Insegurança, Exclusão e Vigilância.

É possível realizar a coleta de dados através da API, permitindo que possam existir ocorrências de ações de coleta automatizada, por agentes externos ao sítio, sem o consentimento prévio do usuário (pois são permissivos novos acordos com empresas parceiras, ao longo do tempo), podendo desencadear atividades vinculadas ao subgrupo Quebra de Sigilo.

Não é transparente ao usuário se a rede social compartilha metadados dos conteúdos multimídia de usuários com empresas parceiras. Caso estes metadados sejam compartilhados, é possível agregá-los com dados pessoais, garantindo a estas empresas parceiras informações

adicionais, como: a localização geoespacial e temporal do conteúdo, atividade que é parte do subgrupo Agregação.

### Twitter

No termo de uso da rede social Twitter, dados pessoais como o nome, o nome de usuário, a foto do perfil e a localização são, por padrão, parte integrante de um catálogo de dados públicos. O serviço não possui um local apropriado para que usuários observem quais, como, onde e por quem estes conjuntos de dados são coletados e com quais outras fontes de dados são agregados, portanto parte integrante de possíveis ocorrências de atividades dos subgrupos Agregação e Apropriação.

O usuário não tem conhecimento pleno sobre quais são as empresas parceiras que acessam suas informações, sendo possível vincular este tipo de atividade ao subgrupo Aumento do Acesso.

Não é possível delimitar de forma transparente quais informações são geradas por empresas parceiras a partir de dados de usuários ou a forma que manipulam e processam estes dados pessoais com outras fontes de dados, o que pode acarretar atividades ligadas aos subgrupos Uso Secundário e Agregação.

As configurações do Twitter estão ajustadas, por padrão, em um modelo que todo o conteúdo publicado por seus usuários torna-se público a qualquer interessado, sem a necessidade de identificação para o acesso e coleta dos dados, sendo possível a realização de atividades de terceiros que prejudiquem a privacidade através da opacidade da divulgação de informações, que são (aparentemente) íntimas, mas que podem ser expostas ou até retransmitidas de forma incompleta. Estas atividades estão vinculadas aos subgrupos Divulgação, Exposição e Uso Secundário.

### 5.2. Resultados da aplicação da Taxonomia da Privacidade

O Gráfico I apresenta a síntese da quantidade de potenciais atividades prejudiciais que foram vinculadas a trechos nos termos de uso, agrupados em cada rede social. A sua estrutura está dividida em três colunas, cada uma delas representando a quantidade de subgrupos identificados para a rede social. Foi considerado que, caso o subgrupo tenha sido vinculado a um trecho ao menos uma vez, é somado o valor 1, caso contrário, 0 – não acumulativo por reincidência.

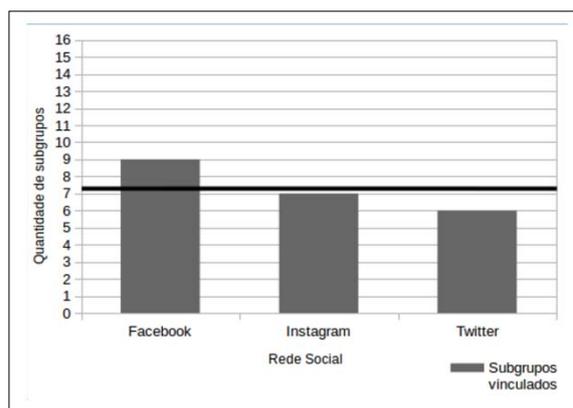


Gráfico I. Quantidade de subgrupos vinculados a trechos de termos de uso, agrupados por rede social

A Taxonomia da Privacidade possui um total de 16 subgrupos, sendo que as redes sociais (colunas) apresentaram totais de vínculos entre 9 e 6, com uma média aproximada de 7,34 (linha horizontal, em evidência) vinculações com subgrupos por rede social.

Os termos de uso apresentaram ao menos uma incidência nos subgrupos Agregação, Uso Secundário, Aumento do Acesso, Vigilância, Insegurança, Exclusão, Quebra de Sigilo, Apropriação, Identificação, Divulgação e Exposição, totalizando 22 (vinte e duas) vinculações com subgrupos, quando somado as três colunas.

O Quadro II sistematiza subgrupos por faixas de incidência, agrupados a partir da quantidade de redes sociais que apresentaram, em seus termos de uso, ao menos um trecho vinculado ao subgrupo.

O quadro está dividido em três colunas: a primeira coluna representa o nome do subgrupo, seguido coluna com a quantidade de redes sociais com termos de uso vinculados com o subgrupo – variando de 0 (quando o subgrupo não fora vinculado a nenhum trecho em nenhum termo de uso) até 3 (quando todos os termos de uso apresentaram, ao menos uma vez, trechos vinculados ao subgrupo).

A terceira coluna, denominada Faixa, estabelece uma divisão por faixas de ocorrências de vinculações destes subgrupos aos termos de uso, onde é delimitado que: a Faixa I concentra os subgrupos com mais incidências; a Faixa II concentra os subgrupos com incidência intermediária; a Faixa III concentra os subgrupos com baixa incidência, e; a Faixa IV concentra os subgrupos com nenhuma incidência para a amostra.

| <i>Subgrupo</i>          | <i>Quantidade de Redes Sociais com vínculo</i> | <i>Faixa</i> |
|--------------------------|--|--------------|
| Agregação                | 3  | Faixa I      |
| Uso Secundário           | 3  |              |
| Aumento do Acesso        | 3  |              |
| Vigilância               | 2  | Faixa II     |
| Insegurança              | 2  |              |
| Exclusão                 | 2  |              |
| Quebra de Sigilo         | 2  |              |
| Apropriação              | 2  |              |
| Identificação            | 1  | Faixa III    |
| Divulgação               | 1  |              |
| Exposição                | 1  |              |
| Interrogatório           | 0  | Faixa IV     |
| Chantagem                | 0  |              |
| Distorção                | 0  |              |
| Intromissão              | 0  |              |
| Interferência Decisional | 0  |              |
| <b>Total</b>             | <b>22</b>                                      | <b>-</b>     |

Quadro II. *Faixas de subgrupos, agrupados a partir da quantidade de redes sociais que apresentaram trechos vinculados.*

O Quadro III (Anexo 2) apresenta a distribuição dos subgrupos de atividades potenciais prejudiciais à privacidade, divididos pelos grupos propostos na Taxonomia da Privacidade, porém com os subgrupos agrupados por faixas de ocorrências em termos de uso das redes sociais. As linhas representam os grupos da Taxonomia da Privacidade e as colunas as faixas de ocorrências do Quadro II (terceira coluna).

A partir do Quadro III, é possível estabelecer que, nos termos de uso das redes sociais, são: **pontos de maior incidência**, elementos que sustentem a possibilidade de atividades prejudiciais à privacidade dos grupos Processamento de Informação e Disseminação de Informação, principalmente sobre questões ligadas aos subgrupos Agregação, Uso Secundário e Aumento do Acesso; **pontos de média incidência**, elementos que sustentem a possibilidade de atividades prejudiciais à privacidade de todos os grupos (Coleta de Informação, Processamento de Informação, Disseminação de Informação e Invasão), principalmente sobre questões ligadas aos subgrupos Vigilância, Exclusão, Insegurança, Quebra de Sigilo e Apropriação, e; **pontos de baixa incidência**, elementos que sustentem a possibilidade de atividades prejudiciais à privacidade dos grupos Processamento de Informação e Disseminação de Informação, principalmente sobre questões ligadas aos subgrupos Identificação, Divulgação e Exposição.

Os subgrupos Interrogatório (do grupo Coleta de Informação); Chantagem e Distorção (do grupo Disseminação da Informação); Intromissão e Interferência Decisional (do grupo Invasão), compõem os subgrupos **sem incidência** e, portanto, não foram vinculados a nenhum termo de uso.

## 6. Considerações Finais

A média aproximada de 7 subgrupos identificados em termos de uso (Gráfico 1), de um total de 16 subgrupos, abre espaço para reflexão sobre a necessidade de um maior aprofundamento nestes termos de uso de serviços privados de redes sociais, principalmente em estudos – guiados por estes eixos temáticos de atividades – para auxiliar a explicitação de dados antes ocultos, como quais são as atividades mais recorrentes ou passíveis de realização, e; quais possíveis lacunas nos termos de uso destas redes, que podem garantir a legalidade da realização de potenciais atividades prejudiciais à privacidade de usuários.

Concluiu-se que a aplicação de uma taxonomia apropriada pode auxiliar no estudo de termos de uso, possibilitando a explicitação de atividades previstas nestes termos, o que permite inclusive, propostas de futuras aplicações desta metodologia em outros contextos.

A aplicação da Taxonomia da Privacidade também permitiu a reorganização destes subgrupos vinculados a trechos dos termos de uso, a partir de uma classificação por faixas de ocorrências (Quadro II) – onde esta classificação apresenta as possíveis atividades mais recorrentes nos termos de uso (Faixa I) até possíveis atividades que não foram verificadas (Faixa IV). Estas faixas podem auxiliar, por exemplo, na elaboração de estratégias de atualização dos termos de uso destes serviços, através da apresentação de focos de atividades prejudiciais mais aderentes às redes sociais.

Como trabalhos futuros, destacam-se as possibilidades de aplicação deste estudo utilizando termos de uso de outras redes sociais; em outros contextos, como *blogs*, mecanismos de busca, entre outros – e a possibilidade de aplicação destes resultados na definição de estratégias de acompanhamento de atualização destes termos de uso.

## Notas

- (1) Tradução da língua inglesa do termo *Online Social Network*.
- (2) Segundo WORLDOMETERS (2015), em 2011 a população mundial superou o número de 7 bi-

lhões, sendo este o valor total adotado neste trabalho.

- (3) Facebook é marca registrada de Facebook Inc.
- (4) A marca Instagram é de propriedade privada de Facebook Inc.
- (5) A marca Twitter é de propriedade privada de Twitter Inc.
- (6) Disponíveis em 20-03-2015.
- (7) As nomenclaturas dos grupos e subgrupos foram traduzidas da língua inglesa para a língua portuguesa pelos autores.
- (8) Quando disponível.
- (9) Tradução da palavra em língua inglesa *decay*.

Solove, D. J. (2006). A Taxonomy of Privacy. *University Of Pennsylvania Law Review*, 154(3), 477. <http://doi.org/10.2307/40041279>

Twitter. (2015a). *Twitter Terms of Service*. Recuperado 10-02-2015, de [https://twitter.com/tos/previous/version\\_8?lang=en](https://twitter.com/tos/previous/version_8?lang=en)

Twitter. (2015b). *Twitter Privacy Policy*. Recuperado 10-02-2015, de [https://twitter.com/privacy/previous/version\\_9?lang=en](https://twitter.com/privacy/previous/version_9?lang=en)

WORLDMETERS. (2015). *Real Time Statistics*. Recuperado 10-01-2015, de <http://www.worldometers.info/>

Young, A. L.; Quan-Haase, A. (2009). Information revelation and internet privacy concerns on social network sites. *Proceedings Of the Fourth International Conference on Communities and Technologies – C&T '09*. <http://doi.org/10.1145/1556460.1556499>

## Referencias

Adamic, L. A.; Adar, E. (2003). Friends and neighbors on the Web. *Social Networks*, 25(3), 211–230. [http://doi.org/10.1016/s0378-8733\(03\)00009-1](http://doi.org/10.1016/s0378-8733(03)00009-1).

BBC. (2014, Dezembro 2014). *Instagram now bigger than Twitter*. Recuperado 10-01-2015, de <http://www.bbc.co.uk/newsbeat/article/30410973/instagram-now-bigger-than-twitter>

Castells, M. (2002). O poder da identidade. São Paulo: Paz e terra.

Castells, M. (2003). A Sociedade em rede. São Paulo: Paz e Terra.

CERN. (2015). *The birth of the web*. Recuperado 12-02-2015, de <http://home.web.cern.ch/topics/birth-web>

Chen, D.; Zhao, H. (2012). Data Security and Privacy Protection Issues in Cloud Computing. *2012 International Conference On Computer Science and Electronics Engineering*. <http://doi.org/10.1109/iccsee.2012.193>

Facebook. (2015a). *Statistics*. Recuperado 10-02-2015, de <http://newsroom.fb.com/company-info/>

Facebook. (2015b). *Terms of Service*. Recuperado 10-02-2015, de <https://www.facebook.com/legal/terms>

Facebook. (2015c). *Data Policy*. Recuperado 10-02-2015, de <https://www.facebook.com/about/privacy/>

Fogel, J.; Nehmad, E. (2009). Internet social network communities: Risk taking, trust, and privacy concerns. *Computers In Human Behavior*, 25(1), 153–160. <http://doi.org/10.1016/j.chb.2008.08.006>

Instagram. (2015a). *Terms of Use*. Recuperado 10-02-2015, de <https://help.instagram.com/478745558852511>

Instagram. (2015b). *API Terms of Use*. Recuperado 10-02-2015, de <https://instagram.com/about/legal/terms/api/>

Instagram. (2015c). *Privacy Policy*. Recuperado 10-02-2015, de <https://help.instagram.com/155833707900388>

Krasnova, H.; Günther, O.; Spiekermann, S.; Koroleva, K. (2009). Privacy concerns and identity in online social networks. *IDIS Identity In the Information Society*, 2(1), 39–63. <http://doi.org/10.1007/s12394-009-0019-1>

Mislove, A.; Marcon, M.; Gummadi, K. P.; Druschel, P.; Bhattacharjee, B. (2007). Measurement and analysis of online social networks. *Proceedings Of the 7th ACM SIGCOMM Conference on Internet Measurement - IMC '07*. <http://doi.org/10.1145/1298306.1298311>

Santana, R. C. G. (2013). Ciclo de Vida dos Dados e o papel da Ciência da Informação. *XIV Encontro Nacional de Pesquisa em Ciência da Informação '13*. ISBN: 978-85-65044-06-6

## Anexo 1

| <i>Grupo</i>                              | <i>Subgrupo</i>          | <i>Conceitos de Atividades</i>   |
|---|--------------------------|--|
| Coleta de Informação<br>(Grupo I)         | Vigilância               | Atividades encadeadas com o propósito de vigiar um indivíduo no seu espaço privado ou em espaço público.   |
|   | Interrogatório           | Atividades de processos de coleta de dados, baseados em interrogatórios e entrevistas.   |
| Processamento de Informação<br>(Grupo II) | Agregação                | Atividades vinculadas ao processo de combinação de dados de múltiplas fontes sobre indivíduos, com o propósito de revelar fatos ocultos, quando eram analisados de forma separada.   |
|   | Identificação            | Atividades a partir do processo de vinculação de dados que permitam a (re)identificação de usuário (e de seus dados pessoais) com suas respectivas pessoas.  |
|   | Insegurança              | Atividades que não perpassam segurança sobre questões de acesso a dados pessoais aos envolvidos.   |
|   | Uso Secundário           | Atividades que envolvem o uso de dados coletados para um determinado propósito e utilizados <i>a posteriori</i> para outras finalidades.   |
|   | Exclusão                 | Atividades que apresentam opacidade ao indivíduo no processo de armazenamento de dados pessoais, no compartilhamento destes dados a terceiros e na ausência ou na inabilidade de participação nas decisões sobre questões envolvendo a coleta, o armazenamento, o uso e o compartilhamento destes dados. |
| Disseminação de Informação<br>(Grupo III) | Quebra de Sigilo         | Atividades em que ocorrem a quebra de confiança entre as partes em manter a confidencialidade das informações sobre indivíduos.  |
|   | Divulgação               | Atividades de divulgação e de disseminação de informações sobre um indivíduo, que acarretam mudanças na maneira que outros indivíduos julgam seu caráter.  |
|   | Exposição                | Atividades vinculadas a exposição para terceiros de atributos emocionais ou físicos de intimidade do indivíduo, tais como a nudez, funções corporais e informações de cunho privado.   |
|   | Aumento do Acesso        | Atividades que visam amplificar o acesso à dados pessoais além do previsto ou do combinado entre as partes.  |
|   | Chantagem                | Atividades de controle, de dominação, de intimidação ou de ameaças a pessoas ou grupos, por terceiros.   |
|   | Apropriação              | Atividades que utilizam dados pessoais de um determinado sujeito em benefício de um terceiro ou para cancelar um serviço ou um produto, sem o pleno consentimento do sujeito.  |
|   | Distorção                | Atividades de disseminação de informações falsas ou interpretadas de maneira dúbia sobre um indivíduo.   |
| Invasão<br>(Grupo IV)                     | Intromissão              | Atividades com o propósito de realizar incursões em assuntos ou em informações de caráter privado.   |
|   | Interferência Decisional | Atividades de envolvimento do Estado em assuntos de caráter privado, alterando decisões em nome do indivíduo.  |

Quadro I. Grupos, subgrupos e conceitos de atividades prejudiciais a privacidade, adaptado de Solove (2006)

---

**Anexo 2**


---

| <i>Grupo/Faixa</i>                        | <i>Faixa I</i>              | <i>Faixa II</i>         | <i>Faixa III</i>        | <i>Faixa IV</i>                         |
|---|-----------------------------|-------------------------|-------------------------|---|
| Coleta de Informação<br>(Grupo I)         |                             | Vigilância              |                         | Interrogatório                          |
| Processamento de Informação<br>(Grupo II) | Agregação<br>Uso Secundário | Exclusão<br>Insegurança | Identificação           |   |
| Disseminação de Informação<br>(Grupo III) | Aumento do Acesso           | Quebra de Sigilo        | Divulgação<br>Exposição | Chantagem<br>Distorção                  |
| Invasão<br>(Grupo IV)                     |                             | Apropriação             |                         | Intromissão<br>Interferência Decisional |

Quadro III. *Subgrupos de atividades prejudiciais a privacidade, divididos pela hierarquia da Taxonomia da Privacidade, agrupados por faixas de ocorrências em termos de uso de redes sociais.*

**Hito 1: Modelos de organización del conocimiento. Modelaje de sistemas y ciclo de vida de la información**

**Hito 2: Entornos y sistemas para la organización del conocimiento. Diseño, normalización e interoperabilidad de sistemas de información abiertos.**

**Hito 3: Instrumentos para el acceso y transparencia de la información: gestión de documentos, plataformas digitales y archivos**

**Hito 4: Usuarios y usos de la información. Patrones vinculados a la transparencia y el gobierno abierto**

**Hito 5: Gobernanza y Auditoría de información en la organización del conocimiento.**

**Hito 6: Representación del conocimiento: procesos y herramientas en sistemas abiertos y multilingües**

**Hito 7: Epistemología en la organización del conocimiento**

**Hito 8: Igualdad, género y diversidad en la organización del conocimiento**

---

# Aplicabilidad del mapa conceptual en la representación del conocimiento registrado en las historias clínicas del paciente

*Applicability of the concept map in the representation of recorded knowledge and medical records.*

---

Virginia Bentes Pinto (1), Odete Máyra Mesquita Sales (2), Henry Pôncio Cruz de Oliveira (3)

(1) Doutora em Ciência da Informação. Universidade Federal do Ceará, vbentes@ufc.br.

(2) Mestra em Ciência da Informação. Universidade Federal do Ceará, mayra.mesquita@gmail.com.

(3) Doutor em Ciência da Informação. Universidade Federal da Paraíba, henry.poncio@gmail.com.

## Resumen

Los mapas conceptuales son como objetos signícos y de la representación del conocimiento. En la Ciencia de la Información y en la Biblioteconomía se utilizan en el tratamiento, representación, organización y recuperación de la información en documentos tradicionales. Pero, en una documentación específica como las historias clínicas del paciente, esta metodología no se aplica. Así presentamos una investigación que problema es: ¿Cómo aplicar la metodología del mapa conceptual para la representación del conocimiento registrado en las evoluciones de las historias clínicas del paciente? El objetivo: construir mapas conceptuales das evoluciones registradas en las historias clínicas del paciente. La investigación tiene características de estudios exploratorios. Las análisis se basaron en el método funcionalista y análisis de contenido. El corpus es compuesto por 10 historias clínicas del paciente de la especialidad de nefrología. Para la construcción del mapa utilizamos el aplicativo *CmapTools*. Los resultados muestran que la macroestructura si constituyen en atributos con las siguientes categorías: evolución clínica, evolución dentro del plan de SOAP, entre otros. Los atributos de la estructura lógica varían de acuerdo con los tipos de enfermedades del paciente y las acciones de atención realizadas por el personal médico y de enfermera.

**Palabras clave:** Representación del conocimiento. Notas de evolución en las historias clínicas del paciente. Evolución del plan de SOAP. Terminología en nota de evolución de los registros de pacientes. Mapa conceptual. SOAP.

## 1. Proceso de publicación

La génesis de los mapas conceptuales está en la Aprendizaje Significativa del psicólogo David Paul Ausbel que ha sido colocada en pirática en los años sesenta e innovadas en 1980 con la colaboración de Joseph Donald Novak e Helen Hanesian. Esa metodología sirve para organizar y representar el conocimiento, principalmente en el contexto de la educación. Mirando esa meto-

## Abstract

The concept maps can be considered as signic objects and representations of knowledge. In Information Science and Librarianship are used in treatment, representation, organization and retrieval of information in traditional documents. However, in a specific documentation such as medical records, this methodology has not been applied. Thus, we present research whose problem is: How to apply the methodology of concept maps in knowledge representation and evolution of medical records? The aim: build a conceptual map of evolutions recorded in the medical records. This research has characteristics of an exploratory study. The analyzes were based on the functionalist method and content analysis. The corpus consists of 10 clinical histories of nephrology specialty. To build the maps we use the *Cmap*. The results show that the macrostructure of evolutions has attributes set to the categories: clinical evolution, evolution within the SOAP plan, among others. Regarding logic structure, attributes varied according to the types of diseases patients and the actions carried out by the doctor and the nurse.

**Keywords:** Representation of the knowledge. Progress notes in the clinical histories. Terminology in progress notes of medical records. Conceptual Map. SOAP.

dología si percebe que ella también puede ser aplicada en las áreas de Ciencia de la Información y de la Biblioteconomía para el tratamiento de la información y del conocimiento registrados, pues si fundamenta en construir estructuras de representación, facto que ya es percibido. Así, si están la introduciendo, en el tratamiento, representación, organización y recuperación de la información de documentos tradicionales. Sin embargo, en una documentación

específica como las historias clínicas del paciente (HCP), esta metodología no todavía ha se aplicada.

En esta perspectiva es que nosotros la colocamos en práctica en una investigación cuya problemática concierne a la pregunta siguiente: ¿De qué manera se puede aplicar la metodología del mapa conceptual para la organización y la representación del conocimiento registrado en las evoluciones de las historias clínicas del paciente, levando en consideración los datos Subjetivo, Objetivo, Evaluación, Plan (SOAP) para mejorar la recuperación de la información en los Servicios de Archivo Médico y de Estadística (SAME)? El objetivo básico de la investigación consiste en: construir un mapa conceptual de las categorías de la representación del conocimiento registrado en las evoluciones descritas en las historias clínicas del paciente visando la recuperación de la información en el SAME.

Los objetivos específicos son:

- a) Estudiar las historias clínicas del paciente visando la identificación de la estructura lógica y física para la especificación de los metadatos;
- b) Mapear los términos utilizados en la redacción de las evoluciones de las historias clínicas del paciente
- c) Construir el mapa conceptual de las evoluciones de las historias clínicas a partir de los metadatos mapeados.

Estos son las cuestiones cruciales que serán tratadas en este artículo teniendo-se la estructura siguiente: la historia clínica del paciente, materiales y métodos de investigación, aplicabilidad del mapa conceptual en las historias clínicas del paciente y conclusiones

## 2. La historia clínica del paciente

Nunca en la historia de la humanidad se tiene hablado tanto en la historia clínica del paciente (HCP). Documento que desde la época de Hipócrates y Imhotep, considerados los padres de la medicina, es marcado como siendo de importancia impar para el contexto de la salud. Porque en ese documento son registradas todas las informaciones y los conocimientos referentes a el estado de salud de una persona y los cuidados que son hechos en prol de suja cura. Talvez este fato sea el reflejo da la propia evolución del ciudadano que cada vez más tiene consciencia de sus derechos. Entonces, en el contexto de la salud, la historia clínica del paciente desarrolla un papel fundamental para que se concretice el flujo de información y de la comunicación entre

los actores involucrados en el proceso de cuidado de los pacientes.

Conforme las reflexiones de Bentes Pinto (2006, 2007) la HCP es un documento transcrito y escrito pela equipe multiprofesional de la salud y es de per si el documento de información, mediación y de comunicación en el ámbito de las organizaciones de salud y por eso su función fundaméntale es aportar las informaciones e los conocimientos de modo a contribuir para el proceso de comunicación con menor ruido y interferencia y así asegurar el mejor cuidado en todos los niveles de atención en la salud: primaria, secundaria y terciaria. La Declaración de Alma Ata (1978) define la atención primaria como siendo

la asistencia sanitaria esencial basada en métodos y tecnologías prácticas, científicamente fundados y socialmente aceptables, puesta al alcance de todos los individuos y familias de la comunidad mediante su plena participación y a un coste que la comunidad y el país puedan soportar en todas y cada una de las etapas de su desarrollo, con un espíritu de auto responsabilidad y autodeterminación. (Declaración de Ama Ata, 1978).

Por lo tanto es un documento cuja esencia es compartir conocimientos e informaciones con todos los sujetos envueltos en el proceso de cura, incluso profesionales que trabajan en otras organizaciones de salud, pero que necesiten consultar la HCP, razón por la cual nosotros le entendemos que él es de carácter interdisciplinaria. Además, el documento da la Agencia National d'Accreditation y de evaluación de santé (1996) ellas también tienen la importancia vital para ayudar la memoria de equipo multiprofesional de salud puesto que en este documento son registradas todas las "informations qu'ils estiment nécessaires pour la suite de la prise en charge, sans se soucier de transcrire fidèlement les actions qu'ils accomplissent". En el Artículo 1º de la Resolución del nº 1.638/2002, del Consejo Federal de Medicina del Brasil está la definición de la historia clínica del paciente como

un documento único constituido por un conjunto de informaciones, señales e imágenes registradas, generadas a partir de hechos, acontecimientos y situaciones sobre la salud del paciente y la asistencia a él prestada, és de carácter legal, sigiloso y científico, utilizado para posibilitar la comunicación entre miembros del equipo multiprofesional y la continuidad de la asistencia prestada al individuo (BRASIL. CFM, 2002).

Desde Hipócrates do Cos e de Imhotep la historia clínica del paciente está en el cimiento de la

asistencia que es prestada a él. Según Arago (19-?, p. 2) los registros, de Hipócrates traerían en su estructura tres partes a saber: descriptio subjectivi, descriptio morbidis y êxitos (sucessos). Para ilustrar este entendimiento se presenta un pasaje en la citación:

Filisco, que vivía cerca de la muralla, se metió en cama. Primer día, fiebre aguda, sudor, la noche fue penosa. Segundo día, exacerbación general; por la tarde, con un lavado, evacuaciones favorables; la noche fue tranquila. Tercer día, por la mañana, hasta el mediodía, pareció quedar apirético; hacia la tarde, fiebre aguda con sudor, sed, la lengua se desecó, orinó negro; la noche fue penosa, no durmió, toda clase de alucinaciones. Cuarto día: agravamiento general, orinas negras; la noche, más soportable; orinas de color más favorable. Quinto día: hacia el mediodía fluyó de su nariz un poco de sangre no mezclada, la orina de vario color, con nubéculas redondeadas, semejantes al esperma y dispersas, no dejaba depósito. Tras la aplicación de un supositorio, deposición escasa y ventosidades. Noche penosa, sueño entrecortado, habló mucho, deliró, extremidades frías en toda su extensión y no era posible recalentarlas, orinó negro; de día durmió un poco, perdió el habla, sudor frío, extremidades lívidas. Murió hacia la mitad del sexto día. En este enfermo la respiración fue hasta el final rara y amplia, como en uno a quien se reanima; el bazo se hinchó y formó un tumor esferoidal; sudores fríos hasta el fin; los accesos, en los días pares.

En este pasaje de texto se puede constatar que desde el siglo V a.C la evolución detallada del estado de salud del paciente ya era muy importante en la historia clínica de Hipócrates. Entonces desde esta época ya había una preocupación por bien registrar los conocimientos y las informaciones sobre la condición de salud de un paciente y otros aspectos que se juzgase importante para su tratamiento. Tal hecho hasta este siglo aún es repetido.

Corroborando con la proposición hipocrática, el médico y Prof. Lawrence L. Weed ha creado, en 1960, la metodología de evolución de datos e informaciones llamada SOAP (Subjetivos, Objetivos, Apreciación y Planos de acciones) para las historias clínicas del paciente. Tal instrumento se fundamenta en las notas de evoluciones hechas por la equipo multiprofesional de salud que escarba en las historias clínicas, por ejemplo, médicos y enfermeros y técnicos de enfermería. Estas notas de evoluciones apuntan el registro de informaciones de salud debe ser hecho en función de cada problema observable

teniendo en cuenta los cuatro tipos de datos e informaciones observables:

- a) Subjetivos (S): conciernen en la percepción relativa a los datos subjetivos recogidos del diálogo entre la equipo multiprofesional y los pacientes.
- b) Objetivos (O): refiere-se las anotaciones de los datos objetivos relativos el examen físico, complementarios e otros jugados importantes en este aspecto.
- c) Apreciación (A): son los discursos que se traducen en los problemas de salud con los cuales los pacientes se confrontan.
- d) Planes (P): corresponde las acciones de la planificación referente a las conductas que se tomarán concernientes a los problemas identificados

En realidad, la metodología de Lawrence Weed tiene características importantes debido su "objetividad, organización, mayor facilidad de acceso a las informaciones para tomada de decisiones y por la descripción sistemática de las evidencias e de las percepciones que apoyan las conclusiones y los planos diagnósticos y terapéuticos durante el acompañamiento del paciente". El entendimiento del *Evidence-Based Medicine Working Group* propone que el SOAP puede contribuir para la pirática de la Medicina Basada en la Evidencia, pues a partir del SOAP es posible si tener un mapa de conceptos relativos a la condición de salud de las paciente e igualmente de las acciones hechas para las soluciones

### 3. Materiales y métodos de investigación

Esta investigación esta sustentada en el método funcionalista pues, creemos que las historias clínicas del paciente tienen funciones sine quann para el tratamiento de los pacientes en las organizaciones de salud. Este método ha ganado su grand desarrollo con los estudios del contexto de la Antropología con las investigaciones de Malinowski que ha traída otras reflexiones sobre el funcionalismo que antes estaba asociado a las investigaciones matemáticas y del campo de la biología. Por lo tanto, en una perspectiva social Malinowski (p. 141) nos dice "[...] A Minha indiferença por certos tipos de Evolucionismo e uma questão de métodos". Ele continua sus reflexiones afirmando que, no sentido etnográfico o valor científico de un documento o una fuente, é "incuestionable siempre que podamos hacer una clara distinción entre (...) lo que son los resultados de la observación directa y las exposiciones e interpretaciones del

indígena". Pero son "[...] las deducciones del autor basadas en su sentido común y capacidad de penetración psicológica" que construí sentido sobre o que observa do objeto de estudio.

Entonces, mismo que nosotros no estamos investigando los indígenas, aun entendimos que en las HCP encontramos las varias observaciones directas a y las exposiciones e interpretaciones del paciente, y las deducciones de los profesionales de la salud, tanto del punto de vista de la objetividad expresada en los discursos de los pacientes, como también del punto de vista de la subjetividad capturada por esos profesionales sobre las condiciones de salud del paciente y que son registradas en las historias clínicas del paciente.

Además, las HCP tienen en funciones impares en las organizaciones de salud puesto que ellas son que hacen con que si efectivo el flujo de la información y de la comunicación entre los sujetos involucrados en las acciones de cuidados implementados visando a cura de una persona enferma.

Para la aplicabilidad del mapa conceptual con la metodología SOAP, nos basamos en el análisis del contenido que fue desarrollada por Bardin (2009, p. 121), observando las siguientes etapas: pré-análise; exploración del material; e, el tratamiento de los resultados (inferencia e interpretación). Diante desde punto de vista, en un plano inicial, en esta primera fase escogemos las historias clínicas a que serían estudiadas o analizadas visando identificar los asuntos incluidos en el SOAP. Entonces, el estudio empírico ha sido hecho sobre un corpus de 10 (diez) historias clínicas de pacientes que hacen tratamiento de enfermedades de nefrología. Todas las HCP están almacenadas en el Servicio de Archivos Médicos e Estadística (SAME) del Hospital Universitario Walter Cantídio (HUWC) de la Universidad Federal del Ceará (UFC). Hicimos la opción por esta especialidad porque ya teníamos una base de datos con tales archivos y además el HUWC es el pionero en trasplantes de riñón del estado de Ceará siendo considerado un hospital de referencia en este asunto y en las investigaciones que a el son concernientes.

Los estudios de Bardin (2009, p.123) nos dicen aún que "Nem todo o material de análise é susceptível de dar lugar a uma amostragem, e, nesse caso, mais vale abstermo-nos e reduzir o próprio universo (e, portanto, o alcance da análise) se este for demasiado importante". Por eso escogemos trabajar solamente con la estructura dos prontuarios referentes as evoluciones sobre las condiciones de salud de los paciente.

En lo que respecta "la exploración del material" nosotros nos debruizamos sobre los discursos registrados en las evoluciones de las HCP para que fuese posible identificar los metadatos de representación de las informaciones y de los conocimientos anotados en las HCP. Finalmente, para el tratamiento de los resultados y las inferencias e interpretaciones fui construido el mapa conceptual.

#### 4. Tratamiento de los datos y discusión de los resultados

Los resultados de la investigación que ha sido realizada en un corpus de 10 (Diez) historias clínicas del paciente, analizando-se solamente la estructura relativa a las evoluciones, han evidenciado que la redacción de informaciones y del conocimiento registrados en estos documentos es marcada por discursos muy particulares de los actores de este proceso- equipo multiprofesional de la salud y pacientes- que se misturan en la perspectiva de capturar los signos presentes en la subjetividad de la acción de cuidar % de cuidar. Entonces después de este análisis, estructuramos el SOAP segundo los problemas observados en las historias clínicas del paciente que han sido estudiadas, conforme si siegue: Para la "S", nosotros nos dedicamos en los datos subjetivos recogidos del diálogo entre el equipo multiprofesional y los pacientes. En la "O" anotamos los datos objetivos relativos el examen físico, complementarios e otros juzgados importantes en esto aspecto. Concerniente la "A" hicimos el mapa de los discursos que se traducen en los problemas de salud con los cuales los pacientes se confrontan. En la "P", se fueran identificadas las acciones de la planificación referente a las conductas que se tomarán concernientes a los problemas identificados. (Vean-se los mapas- 1 a 4 (en Anexo 1).

Como se puede mirar en el Mapa 1, la metodología de identificación de los datos subjetivos de las evoluciones clínicas basadas en el SOAP posibilita espejar la subjetividad de los discursos de los pacientes y también las impresiones percibidas por los profesionales de la salud sobre como el paciente está mejorando. Otra observación de los hallados es que todos estos discursos pueden ser convertidos en "pistas" para a la representación del conocimiento registrado en las HCP y por lo tanto, para la recuperación de información, tanto por parte de los investigadores así que para los que trabajan en el campo de la salud. Además, pueden contribuir para la construcción de redes sociales de información para las personas que tienen las mismas enfermedades y, por lo tanto necesitan discutir los

síntomas que sienten en relación a sus enfermedades.

En que conciernen los discursos objetivos, estos son muchos, pero en este artículo son presentados solamente algunos ejemplos expuestos en el Mapa 2.

Los datos objetivos evidencian la posibilidad de ser hacer cruzamiento de los síntomas con las enfermedades y si tenemos todos estos datos estructurados en una base de datos indizado con estos meta datos, ciertamente que traerá resultados fundamentales para las investigaciones en el contexto de la salud, pues poden ser hechas comparaciones de los resultados de las pruebas con casos semejantes para la medicina basada en evidencia. Luego, en este mapa, igualmente, es posible identificar y seleccionar términos y conceptos representativos del conocimiento registrados en las HCP y por lo tanto, para la recuperación de información en las organizaciones de salud, puesto que se configuran en meta datos del contexto de la salud.

Presentamos, igualmente, en el Mapa 3 algunos resultados del análisis relativas a las evaluaciones anotadas en las historias clínicas del paciente, hechas por la equipe multiprofesional y, principalmente por los médicos.

Nuestras análisis de estos hallados connotan que eles son meta datos de términos del discurso o de la lenguaje de especialidad y que sirven para la indización y la recuperación de la información y del conocimiento registrados en las HCP. Estos meta datos de especialidad pueden traer contribuciones fundamentales el desarrollo de políticas de salud, puesto que se si tiene un mapa de las enfermedades es posibles se planear y ejecutar acciones para la promoción social de salud, principalmente acciones voltadas para la atención primaria.

Otra observación es el uso de abreviaciones del tipo siglas, símbolos y abreviaturas en la redacción de las evaluaciones. Tal recurrencia puede contribuir para los ruidos e interferencias en el proceso de comunicación intra e entre el equipo de salud e entre ella y el paciente.

En el Mapa 4, presentamos os hallados de la investigación y el Plan ejecutado por los médicos con respecto las acciones desarrolladas o a serán colocadas en prácticas visando mejorar la condicione de salud del paciente.

Los resultados de la investigación analítica del registro del plano de las evoluciones de las historias clínicas del paciente, dejan claro que el discurso de los profesionales (en este caso, médicos y fisioterapeutas) es pautado por ac-

ciones muy particulares, para cada paciente, mismo que en algunas manifestaciones si repitan. Entretanto, como en los mapas 2 y 3, el registro evidencia la lenguaje de especialidad, llena de términos técnicos bien como siglas, símbolos y abreviaturas. Contodo, son términos fundamentales para la representación e la recuperación de la información en la salud. Con relación a las abreviaciones, naturalmente que ellas pueden traer ruidos e interferencias en el proceso de comunicación.

## 5. Conclusiones

Las conclusiones nos hacen volver a la cuestión de investigación y los objetivos propuestos en la introducción de este artículo-lo. Así, las categorías de la representación del conocimiento registrado en las evoluciones descritas en las historias clínicas del paciente visando la recuperación de la información en el SAME.

La aplicabilidad de la metodología del mapa conceptual puede muy bien ser aplicada a la organización y la representación del conocimiento registrado en las notas de evoluciones de las historias clínica del paciente, levando en consideración los datos Subjetivo, Objetivos, de Evaluación y del Plano de acciones.

Por consecuente, concluimos que las notas de evoluciones registradas en las HCP, pueden ser utilizadas como instrumentos para la representación y recuperación de información en el contexto de la salud. Todavía, es necesaria que sean estructuradas en formas de conceptos o términos y una herramienta para esta modelaje es el mapeo y la organización en el SOAP.

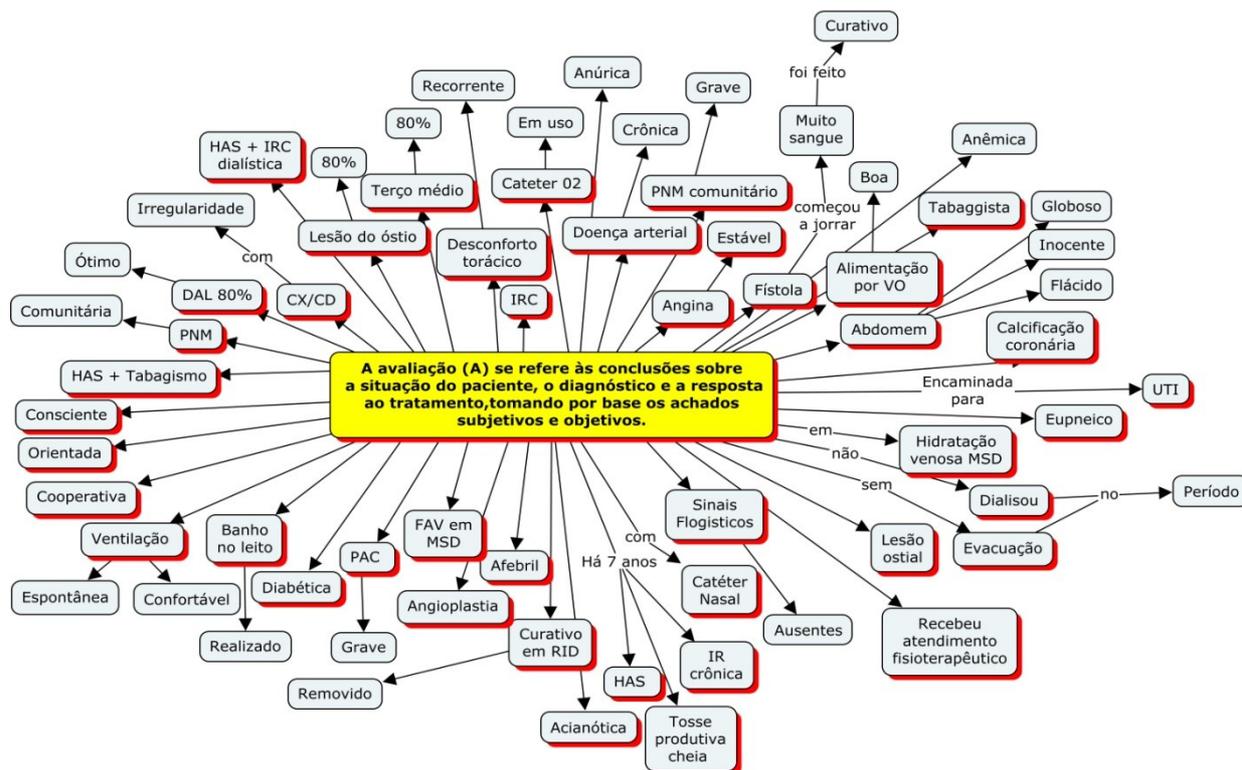
Finalmente, concluimos que las representaciones de los contenidos de las HCP por intermedio del SOAP pueden contribuir para la medicina basada en evidencia

## Referencias

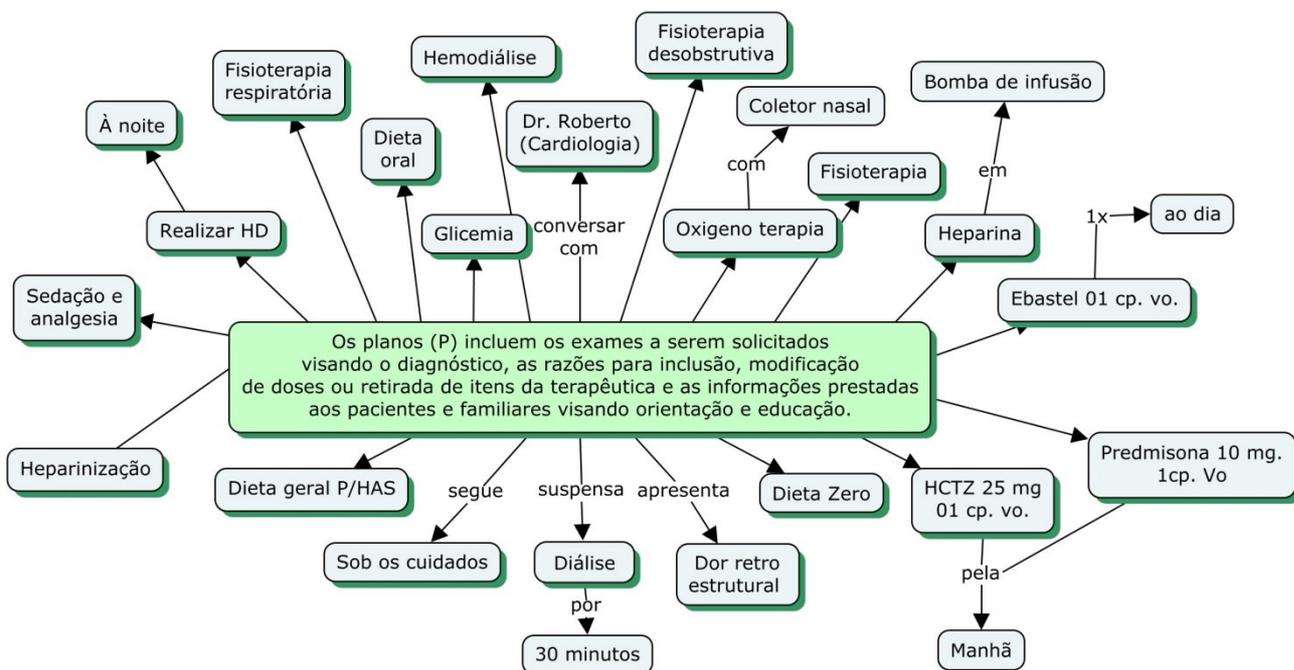
- ARAGO, I. Cuatro lecciones sobre la organización de archivos y bibliotecas médicas : lección primera : la historia clínica. Barcelona: Barcelona: Escuela de Bibliotecarias [19-?].
- AUSUBEL, D. P.; NOVAK, J. D.; HANESIAN, H. Psicología Educacional. Rio de Janeiro: Interamericana, 1980.
- BARDIN, L. Análise de Conteúdo. Lisboa, Portugal; Edições 70, LDA, 2009.
- BENTES PINTO, V. Prontuário eletrônico do paciente: documento técnico de informação e comunicação do domínio da saúde. **Enc. Bibli.** n.21, 1º sem. p. 34-48, 2006
- BRASIL. CONSELHO FEDERAL DE MEDICINA. Resolução nº. 1.638 de 10 de Julho de 2002. Diário Oficial, Brasília, 10 jul. 2002. Seção 1, p. 124-5.

WEED, L.L. Medical records that guide and teach. N Engl J  
Med 1968; 278:593-600  
\_\_\_\_\_. Control de calidad e historia clínica,  
Arch.Intern.Med/vol.127, Jan. 1971.





Mapa 3. Exemple de los datos de la Evaluación de notas de evoluciones anotadas en las HCP, aplicando-se el Plano SOAP.



Mapa 4. Exemple de los datos del Plano de acciones registradas en las notas de evoluciones en las HCP, aplicando-se el Plano SOAP.

---

# Armonización de ISO 25964 y SKOS en software de gestión de tesauros. Propuesta de análisis

*Harmonisation of ISO 25964 and SKOS in thesaurus management tools. Proposal analysis*

---

**María Luisa Alvite Díez (1), M. Mercedes Martínez González (2)**

(1) Área de Biblioteconomía y Documentación. Universidad de León, Campus de Vegazana s/n 24071 León, luisa.alvite@unileon.es.

(2) Departamento de Informática. Edificio T.I.T. Universidad de Valladolid, Campus Miguel Delibes, s/n 47011 Valladolid, mercedes@infor.uva.es

## Resumen

La norma ISO 25964 ha visto la luz en un momento en el que SKOS se ha consolidado como estándar para la representación de vocabularios controlados en la web semántica mediante RDF. Si bien la nueva ISO entiende el tesoro desde un punto de vista conceptual, cercano al enfoque de SKOS, y presenta, entre otras novedades, un modelo de datos, existen dificultades para conciliar la aplicación de ISO 25964 con la publicación de tesauros SKOS. Este trabajo revisa la relación de elementos y estructuras ISO 2788 / ISO 25964 y SKOS/SKOS – XL, así como las reglas de integridad definidas por los tres estándares para garantizar la consistencia del vocabulario. A partir de este examen se presenta una propuesta específica de criterios evaluativos ideados para valorar la adecuación de herramientas de gestión de tesauros a los requisitos de SKOS y de ISO 25964. Nos fijamos específicamente en los elementos, estructuras y requerimientos de integridad que estas aplicaciones deben soportar. Finalmente se realiza un análisis de los cambios relevantes introducidos por la nueva ISO que afectan al instrumento evaluativo propuesto previamente bajo el modelo de ISO 2788. Por el momento los software de gestión de tesauros que utilizan tecnologías semánticas y que fueron sometidos a evaluación en trabajos previos no han introducido las especificaciones de la nueva ISO. Estamos expectantes ante la posible finalización de la extensión de SKOS que aborde por completo el modelo de datos de la nueva ISO.

**Palabras clave:** Herramientas para la gestión de tesauros. ISO 2788. ISO 25964. Metodología de evaluación. SKOS. Tesauros. Web Semántica.

## 1. Introducción

La norma ISO 25964-1:2011 *Information and documentation -- Thesauri and interoperability with other vocabularies -- Part 1: Thesauri for information retrieval* vio la luz en agosto de 2011. En marzo de 2013 se publicó la segunda parte -- *Part 2: Interoperability with other voca-*

## Abstract

ISO 25964 has seen the light when SKOS is accepted as a standard for the representation of vocabularies with RDF on the semantic web. While the new ISO understands the thesaurus from a point of view close to the approach of SKOS, it is a conceptual approach, and the standard presents a data model, among other innovations, there are difficulties to settle the application of ISO 25964 with the publication of SKOS thesauri. This paper reviews the relationship of elements and structures ISO 2788 / ISO 25964 and SKOS/SKOS - XL, as well as integrity rules defined by the three standards to ensure consistency of vocabulary. From this examination, a specific proposal of evaluative criteria designed to assess the adequacy of thesaurus management tools to the requirements of SKOS and ISO 25964 is presented. We focus specifically on the elements, structures and requirements of integrity that must be supported by these thesaurus tools. Finally, an analysis of the relevant changes introduced by the new ISO affecting the evaluative instrument previously proposed under the model of ISO 2788 is carried out. By the time the thesaurus tools using semantic technologies and evaluated in previous works have not introduced the new ISO specifications. We are expectant to the possible completion of SKOS extension to cover the features of the new ISO model.

**Keywords:** Evaluation methodology. ISO 2788. ISO 25964. SKOS. Semantic Web. Thesaurus. Thesaurus tools.

*ularies.* El estándar, entre otras novedades, como subraya García Marco (2014), retoma el peso del concepto en la organización del conocimiento frente al enfoque léxico previo, incrementa el tratamiento del multilingüismo y persigue la interoperabilidad de los tesauros con

otras herramientas conceptuales en el marco de la web semántica y los datos abiertos.

ISO 25964-1 presenta recomendaciones sobre el software de gestión de tesauros, capítulo catorce, define en los capítulos siguientes un modelo de datos que posibilita el desarrollo de aplicaciones de software, los requisitos para la integración de tesauros con otras aplicaciones informáticas así como un esquema XML para la representación de estos vocabularios.

Las aportaciones novedosas de ISO 25964 han aparecido en un momento en el que SKOS (Simple Knowledge Organization Systems) se ha consolidado como estándar de la web semántica, proporcionando un modo de representar vocabularios controlados mediante RDF (Resource Description Framework).

Los tesauros nacieron en la década de los sesenta del pasado siglo como una herramienta de apoyo a la recuperación temática de la información, una herramienta que precisaba evolucionar para adaptarse a las tecnologías semánticas del siglo XXI. SKOS entiende los vocabularios desde un punto de vista conceptual y distingue claramente entre conceptos y etiquetas usadas para representar esos conceptos. En el estándar del W3C las relaciones entre conceptos se asemejan a las relaciones “tradicionales” recomendadas en las normas de tesauros.

Por su parte, los estándares sobre tesauros, se habían centrado fundamentalmente en la construcción de estos vocabularios controlados. ISO 25964 incorpora la definición de un modelo de datos, imprescindible en el entorno actual en el que las máquinas necesitan manipular e interpretar datos y realizar procesos de inferencias. Las aplicaciones precisan un modelo de datos explícito que distinga entre términos y conceptos. La nueva ISO clarifica esta cuestión, señalando con rotundidad que la aplicación principal de un tesoro es la recuperación de información, cuyo objetivo es buscar por conceptos. Por tanto, los conceptos están representados por términos y cada término incluido en un tesoro debería representar a un solo concepto (o unidad de pensamiento). Ahora bien, como subrayan Dextre y Zeng (2012), no debemos olvidar que términos y conceptos se presentan intrínsecamente vinculados.

Señala Pastor Sánchez (2013) cuatro razones fundamentales por las que la nueva ISO 25964 está más cercana a SKOS:

- Define un modelo de datos que facilita el desarrollo de herramientas informáticas y su aplicación en tareas de recuperación de información.

- El modelo establece una distinción clara entre el nivel conceptual y el nivel terminológico.
- La norma define cinco clases de estructuras superiores para representar tesauros: Thesaurus, ThesaurusArray, ThesaurusConcept, ThesaurusTerm y Note.
- Los elementos de anotación son prácticamente idénticos a los de SKOS.

Afortunadamente, la buena comunicación entre los equipos responsables de los desarrollos de SKOS e ISO 25964 han buscado la compatibilidad de ambos modelos y se han desarrollado extensiones de SKOS que resuelven funcionalidades previstas en ISO. Así, con el objetivo de buscar la interoperabilidad entre SKOS e ISO 25964, a lo largo de 2012 y 2013 se completó por parte de Grupo de Trabajo de ISO 25964 y A. Isaac, coeditor de SKOS, un documento en el que se establecen las correspondencias entre los dos estándares (ISO TC46/SC9/WG8 e Isaac, 2013).

Esta extensión ISO-THES se ha actualizado en marzo de 2015 con nuevas notas de aplicación (Isaac y De Smedt, 2015) sobre las correspondencias de ambos estándares y propuestas para extender SKOS de modo que sea capaz de dar cobijo a las posibilidades de modelado de ISO 25964. Entre estos aspectos, destacan los mecanismos para definir relaciones de equivalencia compuesta entre conceptos procedentes de distintos tesauros, los grupos de conceptos, arrays, o las Notas que ISO permite asociar a términos. La propuesta se orienta a reforzar SKOS como estándar en la web semántica.

Teniendo en cuenta este escenario, nuestro trabajo se propone revisar la relación de elementos y estructuras ISO 2788/ISO 25964 y SKOS y presentar una propuesta específica para valorar el grado de adaptación a SKOS y a ISO 25964 por parte de aplicaciones de gestión de tesauros que utilizan tecnologías semánticas. La metodología evaluativa del estudio se asienta en trabajos previos (Martínez González y Alvite Díez, 2014a, 2014b) en los que se formularon criterios evaluativos aplicados a gestores de tesauros atendiendo a los requisitos de ISO 2788 y SKOS. En el presente estudio se revisa y adapta esta propuesta metodológica a la luz de ISO 25964.

## 2. Objetivos y metodología

El objetivo principal de este trabajo se dirige a indagar en el rango de adecuación a ISO 25964 de herramientas de gestión de tesauros ideadas para el tratamiento de vocabularios y estándares.

res surgidos en el contexto de la Web semántica. Nos fijamos específicamente en dos parámetros, la gestión de la integridad que proporcionan las aplicaciones y la correspondencia de las estructuras conceptuales y terminológicas soportadas por estas herramientas y su adecuación a SKOS y a ISO 25964.

Pretendemos verificar la propuesta metodológica de partida, que tomaba en consideración la norma ISO 2788, y adaptarla a la norma ISO 25964.

Para abordar el análisis se ha tenido en cuenta el trabajo realizado sobre SKOS e ISO 25964 de Pastor Sánchez (2013). Las estructuras conceptuales y terminológicas se revisan en función del mapping establecido en Will (2012) y las correspondencias entre el modelo de datos ISO 25964 y el esquema de SKOS (ISO TC46/SC9/WG8 e Isaac, 2013) así como en el documento anexo a este último (Isaac y De Smedt (2015). Como se ha señalado, la metodología evaluativa del estudio se asienta en estudios anteriores (Martínez González y Alvite Díez, 2014a, 2014b) en los que se consideraban los requisitos de ISO 2788.

Los tesauros orientados a conceptos introducen dos niveles, uno conceptual y otro terminológico o léxico (Pastor, 2013) frente al único nivel, terminológico, de los tesauros orientados a términos, de modo que los primeros se acercan al modelo de datos de RDF y adoptado por SKOS, donde la estructura fundamental de un tesoro se define atendiendo a conceptos. Las diferencias relevantes para el tipo de evaluación que nos ocupa afectan a la forma de modelar determinados elementos de un tesoro y a las restricciones de integridad. Hay que tener en cuenta que algunos de estos requisitos varían su formulación según se trate de tesauros léxicos o conceptuales. Pero también es cierto que tanto las relaciones semánticas (TG, TE, TR) como la estructura básica de un tesoro sigue siendo la misma, de modo que es posible plantear la adaptación de un tesoro terminológico en un tesoro de conceptos.

### 2.1. Acotaciones metodológicas

La metodología ha consistido en la revisión de la propuesta evaluativa realizada en los estudios previos tomando ahora como referencia la nueva norma de tesauros, ISO 25964. Así pues, los pasos llevados a cabo siguen el orden en que se presentaba el marco evaluativo en Martínez González y Alvite Díez (2014).

Se podrían haber adoptado alternativas que partiesen del examen de los software de gestión

de tesauros a los que se pretende poder aplicar este conjunto de criterios evaluativos. Sin embargo, la juventud de la norma ISO 25964 y el hecho de disponer de un marco sistemático de valoración previo, que nos proporcionaba un guion de ítems objeto de revisión, nos llevó a considerar que el trabajo con la propia norma era la única garantía para obtener un resultado fiable, es decir, una revisión que realmente tenga en cuenta los cambios introducidos en la norma, se hayan estos trasladado o no a los software de tesauros en el momento actual.

Proponemos centrar el análisis en dos parámetros fundamentales: estructuras conceptuales y terminológicas soportadas y gestión de la integridad.

Los pasos seguidos en este trabajo han sido los siguientes:

- 1) Revisión de la norma ISO 25964, así como de las publicaciones en las que se recogen las principales novedades de la misma con respecto a la ISO 2788.
- 2) Identificación de la lista de estructuras y elementos de un tesoro que se introducen con la nueva ISO o que adquieren una relevancia mayor en esta nueva norma:
  - Conceptos.
  - Términos.
  - Arrays.
  - Tipos de relaciones TG: TGG, Término genérico genérico; TGE, Término genérico enumerativo; TGP, Término genérico partitivo.
  - Grupos de conceptos, para modelar microtesauros o dominios.
  - Equivalencias compuestas, representadas por relaciones USE+.
- 3) Comparación de esta lista con el conjunto de elementos que conformaba la propuesta anterior basada en ISO 2788 y SKOS. Como resultado obtuvimos un listado de elementos actualizado, con las modificaciones pertinentes.
- 4) Revisión de la relación de elementos ISO-SKOS confeccionada con el listado de las correspondencias definidas en ISO TC46/SC9/WG8 e Isaac (2013) y las recientes notas a este documento publicadas en marzo de 2015 (Isaac y De Smedt, 2015). Estos cambios aparecen recogidos en los Test de las Tablas I y II. Los cambios más relevantes competen a celdas cuyo valor es nulo en la columna ISO 2788 y presentan nuevos elementos en la columna ISO 25964.

- 5) Identificación de las reglas de integridad en la nueva ISO. Se revisan las condiciones de integridad recopiladas en el marco evaluativo de la versión anterior atendiendo a estas nuevas reglas, teniendo en cuenta, asimismo, los elementos y estructuras detectadas en los pasos previos.
- 6) Adaptación de las condiciones de integridad ISO 2788 e identificación de las nuevas restricciones que deben considerarse en ISO 25964.
- 7) Revisión paralela de las reglas de integridad en ISO 25964 y en SKOS.
- 8) Examen de los test de integridad y elaboración de los nuevos test recogidos en la Tabla III.
- 9) Análisis de las principales novedades que afectan a la propuesta evaluativa adaptada a ISO 25964.

### 3. Estructuras conceptuales y terminológicas soportadas

Nos fijamos en este epígrafe en los puntos destacables relacionados con los elementos y estructuras soportadas en ISO 25964 y SKOS.

- Tesauros y conceptos. La norma ISO 25964 abandona la idea de base de los descriptores como elemento estructural esencial en un tesoro para alinearse en la visión de tesauros orientados a conceptos, siendo estos últimos los elementos básicos de un tesoro. Los Términos pasan a comportarse como etiquetas ('Labels') de los conceptos.

En SKOS un tesoro se corresponde con un `skos:ConceptScheme` (estructura que recopila y organiza un conjunto de conceptos), y un concepto con un recurso de la clase `"skos:Concept"` que se identifica con un URI. Las etiquetas pueden representar términos preferentes (`skos:prefLabel`), alternativos (`skos:altLabel`) u ocultos (`skos:hiddenLabel`).

- Grupos de conceptos. La nueva ISO introduce la idea de "Grupos de conceptos", en los cuales se engloban los dominios o microtesoros. Tanto ellos como los arrays (ThesaurusArray) son una subclase de `skos:Collection`. Esto es, cualquier array de tesauros o cualquier microtesoro es una `skos:Collection`. Las colecciones SKOS (`skos:Collection`) son grupos de conceptos, que no necesitan estar vinculados por relaciones jerárquicas (TG o TE) (Isaac y Summers, 2009). Los miembros de una colección pueden proceder de diferentes `ConceptSchemes`, lo que significa que una co-

lección puede agrupar conceptos de diferentes tesauros.

En la reciente revisión de la extensión ISO-THES de SKOS se propone la relación "micro-thesaurus of" para indicar que un microtesoro pertenece a un tesoro dado. De modo similar, aunque el anidamiento entre los microtesoros o dominios ha recibido una relación específica para representarlo (`tieneGrupoSuper/tieneGrupoSub`) en ISO 25964, estas relaciones no tienen equivalente en SKOS (ISO TC46/SC9/WG8 e Isaac, 2013), si bien en la reciente revisión de la extensión ISO-THES de SKOS se proponen las relaciones "sub group/super group" para extender SKOS con relaciones equivalentes a las definidas en la ISO 25964.

Arrays (Lista de tesoro). Como ya hemos indicado, al igual que los grupos de conceptos, en (ISO TC46/SC9/WG8 e Isaac, 2013) se propone su representación en SKOS con `skos:Collection`.

- Relaciones de equivalencia. Al pasar a la orientación basada en conceptos, las relaciones semánticas de Descriptores y No Descriptores dejan paso a la idea de un concepto (idea), identificado por una URI unívoca, que tiene asociadas un conjunto de etiquetas (labels), entre las cuales una de ellas es `PreferredLabel`, equivalente a `skos:prefLabel`, y el resto son `NonPreferredLabels`, las cuales tienen su equivalencia SKOS en `skos:altLabel`. En SKOS es posible optar por dos soluciones. Con SKOS-XL, los términos se relacionan con los conceptos a través de las relaciones `PreferredLabel`, `NonPreferredLabel`. Con SKOS Core se asigna directamente un literal, que es la opción básica y más sencilla.

Además, en la nueva ISO, se introduce la equivalencia compleja (compound equivalence), relaciones de equivalencia entre términos USE/USE+ y UP/UP+ que pueden definirse entre conceptos de tesauros diferentes. En ciertos casos el nuevo concepto es la "unión" de los subconceptos, en cuyo caso se recomienda modelarlo con relaciones NT, e incluir el concepto compuesto como un concepto más del tesoro (Will, 2012). La solución para representar esto con SKOS está pendiente, aunque en el anexo a la Extensión SKOS-XL (Isaac y De Smedt, 2015) se aborda una posible extensión de SKOS con este fin. Pastor Sánchez (2013) subraya la especial complejidad de la norma ISO en este sentido al definir relaciones que complementan las que se establecen entre términos

y conceptos (“isNonPreferredLabelFor”, “isPreferredLabelFor”).

Asimismo, conviene advertir que la noción de Término en ISO 25964 viene a encontrar su equivalencia en SKOS-XL, la extensión de SKOS donde los términos o palabras son elevados a la categoría de objetos y tratados en consecuencia como ítems que pueden tener sus propias relaciones, algo que una ‘label’ no puede tener.

Todo lo señalado nos sitúa en una cierta ‘flexibilidad’ de modelado que previsiblemente será abordada por los diseñadores de tesauros de modos variados. Si añadimos a esto que este tipo de relaciones no son frecuentes en los tesauros, es de prever que las herramientas de tesauros no proporcionen soporte para este tipo de relaciones, del mismo modo que venía ocurriendo con las jerarquías de microtesauros (ver la evaluación realizada en Martínez y Alvite, 2014b).

- Relaciones jerárquicas y transitividad: expresan la subordinación entre conceptos: TG (Término Genérico) y TE (Término específico).

Se mantienen en la nueva norma, y se mantienen los símbolos tradicionales por coherencia con la práctica habitual, aunque ahora deben leerse como “concepto genérico/concepto específico” en lugar de “término genérico/término específico”. Cabe destacar la introducción de un ‘rol’ para distinguir tres posibles tipos de relaciones jerárquicas: genérica (kind of), partitiva (part of) y enumerativa o instanciativa (instance of). El efecto de esta distinción sobre la posible aplicación de la transitividad en los razonadores vinculados a las ontologías y que utilizan la subsumpción como eje de la clasificación es crucial: de hecho, resulta imposible aplicar subsumpción en casos de relaciones partitivas (por ejemplo, inferir a partir de “rueda de bicicleta”, que está relacionada con “bicicleta” por una relación ‘NT part of’ que “rueda de bicicleta” es a su vez una “bicicleta” resultaría en realidad una inconsistencia). Esta distinción entre relaciones genéricas (TGG), partitivas (TGP) y enumerativas (TGE) ha planteado la cuestión de una posible extensión de SKOS para representar estos subtipos y nos provoca algunas consideraciones sobre el razonamiento basado en la transitividad que se abordan en la sección de Análisis.

- Relaciones asociativas, entre dos términos o conceptos que no pertenecen a la misma estructura jerárquica aunque sean semántica o

contextualmente similares. TR (Término relacionado).

Se mantiene, pero vincula conceptos. Su equivalente en SKOS es directo: skos:related. Ahora se permite especificar la naturaleza (esto es, semántica) de la relación, para aproximarlo a las ontologías donde las relaciones tienen semántica.

Notas: notas de alcance, de definición, notas de historia, notas del editor del tesoro, notas de uso (nota de cualquier otra clase en relación a un concepto).

La nueva ISO aclara que cuando en una nota de alcance se hace referencia a otros conceptos, se debería proporcionar una nota de alcance recíproca para cada uno de los conceptos mencionados.

### 3.1. Test comparativo ISO 2788 /ISO 25964

Para redefinir los test ha sido necesario prestar atención a los cambios que se comentan en el análisis de las modificaciones en la comprobación de los elementos estructurales soportados que se detalla en la Sección 5.

En las tablas siguientes se indica cómo se definía el test en virtud de la norma ISO 2788 y cómo se redefine en virtud de la norma ISO 25964. Se ha distinguido entre los test que permiten crear los constructores básicos de un tesoro y aquellos que suponen la inclusión de relaciones.

| Test | Def. ISO 2788              | Def. ISO 25964   |
|------|----------------------------|--|
| 1    | Crear tesoro               | Ídem   |
| 2    | Crear dominio              | Crear grupo de conceptos   |
| 3    | Crear Término Preferido    | 3.1 Crear un concepto<br>3.2 Asociar una PreferredLabel al concepto creado |
| 4    | Crear Término No Preferido | Asociar NonPreferredLabel a un concepto                                    |
| 5    | Crear nota                 | Ídem   |
| 6    | Crear array                | Crear array  |
| 7    |                            | Crear un Término   |

Tabla I. Test para constructores

| Test | Def. ISO 2788  | Def. ISO 25964   |
|------|--|--|
| 8    | Crear relación de inclusión entre dominios (subdominios) | Crear una relación tieneGruposSuper/tieneGrupoSub                            |
| 9    | Incluir un concepto en un dominio                        | Incluir un concepto en un grupo de conceptos (crear relación ConceptoTesoro) |

|    |   |   |
|----|---|---|
| 10 | Crear relación TG entre dos Términos Preferentes relacionados previamente                 | Crear relación TG entre dos conceptos no relacionados previamente:<br>a) de tipo genérico<br>b) de tipo enumerativo<br>c) de tipo partitivo |
| 11 | Crear relación TE entre dos Términos Preferentes relacionados previamente                 | Crear relación TE entre dos conceptos no relacionados previamente:<br>a) de tipo genérico<br>b) de tipo enumerativo<br>c) de tipo partitivo |
| 12 | Crear dos relaciones entre un Término Preferente y dos de sus antecesores (polijerarquía) | Crear dos relaciones TE entre un concepto y dos de sus antecesores  |
| 13 | Crear relación TR entre dos Términos Preferentes relacionados previamente                 | Crear relación TR entre dos conceptos no relacionados previamente   |
| 14 | Crear un Término Cabecera   | Crear un Concepto Cabecera (Top Concept)  |
| 15 |   | Incluir un concepto en un array   |

Tabla II. Test para inclusión de relaciones

Se comentan a continuación algunas de las modificaciones más significativas respecto a la propuesta presentada en (Martínez González, Alvite Díez, 2014a, 2014b). En línea con la introducción de los dos niveles, conceptos y tesauros, que supone la evolución de la ISO 2788 a la ISO 25964, se hace necesario introducir un nuevo test, número 7, para crear Términos en el tesauro, estos Términos serán asociados a los conceptos a través de las relaciones *hasPreferredLabelFor*, etc. Es relevante resaltar que esto no ocurría así en la versión evaluativa anterior, dado que, aunque se había considerado la correspondencia con SKOS (Martínez González, Alvite Díez, 2014a, 2014b), en SKOS Core las etiquetas asociadas a los conceptos no pueden relacionarse entre sí, con lo cual se establecía una relativa equivalencia entre un Término Preferente y un Concepto SKOS en cuanto a su importancia en el tesauro (eran estos constructores respectivamente los que podían participar en las relaciones básicas de un tesauro). Ahora las relaciones básicas se establecen tanto en ISO 25964 como en SKOS entre conceptos.

Dado el énfasis mostrado en la ISO 25964 en la distinción entre los tres tipos de relaciones TG/TE -TGG, TGE, TGP- se plantea

redefinir el test (10/11) que previamente se había planteado, de modo que se pueda comprobar si las aplicaciones soportan la creación de las relaciones jerárquicas, pero también si se permite distinguir el rol de la relación jerárquica.

Por razones similares, se han integrado nuevos test, para comprobar el soporte de arrays. En concreto los test que se ocupan de crear un array (número 6), y el que comprueba la inclusión de conceptos en los arrays (número 15).

Sin embargo, queremos resaltar que el soporte de microtesauros o dominios ya estaba contemplado en la versión anterior de este entorno de evaluación.

#### 4. Gestión de la integridad

ISO 25964 dedica un capítulo específico a las recomendaciones sobre el software de gestión de tesauros y epígrafes concretos sobre los requisitos de las aplicaciones para gestionar las relaciones entre términos y conceptos, las notas y los códigos y notaciones, entre otros. Del análisis de estas recomendaciones extraemos las condiciones de integridad que se presentan en este apartado y se obtienen a partir de ellas las reglas que permiten la definición de los test correspondientes.

La nueva norma conserva la esencia de las relaciones de integridad establecidas en ISO 2788. Esto significa que el grueso de las cuestiones planteadas con respecto a la consistencia de un tesauro se mantiene, si bien deben adaptarse en algunos casos a los cambios que supone introducir la idea de concepto como elemento fundamental de un tesauro. Así pues, es necesario asegurar las siguientes condiciones:

- 1) Unicidad. No puede haber elementos duplicados en un tesauro. Los conceptos, grupos de conceptos, o arrays no pueden estar duplicados. Asimismo, los términos, *ThesaurusTerm*, tampoco deben estar duplicados (según se refleja en la cláusula 14.3.a de la norma). A esto debe añadirse que un elemento de un tesauro no puede pertenecer simultáneamente a más de una de estas dos clases. Esto es, no puede haber conceptos que sean simultáneamente grupos de conceptos o términos, etc. Este tipo de restricciones son equivalentes a las que anteriormente venían expresadas como que un Término Preferente sólo podía aparecer una vez, que los Términos Preferentes no podían ser simultáneamente Dominios o Términos no Preferentes.
- 2) Adicionalmente, se mantiene la exigencia de que un concepto tenga un único Término

Preferente en cada lengua del tesoro. En la nueva norma esto significa que sólo puede haber una relación `hasPreferredLabel` para cada concepto. Ahora son los conceptos quienes pueden participar en las relaciones semánticas (antes eran los Términos Preferentes). Esto se aproxima al modelo de SKOS y significa que debe controlarse que no se establecen relaciones básicas entre Términos. Los términos no preferentes (con una relación `USE` o `USE+` con otro término) no puede tener relaciones `TG`, `TE`, `TR` o `UP`.

- 3) Las relaciones básicas son incompatibles, tal y como establecía ISO 2788. Se debe seguir previniendo la introducción de relaciones inconsistentes. Si dos conceptos ya tienen una de las relaciones básicas, no es admisible otra relación básica entre los mismos conceptos. Esto significa que dos conceptos no pueden estar relacionados simultáneamente por una relación `TE/TG` y `TR`.
- 4) Los ciclos siguen estando prohibidos. Si un concepto A es `TG` de otro concepto B, B no puede ser a su vez `TG` de A. Esto es extensible a la jerarquía de modo transitivo. Tampoco puede darse que un Top Concept sea `TE` de otro concepto. No es admisible la relación de un término o un concepto consigo mismo, es decir, todas las relaciones básicas son irreflexivas. Estas restricciones ya se habían contemplado en la propuesta anterior.
- 5) Las relaciones inversas deben introducirse automáticamente, como ya se venía haciendo, esto es, la aplicación debe soportar los requisitos de reciprocidad en las relaciones `TE/TG`, `TR` y `USE/UP`. La nueva norma recomienda que el software se encargue de hacer disponibles automáticamente las relaciones recíprocas en respuesta a la inserción por el editor de cualquiera de ellas, pero como mínimo el software debería advertirle si se produce cualquier relación no recíproca. Además, debería ser posible establecer relaciones recíprocas definidas por el usuario, por ejemplo para distinguir entre diferentes tipos de `TG/TE` o diferentes tipos de relaciones asociativas.
- 6) La eliminación o enmienda de un concepto o término debe conllevar la eliminación de todas las relaciones en las que aparece. También estaba contemplado este requisito en la versión previa de este entorno de evaluación.

Añade ahora la ISO 25964 que, sin embargo, si la eliminación de un término, concepto o relación dejara a cualquier concepto sin por lo menos un término/concepto genérico, o a cualquier término no preferente sin un tér-

mino preferente, la aplicación debería advertirlo.

Sin embargo, con la introducción de los dos niveles, conceptos y términos (Pastor, 2013), aparecen nuevas restricciones que deben incorporarse a las ya contempladas:

- 7) Debe soportarse la posibilidad de asignar el tipo de una relación `TG/TE`. Cada una de las relaciones debe ser de un único rol.
- 8) En caso de que se utilice la relación `USE+`, el software debe comprobar que la relación es por lo menos ternaria.

### Restricciones de integridad en SKOS

Por su parte SKOS define sus propias restricciones de integridad (W3C, 2009), expresadas a través de un conjunto de axiomas:

S9 - `skos:ConceptScheme` is disjoint with `skos:Concept`.

S13 - `skos:prefLabel`, `skos:altLabel` and `skos:hiddenLabel` are pairwise disjoint properties.

S14 - A resource has no more than one value of `skos:prefLabel` per language tag.

S19 - The `rdfs:domain` of `skos:semanticRelation` is the class `skos:Concept`.

S20 - The `rdfs:range` of `skos:semanticRelation` is the class `skos:Concept`.

S27 - `skos:related` is disjoint with the property `skos:broaderTransitive`.

S37 - `skos:Collection` is disjoint with each of `skos:Concept` and `skos:ConceptScheme`.

### Comparación ISO - SKOS

En este apartado se plantea una comparación de las restricciones de integridad entre la norma ISO y SKOS que no figura en los documentos de trabajo utilizados por el grupo de trabajo del W3C, pero que ayuda a prever qué restricciones de integridad de un tesoro pueden estar previsiblemente soportadas por una herramienta diseñada para trabajar con SKOS y cuáles no.

Algunas de estas restricciones se corresponden con las reglas de integridad para los tesauros. Esto supone que una herramienta SKOS, esté o no específicamente diseñada para tesauros, debería aportar la garantía de esas restricciones. Sin embargo, en los casos en que una restricción presente en tesauros no esté incluida en SKOS sí será importante prestar atención al hecho de que la herramienta esté creada para manipular tesauros, puesto que recae sobre la

responsabilidad de sus creadores aportar ese plus de comprobación.

- Los axiomas S9 y S37 garantizan que un concepto no puede ser simultáneamente un grupo de conceptos. Esto es así tanto en los casos, como el del tesoro UKAT, donde los microtesoros se han representado con `skos:ConceptScheme`, como en aquellos donde se ha seguido la propuesta hecha en ISO TC46/SC9/WG8 e Isaac (2013) de utilizar `skos:Collections` para representar este tipo de estructura.
1. El axioma S14 garantiza que no hay más de un Término Preferente (`hasPreferredLabel`) para un tesoro.
  2. El axioma S13 asegura que los términos elegidos con Preferred Labels no son utilizados simultáneamente como Non Preferred Labels.
  3. Los axiomas S19 y S20 garantizan que son los conceptos quienes participan en las relaciones semánticas.
  4. El axioma S27 asegura la incompatibilidad de relaciones básicas, esto es, que dos conceptos no pueden estar relacionados simultáneamente por una relación TG/TE y TR.

No obstante, es interesante poner atención en las restricciones de la norma ISO que no están cubiertas por SKOS, ya que son las que con mayor probabilidad van a condicionar las diferencias entre una aplicación SKOS y una aplicación de tesoros. Además, con la introducción en la nueva ISO de las relaciones entre términos para representar conceptos complejos (relación USE+), cuya equivalencia en SKOS aún no está resuelta (ISO TC46/SC9/WG8 e Isaac, 2013) se añade un reto adicional para estas herramientas. Se resumen a continuación las restricciones ISO que no encuentran su equivalencia en restricciones SKOS:

5. En SKOS, la duplicidad de etiquetas, aunque no está recomendada, no está estrictamente prohibida (ver en la Recomendación SKOS, Sección 2.2.1, Appendix). Esto significa que dos `skos:Concept` pueden compartir la misma Preferred Label sin que el grafo sea inconsistente. Sin embargo, no debería ser así en un tesoro. Cada concepto del tesoro está representado por un término preferente por lengua. Aunque esto se aborda en algunas herramientas y también en la validación de SKOS, no está incluido en la recomendación y, por tanto, no debería asumirse que cualquier herramienta SKOS detectará un

caso semejante como un conflicto que necesariamente deba resolverse.

6. En SKOS, los ciclos sí están permitidos, a diferencia de lo que ocurre en los tesoros. Por ejemplo, es consistente en SKOS que `A skos:broader B` y `B skos:broader A` (ver sección 8.6.8 de la Recomendación SKOS).
7. SKOS no establece ninguna restricción de integridad en relación con el hecho de que `skos:broader` y `skos:narrower` son relaciones inversas. Aunque se pueden establecer dependencias que permitan inferir la una a partir de la otra, es consistente que sólo aparezca una de ellas. Así pues, nuevamente estamos ante una restricción más fuerte en los tesoros que en SKOS.
8. En SKOS Core solamente los conceptos participan en relaciones. En SKOS-XL sí se pueden relacionar las 'labels'. De hecho, es esta versión extendida de SKOS la que mejor acoge la posibilidad de modelar los términos y las relaciones entre ellos, si bien tiene limitaciones para ofrecer toda la potencialidad que requiere la norma ISO 25964 (Pastor, 2013) y actualmente se sigue trabajando en las extensiones o modificaciones a SKOS que permitan hacerlo (Isaac y Smedt, 2015).

#### 4.1. Test comparativo ISO 2788/ISO 25964

Para comprobar las condiciones de consistencia del vocabulario se ha definido un conjunto de test para las restricciones de integridad.

| Test | Def. ISO 2788  | Def. ISO 25964   |
|------|--|--|
| 1    | Relacionar un Término No Preferente con más de un Término Preferente | Asignar más de un Término Preferente a un Concepto   |
| 2    | Crear un dominio igual a otro existente                              | Crear un grupo de conceptos equivalente a otro existente (usar el mismo Término Preferente para dos grupos de conceptos) |
| 3    | Crear un Término Preferente igual a otro ya existente                | Crear un concepto equivalente a otro existente (usar el mismo Término Preferente para dos conceptos)                     |
| 4    | Crear un Término Preferente igual a un dominio                       | Usar el mismo Término Preferente para un concepto y para un grupo de conceptos   |
| 5    | Crear un Término Preferente igual a un Término no Preferente         | Usar el mismo Término Preferente como PreferredLabel y como NonPreferredLabel  |

|    |  |  |
|----|--|--|
| 6a | Crear una relación básica (TR, TG, TE) entre dos Términos no Preferentes                       | Crear una relación básica en la que participe un Término                                   |
| 6b | Crear una relación básica entre un Término Preferente y un Término no Preferente               | Crear una relación básica en la que participe un Término                                   |
| 7  | Crear una relación TE entre dos términos que ya están relacionados por una relación TG         | Crear una relación TE entre dos conceptos que ya están relacionados por una relación TG    |
| 8  | Crear una relación RT entre dos términos que ya están relacionados por una relación TE/TG      | Crear una relación TR entre dos conceptos que ya están relacionados por una relación TE/TG |
| 9  | Hacer un Término Cabecera TE de otro término   | Relacionar un TopConcept con otro concepto por una relación jerárquica                     |
| 10 | Crear una relación TG/TE que cierra un ciclo de tipo (TG/TE, TR) (ej.: A TG B, B TG C, C TR A) | Ídem   |
| 11 | Crear una relación TG/TE que cierra un ciclo TG/TE (ej.: A TE B, B TE C, C TE A)               | Ídem   |
| 12 | Eliminar una relación TG/TE  | Ídem   |
| 13 | Eliminar una relación TR   | Ídem   |
| 14 | Eliminar un Término Preferente   | Eliminar un concepto   |
| 15 |  | Eliminar un término  |
| 16 | Eliminar un dominio  | Eliminar un grupo de conceptos   |
| 17 | Eliminar un Término Cabecera   | Eliminar un concepto cabecera  |
| 18 |  | Eliminar un array  |
| 19 |  | Eliminar una relación de equivalencia de Término compuesto                                 |
| 20 | Eliminar una relación de 'subdominio'  | Eliminar una relación SubGroupOf/SuperGroupOf entre dos grupos de conceptos                |

Tabla III. *Test de restricciones de integridad*

Para definir los test se han examinado en cada una de las condiciones de integridad los posibles casos de quiebra de la restricción. Por ejemplo, la Unicidad puede romperse por varias vías, lo cual se traduce en los test 1 a 5.

En el conjunto de test relacionados con la integridad son relevantes los siguientes cambios:

9. Se introducen test para comprobar la eliminación de los constructores cuya creación deviene relevante en la nueva ISO: arrays, equivalencias compuestas y términos.
10. Se adaptan los test que comprueban la unicidad (esto es, que no se duplican ítems en el tesoro) de modo que ya no se comprueba si se trata del mismo Término Preferente sino que no se asigne el mismo Término como 'label' de dos conceptos diferentes.
11. Se ha hecho necesario adaptar los test que verificaban el tipo de estructuras que pueden participar en las relaciones básicas. Ahora son conceptos, en los test, donde antes se preveía la participación de Términos No Preferentes, ahora debe prevenirse la participación de Términos. Los test afectados son los 6.a y 6.b, que se fusionan ahora en un único test (fijarse en que la redacción dada actualmente es la misma en ambos casos).
12. Asimismo, aunque no esté expresada como tal en la norma ISO 25964, se ha determinado, por considerarla lógica desde un punto de vista de la semántica subyacente, que dos conceptos, A y B, no pueden estar relacionados por dos relaciones TG/TE, aunque tengan roles diferentes. Esto es, no aceptamos situaciones como *A TGG B* y *A TGP B*. En consecuencia se han mantenido los test 7 y 8, e incluso se está considerando la conveniencia de un test adicional expresado como *'Crear una relación básica, TG/TE/TR, entre dos conceptos que ya están vinculados por una relación del mismo tipo'*.

## 5. Análisis

En este epígrafe se proyectaba recoger los resultados de la aplicación de los nuevos test sobre las dos herramientas de gestión de tesoros que utilizan tecnologías semánticas y que fueron empleadas en los trabajos evaluativos previos sustentados en ISO 2788: PoolParty (1) y SKOSEd (2).

Para contrastar la metodología, en las pruebas proponíamos emplear los cuatro tesoros utilizados en los estudios anteriores:

13. EuroVoc (3), versión del Tesoro EuroVoc representada con SKOS.
14. NAL -National Agricultural Library thesaurus- (4). Representación en SKOS/RDF.
15. UKAT Thesaurus (4), representación en SKOS/RDF. Cabe destacar que en la fecha de redacción de este trabajo (junio de 2015) la versión disponible para descarga sigue siendo de 2009. En la web de UKAT se in-

cluye como aspecto novedoso la posibilidad de acceso a una versión UKAT 'Linked Data', pero cuando se ha intentado este acceso ha sido imposible conseguir conexión con el servidor, con lo cual no se ha podido comprobar si se mantiene la representación de microtesauros como skos:ConceptScheme o se ha modificado a tenor de lo expresado en las correspondencias ISO - SKOS (ISO TC46/SC9/WG8 e Isaac, 2013).

16. Tesoro ad-hoc, un mini-tesoro construido a partir de EuroVoc con 70 términos, 42 descriptores, 24 no descriptores y 2 notas de alcance.

Queremos advertir que la ejecución de las pruebas que planteábamos incluir en el trabajo actual no ha podido llevarse a efecto debido a las especificidades que presentan estos software en el momento actual.

PoolParty es un sistema de gestión de tesauros dirigido a la Web Semántica (Nagy et al., 2011). Se trata de un software comercial que cuenta con una versión de evaluación en la que algunas de las utilidades del programa no están incluidas. PoolParty utiliza SKOS y trabaja con las especificaciones de este estándar. Cuenta con una interfaz de usuario sencilla y permite crear y mantener tesauros multilingües. Desde que se realizaron los pruebas presentadas en (Martínez González, Alvite Díez, 2014a, 2014b), PoolParty ha seguido evolucionando (actualmente en la versión 5.1) incluyendo más funcionalidades, destacando, entre otras, la extracción automática de conceptos a partir de textos. Sin embargo, en lo que nos afecta (elementos estructurales soportados en los tesauros y restricciones de integridad) el cambio más relevante es que, según expresan en su propia web, esta herramienta es ahora un "editor de ontologías", lo cual nos genera ciertas dudas sobre la intención o urgencia de sus desarrolladores por incorporar las modificaciones introducidas en la nueva ISO para tesauros.

SKOS Ed es un plugin del editor de ontologías OWL, Protégé. Esta aplicación permite crear y editar vocabularios representados en SKOS. Introduce ciertas variaciones con respecto a las notaciones comunes en estos lenguajes normalizados, así aparece la relación "Narrower Transitive" y "Broader Transitive". En cierto sentido, el usuario puede sentirse desbordado por la aplicación si no está familiarizado con las funcionalidades de Protégé o con editores de ontologías. En la propuesta que ahora revisamos se había evaluado la versión 1.0, siendo la última disponible la versión 2.0, de agosto de 2012. Si bien en esa fecha la Parte 1 de la norma ISO

25964 ya se había publicado, no es de esperar que SKOS Ed soporte los cambios introducidos en ella puesto que, como ya se indicaba en su momento y se insiste en este trabajo, esta herramienta es un editor SKOS y no un editor de tesauros, razón por la cual, no habiendo sido, entre tanto, modificada la recomendación SKOS, no existe exigencia alguna para soportar las nuevas estructuras y capacidades incorporadas en ISO 25964.

A la espera de que las extensiones de SKOS que buscan la compatibilidad con ISO 25964 puedan incorporarse a las aplicaciones de gestión de tesauros y sea viable el correspondiente testeo de los criterios evaluativos propuestos en este estudio, procedemos a analizar a continuación las principales modificaciones introducidas en la propuesta evaluativa.

### **Análisis de las modificaciones en la comprobación de elementos y estructuras soportadas**

En lo que respecta a los cambios relevantes introducidos por la nueva ISO que afectan al instrumento evaluativo propuesto previamente bajo el modelo de ISO 2788, destacamos los puntos siguientes:

17. En lugar de términos preferentes o descriptores como estructuras básicas de un tesoro, se pasa ahora a usar Conceptos, tanto en ISO 25964 como en SKOS, como estructura básica de un tesoro. Los tesauros (en ISO 2788 e ISO 25964) se corresponden en SKOS con esquemas de conceptos (skos:ConceptScheme).
18. Los descriptores de ISO 2788 deben corresponderse con las PreferredLabels de ISO 25964, y los No-descriptores con las NonPreferredLabels.
19. Si bien no es recomendable, es posible en ISO 25964 y en SKOS que dos PreferredLabels coincidan (en rigor, los conceptos se identifican por su URI y no por sus Labels). Sin embargo, en ISO 2788 no podía haber dos Descriptores iguales, ya que originaba un problema de identificación.
20. Aunque la relación hasTopConcept/isTopConceptOf de ISO 2788 o TopConcepts de ISO 25964 no tienen una equivalencia directa en SKOS según se indica en ISO TC46/SC9/WG8 e Isaac (2013), es posible indicar que un concepto es Top Concept de un esquema de conceptos usando la propiedad skos:topConceptOf, con lo cual es posible representar los Top Terms o Top Concepts de las respectivas normas ISO en SKOS.

21. Las relaciones TG, TE y TR se corresponden con relaciones equivalentes, que ahora relacionan conceptos en lugar de términos.

ISO 25964 introduce un 'rol' con tres valores posibles en las relaciones TG/TE: genérica (kind of), partitiva (part of) y enumerativa o instanciativa (instance of). Esto afecta a las posibilidades de inferencia en base a la transitividad de estas relaciones (Will, 2012), que pueden no ser cierta cuando se trata de un rol 'partitivo'. En este sentido, nos gustaría llamar la atención sobre un aspecto de SKOS que resulta especialmente útil ante esta nueva posibilidad: contra lo que en principio se podría pensar sin hacer una lectura detallada de la recomendación, en SKOS las relaciones `skos:broader/skos:narrower` no son transitivas. Esto es, un razonador no debería inferir nunca ante la presencia de las aserciones 'A `skos:broader` B' y 'B `skos:broader` C' que es cierto 'A `skos:broader` C'. Este es exactamente el caso que ocurriría en la situación de una relación de esta naturaleza con un rol 'partitivo'. Pero es posible representar y utilizar la transitividad en SKOS para razonar en consecuencia, utilizando para ello las relaciones expresamente creadas al efecto: `skos:broaderTransitive/skos:narrowerTransitive`. Estas son, pues, las que deberían utilizarse en caso de que el rol permita razonar en base a la transitividad a la hora de hacer la equivalencia de las normas ISO a la representación con SKOS. Igualmente, relacionado con esta puntualización, conviene señalar que la propiedad `skos:related` no es transitiva en SKOS, ni tiene ninguna alternativa similar a las que hemos indicado para la relación jerárquica que permita razonar en base a la transitividad. No obstante, dado que las extensiones de SKOS para soportar esta distinción -a través de subpropiedades de `skos:broader` y `skos:narrower` según se plantea en ISO TC46/SC9/WG8 e Isaac, (2013)- aún no están disponibles, nuestra previsión es que la mayoría de las herramientas den resultado negativo en estos tres nuevos test específicos introducidos.

22. Las notas se corresponden con notas equivalentes en ISO 25964 y/o SKOS.

23. Los microtesauros o dominios se representan con Grupos de conceptos (ISO 25964), los cuales se corresponden a su vez con colecciones en SKOS (`skos:collection`). Esto supone un cambio respecto a la situación anterior, donde quedaba sin determinar esta equivalencia, de modo que en algunos tesau-

ros utilizados en nuestra evaluación anterior, por ejemplo el tesoro UKAT, los microtesauros se habían representado como `skos:ConceptSchemes`. Otra alternativa habría sido recurrir a las colecciones (Pastor, 2013). Ahora, sin embargo, parece más claro que la nueva norma ha optado por recurrir a `skos:Collection` para representar este tipo de estructura en SKOS.

24. Los anidamientos entre microtesauros o dominios se representan en ISO 25964 por la relación `hasSuperGroup/hasSubGroup`, si bien en SKOS no está resuelto cómo representar estas relaciones, lo cual mantiene una incertidumbre que, en nuestra opinión, difícilmente animará por el momento a los desarrolladores de software de tesauros a incorporar esta potencialidad que hasta ahora venían descartando.

#### **Análisis de las modificaciones con respecto a los requerimientos de integridad**

Con respecto a las restricciones de integridad subrayamos los puntos siguientes:

25. A la luz de ISO 25964 es necesario comprobar la eliminación de los constructores cuya creación deviene relevante en la nueva norma: arrays, equivalencias compuestas y términos.

26. En lo que se refiere a la unicidad (esto es, que no se dupliquen ítems en el tesoro) ya no se requiere verificar si se trata del mismo Término Preferente sino comprobar automáticamente que no se asigne el mismo Término como 'label' de dos conceptos diferentes.

27. Ni los conceptos ni los términos pueden estar duplicados. Las URI son siempre únicas, lo cual garantiza la unicidad en sentido estricto. En consecuencia, lo que debe garantizarse es que el mismo término no puede usarse como etiqueta preferente para dos conceptos distintos, que sería el error que previsiblemente podría cometer un editor de tesauros durante su creación.

28. La especificación de dos niveles, conceptos y términos, introduce la posibilidad de que un término quede flotante en el tesoro, sin relación con ningún concepto. Esto afecta a los resultados esperados de los test de integridad relacionados con la eliminación de elementos del tesoro, que ahora son más exigentes.

29. La comprobación de los ciclos sigue siendo necesaria, independientemente de la introducción de roles en las relaciones TG/TE.

Las consideraciones sobre la comparación con SKOS apuntan a que este es un conjunto de test en los que se debe prestar especial atención a los resultados que se obtengan con las herramientas.

30. Algo similar ocurre con los test que comprueban la inserción automática de relaciones inversas en un tesoro.
31. Donde en ISO 2788 se preveía la participación de Términos No Preferentes, ahora debe prevenirse la participación de Términos. Los test afectados son los 6a y 6b, que se fusionan ahora en un único test en el instrumento evaluativo que hemos diseñado (Tabla III).
32. Se decide, aunque no esté expresada como tal en la norma ISO 25964, incluir una restricción que prevenga que dos conceptos A y B, no pueden estar relacionados por dos relaciones TG/TE, aunque tengan roles diferentes.

## 6. Consideraciones finales

El trabajo nos ha permitido identificar los cambios introducidos por ISO 25964 que afectan al marco evaluativo previo propuesto para la valoración de herramientas de gestión de tesauros con ISO 2788.

Aplazamos para una segunda fase el estudio de las relaciones entre Términos para expresar equivalencias de términos compuestos y otras relaciones terminológicas complejas ("USE/USE+", "UF/UF+"), así como los Arrays.

Estamos expectantes ante las conclusiones finales del trabajo que se está realizando para establecer las correspondencias entre SKOS y la norma ISO 25964. Resultará crucial que este trabajo derive finalmente en una extensión de la recomendación SKOS, como parece que se está apuntando desde el grupo de trabajo encargado de esta revisión. En ese momento será necesario examinar las consideraciones realizadas en nuestra comparación de SKOS y el modelo de datos de la nueva norma, en particular las que afectan a las restricciones de integridad, para verificar si estos dos estándares se acercan más en este aspecto o, por el contrario, se mantienen las diferencias más significativas detectadas en la actualidad.

Por ahora parece que las aplicaciones que vienen trabajando con SKOS para la gestión de tesauros todavía no han implementado las funcionalidades derivadas del modelo de datos de ISO 25964. Con todo, consideramos relevante conocer en qué medida se está llevando a cabo la armonización de estos estándares dentro de

las aplicaciones ya ideadas para la gestión de vocabularios en la web semántica y nos planteamos ampliar el análisis a herramientas de gestión de tesauros "tradicionales" que adecúen sus aplicaciones a las exigencias de la nueva normativa y, en consecuencia, al entorno de interoperabilidad entre vocabularios que abandera SKOS.

## Notas

- (1) <http://www.poolparty.biz/>
- (2) <https://code.google.com/p/skoseditor/>
- (3) <http://eurovoc.europa.eu/>
- (4) <http://agclass.nal.usda.gov/>
- (5) <http://www.ukat.org.uk/thesaurus/>

## Referencias

- Dextre Clarke, S.G.; Zeng, M.L. (2012). From ISO 2788 to ISO 25964: The evolution of thesaurus Standards towards interoperability and data modeling. *Information Standards Quarterly (ISQ)*. 24(1), 20-26. Recuperado 27-05-2015 de [http://www.niso.org/apps/group\\_public/download.php/8285/isqv24no1.pdf](http://www.niso.org/apps/group_public/download.php/8285/isqv24no1.pdf)
- García Marco, F. J. (2014). La nueva norma de tesauros UNE ISO 25964-1. *Anuario ThinkEPI*. 8, 159-163.
- Isaac, A.; De Smedt, J. (2015). *Notes apply to the published mapping between the ISO 25964 data model and the SKOS schema*. Recuperado 27-05-2015 de <http://pub.tenforce.com/schemas/iso25964/skos-thes/>
- Isaac, A.; Summers, E. (2009). *SKOS Simple Knowledge Organization System Primer. W3C Working Group, Note 18 August 2009*. Recuperado 29-05-2015 de <http://www.w3.org/TR/skos-primer>
- ISO (2011). *ISO 25964-1:2011 Information and documentation -- Thesauri and interoperability with other vocabularies -- Part 1: Thesauri for information retrieval*. Geneva: International Organization for Standardization.
- ISO (2013). *ISO 25964-1:2011 Information and documentation -- Thesauri and interoperability with other vocabularies -- Part 2: Interoperability with other vocabularies*. Geneva: International Organization for Standardization.
- ISO TC46/SC9/WG8 ; Isaac, A. (2013). *Correspondence between ISO 25964 and SKOS/SKOS-XL models*. Recuperado 27-05-2015 de [http://www.niso.org/apps/group\\_public/download.php/12351/Correspondence%20ISO25964-SKOSXL-MADS-2013-12-11.pdf](http://www.niso.org/apps/group_public/download.php/12351/Correspondence%20ISO25964-SKOSXL-MADS-2013-12-11.pdf)
- Martínez González, M. M.; Alvite Díez, M. L. (2014a). On the evaluation of thesaurus tools compatible with the Semantic Web. *Journal of Information Science*. 40(6), 711-722. doi: 10.1177/0165551514545603
- Martínez González, M. M.; Alvite Díez, M. L. (2014b). Propuesta metodológica de evaluación de gestores de tesauros compatibles con la web semántica. *Anales de Documentación*. 17(1). Recuperado 27-05-2015 de <http://revistas.um.es/analesdoc/article/view/analesdoc.17.1.186271>
- Nagy, H.; Pellegrini, T.; Schandl, T.; Blumauer, A.; Mader, C. (2011). La realización de los casos de uso basados en tesauros con el paquete PoolParty. *BiD: textos universitarios de biblioteconomía i documentació*. (27). Recuperado 27-05-2015 de <http://bid.ub.edu/27/nagy2.htm>

- Pastor Sánchez, J. A. (2013). ISO-THES: ampliando Skos a partir de la norma de tesauros ISO 25964. *Anuario ThinkEPI*, 7, 189-193.
- Pastor Sánchez, J. A.; Martínez Méndez, F. J.; Rodríguez Muñoz, J. V. (2012). Aplicación de SKOS para la interoperabilidad de vocabularios controlados en el entorno de linked open data. *El profesional de la información*, 21(3), 245-253.
- Sánchez Cuadrado, S.; Colmenero Ruiz, M. J.; Moreiro, J. A. (2012). Tesauros: estándares y recomendaciones. *El profesional de la información*, 21(3), 229-235.
- Will, L. (2012). The ISO 25964 data model for the structure of an information retrieval thesaurus. *Bulletin of the American Society for Information Science and Technology*, 38(4), 48-51. Recuperado 27-05-2015 de [https://www.asis.org/Bulletin/Apr-12/AprMay12\\_Will.pdf](https://www.asis.org/Bulletin/Apr-12/AprMay12_Will.pdf)
- World Wide Web Consortium (W3C). (2009). *SKOS Simple Knowledge Organization System Reference*. W3C Recommendation 18 August 2009. W3C. Recuperado 27-05-2015 de <http://www.w3.org/TR/2009/REC-skos-reference-20090818>

---

# Organização do conhecimento e tratamento da informação: analisando a subjetividade no processo de indexação e implicações na recuperação da informação

*Knowledge organization and information treatment: analyzing subjectivity in the process of indexing and implications in information retrieval*

---

**Eliane Pawlowski Oliveira Araújo (1), Claudio Paixão Anastácio de Paula (2), Jacqueline Pawlowski Oliveira (3), Simone Torres (4)**

- (1) Universidade Federal de Minas Gerais Av. Antônio Carlos 6627 – Programa de Pós-Graduação da Escola de Ciência da Informação. Belo Horizonte / Minas Gerais / Brasil. 31270-901, elianepaw@yahoo.com.br;  
(2) claudiopap@hotmail.com,  
(3) jackiepaw@gmail.com,  
(4) simone.torres@almg.gov.br

## Resumo

Muito se tem discutido sobre como o desenvolvimento das tecnologias de informação e comunicação alterou o cenário e os comportamentos informacionais, principalmente a partir dos primeiros anos do Século XXI. Entretanto, nesse contexto, é interessante verificar que, para além da preocupação inicial de disponibilizar ao usuário o máximo de informações, vai se impondo um propósito mais efetivo do que priorizar a mera disseminação da informação: a necessidade de recuperar aquilo que possa atender as demandas dos indivíduos de forma eficaz, tendência que vem se apresentando como uma nova perspectiva no que tange à organização do conhecimento e como um desafio aos sistemas informacionais. Considera-se que um dos pontos chave para que o processo de recuperação da informação ocorra de forma adequada se encontra na etapa de tratamento informacional, processo que compreende, dentre várias tarefas, a indexação. Esta atividade, que apresenta o desafio de atribuir termos que representem fielmente o objeto analisado, tem sua qualidade condicionada à qualificação do indexador e sua "imparcialidade" no processo. Contudo, sabe-se ser quase inevitável que alguns elementos de julgamento subjetivo não afetem a realização dessa atividade. O presente trabalho pretende apresentar o resultado parcial de estudo realizado em Bibliotecas Universitárias que teve como objetivos: identificar como as perspectivas individuais permeiam o processo de tratamento informacional; verificar a influência da subjetividade na atividade de indexação impactando consequentemente o processo de recuperação de informação; e validar o uso de métodos alternativos para investigação de comportamentos informacionais e estudos de usuários em Ciência da Informação. A pesquisa teve como eixo norteador a Abordagem Clínica da Informação,

que é uma perspectiva que apresenta a possibilidade de investigar o comportamento informacional considerando a influência de elementos culturais, simbólicos, cognitivos e afetivos, bem como fatores psicodinâmicos – conscientes e inconscientes. Como métodos constituintes dessa abordagem ressalta-se a entrevista, a análise de tarefa, a técnica do incidente crítico, o protocolo verbal, com destaque para a utilização do Teste Arquetípico de Nove Elementos para analisar os aspectos subjetivos por meio da via simbólica. Os resultados obtidos possibilitaram compreender como a subjetividade se integra às competências individuais para influenciar o tratamento informacional repercutindo na recuperação de informação pelos usuários. As constatações possibilitadas pela pesquisa demonstram a necessidade de caminhar na busca pelo estabelecimento de diretrizes mais precisas de modo a possibilitar a definição de parâmetros mais eficazes para orientar a execução dessa atividade.

**Palavras-chave:** Recuperação da Informação. Tratamento da informação. Análise de assunto. Indexação. Catalogação. Subjetividade

## Abstract

There has been much debate about how the development of information technology and communication has changed the information scenario and behaviours, especially since the beginning of the 21<sup>st</sup> century. However, it is interesting to note, in this context, that besides the initial concern of making available the greatest amount of information for the user, and prioritizing the dissemination of information, a greater purpose imposes itself: the need to retrieve what can attend the user's demand efficiently. This is a new perspective regarding knowledge organization and a challenge to information systems. One of the

key activities for the process of adequate information retrieval is information treatment, which comprehends, among various tasks, indexing. This activity presents the challenge of attributing terms that represent the analyzed object faithfully. The quality of the process of indexing is conditioned by how qualified the indexer is, and his or her impartiality in the process. However, it is known that it is inevitable that some elements of subjective judgment affect this activity. This article aims to present the partial results of a study carried out in university libraries, with the objective of: identifying how the individual perspectives permeate information processing; verifying the influence of subjectivity in indexing, which affects information retrieval; and validating the use of alternative methods to investigate information behaviour and to carry out user studies in Information Science. This research had the Clinical Approach to Information as a guiding principle, a perspective that presents the possibility of

investigating information behaviour considering the influence of cultural, symbolic, cognitive and affective elements as well as psychodynamic factors – conscious and unconscious. The methods that constitute this approach are interviews, task analysis, the critical incident technique, the verbal protocol, and the use of the Archetypal Test with Nine Elements. The results obtained enabled the understanding of how subjective aspects integrate themselves to individual competences influencing the process of information processing, having an impact on information retrieval by users. The results of the study show the need to establish more precise guidelines to enable the definition of more effective parameters to orient this activity.

**Keywords:** Information retrieval. Information treatment. Subject analysis. Indexing. Cataloging. Subjectivity.

## 1. Introdução

A recuperação da informação é uma preocupação antiga. Essa afirmação pode ser confirmada em algumas iniciativas como o “*mundaneum*” de Paul Otlet que, já no Século XIX, objetivava tornar acessível o conhecimento produzido, assim como o “*memex*”, de Vannevar Bush que visava facilitar a recuperação da informação nos idos da década de 1940 (Garcia, 2007).

A recuperação da informação, em sistemas de gestão de bibliotecas, é um dos aspectos mais importantes decorrentes do tratamento informacional. De acordo com Araújo (2013), “tratar” a informação implica a execução de várias atividades - dentre as quais a indexação - todas visando possibilitar ao usuário a recuperação da informação desejada com a maior precisão e rapidez possíveis. Entretanto, apesar da existência de metodologias que orientam a prática da indexação existe um vazio teórico e metodológico na elaboração de informações documentárias cujas regras não apresentam indicações objetivas sobre o modo de obter os produtos desejados ficando o êxito do trabalho realizado, segundo Kobashi (1994), relacionado ao bom senso e experiência do indexador.

Esse fato nos leva a atentar para a influência dos aspectos subjetivos na realização do processo de indexação e, em especial, na atividade de análise de assunto. Desta forma, compreender como se integram e interrelacionam a subjetividade e os comportamentos vinculados ao tratamento da informação pode fornecer indícios de como minimizar a interferência desse aspecto no processo favorecendo, conseqüentemente, que a recuperação possa ocorrer da forma mais eficaz pelo usuário.

## 2. Aspectos teóricos

Para descortinar esse universo foi realizada uma pesquisa em Bibliotecas Universitárias com bibliotecários com experiência na execução da atividade de análise de assunto em processos de catalogação/indexação visando obter, pela via simbólica, percepções sobre os aspectos intrínsecos presentes nesta atividade.

Utilizou-se como eixo norteador para verificação da influência da subjetividade no processo de tratamento da informação a “Abordagem Clínica da Informação”, perspectiva proposta por Paula (2011). Essa abordagem apresenta a possibilidade de investigar o comportamento informacional considerando a influência de elementos culturais, simbólicos, cognitivos e afetivos, bem como fatores psicodinâmicos – conscientes e inconscien-

tes. Inspirada na designação francesa *approche clinique*, tem por característica um olhar profundo do fenômeno da informação utilizando-se de uma perspectiva clínica (sem o viés psicopatológico) para atingir níveis de análise não usuais nos estudos comportamentais e cognitivistas tradicionais.

Essa ação, pressupõe-se, irá permitir a descrição de fenômenos e possibilitar tecer diagnósticos, prognósticos ou prescrever intervenções, pois os sujeitos do estudo são compreendidos em suas interações com o contexto que os rodeia e com seus elementos intrínsecos. Segundo Paula (2011), a adoção desta prática possibilita o “sair de cena” de uma postura mais funcionalista da relação com a informação e adentrar ao palco das ações numa busca intensa pelos “comos” e os “porquês” das ações, consideradas subjetivas e dotadas de significados.

Dentre os métodos e técnicas utilizados na pesquisa para compor esta abordagem foram selecionadas metodologias que possuem como fundamento os aspectos simbólicos e o imaginário, pois, segundo Gilbert Durand (1997) o imaginário é considerado como o alicerce sobre o qual se constroem as concepções de homem e de mundo. Os estudos deste autor, segundo Oliveira e Maia (2008, p.1), procuraram colocar “a imagem, a imaginação e o imaginário no cenário dos estudos acadêmicos” e se consolidaram em uma teoria que se organiza sob o método da convergência, isto é, os símbolos se (re)agrupam em torno de núcleos organizadores, indicando que há estreita relação entre os gestos do corpo e as representações simbólicas (Cemim *et al*, 2001).

Os estudos de Gilbert Durand foram sistematizados pelo psicólogo Yves Durand (1988), compondo o Teste Arquetípico dos Nove Elementos – AT-9. O Teste, conforme explicitado por Cemin *et al* (2001), permite entender como o sujeito age frente às suas angústias e enfrentamentos mais elementares. De acordo com Paula (2012), a intenção é utilizar estímulos arquetípicos “que têm o papel de colocar o problema trabalhado numa perspectiva de tempo, ameaça e finitude” para construir modos de enfrentamento de um problema.

No contexto da pesquisa realizada, o “problema” em questão é a decisão sobre a atribuição de termos na execução da atividade de análise de assunto. Essa atividade, que constitui-se como a etapa do tratamento temático em que um documento é analisado, visa a extração de conceitos que possam traduzir sua essência. Tem como um de seus pontos relevantes o fato

de que a correta seleção de termos facilita o atendimento da demanda informacional por parte dos usuários em sistemas informatizados.

Naves (1996) define a análise de assunto como a operação base para os procedimentos de recuperação da informação. Ampara-se em Harris (1970) ao afirmar que a ideia de que esta atividade constitui-se uma área aparentemente simples “ocorre por absoluto desconhecimento da complexidade do processo que exige esforços [...] no sentido de seguir uma metodologia adequada para obter resultados satisfatórios” (Naves, 1996, p.217). A autora também aponta a influência da pessoa que executa a atividade, pois “não há dúvidas de que o indexador interpõe suas próprias ideias e preconceitos na sua atuação de intermediário entre autores e usuários” (Naves, 1996, p.221).

O que torna esse processo complexo, de acordo com Ferneda (2003), é que a necessidade de informação do usuário será representada por meio de uma expressão de busca e prever os termos que foram usados para representar os documentos pelo indexador tem sido um desafio considerando a quantidade de documentos disponibilizados nos sistemas de informação.

Cabe destacar que a ação de descrever um documento de acordo com seu assunto – definida pelo UNISIST (1981) como indexação – possui terminologia e conceitos próprios no contexto das bibliotecas, sendo considerada como uma das atividades inerentes ao processo de catalogação. Apesar de alguns estudos como o de Fujita (2003), Silva e Fujita (2004), Lancaster (2004) e Fujita *et al* (2010) apresentarem diferentes perspectivas teóricas e metodológicas sobre a indexação e a catalogação, no presente trabalho, as diferenças terminológicas dessas atividades não foram abordadas. Optou-se por analisar o ponto de equivalência entre ambas que é a atividade de análise de assunto, tarefa que busca representar documentos por meio de termos visando sua posterior recuperação, procedimento inerente ao exercício do bibliotecário na qualidade de catalogador. Conforme afirmam Silva e Fujita (2004, p.142) “Apesar das divergências sobre semelhanças e diferenças entre os termos, a indexação alfabética de assuntos e a catalogação de assuntos são equivalentes porque são resultados de um mesmo processo: a análise de assunto”.

### 3. Aspectos metodológicos

A pesquisa foi conduzida em uma biblioteca especializada integrante de um Sistema de Bibliotecas de uma instituição de ensino superior

cujos acervo bibliográfico é um dos maiores do Sistema. Esse contexto foi selecionado considerando que, nas bibliotecas com maiores acervos, o volume de entrada de exemplares e consequente tratamento de informação para compor o catálogo bibliográfico também são maiores. O quantitativo de sujeitos da amostra foi definido considerando o caráter da pesquisa que permitiu utilizar uma amostragem não probabilística, critério referenciado por Oliveira (2001) quando se trata de uma população homogênea, sendo possível a escolha de *experts* nas amostras definidas por intenção ou julgamento. Desta forma, os sujeitos da pesquisa foram três bibliotecários (S1, S2 e S3) cujo critério de seleção foi a experiência na atividade de catalogação.

Como métodos utilizados relacionam-se a entrevista semiestruturada e a atividade de análise de tarefa, aos quais foram incorporadas algumas técnicas para compor a coleta de dados. A entrevista teve com um de seus objetivos conhecer a trajetória profissional dos sujeitos da pesquisa. Foram entremeados aspectos simbólicos e afetivos em sua condução visando reunir elementos para compor também um cenário de sentimentos, percepções e construções afetivas relacionadas ao exercício profissional de cada sujeito. Nesta atividade foi incorporada a Técnica do Incidente Crítico (TIC) definido por Flanagan (1973, p.100) como “qualquer atividade humana observável que seja suficientemente completa em si mesma para permitir inferências e previsões a respeito da pessoa que executa o ato”. O incidente analisado nesta pesquisa foi relacionado a uma situação de tomada de decisão na execução da atividade de análise de assunto que o entrevistado tenha considerado relevante destacar.

A análise de tarefa realizada contemplou a execução da atividade de análise de assunto em três livros pré-selecionados com foco na atribuição de termos. A seleção dos materiais considerou apenas que os livros não pertencessem ao acervo da Universidade e que os temas fossem genéricos. Na execução desta atividade foi incorporado o preenchimento de uma ficha baseada em Coutinho e Araújo (2010) com vistas a identificar os procedimentos realizados. Após a execução da atividade foi aplicada a técnica Protocolo Verbal, que consiste na verbalização consciente dos pensamentos dos indivíduos diante da execução de uma atividade com vistas a fornecer informações de seus processos mentais a partir da exteriorização por meio da fala. Ao término da realização da tarefa proposta aplicou-se o teste AT-9.

## 4. Resultados

A apresentação dos resultados obtidos é segmentada pelos tipos de análises efetuadas conforme abaixo:

### 4.1 Análise dos procedimentos

Verificou-se que a atuação dos bibliotecários não foi uniforme (Quadros 1 a 3) variando os procedimentos executados conforme o livro analisado. Apenas a leitura do título e subtítulo foi o procedimento realizado por todos os bibliotecários em todos os livros.

| <i>Procedimento</i>  | <i>Livro1</i> | <i>Livro2</i> | <i>Livro3</i> |
|--|---------------|---------------|---------------|
| Leitura do título e subtítulo  | X             | X             | X             |
| Nome do autor  | X             | X             | X             |
| Lombada  |               |               |               |
| Área do livro  | X             | X             | X             |
| Leitura do resumo  |               |               |               |
| Leitura do índice  |               |               |               |
| Leitura do sumário   | X             | X             | X             |
| Leitura da introdução  | X             | X             | X             |
| Leitura dos capítulos  |               |               |               |
| Leitura da orelha do livro   | X             | X             |               |
| Leitura dinâmica   |               | X             |               |
| Folheada geral   |               | X             | X             |
| Ilustrações, diagramas, tabelas e seus títulos explicativos                      |               |               | X             |
| Leitura de palavras em destaque (sublinhadas, impressas em tipos diferentes etc) | X             |               |               |
| Exame das referências bibliográficas   | X             |               | X             |
| Material adicional   |               |               |               |
| Catálogo na fonte  | X             | X             | X             |
| Consulta outro bibliotecário   | X             | X             | X             |
| Outras (inseridas pelos bibliotecários)  |               |               |               |
| Consulta a outras bases de dados   | X             | X             | X             |

Quadro 1. *Procedimentos realizados para definir a escolha dos termos – S1*

| <i>Procedimento</i>  | <i>Livro1</i> | <i>Livro2</i> | <i>Livro3</i> |
|--|---------------|---------------|---------------|
| Leitura do título e subtítulo  | X             | X             | X             |
| Nome do autor  |               |               |               |
| Lombada  | X             |               |               |
| Área do livro  |               |               |               |
| Leitura do resumo  | X             | X             |               |
| Leitura do índice  | X             |               |               |
| Leitura do sumário   |               |               |               |
| Leitura da introdução  | X             | X             |               |
| Leitura dos capítulos  |               | X             |               |
| Leitura da orelha do livro   |               | X             |               |
| Leitura dinâmica   |               | X             |               |
| Folheada geral   | X             | X             | X             |
| Ilustrações, diagramas, tabelas e seus títulos explicativos                      |               |               |               |
| Leitura de palavras em destaque (sublinhadas, impressas em tipos diferentes etc) |               |               |               |
| Exame das referências bibliográficas   |               |               |               |
| Material adicional   |               |               |               |
| Catálogo na fonte  |               |               |               |
| Consulta outro bibliotecário   |               |               |               |

Quadro 2. *Procedimentos realizados para definir a escolha dos termos – S2*

| <i>Procedimento</i>  | <i>Livro1</i> | <i>Livro2</i> | <i>Livro3</i> |
|--|---------------|---------------|---------------|
| Leitura do título e subtítulo  | X             | X             | X             |
| Nome do autor  |               | X             | X             |
| Lombada  |               |               |               |
| Área do livro  | X             | X             | X             |
| Leitura do resumo  |               |               | X             |
| Leitura do índice  |               |               |               |
| Leitura do sumário   | X             | X             | X             |
| Leitura da introdução  |               |               |               |
| Leitura dos capítulos  |               |               |               |
| Leitura da orelha do livro   | X             | X             |               |
| Leitura dinâmica   |               |               |               |
| Folheada geral   | X             |               |               |
| Ilustrações, diagramas, tabelas e seus títulos explicativos                      |               |               |               |
| Leitura de palavras em destaque (sublinhadas, impressas em tipos diferentes etc) |               |               |               |
| Exame das referências bibliográficas   |               |               |               |
| Material adicional   |               |               |               |
| Catálogo na fonte  |               | X             | X             |
| Consulta outro bibliotecário   |               |               |               |
| Outras (inseridas pelos bibliotecários)  |               |               |               |
| Rede Pergamum  | X             |               |               |
| Catálogo da Universidade que publicou  |               |               | X             |

Quadro 3. *Procedimentos realizados para definir a escolha dos termos – S3*

#### 4.2 Análise dos termos atribuídos

A análise dos termos atribuídos pelos bibliotecários na tarefa executada revelou também resultados diferenciados:

- Livro 1 - houve certa homogeneidade na definição dos termos com a adoção das expressões “extensão universitária” e “ensino superior” pelos três bibliotecários;
- Livro 2 - a seleção de termos girou em torno de um eixo comum – Mudanças climáticas – apesar dos termos escolhidos pelos pesquisados terem sido variados como: Brasil - Clima (S1); Tempo (S2); Climatologia (S3);
- Livro 3 - os termos foram diversificados, não oferecendo um ponto comum que pudesse ser usado na recuperação deste livro. Foram apresentados termos como: Informática – estudo e ensino (S1), Programação de computadores – computação (S2) e Armazenamento de dados (S3).

#### 4.3 Análise simbólica

A partir das inferências sobre os incidentes críticos relatados foi possível identificar os seguintes pontos centrais na situação apresentada por cada entrevistado (Quadro 4):

|    | <i>Ponto central</i>   | <i>Aspecto crítico</i>                                      |
|----|--|---|
| S1 | Dúvida quanto ao termo a ser usado. Não existia o termo autorizado | O usuário não vai achar o livro                             |
| S2 | Divergência com o usuário  | Conflito de autoridade                                      |
| S3 | Diversidade de definições de um termo                              | O termo pode ter vários significados dependendo do contexto |

Quadro 4. *Incidentes críticos*

O uso desta técnica possibilitou identificar eventos relatados na perspectiva dos entrevistados de modo a compor a análise dos aspectos subjetivos numa perspectiva holística na qual o indivíduo não é considerado isolado de seu contexto e história. Esta condição permitiu um entrelaçamento entre fatos presentes e passados possibilitando um olhar profundo sobre o fenômeno estudado.

Em relação aos aspectos simbólicos, a consolidação das imagens características do exercício da catalogação e representativas do incidente crítico constam dos Quadros 5 e 6.

|    | <i>Imagem</i>      | <i>Explicação</i>  |
|----|--------------------|--|
| S1 | Livro aberto       | “Sempre que eu penso em uma coisa ... boa, eu penso em um livro aberto..., sempre.”  |
| S2 | Árvore             | “Porque a árvore está crescendo, dá frutos, dá folha; eu vejo como uma árvore mesmo, uma coisa produtiva.”   |
| S3 | Globo em movimento | “...porque é movimento, é movimento ... de ideias, de assunto, de saberes, de...então... movimento, uma coisa que representasse o movimento, um globo” |

Quadro 5. *Imagens representativas da atividade de catalogação*

|    | <i>Imagem</i>   | <i>Explicação</i>   |
|----|---|---|
| S1 | “Ah, barreira, eu acho que um muro. rsrs, É, uma barreira, assim bem, sabe...”  | “Ah, porque é, porque é uma barreira mesmo, sabe, eu acho que é, é a dificuldade que o usuário tem de localizar e a gente tem de disponibilizar desta forma |
| S2 | “um menino correndo, correndo e você não conseguindo pegar ele ... um menino bem levadinho, aquele que você dá uns tapinhas nele e ele..., sai correndo desesperado.” | “Aí ... mais ou menos isso, porque nós não entramos em... num diálogo né?”  |

|    |   |  |
|----|---|--|
| S3 | <i>“hum... imagem que vem para mim é muito de... antena,”</i> | <i>“A antena com essas... sempre... nunca sozinha... sempre com alguma coisa junto...”</i> |
|----|---|--|

Quadro 6. *Imagens representativas do incidente crítico*

Segundo Chevalier e Gheerbrant (2008, p. XII), os símbolos constituem o cerne da vida imaginativa, revelando segredos do inconsciente e traduzindo o esforço dos indivíduos em “decifrar e subjugar um destino que lhe escapa através das obscuridades que o rodeiam”. Esta afirmação pode ser observada nas significações que se “corporificaram” por meio da análise do comportamento de cada entrevistado apresentadas conforme demonstrado a seguir:

Para S1 a atividade de catalogação é representada como um “livro aberto”. Cirlot (1984) trata do simbolismo do livro dizendo que os chineses consideram emblemas com livros como símbolos do poder para afastar espíritos malignos; já Chevalier e Gheerbrant (2008) destacam que os livros silibinos eram consultados pelos romanos em situações excepcionais para encontrar neles as respostas divinas para suas angústias. Os mesmos autores destacam que o livro dos mortos egípcio era utilizado para implorar aos deuses uma boa travessia dos infernos e a consequente chegada ao sol eterno.

Ao representar o seu incidente crítico através da imagem de uma barreira a ser transposta, S1 cria uma metáfora que se constela com o simbolismo do livro. Uma vez que a barreira (ou o muro) é descrita por Cirlot (1984) como uma incapacidade de avançar pelo caminho pretendido, o livro, do modo como foi citado parece evocar um salvo-conduto ou proteção para afastar os “espíritos malignos” ou, ainda, como fonte de “respostas divinas para as suas angústias” e prescrição para implorar uma “passagem pelos infernos” e o acesso ao “sol eterno”. Desse modo, parecem ficar bastante claras as associações, não somente do livro como um lenitivo para a angústia diante do desconhecido e do muro com uma barreira – um elemento que tem em sua significação mais fundamental o sentido de separação – como também entre essas duas imagens (livro e barreira/muro) na forma de Nêmesis um do outro. Isso revela muito da visão de S1 sobre o processo de catalogação e seus percalços. Para S1, parece tratar-se de uma atividade tranquila, que, porém, pode tornar-se turbulenta quando a impossibilidade de acesso se configura como uma separação, uma “comunicação cortada”, frustrando que o objetivo maior do entrevistado – que é disponibilizar

para o usuário o acesso a informação da melhor forma possível – se concretize conforme seu intento.

Para S2, a catalogação é representada como uma árvore, cujas interpretações principais de seu simbolismo articulam-se em torno da ideia de um cosmo vivo em perpétua regeneração (Chevalier e Gheerbrant, 2008). Como vida inesgotável, equivale à imortalidade e esse conceito de vida sem morte se traduz em realidade absoluta, tornando-se a árvore o centro do mundo (Cirlot, 1984). Tida como um dos temas simbólicos mais ricos e difundidos, a árvore consolida o “aspecto cíclico da evolução cósmica: morte e regeneração”, significação que é referenciada por Mircea Eliade assim retratada por Chevalier e Gheerbrant (2008, p.85): “é por ser vertical, é porque cresce, pende suas folhas e torna a recuperá-las e porque, conseqüentemente, se regenera: morre e renasce inúmeras vezes.”

No imaginário de S2 este sentido justifica a sua escolha pelo símbolo da árvore ao percebê-lo como algo que “*está crescendo, dá fruto, dá folha [...] uma coisa produtiva*”. Essa imagem de morte e regeneração relacionada à catalogação relembra seus processos em que o texto analisado “morre”, se “regenerando” nos termos atribuídos para representá-lo no final do processo de análise de assunto, numa atividade que abriga o sentido de produção e transformação. Também o verticalismo da árvore, que tem suas raízes fincadas no solo e seus galhos elevados para o céu evoca-a como símbolo das relações entre o céu e a terra, interpretação vista nos mulçumanos xiitas de rito ismaelita, no qual a árvore simboliza *hakikal*, um estado de beatitude no qual o mítico, ao ultrapassar a dualidade das aparências, encontra a realidade suprema (Chevalier e Gheerbrant, 2008).

Pode-se inferir a existência de uma postura de “supremacia” no posicionamento que S2 adota em relação ao resultado de seu trabalho. Encarnado como uma “verdade absoluta”, que é refletido no sentido da árvore como “centro do mundo”, essa noção indica que qualquer questionamento ou interferência externa não será vista com naturalidade. Essa constatação aparece explicitada na justificativa da evocação da imagem do incidente crítico na qual o questionamento a uma posição adotada por S2 e ausência de concordância com seu ponto de vista são considerados como fator de impasse. Isso se dá em virtude de chocar-se com o sentido de “eixo do mundo”, “pilar central”, simbolizado pela árvore, que constitui o significado da ativi-

dade de catalogação para o entrevistado e em torno do qual circundam seus comportamentos

O reforço a esta interpretação pode ser associado na imagem representativa do incidente crítico no qual o confronto de opiniões – que gerou um conflito de autoridade e um questionamento da “verdade suprema”, que é a decisão tomada – configura-se como um ato de rebeldia e não aceitação de uma realidade adversa. Segundo Cirlot (1984, p.378), a imagem do menino representa “o produto da *coniunctio* entre inconsciente e o consciente, símbolo alquímico de união de substâncias desiguais. A corrida remete ao sentido de fuga, o que pode caracterizar a resistência em proceder a união das “ideias opostas”, desiguais, fruto de pressão exercida pelo contexto, ao qual S2 demonstra resistência e do qual gostaria de fugir. A imagem do menino, nas colocações de Chevalier e Gheerbrant (2008, p.302), remete à “vitória sobre a complexidade e a ansiedade e a conquista da paz interior e da autoconfiança” e sua fuga, na representação do incidente crítico, busca preservar esse estado de completude que está sendo ameaçado.

Para S3, a atividade de catalogação foi representada pela imagem de um globo em movimento. Este símbolo evoca em seu significado o sentido de poder, “o domínio ou o território sobre o qual se estende a autoridade do soberano e o caráter *totalitário* dessa autoridade” (Chevalier e Gheerbrant, 2008, p.472). Outra associação importante desta imagem relaciona-se a sua forma esférica que, corresponde à perfeição e felicidade, pois a ausência de arestas (quinas) remete à falta de inconvenientes, estorvos e contrariedades. Articula-se também ao símbolo a ideia de esfericidade relacionada a forma circular do globo que, inspirado nos emblemas chineses, representa atividade, símbolo do princípio masculino (*Yang*) representado por um círculo branco (Cirlot, 1984).

S3 justifica a escolha dessa imagem associando a ela o conceito de movimento (“*de ideias, de assunto, de saberes*”...), termo que reforça o simbolismo da perfeição como citado por Chevalier e Gheerbrant (2008, p.389): “A noção de esfera e de movimento orbicular é dominante e exprime a perfeição. Se um ser for concebido como perfeito, ele será simbolicamente imaginado como uma esfera”.

A representação da atividade de classificação por S3 como um globo em movimento sugere que o pesquisado vê seu trabalho como uma atividade “perfeita”, sem arestas, o que possibilita compreender como a antena – imagem associada ao incidente crítico – pode ser responsável por desarticular esse mundo considerado

pleno. A simbolização da antena neste contexto está associada ao inconstante, ao múltiplo, visto que pode absorver vários significados dependendo do cenário. Essa faceta que a antena representou constitui uma ameaça ao universo perfeito e controlado de S3 porque, apesar deste representar movimento, esse movimento é orbital e gira em torno de si mesmo e a dinamicidade representada pela antena não é linear, trazendo em sua representação o sentido do “imperfeito” e imprevisível.

Reunindo as associações feitas por S1, S2 e S3 pode-se perceber uma conexão simbólica estreita entre essas metáforas. Quando S2 e S3 evocam como representação da atividade de catalogação, respectivamente, uma árvore e o globo, verifica-se como o conceito de plenitude – traduzido na “verdade absoluta” e na “perfeição” – encontra repercussão no imaginário de ambos entrevistados, sugerindo que o produto do trabalho realizado é completo em si mesmo. A mesma similaridade é vista nas imagens evocadas nos incidentes críticos caracterizadas pela quebra dessa hegemonia expressa tanto na figura de outra opinião, quanto na dinamicidade de significados que um termo pode assumir. As duas situações representam um conflito que ameaça esse conceito de “supremacia” que permeia a visão dos entrevistados sobre a realização de suas atividades. O que ameaça a ambos é algo que conteste ou abale suas certezas.

Cabe destacar que o livro, metáfora atribuída à catalogação por S1, tem, dentre seus significados, o relacionado ao “Livro da Vida do Apocalipse” onde este se identifica com a Árvore da Vida sendo “as folhas da árvore como os caracteres do livro” representando “a totalidade dos decretos divinos” (Chevalier e Gheerbrant, 2008, p.555). Desta forma, as associações entre os diversos sentidos atribuídos aos símbolos permitiram construir uma rede de significados (Figura 1) cuja sutura se dá pela interpretação da catalogação como uma atividade que tem em si um aspecto de finitude, sendo seu resultado tido como um “produto fechado”, total e completo que, apesar de ser construído sob os olhares e caminhos diferentes por cada um dos entrevistados, para cada um deles o resultado de seus trabalhos adquire um caráter de verdade “incontestável”.

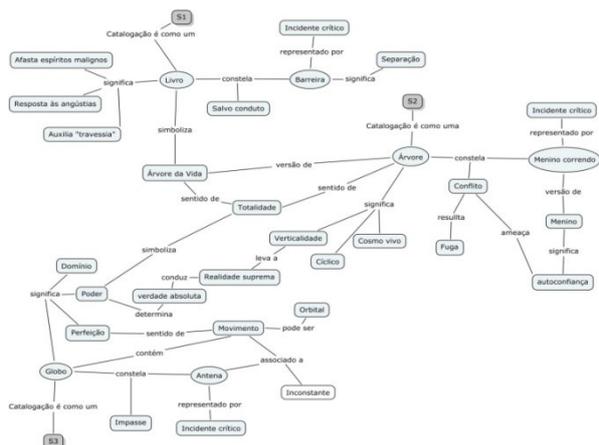


Figura 1. Rede interconectada de símbolos e significados

A doutrina de *Mohyiddin ibn Arabi*, de acordo com Cirlot (1984), propõe o universo como uma condensação do imenso livro escrito pela pena divina e que o universo transcendente do livro descende para tomar a forma do universo manifesto onde o homem vive. Esse simbolismo, destaca o autor, é o mesmo que os Rosacruzes chamavam de *Liber Mundi* e o Apocalipse de *Liber Vitae*. Já Chevalier e Gheerbrant (2008) fazem referência ao livro (*Liber Mundi*) não somente como uma simbolização da relação/correspondência, via mensagem divina, do macrocosmo com o microcosmo, como também faz uma referência a versões da busca do Graal onde o livro seria uma representação do próprio Graal cuja busca seria a “procura da palavra perdida”, da sabedoria tornada acessível ao comum dos mortais. Neste caso, a conexão simbólica com a atividade de catalogação parece inquestionável. Os autores destacam ainda que enquanto o livro fechado representa a matéria virgem, o “livro aberto” representa a matéria fecundada. Desse modo, enquanto o livro permanece fechado, o mesmo conserva seus segredos, já quando é aberto seu conteúdo é tomado por quem o investiga. Não seria essa a essência da atividade de catalogação? Tornar abertos os livros para permitir que os investigadores tenham acesso às suas mensagens? Ainda fazendo referência aos autores retro mencionados, cabe à catalogação repetir o trabalho dos alquimistas que, através de suas operações sucessivas, abrem os selos e as fitas que fecham o *Grande Livro da Natureza* e permitem que a matéria prima seja extraída da mina e trabalhada. Para que isso ocorra é necessário que, dos conteúdos intrínsecos dos livros, seja retirada sua essência de modo que possam ser representados fielmente.

É possível verificar uma analogia no sentido de verticalidade da árvore que coloca em comunicação os três níveis do cosmo: o subterrâneo, a superfície e as alturas (Chevalier e Gheerbrant, 2008). No processo de catalogação, esses níveis podem ser refletidos, respectivamente, no texto do livro, nos termos atribuídos e no usuário da informação. Por meio das raízes se acessa as profundezas onde se enterram as essências; na superfície da terra o tronco, representação visível da estrutura, vai levar aos galhos e folhas a seiva (informação) necessária para saciar e atender a necessidade daqueles que estão no final do processo. “Ela estabelece, assim, uma relação entre o mundo ctoniano e o mundo uraniano” e reúne todos os elementos (Chevalier e Gheerbrant, 2008, p.84).

Finalizando esse ciclo verifica-se o significado contido no globo que, por sua forma esférica, corresponde à plenitude da realização. Seu aspecto circular é associado às propriedades simbólicas da perfeição e homogeneidade, atributos que se espera que a atividade de catalogação compreenda que é a perfeita representação do livro e uma codificação que traduza de maneira homogênea e fiel seu conteúdo. Na tradição grega, em especial em Parmênides e nos textos órficos, as esferas são utilizadas para representar dois mundos: o mundo terrestre e o Outro-Mundo, sendo a morte responsável por fazer passar de uma esfera a outra. Também nesse sentido a catalogação pode ser vista como representativa da passagem de um mundo do conteúdo para um mundo de conceitos, no qual o primeiro é “abandonado/eliminado” após a inserção no novo mundo que passará a representar o mundo desejável.

Importante destacar que essa análise não se esgota na mera interpretação dos significados latentes, pois as palavras, segundo Jean Chevalier, são incapazes de expressar todo o valor de um símbolo (Chevalier e Gheerbrant, 2008, p. XIII). Cabe nessas reflexões o cuidado de não encerrar em estreitos limites todas as dimensões descortinadas pela interpretação simbólica, mas ampliá-la em outras metodologias que possam expandir seus significados.

Essa proposta foi complementada nesta pesquisa com o uso do AT-9 por meio do qual se buscou compreender o que permeia as ações dos sujeitos em situações de angústia (Cardoso, 2005) representada neste estudo pela decisão do termo adequado na análise de assunto.

#### 4.4 Consolidação das análises

O uso do Teste Arquetípico de Nove Elementos (AT-9) e das outras técnicas complementares

possibilitou explorar os aspectos subjetivos utilizando-se as dimensões simbólico-afetivas. A compilação dos dados obtidos foi realizada sob o crivo das estruturas antropológicas do imaginário de Gilbert Durand (1997), consolidada na metodologia desenvolvida por Yves Durand (1988) com vistas a identificar indícios que revelassem como se deu o processo informacional, subjetivo e afetivo envolvido na atividade executada.

Como resultado das análises efetuadas foi possível identificar os microuniversos míticos de cada indivíduo e suas relações com os comportamentos informacionais adotados, o que possibilitou compreender o porquê dos resultados diferenciados na execução da atividade de análise de assunto.

O microuniverso mítico de S1, denominado Místico Integrado, o caracteriza como um indivíduo que procura disfuncionalizar a ameaça trazendo-a para seu universo controlado e submetendo-a a sua experiência de forma a neutralizar a angústia que ela encerra. Esse perfil direciona suas atitudes no processo de tratamento informacional e os termos selecionados na atividade de análise de assunto fecham um ciclo “perfeito”, no entender do pesquisado, que é atender ao usuário (Figura 2).

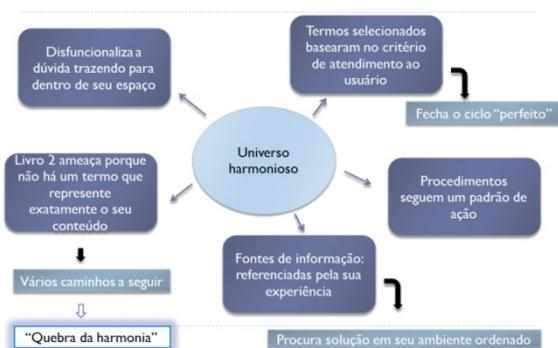


Figura 2. Microuniverso de S1 e ações decorrentes

O Microuniverso mítico de S2 denominado Sintético Simbólico de forma Diacrônica de Evolução Cíclica - no qual o movimento cíclico é traduzido pelo “eterno retorno” - permitiu vislumbrar que o que caracteriza o pesquisado é sua postura fechada para decidir. Esta atitude torna seu processo decisório bloqueado a interferências externas, sendo esse modo de se proteger em si mesmo sua forma de lidar com os desafios da decisão (Figura 3). Os termos selecionados na atividade de análise de assunto seguiram o perfil de S2 de procurar em um ambiente interno – ou seja, no próprio livro – a resposta à atividade, mantendo nessa postura o ambiente hermético.



Figura 3. Microuniverso de S2 e ações decorrentes

O microuniverso mítico de S3 - Sintético Existencial Diacrônico - o caracteriza como um indivíduo que busca lidar com a decisão sob duas perspectivas: volta-se para si para encontrar respostas e, em seguida, parte para uma postura de enfrentamento utilizando os instrumentos de que dispõe para enfrentar a dúvida e decidir (FIG.4). Na atribuição de termos, o pesquisado trouxe para si o processo de análise, mas, na hora de decidir, utilizou um instrumento externo capaz de referendar de forma pragmática sua decisão.



Figura 4. Microuniverso de S3 e ações decorrentes

## 5. Conclusão

As técnicas e métodos adotados permitiram concluir que os aspectos subjetivos, mesmo em um processo que tem uma metodologia estruturada - como é o caso da análise de assunto – que apresenta termos de vocabulários controlados, normatizações de procedimentos e fontes de informação padronizadas – são responsáveis por resultados diferenciados. Isto ocorre porque o processo de análise, que é uma atividade de tomada de decisão, não carrega em si apenas aspectos racionais, mas é permeada por toda uma história de vida, experiências profissionais, preferências, estruturações mentais e perfis psicológicos. Esses aspectos fazem desse processo o resultado de toda uma confluência de

fatores que não se situam apenas em nível consciente, estando influenciado por circunstâncias que não estão palpáveis ou explícitas no comportamento do indivíduo.

Foi possível compreender a complexidade do processo de tratamento da informação e apresentar o entendimento da necessidade de se buscar o estabelecimento de diretrizes mais precisas e de parâmetros mais eficazes para orientar o processo de indexação. Acredita-se que o reflexo de um processo de tratamento informacional balizado em critérios mais efetivos e que observem a subjetividade como fator intrínseco à atividade irá culminar em uma recuperação de informação mais precisa contribuindo para que o usuário tenha atendida sua demanda de forma satisfatória.

Cabe ressaltar que a recuperação da informação, conforme destacado por Fernalda (2003), implica em operar seletivamente um estoque de informação, o que envolve processos cognitivos por parte dos usuários, bem como a capacidade de abstração, apreensão e representação do significado de forma contextual por parte dos indexadores. Essas particularidades, segundo o autor, dificilmente podem ser formalizadas por meio de um algoritmo, devendo, portanto, ser observadas sob outras perspectivas visando resultados mais efetivos no processo.

## Referências

- Araujo, E. P. O. (2013). *Tomada de decisão organizacional e subjetividade: análise das dimensões simbólico-afetivas no uso da informação em processos decisórios*. 2013. 162 p. Dissertação (Mestrado em Ciência da Informação) - Escola de Ciência da Informação, Universidade Federal de Minas Gerais. Belo Horizonte..
- Cemin, A. B. et al. (2001) Gênero e imaginário. *Revista eletrônica do Centro de Estudos do Imaginário*. Ano I, n.3, out-dez. Recuperado 18-5-2014 de <http://www.cei.unir.br/artigo32.html>.
- Chevalier, J.; Gheerbrant, A. (2008). *Dicionário de símbolos*. 22 ed. Rio de Janeiro: José Olympio,
- Cirlot, J. E. (1984) *Dicionário de símbolos*. São Paulo: Editora Moraes,
- Coutinho. L. F.; Araujo, C. A. A. (2010) *A indexação nas áreas do conhecimento: uma comparação das áreas de ciências exatas e da terra, das ciências humanas e da linguística, letras e artes*. Rio de Janeiro. 2010 Anais. Rio de Janeiro. XI ENANCIB.
- Durand G. (1997). *As estruturas antropológicas do imaginário*. Martins Fontes. São Paulo.
- Durand, Y. (1988) *L'exploration de L'imaginaire: Introduction à la modélisation des Univers Mythiques*. L'espace bleu. Paris.
- Estrada, A. A.. (2002) O teste AT-9 na escola: considerações preliminares acerca do universo da angústia. *Educere. Revista da Educação*. V.2, n.1, p. 25-38, jan/jun
- Flanagan, J. C. (1973) A técnica do incidente crítico. *Arquivos Brasileiros de Psicologia Aplicada*, v.25, n.2, abr/jun
- Fernalda, E. (2003) *Recuperação de informação: análise sobre a contribuição da Ciência da Computação para a Ciência da Informação*. Tese (Doutorado em Ciências da Comunicação). Escola de Comunicação e Artes. Universidade de São Paulo. São Paulo.
- Fujita, M. S. L. (2003). *A leitura documentária do indexador: aspectos cognitivos e linguísticos influentes na formação do leitor profissional*. Tese (Livre-docência em Análise Documentária e Linguagens Documentárias Alfabéticas). Universidade Estadual Paulista. Marília.
- Fujita, M. S. L et al. (2010). O contexto da indexação para a catalogação de livros em abordagem sociocognitiva. *Brazilian Journal of Information Science*, Marília (SP), v.4, n.2, p.22-40, jul/dez 2010. Recuperado em 08-5-2014 de <http://www2.marilia.unesp.br/revista/index.php/bjis/index>.
- Garcia. J. C. R. (2007) Recuperação da informação. In: *DataGramaZero - Revista de Ciência da Informação*, v.8 n.6. Recuperado em 20-5-2014 de [http://www.dgz.org.br/dez07/F\\_I\\_com.htm](http://www.dgz.org.br/dez07/F_I_com.htm).
- Harris S, J. L. (1970) *Subject analysis: computer implications of rigorous definition*. Metuchen, NJ: Scarecrow Press,
- Kobashi, N. Y. (1994) *Elaboração de informações documentárias: em busca de uma metodologia*. Tese (Doutorado em Ciências da Comunicação). Escola de Comunicações e Arte, Universidade de São Paulo. São Paulo.
- Lancaster, F.W. (2004) *Indexação e resumos: teoria e prática*. 2 ed. Briquet de Lemos. Brasília.
- Naves, M. M. L. (1996) Análise de assunto: concepções. *Revista Biblioteconomia*. Brasília, v.20, n.2, p.215-226, jul/dez..
- Oliveira. T. M. V. (2001) Amostragem não probabilística: adequação de situações para uso e limitações de amostra por conveniência, julgamento e quotas. *Administração online*, v.2, n.3, jul/ago/set. Recuperado em 15-5-2014 de [http://www.fecap.br/adm\\_online/art23/tania2.htm](http://www.fecap.br/adm_online/art23/tania2.htm)
- Oliveira, G. P.; MAIA, L. S. L. (2008) Estudo do universo imaginário de professores de matemática: uma análise a partir da teoria de Gilbert Durand.. Recuperado em 20-5-2014 em: <http://31reuniao.anped.org.br/1trabalho/GT19-4798-Int.pdf>.
- Paula, C. P. A. (2011) *Dimensões simbólicas e afetivas do uso da informação: uma análise das comunicações entre professores do departamento de psicologia de uma instituição de ensino superior pública brasileira*. Brasília. 2011. Anais. Brasília. XII ENANCIB..
- Paula, C. P. A. (2012) *Proposta de metodologia para a investigação do comportamento de busca informacional e do processo de tomada de decisão dos líderes nas organizações: introduzindo uma abordagem clínica na informação*. Rio de Janeiro. 2012. Anais. Rio de Janeiro. XIII ENANCIB.
- Silva, M.R. et al. (2004). A prática de indexação: análise da evolução de tendências teóricas e metodológicas. *Transinformação*, v.16, n.2, p-133-161, mai/ago. Recuperado em 09-5-2014 de <http://revistas.puc-campinas.edu.br/transinfo/viewarticle.php?id=65>.
- UNISIST. (1981) Princípios de indexação. *Revista da Escola de Biblioteconomia da UFMG*, Belo Horizonte, v.10, n.1, p. 83-94, mar.

---

# Metadatos en noticias: un análisis internacional para la representación de contenidos en periódicos

*Metadata in news: an international review for representing the content of newspapers*

---

**María-José Baños-Moreno (1), Eduardo R. Felipe (2), Juan Antonio Pastor-Sánchez (3), Rodrigo Martínez-Béjar (4), Gercina Lima (5)**

(1) Universidad de Murcia, Campus de Espinardo, C.P. 30100, mbm41963@um.es,

(2) Universidad Federal de Minas Gerais, erfelipe@hotmail.com,

(3) Universidad de Murcia, Campus de Espinardo, C.P. 30100, pastor@um.es,

(4) rodrigo@um.es,

(5) Universidad Federal de Minas Gerais, glima@eci.ufmg.br

## Resumen

Los periódicos trabajan con un gran volumen de información que necesita ser descrita adecuadamente. Para ello, las etiquetas "title", "keywords" y "description" son muy utilizadas en el código fuente de las noticias online. Sin embargo, estas no resultan suficientemente descriptivas. Así, surgen estándares de metadatos, con el fin de facilitar la interoperabilidad y profundizar en la descripción. Actualmente, las etiquetas HTML y diversos estándares conviven en el sector periodístico, con diversos grados de implantación. Se analiza el código fuente de diarios internacionales de información general y se lleva a cabo una profunda revisión bibliográfica sobre estándares de metadatos. El propósito es conocer qué estándares existen, y evaluar su uso en los códigos fuente de una muestra de periódicos. Para ello se identifican los metadatos de contenido semántico de los códigos fuente. Además se desarrolla el software MetadadosHTML. Como conclusiones destacan la gran distancia entre los estándares recogidos en la bibliografía y los mostrados en los código fuente. En el primer caso, los más referenciados son los formatos NewsML y NITF, implementados por algunos medios y agencias de prensa, al menos a nivel interno. Por el contrario, en el código fuente los más habituales son schema.org y dos esquemas para mostrar información en redes sociales, Open Graph Protocol (usado por Facebook) y Twitter Cards. Esto, evidencia la convivencia de diversos estándares de metadatos en el ámbito de los medios de comunicación y pone de relieve la falta de uniformidad en su uso. Para alcanzar el ideal de interoperabilidad de los contenidos, es preciso utilizar tecnologías de la Web Semántica. En este sentido, se debería tender a definir ontologías o vocabularios RDF para las diferentes propuestas analizadas.

**Palabras clave:** Periódicos, metadatos, schema.org, Open Graph Protocol, NITF, NewsML

## Abstract

Newspapers use a large volume of information that must be described adequately. To do this, the tags "title", "keywords" and "description" are widely used in the source code of online-news items. However, these are not descriptive enough for the news items. So, metadata standards are created in order to facilitate the interoperability and make a deeper description of them. Currently, HTML tags and several standards live together in the newspaper industry, with different degrees of implementation. In this paper, we analyze the source code of international general-information newspapers. Moreover, we conducted an in-depth literature review on metadata standards. The aim is to analyze what standards exist and how they are used in a sample of newspapers' source codes. For that, we identify the semantic metadata in the source codes and design the software MetadadosHTML. As conclusions, the great distance between metadata standards identified in the literature review and those in the source codes is clearly shown. In the former, the most cited metadata are NewsML and NITF formats, implemented at least in an internal level by some media and press agencies. By the contrary, schema.org and two social media schemas (Open Graph Protocol for Facebook and Twitter Cards) are the most common in the latter one. The coexistence of different metadata standards in the media sector is exhibited, stressing the lack of uniformity in their use. To achieve the ideal of interoperability between contents, the use of Web Semantic technologies is needed. In this regard, to define ontologies and RDF vocabularies for the different proposals analyzed should be a tendency.

**Keywords:** Newspapers, metadata, schema.org, Open Graph Protocol, NITF, NewsML

## Agradecimientos

Este trabajo ha sido posible gracias a la financiación de la Universidad de Murcia a través del programa de becas predoctorales con resolución R-406/2011. Agradecimientos ao apoio das Agências de fomento FAPEMIG e CNPq/Brasil

## 1. Introducción

Internet, que ha propiciado un profundo cambio cultural, ha puesto al alcance de los ciudadanos el mayor volumen de información de todos los tiempos, aunque gran parte de forma poco estructurada. Esto se traduce en la dificultad para descubrir conocimiento relevante (Paepen et al., 2002; Pereira y Baptista., 2003a; Castells et al., 2006; Silva et al., 2008; Kallipolitis et al., 2012; Nies et al., 2012; Díaz Nosty, 2013; Mannens et al., 2013), afectando de forma considerable a la prensa escrita, el dominio más rico en contenidos en la Web (Saleh y Al-Khalifa, 2009, Abadal et al., 2014). Este crecimiento y disponibilidad contribuye a la necesidad de agrupar la información a nivel semántico (Paepen et al., 2002; Pereira y Baptista, 2003a; Pereira y Baptista, 2004). Buena parte de los métodos de recuperación de información, fundamentados en algoritmos de localización de términos o cálculo de frecuencias, entre otros, son útiles, pero insuficientes para lograr un resultado relevante y preciso. Los metadatos, un conjunto común de elementos que representan un documento (Yaginuma et al., 2003a; Shoval et al., 2008) pueden suplir las carencias de estos métodos. Los metadatos pueden concebirse como una forma efectiva de gestionar, de forma estandarizada, enormes cantidades de información (Baptista y Machado, 2001; Paepen et al., 2002; Martínez-Fernández et al., 2004; Wong et al., 2010).

Si bien el análisis e identificación de los elementos formales de una noticia están resueltos, el empuje de la Web Semántica en el ámbito periodístico vendría del mercado del contenido de los artículos. Desde principios de siglo, las principales líneas de investigación para este propósito se centran en la indización humana y automática, procesamiento de lenguaje natural (PLN), métodos estadísticos, etc. Sin embargo, los problemas de interoperabilidad y la falta de un estándar de representación comúnmente aceptado (Castells et al., 2006; Troncy, 2008) persisten en tanto no se afianzan las técnicas y herramientas de anotación semántica. La utilización de metadatos de presentación del contenido es un paso intermedio.

Con la llegada de los medios de comunicación a Internet, muchas noticias se publican en línea y en diferentes formatos (Paepen et al., 2002;

Agarwal et al., 2012). Los periódicos se transforman aprovechando las posibilidades que las nuevas tecnologías les brindan. Por esto, no es extraño encontrar proyectos que aplican tecnologías de la Web Semántica en los procesos de producción y difusión de información de actualidad, con experiencias basadas en una cierta automatización asistida por personal del periódico, como Rubio Lacoba (2012), Fernández et al. (2006, 2007a, 2007b, 2010 y 2012) ó Castells et al. (2006).

Una de las bases de la Web Semántica es la reutilización de productos previos. En este sentido, es posible aprovechar aquellos metadatos definidos por el propio medio. En función de las necesidades y posibilidades de los periódicos, existen diferentes opciones para expresar estos elementos. Además, hay que tener en cuenta los estándares propios del estado del arte, tales como NITF o NewsML. Estos son utilizados por agencias y periódicos para hacer accesibles sus contenidos y pueden ser fuentes para una ontología de dominio (Fernández et al., 2007; 2010). De hecho, este trabajo se enmarca en un proyecto de modelado de una ontología de dominio en política y economía. La ventaja de los metadatos es su potencialidad para la recuperación de información utilizando los principios de interoperabilidad semántica.

El objetivo de este trabajo es conocer qué estándares de metadatos se utilizan para representar el contenido de una noticia, tanto en periódicos como en la bibliografía especializada. En este sentido y dentro del contexto de los medios de comunicación escrita, se entiende noticia como un texto que relata objetivamente un hecho (noticioso) susceptible de ser de interés público, por ser algo novedoso, destacable y/o actual. La noticia es el género informativo por excelencia y el de uso más extendido (García Gutiérrez, 2014). Para ello, se ha realizado una búsqueda en la literatura científica sobre este tema, al tiempo que se ha utilizado una muestra internacional de periódicos de información general para el análisis del código fuente y conocer los metadatos utilizados.

El trabajo se organiza de la siguiente manera: primero se muestra qué metadatos para prensa son referenciados en la bibliografía; a continuación se describe la metodología para la selección de diarios de información general, la toma de datos y la comparación entre los metadatos identificados en los trabajos previos y los presentes en los códigos fuente de noticias de diarios de la muestra; después se plantean algunos de los resultados obtenidos; finalmente se exponen conclusiones y propuestas de trabajos

futuros. Entre las conclusiones, destaca la gran distancia entre los estándares recogidos en la bibliografía y los mostrados el código fuente de los periódicos. En el primer caso, los más referenciados son los formatos NewsML y NITF, implementados por algunos medios y agencias de prensa a nivel interno. En proyectos desarrollados en el ámbito académico para la gestión de noticias como Omnipaper, ePaper o News, también son frecuentes. Por el contrario, en el código fuente los metadatos más habituales son Schema.org y los esquemas para visualizar información en redes sociales, *Open Graph Protocol* (usado por Facebook) y *Twitter Cards*. Esto, evidencia, por un lado, la convivencia de diversos estándares de metadatos en el ámbito de los medios de comunicación y pone de relieve la falta de uniformidad en el uso de estos. Cabe preguntarse el porqué del salto entre lo referenciado y lo expuesto en los códigos fuente. Es posible que se deba a ciertas reticencias a una exposición pública de los mecanismos de representación de noticias ante un potencial uso por otros agentes, o simplemente a un modelo de gestión que distingue claramente entre lo interno y lo externo.

## 2. Metadatos en prensa

Los metadatos son el conjunto de atributos o elementos que representan un recurso, facilitando su identificación y distinguiendo documentos relevantes de los no relevantes (Yaginuma et al., 2003b; Hillmann, 2005; Abbud Grácio y Fadel, 2010, p. 10-11; Kallipolitis et al., 2012; Asociación Española de Normalización y Certificación, UNE-ISO/TR 23081-3:2012). Más específicamente:

“Los metadatos son descripciones estructuradas y codificadas que describen características y propiedades de objetos y recursos para facilitar su localización, recuperación, valoración, administración, persistencia e interoperabilidad” (Pastor-Sánchez, 2013, p. 22).

Centrándonos en los que tienen por objeto la representación del contenido de un documento, podemos definirlos como el “conjunto de propiedades asociadas a la faceta descriptiva del contenido de una noticia, esto es, datos que se pueden deducir desde el consumo (lectura, escucha o visionado) de la noticias (IPTC, 2009, pp. 20). Se excluyen, por tanto, aquellos datos formales de la producción o difusión de la noticia, como autor, medio, fecha de publicación [que no responden] al análisis estructural de contenido que nos ocupa, pues pertenecen a otro plano de descripción” (García-Gutiérrez, 2014). En agencias y medios de comunicación los metadatos semánticos se han ido añadiendo

manualmente para, en las búsquedas, filtrar mediante palabras clave y categorías (Fernández et al., 2006). Además del código fuente, estos metadatos pueden encontrarse en un documento independiente (al que se referencia), o bien almacenados en bases de datos externas. Este meta-contenido, inicialmente se refleja en su código fuente mediante las etiquetas <meta> “title”, “keywords” y “description” (1) (Alesso y Smith, 2004, pp. 60):

- “Title” provee información general sobre el documento, es usada para indexar la página basándose en palabras clave del título
- “Description” es el resumen de la página web que el motor de búsqueda mostrará en la página de resultados
- “Keywords” permite listar explícitamente las palabras clave que se asocian con la página web, es usada por los motores de búsqueda para la indexación. Desde 2009, esta etiqueta dejó de ser utilizada por Google (Cutts, 2009) siendo reemplazada por <news\_keywords> para las noticias (Galfi, 2012; McGee, 2012).

Con el tiempo fue evidente la necesidad de definir conjuntos de elementos comunes (Yaginuma et al., 2003a; Pereira y Baptista, 2004) que facilitaran el intercambio de información a nivel semántico, profundizando en la descripción de los recursos. Así, surgen estándares para metadatos como *Dublin Core Metadata Initiative* (DCMI) y otras opciones más específicas para las necesidades concretas de un área de conocimiento, una tipología documental, una organización, etc. En el ámbito periodístico existen diversas posibilidades. En este sentido, las entidades más relevantes son la *Newspaper Association of America* (en adelante NAA) y la *International Press Telecommunications Council* (en adelante IPTC), un consorcio internacional de agencias de noticias, editores y distribuidores de periódicos (Pellegrini, 2012).

La forma de indicar las propiedades de los recursos, en este caso noticias, se basa en una idea similar a las sentencias RDF (*Resource Description Framework*) en las que los recursos son descritos a partir de tripletas compuestas por un sujeto, un predicado y un objeto (Pastor Sánchez, 2013, p. 56). Con los estándares de metadatos, aparece un elemento intermedio que busca sistematizar y ahondar en las posibilidades de descripción, como se muestra a continuación:

| <i>Sujeto</i> | <i>Predicado</i>                  | <i>Objeto</i>              |
|---------------|-----------------------------------|----------------------------|
| Noticia       | Elemento de meta-dato (propiedad) | Valor que toma el elemento |
| Noticia       | Keywords                          | Keyword1, Keyword2, etc.   |

↓

| <i>Sujeto</i> | <i>Esquema metadatos</i> | <i>Elemento metadato</i> | <i>Valor</i>             |
|---------------|--------------------------|--------------------------|--------------------------|
| Noticia       | p.e. Dublin Core         | Keywords                 | Keyword1, Keyword2, etc. |

Tabla 1. Formas de describir una noticia en el código fuente

Se indican, por orden alfabético, los metadatos para representar noticias en medios de comunicación de acuerdo con la bibliografía. En el anexo se muestra las referencias que citan estas especificaciones:

ANPA 89-3 (The ANPA Wire Service Guidelines for 1200 Baud transmission). Estas guías, también conocidas como ANPA 1312 y NAA 89-3, fueron desarrolladas al amparo de NAA como formato para la codificación y transmisión de información en los servicios de noticias y los métodos por los cuáles los periódicos podían utilizarlos. Su última actualización es de 1989.

Dublin Core (DC) es un simple pero efectivo conjunto de metadatos para la descripción de un amplio rango de recursos online. Este estándar es creado y mantenido por un grupo internacional e interdisciplinar de profesionales de diversas áreas. DC trabaja a dos niveles: simple, con 15 elementos básicos de descripción; y cualificado, que incluye otros tres elementos y una serie de cualificadores. Algunas iniciativas se han centrado en la adaptación de DC a sectores concretos como DC-Lib, para bibliotecas. Allen et al. (2007), Fernández et al., (2006), Castells et al. (2006) o Yaginuma et al. (2003) son sólo algunos de los trabajos que utilizan este esquema, combinado con otros como NITF, para proyectos como NEWS, Neptuno u Omnipaper, por ejemplo.

Information Interchange Model (IMM) fue definido por IPTC para el intercambio de contenido binario de noticias, principalmente fotografías (Fernández et al, 2007a).

IPTC 7901 fue la primera recomendación de la IPTC para el intercambio de mensajes sobre noticias entre periódicos y agencias de noticias. Se desarrolló con el propósito de facilitar la comunicación, superando la rigidez del ANPA 89-3, en la que se inspira. Fue actualizado por última vez en 1995. Según IPTC todavía es el for-

mato más utilizado por las agencias para enviar artículos de texto a clientes junto con ANPA 1312 (en América).

Los microdatos son frameworks para introducir contenido semántico en el código HTML, un grupo de pares nombre-valor (Hickson, 2013) donde se identifica qué esquema se utiliza, qué recurso se describe (itemscope itemtype) y sus características (itemprop). Entre los referenciados por la bibliografía destacan: rNews, un modelo compuesto de clases conceptuales con un nivel suficiente de expresividad semántica en términos de vocabulario y tipos de datos que cubren los atributos más importantes de la noticia para su marcado semántico (Pellegrini, 2012; Heravi y McGinnis, 2013); vCard es un modelo para el intercambio de información de datos de identificación y localización de una persona o empresa, tales como nombre, apellidos, domicilio o email. La última versión de este modelo, RFC6350 se ha adaptado a la Web Semántica mediante su representación en una ontología por el W3C (Iannella y McKinney, 2013).

Microformatos “son pequeños patrones en HTML para representar cosas habitualmente publicadas como personas, eventos, posts de blogs, revisiones y etiquetas en la web” (<http://microformats.org/wiki/introduction>). Estos patrones, embebidos en el código HTML, permiten marcar semánticamente cierto contenido de la web. En esencia se utiliza el atributo “class” para identificar vocabularios, clases y propiedades (Pastor-Sánchez et al., 2013). Existen varios tipos, dependiendo de la información a describir. En la literatura se hace mención de hNews, un microformato proporcionado por la agencia de noticias Associated Press (Pellegrini, 2012) que establece un número de campos para describir completamente el trabajo periodístico. En la última modificación de este estándar también participaron autores de la Media Standards Trust, una organización independiente para fomentar la calidad, transparencia y rendición de cuentas en la era digital. Para información más general es necesario recurrir a un microformato más general, hEntry, que sustituye a hAtom.

NewsForm es presentado por Mueller (2000) como un nuevo formato para representar información sobre eventos en noticias, pues NITF y XMLNews eran demasiado generales. Este formato es compatible con los dos anteriores.

News Industry Text Format (NITF): desarrollado por la IPTC y la NAA (Mueller, 2000), es un formato estándar abierto y público, en XML, que define la estructura de artículos de noticias, facilitando la reutilización de información. Tam-

bién permite la anotación de entidades que aparezcan en el contenido de los artículos (inline metadata). Entre otros, según IPTC ([2014]) es utilizado por The Associated Press, Dow Jones, The New York Times y Deutsche Presse-Agentur (2). Es el primer estándar de la nueva generación tras ANPA 89-3, IPTC 7901 e IIM (Fernández et al., 2007a). La última versión fue liberada en junio de 2012.

NewsML Architecture (NAR): también diseñado por la IPTC, es un estándar de intercambio para noticias en general utilizado para representar noticias como paquetes multimedia (García et al., 2006). NewsML apoya la expresión de metadatos e información de noticias electrónicas y las relaciones entre datos de la noticia (Kodama et al., 2008), facilitando el reaprovechamiento de contenidos descritos a partir de metadatos. La última versión desarrollada es NewsML-G2 2.19. Se complementa con NITF, ya que presentan diferentes opciones y posibilidades. Este formato es utilizado, por ejemplo, por Reuters (Tenenboim et al., 2008) y en el proyecto ePaper como una de las opciones, junto a RSS. Entre otros NewsML ofrece soporte para el contenido de noticias en cualquier formato y metadatos para la descripción de una noticia como un todo. EventsML, diseñado por IPTC es un estándar para transmitir información sobre eventos en la industria informativa (Biyun et al., 2009). Actualmente se integra en NewsML.

News Value Markup Language (NVML) es un lenguaje desarrollado por Kodama et al. (2008) para la descripción de metadatos para la transmisión de documentos con NewsML de forma independiente a este estándar, permitiendo el control de la distribución y gestión estos documentos. Entre los metadatos definidos para NVML en el cuerpo del documento se incluye nvValue para el contenido de la noticia, que incluye, entre otros, resumen y categoría. El sistema de edición permite la clasificación automática de las noticias en categorías de IPTC SubjectCodes en función de la relevancia de cada palabra del texto y de las noticias ya almacenadas en el sistema.

Publishing Requirements for Industry Standard Metadata (PRISM) es una iniciativa de PRISM Working Group y nextPub/PRISM Source Vocabulary Working Group que aúna un conjunto de esquemas y, sobre todo, vocabularios, que los publicadores pueden utilizar para la descripción de contenidos a descargar o intercambiar mediante sindicación, agregación, etc., tales como noticias, blogs, libros, revistas, newsletter, entre otros. Para ello, usa como base Dublin Core (DC) y Dublin Core Metadata Initiative (DCMI).

Uno de los usos más frecuentes de PRISM es la codificación tanto de metadatos como del texto de un artículo en XML, usando para ello PRISM Aggregator Message (PAM) para enviar contenidos entre miembros de la cadena de suministro de información. Para el marcado en el propio texto, también cuenta con PRISM Inline Markup (PIM). La última versión (3.1.) fue publicada a finales de 2014.

Really Simple Syndication (RSS) es un formato para la sindicación de contenidos y metadatos de descripción. “La sindicación de contenidos ayuda al usuario a estar actualizado sobre nuevos contenidos publicados en diferentes y distribuidas fuentes de información y mejora la visibilidad dichos contenidos, garantizando que los usuarios serán advertidos cuando un nuevo contenido sea publicado” (Pereira y Baptista, 2009). Es utilizado, entre otros, por Saias et al. (2006) para obtener una primera clasificación de las noticias de última hora recibidas en un medio de comunicación.

XMLNews fue creado por WavePhore y mantenido por Megginson technologies Ltd., se dividía en dos especificaciones. XMLNews-Story para marcar el contenido de noticias en formato texto mediante un subconjunto de etiquetas de NITF. XMLNews-Meta era un formato de metadatos que podía aplicarse sobre una noticia o cualquier otro recurso, sobre el que aportaba información. Las últimas actualizaciones se realizaron en 1999.

Algunas de estas iniciativas se solapan en ciertos aspectos, aunque sus propósitos son diferentes. En otras ocasiones, se complementan. De hecho, los grupos de trabajo de PRISM y de IPTC están trabajando juntos para definir un formato y vocabulario de metadatos comunes. Por ejemplo, PRISM y IPTC NewsCodes pueden ser utilizados en NewsML.

Merece la pena destacar NewsCodes, un conjunto de vocabularios diseñados por la IPTC, una jerarquía de clasificación de materias con 3 niveles y 17 categorías en su primer nivel (García et al. 2006; Troncy, 2008). Anteriormente fue conocida como Topic Sets (Fernández et al., 2007). Son utilizados, entre otros, por Pereira et al. (2003), Pereira y Baptista (2003) y Schranz (2005) en el proyecto Omnipaper; Fernández et al. (2006), Fernández et al. (2007b) y Fernández et al. (2010) en NEWS; Tenenboim et al. (2008) en ePaper; y Kallipolitis et al. (2012) en World News Finder (WNF). En todos los casos, esta taxonomía fue utilizada para la construcción de las ontologías que permitían la categorización de noticias. NewsCodes están compuestos por términos de diversos dominios, clasificados

jerárquicamente en tres niveles, de más general a más específico: subject, subject matter, subject detail. Estos vocabularios aportan valores para los elementos y atributos, por ejemplo, de documentos NewsML y NITF. Algunas taxonomías que destacan de NewsCodes son:

- Subject Reference System (SRS), creadas por IPTC y NAA. Ofrecen un esquema en el que se define valores para clase de objeto, atributo, referencia de materia (subject reference), sinónimos y cualificadores. NITF, NewsML e IIM fueron concebidos en 2003 como extensiones de SRS. Está compuesto por Subject Code, Subject Qualifier, Media Type, NewItem y Genre
- SubjectCodes es una taxonomía centrada en el texto, desarrollada desde 1990 hasta 2010, cuando es reemplazada por una nueva extensión, Media Topics, actualizada recientemente

Aunque la mayor parte de la bibliografía y la muestra seleccionada se centra en periódicos, los metadatos también son utilizados en agencias de noticias. AFP, por ejemplo, según Guerrillot (2006) implementa el formato NewsML y hasta aboga por la implicación de la agencia en el desarrollo de NewsML-G2. También EFE y ANSA (Sánchez-Fernández et al., 2005; Fernández et al., 2010). La Agencia EFE, de hecho, muestra en su sitio web algunos documentos sobre cómo utilizar NewsML y NITF en sus artículos.

Algunos autores han reutilizado parte de estos metadatos para el desarrollo de proyectos de gestión de artículos periodísticos. ePaper (Tenenboim, 2008; Shoval et al., 2008) es un periódico personalizado basado en la categorización de noticias y perfiles de usuarios usando una ontología; Omnipaper (Yaginuma et al., 2003a, 2003b, 2004; Pereira y Baptista, 2003a, 2003b, 2004; Ariza Ávila y Baptista, 2004; Schranz, 2005), fue un proyecto desarrollado en el seno de la Unión Europea para la creación de un perfil de aplicación que resultara útil para el acceso a artículos periodísticos de diferentes medios europeos; Neptuno (Castells et al., 2006) tiene por objetivo desarrollar un archivo semántico de alta calidad para Diari SEGRE facilitando la descripción y anotación de artículos y la búsqueda de información integrando una ontología en el sistema; NEWS (*News Engine Web Services*) tiene como propósito desarrollar herramientas que asistan a las agencias de noticias en sus procesos de producción y distribución de artículos. También usa una ontología como herramienta principal del proyecto (Fernández-García et al, 2007, 2010); WNF (World

News Finder) vincula una herramienta de extracción de conocimiento con una ontología para la extracción de información a partir de documentos de noticias del mundo en HTML y extrae conocimiento que coincide con la estructura de clasificación dada (Kallipolitis et al., 2012).

### 3. Metodología

Para configurar la muestra se llevó a cabo una selección continua de zonas geopolíticas, países y periódicos, siguiendo un proceso de muestreo no aleatorio intencional. Así, se dividió el mundo en 9 zonas, de acuerdo a sus competencias geopolíticas, históricas y socioeconómicas. A continuación, en cada área se seleccionaron aquellos países con mayor Producto Interior Bruto (PIB). Después, se escogió el diario generalista de cobertura nacional o internacional más leído, de acuerdo a los datos de accesos Web y compra de edición impresa de 4International Media & Newspaper (2012).

Para esta selección, se siguieron los siguientes criterios: Edición en inglés, español, portugués o francés. En Alemania no se halló ningún medio representativo escrito en un idioma distinto al alemán; Los periódicos deben contemplar de manera habitual noticias de política y economía, la temática de los artículos que queremos componer como muestra. De hecho, se han incluido en la muestra los tres diarios económicos más importantes a nivel internacional: Financial Times, The Economist (Reino Unido) y The Wall Street Journal (Estados Unidos); Los periódicos deben ser de tipo generalista, ya que facilita la comprensión de las noticias, con la excepción de los tres especializados ya indicados. En el caso particular de Australia, la ausencia de medios generalistas de referencia, hizo que se optara por un medio de corte económico.

La siguiente tabla muestra los veintiún periódicos analizados:

| Parte del mundo      | País           | Periódico   |
|----------------------|----------------|---|
| UE y Norte de Europa | Alemania       | Süddeutsche Zeitung                                       |
|                      | Francia        | Le Monde  |
|                      | Reino Unido    | The Daily Telegraph;<br>Financial Times;<br>The Economist |
| Norte de América     | Estados Unidos | The New York Times;<br>The Wall Street Journal            |
| Sudamérica           | Brasil         | O Globo   |
|                      | México         | El Universal  |
| ExRepúblicas URSS    | Rusia          | Pravda  |
| África               | Nigeria        | Nigerian Tribune  |

|                 |                        |                                 |
|-----------------|------------------------|---------------------------------|
| subsahariana    | Sudáfrica              | Independent Online              |
| Asia            | China                  | China Daily                     |
|                 | India                  | The Times of India              |
|                 | Japón                  | Asahi Shimbun                   |
| Oriente Medio   | Arabia Saudí           | Arab News                       |
|                 | Emiratos Árabes Unidos | Gulf News                       |
|                 | Israel                 | Yedioth Aharonot                |
|                 | Turquía                | Today's Zaman                   |
| Oceanía         | Australia              | The Australian Financial Review |
| Norte de África | Egipto                 | The Daily News Egypt            |

Tabla 2. Muestra de periódicos

La toma de datos se basó en el análisis de los códigos fuente de tres noticias de cada uno de los diarios de la muestra. En diciembre de 2014, febrero y abril de 2015 se accedió a una noticia, seleccionada al azar, de la portada de cada periódico. Posteriormente se analizaron los metadatos de descripción de contenidos comprendidos entre las etiquetas `<head>` y `</head>`, esto es, etiquetas meta. También se tuvo en cuenta los metadatos utilizados para anotar el texto en el cuerpo de la propia noticia. Finalmente todos los datos obtenidos se recogieron en un documento Excel que facilitó la comparación entre lo mostrado por la bibliografía y los códigos fuente.

Finalmente, el software MetadadosHTML, una aplicación creada ad hoc y autoejecutable en Object Pascal permitió la descarga del código fuente, en HTML, de las 5582 noticias. Después se extrajeron diversos metadatos, a partir de los esquemas y patrones detectados.

## 4. Resultados

### 4.1. Aclaraciones terminológicas

Con una estructura jerárquica de metadatos es posible representación información en varios grados de detalle (Jokela et al., 2001) y con diferentes funciones que conviene aclarar: Metadatos en HTML, que aparecen en el código fuente de las noticias, mediante las etiquetas "title", "keywords", "news\_keywords" y "description" o utilizando microformatos para la descripción semántica de información muy concreta. Fueron los primeros en aparecer y aún son muy utilizados; Vocabularios, tales como taxonomías, tesauros y ontologías, permiten indizar el contenido de una noticia a partir de una selec-

ción previa de conceptos y relaciones entre los mismos, como NewsCodes; Espacios de nombre ("namespace"), compuestos por un grupo de elementos y atributos. Un elemento existe en un espacio de nombre particular y debe ser validado contra el espacio de nombre referenciado. Así identificamos de dónde procede cada elemento y evitamos colisiones por el uso de elementos comunes (Alesso y Smith, 2004, pp. 70 y 200). Son imprescindibles para conocer la procedencia de esquemas y perfiles de metadatos; Esquemas y perfiles de metadatos, que definen elementos concretos de descripción de contenido de las noticias. Son insertados en el código fuente mediante espacios de nombre; Formatos de marcado semántico, como los que enumeran, Pastor-Sánchez et al. (2013) y Sporny et al., 2014: microformatos, RDFa y microdatos. En la figura 1, en anexo, se puede observar cómo se utilizan en una noticia seleccionada al azar.

### 4.2. Metadatos contenidos en el código fuente

Los estándares identificados fueron Dublin Core (descrito en el apartado 2), Schema.org, Open Graph Protocol (para distribuir información en Facebook), Twitter Cards. Además de las etiquetas meta "title", "keywords" y "description" y los microformatos de HTML, "news\_keywords" resultó ser muy utilizada.

Schema.org es una colección de esquemas para marcar páginas HTML. Es usado por motores de búsqueda como Bing, Google Yahoo! y Yandex para mejorar la visualización y recuperación de recursos. El esquema más general es "Thing", que se aplica a cualquier tipo de recurso, pero lo habitual es utilizar uno más específico. Entre los diarios de la muestra se utilizan dos esquemas, Breadcrumb (6 diarios) y News Article (9 diarios, 2 de ellos hasta el nivel más específico). Breadcrumbs permite definir jerarquías temáticas o secciones. Funciona mediante la definición de enlaces, de más general a más específico; Para noticias, de más genérico a específico, encontramos los esquemas Thing > Creative Work > Article > News Article. Las instancias de News Article deben aparecer como valores de la propiedad associated Article del tipo Media Object. Esta clase contiene derivados de las propiedades de rNews, de IPTC. Para contenidos de noticias, a nivel del esquema Creative Work, la etiqueta `<meta itemprop=keywords>`, muestra las palabras clave, delimitadas por comas.

El Protocolo Open Graph permite que cualquier página web se convierta en un objeto rico, en un gráfico social que, entre sus objetivos, persigue

facilitar la visualización de sus contenidos en Facebook. Creado originalmente por los desarrolladores de esta red social, se inspira en Dublin Core, link-rel canonical, Microformatos y RDFa y actualmente es mantenida por la OWF (Open Web Foundation, 2014). Para la descripción de un recurso, cuenta con cuatro propiedades básicas (title, type, url e image), y siete opcionales (por ejemplo, site\_name o description). Se definen en la cabera del código fuente “head” mediante la etiqueta <meta property=“og:propiedad” >.

Con Twitter Cards, los desarrolladores de esta red de microblogging han definido algunos metadatos para visualizar información relativa a los recursos multimedia adjuntos en los tuits. Los veintiocho metadatos se organizan en siete subclases entre las que “Summary” se utiliza por defecto, como un resumen, y emplea cinco propiedades; “Summary Card with Large Image” es similar a la anterior, pero destaca una imagen; “Product card” se utiliza para productos informativos, con un total de 11 propiedades. Algunos tags a resaltar, por ejemplo, son <twitter:title>, <twitter:description>. Se definen en la cabera del código fuente “head” mediante la etiqueta <meta property=“twitter:propiedad” >.

Google News utiliza el esquema sitemaps para marcar noticias y proporcionar metadatos a Google sobre ésta (Google Inc., 2014b). <news:keywords> consiste en “una lista separada por comas de palabras clave que describan el tema del artículo. Las palabras claves se pueden extraer de la lista de palabras clave existente de Google Noticias, aunque no es la única fuente de donde se puede obtener (Google Inc., 2014a)” (4)

#### 4.3. Comparación bibliografía – códigos fuente

Los datos, de acuerdo con el siguiente gráfico, muestran la escasa coincidencia entre los metadatos referenciados en el estado del arte y los utilizados en el código fuente de las noticias de los diarios que componen la muestra.

Respecto a la bibliografía, destacan NewsML y NITF, ambos creados por IPTC como los estándares más referenciados. También Dublin Core es citado en un buen número de trabajos, la mayoría de ellos vinculados al proyecto Omnipaper. Es utilizado por un par de periódicos de la muestra. En contraste, en todos los códigos fuente aún se utilizan etiquetas meta para reflejar parte del contenido de las noticias (título, resumen y palabras clave). Open Graph Protocol, el esquema empleado para Facebook, y Twitter Card, son también muy utilizados. En este caso, el fin es facilitar la visualización de

parte de las noticias en estas redes sociales. schema.org, que cuenta con un esquema propio para noticias (News Article) es aplicado por más de la mitad de los diarios de la muestra.

Los datos extraídos a partir del software MetadatosHTML facilitó el análisis de esta información. Se comprobó que título, resumen y palabras clave, los datos más comunes entre los estándares, coincidían. Esta información es utilizada en un trabajo posterior.

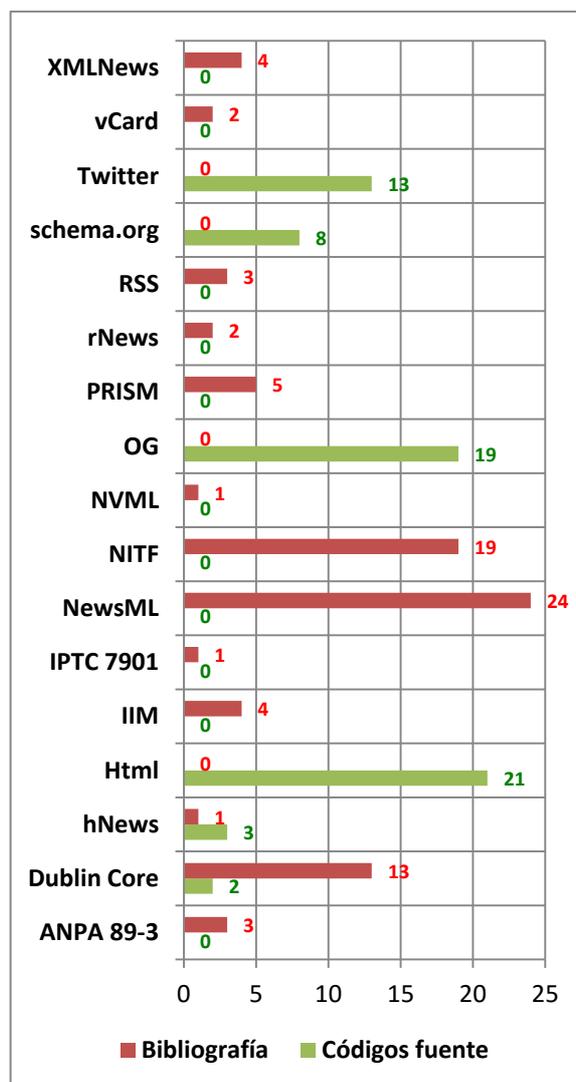


Gráfico 1. Comparación entre metadatos en bibliografía y códigos fuente de noticias

## 5. Conclusiones y propuestas de futuro

En el ámbito periodístico existen diversas formas de representar metadatos para el contenido, desde las más generales (las propias etiquetas de HTML, Dublin Core o los microdatos) a otras específicas y complejas como NITF y NewsML. La mayor parte de los estándares específicos, recogidos en la bibliografía, perma-

necen ocultos al usuario final. El mapeado del código fuente de noticias, muestra que estos metadatos no se muestran en la parte de visible de las webs de los diarios. Por otro lado, aquellos metadatos utilizados en las webs apenas aparecen referenciados en el estado del arte.

Además, la diversidad de esquemas y tecnologías dificultan un mayor aprovechamiento y pueden hasta ir en contra de la estandarización u homogeneización de las formas de representar un documento. Si bien algunos esquemas generales, especialmente los basados en la difusión de información en redes sociales, se han propagado en todos los medios, la escasa presencia de esquemas más específicos en el código fuente de los 21 diarios, en los que se identifique algo más que palabras clave y resúmenes, apunta hacia dos posibles causas: no los utilizan y/o no los visibilizan. Esta conclusión coincide con lo indicado por Pellegrini (2012), que confiaba en el impulso de rNews para solventarlo. En cualquier caso, desde el punto de vista de la reutilización de la información sería deseable que, de entre los sets de metadatos para la representación de noticias, se alzara uno que fuera empleado por todos los medios de comunicación (García et al, 2006). Un modelo estándar o, como mínimo, consensuado por usuarios y desarrolladores (Pastor-Sánchez, 2013, p. 21) facilitaría un mayor intercambio y reutilización de información. Precisamente, González Cristóbal et al. (2002) advertían de la imposibilidad de “definir un recurso accesible de forma unificada debido a la heterogeneidad en los formatos de almacenamiento de noticias y en la meta-información que las describe”. Por otro lado, para alcanzar el ideal de interoperabilidad de los contenidos, es preciso utilizar tecnologías de la Web Semántica. En este sentido, se debería tender a definir ontologías o vocabularios RDF para las diferentes propuestas analizadas en este trabajo como, por ejemplo, vCard.

Este trabajo representa un avance en el modelado de una ontología de dominio de política y economía, utilizando como base noticias. La principal ventaja de la información de actualidad es su abundancia en Internet y permanente actualización, lo que asegura una continua renovación de términos, conceptos y relaciones (Baños-Moreno et al., 2013). Esta idea no es nueva, pues ya existen ontologías que son punto de partida y también producto de noticias de prensa, tales como Saleh y Al-Khalifa (2009) y Fernández et al. (2006, 2007b, 2010 y 2012). El siguiente paso es la extracción de la información recogida por los metadatos del código fuente. Para ello, el software metadatosHTML, permite obtener un listado de términos sobre política y

economía, de acuerdo con el contenido de las noticias. El análisis en profundidad de algunos de los esquemas de la bibliografía facilitará la identificación de entidades. Esta ontología cubriría “el hueco que existe entre la necesidad de los usuarios para la selección personalizada de contenidos y lo que la industria mediática puede ofrecer” (Fernández et al., 2007b), entre otros. En muchos casos, además, se reutilizan herramientas como IPTC Subject Codes o tesauros de una agencia o medio de comunicación. Castells et al. (2006), por ejemplo, parten de IPTC Subject Codes y aprovechan la experiencia de documentalistas y periodistas del Diari SEGRE; Tenenboim (2008) usa IPTC Subject Codes para la ontología implementada en ePaper; Biyun et al. (2009) parten de EventsML, NewsML y NITF con el fin de desarrollar una ontología para el análisis de noticias e implementar algoritmos para alimentar dicha ontología con nuevos elementos.

Una limitación importante de este trabajo es que, más allá de la identificación de espacios de nombre, prefijos y elementos visibles en el código fuente, no se alcanza a conocer qué estándares están siendo utilizados a nivel interno en cada medio de comunicación o para el intercambio con agencias de noticias. Entrevistas o encuestas a conocedores de los métodos de trabajo y de difusión de noticias de los diarios podría solventar esta cuestión.

Otra posibilidad habría sido utilizar las fuentes RSS de los diarios, o bien aquellos contenidos producidos por los lectores, ya convertidos en prosumidores (Toffler, 1980). La tendencia observada es la asunción de las posibilidades de la web 2.0. Facebook, de hecho, se ha convertido en un contenedor más de información, en un canal de comunicación entre el medio y los lectores. Tanto es así que, recientemente, The New York Times, The Guardian y la BBC, entre otros, han comenzado a publicar algunos de sus contenidos directamente en la red social (Jiménez Cano y Abad Liñán, 2015)

## Notas

- (1) También se utiliza “http-equiv” en lugar de “name”, que “permite que los servidores que funcionan con el protocolo de transferencia de hipertexto (HTTP), recopilen la información para ofrecer los encabezados del mensaje de respuesta” (Lamarca Lapuente, 2013).
- (2) Más información de organizaciones que usan NITF en [https://www.iptc.org/site/News\\_Exchange\\_Formats/NITF/Who's\\_using\\_it/](https://www.iptc.org/site/News_Exchange_Formats/NITF/Who's_using_it/)

- (3) El recurso de Google Testing-tool permite identificar el esquema utilizado y ver a qué nivel del esquema se ha implementado.
- (4) Este servicio ha desaparecido en España, para diarios españoles, tras imponer el pago a editores en concepto de derechos de autor por la inclusión de artículos en este servicio desde el 16 de enero de 2014.

## Referencias

- 4IMN. (2012). 4International Media & Newspaper. Retrieved October 11, 2012, from <http://www.4imn.com>
- Abadal, E., Guallar, J., & Codina, L. (2014). Sistemi di documentazione della stampa periodica: quali sono e come valutarli? *AIB studi*, 54(1), 75–86. <http://doi.org/10.2426/aibstudi-9486>
- Abbud Grácio, J. C., & Fadel, B. (2010). Estratégias de preservação digital. En *Gestão, mediação e uso da informação* (pp. 58–83). São Paulo: Editora UNESP; Cultura Acadêmica. Recuperado de <http://books.scielo.org/id/j4gkh/pdf/valentim-9788579831171-04.pdf>
- Agarwal, S., Singhal, A., & Bedi, P. (2012). Classification of RSS Feed News Items Using Ontology. En A. Abraham, A. Zomaya, S. Ventura, R. Yager, V. Snasel, A. K. Muda, & P. Samuel (Eds.), *International Conference on Intelligent Systems Design and Applications, ISDA* (pp. 491–496). New York: IEEE. <http://doi.org/10.1109/ISDA.2012.6416587>
- Alesso, H. P., & Smith, C. F. (2004). *Developing Semantic Web Services*. Natick, Mass: A K Peters/CRC Press.
- Allen, R. B., Japzon, A., Achananuparp, P., & Lee, K. J. (2007). A framework for text processing and supporting access to collections of digitized historical newspapers. En M. J. Smith & G. Salvendy (Eds.), *Human Interface and the Management of Information: Interacting En Information Environments, Pt 2, Proceedings* (Vol. 4558, pp. 235–244). Berlin: Springer-Verlag Berlin. Recuperado de <https://scholarworks.iupui.edu/bitstream/handle/1805/4552/allen-2007-framework.pdf?sequence=1&isAllowed=y>
- Ariza Ávila, C. E., & Baptista, A. A. (2004). Uso de RDF y bases de datos de metadatos nativas dentro del proyecto Ominpaper. En *XATA 2004: actas da 2ª conferência nacional* (Vol. 2, pp. 166–169). Porto: FEUP. Recuperado de <http://repositorium.sdum.uminho.pt/handle/1822/2249>
- Asociación Española de Normalización y Certificación (2012). Información y documentación. Metadatos para la gestión de documentos. Parte 3: Método de auto-evaluación. UNE-ISO/TR 23081-3. Madrid: AENOR
- Baños-Moreno, M.-J., Pastor-Sánchez, J.-A., & Martínez-Béjar, R. (2013). Propuesta de actualización de macrotesoros a partir de noticias de divulgación científico-tecnológica. En *Informação e/ou Conhecimento: as duas faces de Jano* (pp. 99–112). Porto (Portugal): Faculdade de Letras da Universidade do Porto / CETAC.MEDIA. Recuperado de <http://hdl.handle.net/10760/20684>
- Baptista, A. A., & Machado, A. (2001). Metadata Usage En an Online Journal - An application Profile. En A. Hübler, P. Linde, & J. W. . Smith (Eds.), *Electronic Publishing '01 - 2001 En the Digital Publishing Odyssey* (pp. 59–64). Kenterbury, UK: University of Kent. Recuperado de <http://elpub.scix.net/cgi-bin/works/Show?200106>
- Biyun, H., Jun, W., & Yiming, Z. (2009). Ontology design for online news analysis (Vol. 4, pp. 202–206). <http://doi.org/10.1109/GCIS.2009.78>
- Castells, P., Perdrix, F., Pulido, E., Rico, M., Fuentes, J. M., Benjamins, R., ... Granollers, T. (2006). Newspaper Archives on the Semantic Web. En R. Navarro-Prieto & J. L. Vidal (Eds.), *HCI related papers of Interacción 2004* (pp. 267–276). Springer Netherlands. Recuperado de [http://link.springer.com/chapter/10.1007/1-4020-4205-1\\_22](http://link.springer.com/chapter/10.1007/1-4020-4205-1_22)
- Cutts, M. (2009, September 21). Google does not use the keywords meta tag En web ranking. Recuperado de <http://googlewebmastercentral.blogspot.com/2009/09/google-does-not-use-keywords-meta-tag.html>
- Díaz Nosty, B. (2013). *La prensa en el nuevo ecosistema informativo. «¡Que paren las rotativas!»*. La transición al medio continuo. Barcelona: Ariel, Fundación Telefónica, Planeta. Recuperado de [http://www.fundacion.telefonica.com/es/arte\\_cultura/publicaciones/detalle/238](http://www.fundacion.telefonica.com/es/arte_cultura/publicaciones/detalle/238)
- Fernández, N., Arias Fisteus, J., Sánchez, L., & López, G. (2012). IdentityRank: Named Entity Disambiguation En the News Domain. *Expert Syst. Appl.*, 39(10), 9207–9221. <http://doi.org/10.1016/j.eswa.2012.02.084>
- Fernández, N., Blázquez, J. M., Fisteus, J. A., Sánchez, L., Sintek, M., Bernardi, A., ... Ben-Asher, Z. (2006). NEWS: Bringing Semantic Web Technologies into News Agencies. En I. Cruz, S. Decker, D. Allemang, C. Preist, D. Schwabe, P. Mika, ... L. M. Aroyo (Eds.), *The Semantic Web - ISWC 2006* (pp. 778–791). Springer Berlin Heidelberg. Recuperado de [http://link.springer.com/chapter/10.1007/11926078\\_56](http://link.springer.com/chapter/10.1007/11926078_56)
- Fernández, N., Blázquez, J. M., Sánchez, L., & Bernardi, A. (2007). IdentityRank: Named entity disambiguation En the context of the NEWS project (Vol. 4519 LNCS, pp. 640–654). Recuperado de <http://www.scopus.com/inward/record.url?eid=2-s2.0-34548061897&partnerID=40&md5=09bf835b0eb41d2e206ff59b97ec5ca3>
- Fernández, N., Fuentes, D., Sánchez, L., & Fisteus, J. A. (2010). The NEWS ontology: Design and applications. *Expert Systems with Applications*, 37(12), 8694–8704. <http://doi.org/10.1016/j.eswa.2010.06.055>
- Fernández, N., Sánchez-Fernández, L., Blázquez-del-Toro, J. M., & Villamor-Lugo, J. (2007). The News Ontology for Professional Journalism Applications. En R. Sharman, R. Kishore, & R. Ramesh (Eds.), *Ontologies* (pp. 887–919). Springer US. Recuperado de [http://link.springer.com/chapter/10.1007/978-0-387-37022-4\\_32](http://link.springer.com/chapter/10.1007/978-0-387-37022-4_32)
- Galfi, R. (2012, September 19). Google News Blog: A newly hatched way to tag your news articles. Recuperado de <http://googlenewsblog.blogspot.com.es/2012/09/a-newly-hatched-way-to-tag-your-news.html>
- García Gutiérrez, A. (2014). Análisis documental de noticias de prensa en sistemas de información factual. *Revista Española de Documentación Científica*, 37(2), e046. <http://doi.org/10.3989/redc.2014.2.1094>
- García, R., Perdrix, F., & Gil, R. (2006). Ontological Infrastructure for a Semantic Newspaper. En *In "Semantic Web Annotations for Multimedia Workshop, SWAMM 2006". 15th World Wide Web Conference*.
- González Cristóbal, J. C., Villena Román, J., Bueno Carrillo, F. J., García Serrano, A. M., Ruiz Cristina, A., & Martínez Fernández, P. (2002). OmniPaper: acceso inteligente a periódicos europeos. Recuperado de <http://rua.ua.es/dspace/handle/10045/1752>
- Google Inc. (2014a). Creating a Google News Sitemap. Retrieved November 30, 2014, from <https://support.google.com/news/publisher/answer/74288?hl=en>
- Google Inc. (2014b). Learn about sitemaps. Retrieved February 12, 2014, from <https://support.google.com/webmasters/answer/156184?hl=en>

- Guerrillot, S. (2006). Use of semantic technologies at Agence France-Presse (AFP). Presented at the Semantic Technology Conference, San José. Recuperado de <http://ceur-ws.org/Vol-194/paper8.pdf>
- Heravi, B. R., & McGinnis, H. (2013). A Framework for Social Semantic Journalism. En *First International IFIP Working Conference on Value-Driven Social & Semantic Collective Intelligence (VaSCO)*. Paris, France. Recuperado de <http://members.deri.ie/~bahher/Publications/A%20Framework%20for%20Social%20Semantic%20Journalism%20Final.pdf>
- Hickson, I. (2013, October 29). HTML Microdata: W3C Working Group Note 29 October 2013. Recuperado de [www.w3c-prg/TR/microdata/](http://www.w3c-prg/TR/microdata/)
- Hillman, D. (2005, November 7). Using Dublin Core (DCMI Recommended Resource). Recuperado de <http://dublincore.org/documents/usageguide/>
- Iannella, R., & McKinney, J. (2014). vCard Ontology - for describing People and Organizations. W3C Interest Group Note 22 May 2014. Recuperado de <http://www.w3.org/TR/vcard-rdf>
- IPTC (International Press Telecommunications Council). (2009). *NewsML G2. Specification Version 2.4. Power Conformance Level* (No. Document Revision 1). Recuperado de [https://www.iptc.org/std/NewsML-G2/2.4/specification/NewsML-G2\\_2.4-spec-PCL.pdf](https://www.iptc.org/std/NewsML-G2/2.4/specification/NewsML-G2_2.4-spec-PCL.pdf)
- IPTC (International Press Telecommunications Council). (2014). NITF News Industry Text Format. Recuperado December 17, 2014, de [https://www.iptc.org/site/News\\_Exchange\\_Formats/NITF/](https://www.iptc.org/site/News_Exchange_Formats/NITF/)
- Jiménez Cano, R., & Abad Liñán, J. M. (2015, May 13). "The New York Times", "The Guardian" y la BBC publican directamente en Facebook desde hoy. *El País*. San Francisco / Madrid. Recuperado de [http://tecnologia.elpais.com/tecnologia/2015/05/13/actualidad/1431490102\\_473389.html](http://tecnologia.elpais.com/tecnologia/2015/05/13/actualidad/1431490102_473389.html)
- Jokela, S., Turpeinen, M., Kurki, T., Savia, E., & Sulonen, R. (2001). The role of structured content En a personalized news service. *Acta Polytechnica Scandinavica Mathematics and Computing Series*, (114), XX–XXI. Recuperado de <http://www.computer.org/csdl/proceedings/hicss/2001/0981/07/09817044.pdf>
- Kallipolitis, L., Karpis, V., & Karali, I. (2012). Semantic search En the World News domain using automatically extracted metadata files. *Knowledge-Based Systems*, 27, 38–50. <http://doi.org/10.1016/j.knosys.2011.12.007>
- Kodama, M., Ozono, T., Shintani, T., & Aosaki, Y. (2008). Realizing a news value markup language for news management systems using newsML. En F. Xhafa & L. Barolli (Eds.), *2nd International Conference on Complex, Intelligent and Software Intensive Systems* (pp. 249–255). Los Alamitos: IEEE Computer Soc. <http://doi.org/10.1109/CISIS.2008.70>
- Lamarca Lapuente, M. J. (2013, August 12). *Hipertexto, el nuevo concepto de documento en la cultura de la imagen: Metadatos en HTML* (tesis). Universidad Complutense de Madrid, Madrid. Recuperado de [http://www.hipertexto.info/documentos/meta\\_html.htm](http://www.hipertexto.info/documentos/meta_html.htm)
- Mannens, E., Coppens, S., De Pessemier, T., Dacquin, H., Van Deursen, D., De Sutter, R., & Van de Walle, R. (2013). Automatic news recommendations via aggregated profiling. *Multimedia Tools and Applications*, 63(2), 407–425. <http://doi.org/10.1007/s11042-011-0844-8>
- Mannens, E., Troncy, R., Braeckman, K., Van Deursen, R. D., Van Lancker, V. W., De Sutter, R., & Van De Walle, R. (2009). Automatic metadata enrichment En news production. En *10th International Workshop on Image Analysis for Multimedia Interactive Services* (pp. 61–64). New York. <http://doi.org/10.1109/WIAMIS.2009.5031432>
- Martínez-Fernández, J. L., García-Serrano, A., Martínez, P., & Villena, J. (2004). Automatic Keyword Extraction for News Finder. En A. Nürnberger & M. Detyniecki (Eds.), *Adaptive Multimedia Retrieval* (pp. 99–119). Springer Berlin Heidelberg. Recuperado de [http://link.springer.com/chapter/10.1007/978-3-540-25981-7\\_7](http://link.springer.com/chapter/10.1007/978-3-540-25981-7_7)
- McGee, M. (2012, September 19). Google Announces News Keywords Meta Tag For Publishers. Recuperado de <http://searchengineland.com/google-announces-news-keywords-metatag-133759>
- Mueller, E. T. (2000). Making news understandable to computers. *arXiv*. Recuperado de <http://arxiv.org/html/cs/0003001>
- Nies, T. de, D'heer, E., Coppens, S., Van Deursen, D., Mannens, E., & Van de Walle, R. (2012). Bringing newsworthiness into the 21st century. En *Web of Linked Entities, Workshop proceedings* (pp. 106–117). Recuperado de <http://ceur-ws.org/Vol-906/paper11.pdf>
- Open Web Foundation (OWF). (2014, October 20). The Open Graph protocol. Retrieved December 3, 2014, from [ogp.me](http://ogp.me)
- Paepen, B. E. (2002). Omnipaper: Bringing electronic news publishing to a next level using XML and Artificial Intelligence. En J. A. Carvalho, A. Hübler, & A. A. Baptista (Eds.), *Proceedings of the 6th International ICC/IFIP Conference on Electronic Publishing*. Karlovy Vary, Czech Republic: VWF Berlin. Recuperado de <http://elpub.scix.net/cgi-bin/works/Show?02-29>
- Pastor-Sánchez, J. A. (2011). *Tecnologías de la Web Semántica* (Edición: 1). Barcelona: Editorial UOC, S.L.
- Pastor-Sánchez, J. A., Orduña-Malea, E., & Saorín Pérez, T. (2013). Marcado semántico automático en gestores de contenidos: integración y cuantificación. *El Profesional de la Información*, 22(5), 381 – 391. <http://doi.org/10.3145/epi.2013.sep.02>
- Pellegrini, T. (2012). Semantic metadata En the news production process: achievements and challenges. En *Proceedings of the 16th International Academic MindTrek Conference 2012* (pp. 125–133). Tampere; Finland: ACM Press. <http://doi.org/10.1145/2393132.2393158>
- Pereira, T., & Baptista, A. A. (2003). Omnipaper: descrição de recursos de notícias digitais em RDF. En J. C. Ramalho, Pedro Rangel Henriques, G. R. Librelotto, & G. V. Arnold (Eds.), *XML, aplicações e tecnologias associadas*. Braga, Portugal: Universidade do Minho. Recuperado de <http://repositorium.sdum.uminho.pt/handle/1822/283>
- Pereira, T. S. M., & Baptista, A. A. (2009). The instantiation of OmniPaper RDF prototype En the context of scientific publications. *Electronic Library*, 27(5), 767–778. <http://doi.org/10.1108/02640470910998506>
- Rubio Lacoba, M. (2012). Nuevas destrezas documentales para periodistas: el vocabulario colaborativo del diario El País. *Trípodos*, 31, 65–78. Recuperado de <http://www.raco.cat/index.php/Tripodos/article/view/262073/349255>
- Saias, J., & Quaresma, P. (2006). A proposal for an ontology supported news reader and question-answer system. En S. Oliveira Rezende (Ed.), *Proceedings of International Joint Conference, 10th IBERAMIA, ICMC-USP*. Ribeirao Preto, Brazil. Recuperado de <http://ceur-ws.org/Vol-199/wonto-05.pdf>
- Saleh, L. M. B., & Al-Khalifa, H. S. (2009). AraTation: An Arabic semantic annotation tool (pp. 447–451). <http://doi.org/10.1145/1806338.1806421>

- Sánchez-fernández, L., Bernardi, A., & Fuentes, M. (2005). An experience with Semantic Web technologies En the news domain. En *Integration of Knowledge, Semantics and Digital Media Technology, 2005. EWIMT 2005. The 2nd European Workshop on the (Ref. No. 2005/11099)* (pp. 455–459). London.
- Schranz, M., Dustdar, S., & Platzer, C. (2005). Building an Integrated Pan-European News Distribution Network. En L. M. Camarinha-Matos, H. Afsarmanesh, & A. Ortiz (Eds.), *Collaborative Networks and Their Breeding Environments* (pp. 587–596). Springer US. Recuperado de [http://link.springer.com/chapter/10.1007/0-387-29360-4\\_62](http://link.springer.com/chapter/10.1007/0-387-29360-4_62)
- Shoval, P., Maidel, V., & Shapira, B. (2008). An Ontology-Content-based Filtering Method. *International Journal ITA*, 15(4), 303–314. Recuperado de <http://sci-gems.math.bas.bg:8080/jspui/handle/10525/88>
- Silva, D. L. da, Souza, R. R., & Almeida, M. B. (2008). Ontologias e vocabulários controlados: comparação de metodologias para construção. *Ciência da Informação*, 37(3), 60–75. Recuperado de [http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S0100-19652008000300005](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-19652008000300005)
- Sporny, M., Longley, D., Kellogg, G., Lanthaler, M., & Lindström, N. (2014, January 16). JSON-LD 1.0: A JSON-based Serialization for Linked Data. Recuperado de <http://www.w3.org/TR/json-ld/>
- Tenenboim, L., Shapira, B., & Shoval, P. (2008). Ontology-based classification of news En an electronic newspaper (pp. 89–97). Presented at the INFOS 2008: Intelligent Information and Engineering Systems, Varna, Bulgaria. Recuperado de <http://sci-gems.math.bas.bg/jspui/bitstream/10525/1035/1/IBS-02-p12.pdf>
- Toffler, A. (1980). *The third wave: The classic study of tomorrow*. New York, NY, USA: Bantam.
- Troncy, R. (2008). Bringing the IPTC News Architecture into the Semantic Web. En A. Sheth, S. Staab, M. Paolucci, D. Maynard, T. Finin, & T. Krishnaprasad (Eds.), *Semantic Web - Iswc 2008* (Vol. 5318, pp. 483–498). Berlin: Springer-Verlag Berlin. <http://doi.org/10.1007/978-3-540-88564-1-31>
- Wong, W., Liu, W., & Bennamoun, M. (2010). An ontology-based interface for improving information exploration. En *Proceedings of the first international workshop on Intelligent visual interfaces for text analysis* (pp. 29–32). New York, USA: ACM. <http://doi.org/10.1145/2002353.2002364>
- Yaginuma, T., Pereira, T., & Baptista, A. A. (2003a). Design of metadata elements for digital news articles En the omnipaper project. En S. M. de Souza Costa, J. A. Carvalho, A. A. Baptista, & A. C. Santos Moreira (Eds.), *From information to knowledge: 7th ICCC/IFIP International Conference on Electronic Publishing* (pp. 132–139). Minho, Portugal: Universidade do Minho. Recuperado de <http://repositorium.sdum.uminho.pt/handle/1822/170>
- Yaginuma, T., Pereira, T., & Baptista, A. A. (2003b). Metadata elements for digital news resource description. En *Proceedings CLME'2003 - 3º Congresso Luso-Moçambicano de Engenharia* (pp. 1317–1326). Maputo. Recuperado de <http://repositorium.sdum.uminho.pt/handle/1822/279>

## Anexo 1

| <i>Especificaciones</i>           | <i>Citados en bibliografía</i>   | Nº |
|-----------------------------------|--|----|
| ANPA 89-3                         | Sánchez-Fernández et al., 2005; Fernández-García et al., 2007; Pellegrini, 2012  | 3  |
| Dublin Core                       | Yaginuma, Pereira y Baptista, 2003; Pereira y Baptista, 2003a; Pereira y Baptista, 2003b; Pereira et al., 2003; Ariza y Baptista, 2004; Martínez-Fernández et al., 2004; Schranz, 2005; Fernández et al., 2006; Allen et al., 2007; Fernández-García et al., 2007; Fernández et al., 2010; Moreira et al., 2012; Pellegrini, 2012  | 13 |
| hNews                             | Pellegrini, 2012   | 1  |
| IIM                               | García et al., 2006; Fernández-García et al., 2007; Tse-Yi y Teong, 2008; Fernández et al., 2010   | 4  |
| IPTC 7901                         | Fernández-García et al., 2007  | 1  |
| NewsML (NAR)                      | Pereira y Baptista, 2003a; Pereira y Baptista, 2003b; Pereira et al., 2003; Yaginuma, Pereira y Baptista, 2003; Martínez-Fernández et al., 2004; Yaginuma, Pereira y Baptista, 2004; Sánchez-Fernández et al., 2005; Schranz, 2005; Schranz et al., 2005; Fernández et al., 2006; García et al., 2006; Guerrillot, 2006; Fernández-García et al., 2007; García et al., 2008; Kasper et al., 2008; Kodama et al., 2008; Tenenboim et al., 2008; Troncy, 2008; Tse-Yi y Teong, 2008; Biyun et al., 2009; Mannens et al., 2009; Fernández et al., 2010; Kallipolitis et al., 2012; Mannens et al., 2013 | 24 |
| News Value Markup Language (NVML) | Kodama et al., 2008  | 1  |
| NewsCodes (SRS)                   | Pereira y Baptista, 2004; Sánchez-Fernández et al., 2005; Fernández et al., 2006; García et al., 2006; Fernández-García et al., 2007; García et al., 2008; Kasper et al., 2008; Shoal et al., 2008; Tenenboim et al., 2008; Troncy, 2008; Tse-Yi y Teong, 2008; Heravi y McGinnis., 2013; Kallipolitis et al., 2012; Heravi et al., 2012; Chy et al., 2014   | 15 |
| NITF                              | Mueller, 2000; Pereira y Baptista, 2003a; Pereira y Baptista, 2003b; Pereira et al., 2003; Yaginuma, Pereira y Baptista, 2003; Ariza y Baptista, 2004; Martínez-Fernández et al., 2004; Yaginuma, Pereira y Baptista, 2004; Sánchez-Fernández et al., 2005; Schranz, 2005; Schranz et al., 2005; Fernández et al., 2006; García et al., 2006; Fernández-García et al., 2007; Fernández et al., 2010; García et al., 2008; Tse-Yi y Teong, 2008; Biyun et al., 2009; Pellegrini, 2012   | 19 |
| PRISM                             | Fernández-García and Sánchez-Fernández, 2004; Fernández et al., 2006; Fernández-García et al., 2007; Tse-Yi y Teong, 2008; Fernández et al., 2010  | 5  |
| rNews                             | Heravi et al., 2012; Pellegrini, 2012; Heravi y McGinnis, 2013   | 2  |
| RSS                               | Tenenboim et al., 2008; Fransincar et al., 2009; Pereira y Baptista, 2009  | 3  |
| vCard                             | Pereira y Baptista, 2003a; Pereira y Baptista, 2003b   | 2  |
| XMLNews                           | Mueller, 2000; Yaginuma, Pereira y Baptista, 2003; Yaginuma, Pereira y Baptista, 2004; Fernández-García et al., 2007   | 4  |

Tabla 3. Especificaciones de metadatos referencias por la bibliografía analizada

| <i>Especificaciones</i> | <i>Referencias</i>   | Nº |
|-------------------------|--|----|
| Etiquetas meta (html)   | Süddeutsche Zeitung; Le Monde; The Daily Telegraph; Financial Times; The Economist; The New York Times; The Wall Street Journal; O Globo; El Universal; Pravda; Nigerian Tribune; Independent Online; China Daily; The Times of India; Asahi Shimbun; Arab News; Gulf News; Yedioth Aharonot; Today's Zaman; The Australian Financial Review; The Daily News Egypt | 21 |
| Dublin Core             | Le Monde; The Telegraph  | 2  |
| hNews                   | Nigerian Tribune; Independent Online; Asahi Shimbun  | 3  |
| schema.org              | Süddeutsche Zeitung; Le Monde; The Daily Telegraph; The New York Times; The Wall Street Journal; O Globo; Today's Zaman; The Daily News Egypt  | 8  |
| Open Graph Protocol     | Süddeutsche Zeitung; Le Monde; The Daily Telegraph; Financial Times; The Economist; The New York Times; The Wall Street Journal; O Globo; El Universal; Pravda; Nigerian Tribune; The Times of India; Asahi Shimbun; Arab News; Gulf News; Yedioth Aharonot; Today's Zaman; The Australian Financial Review; The Daily News Egypt                                  | 19 |
| Twitter Card            | Süddeutsche Zeitung; Le Monde; The Daily Telegraph; The Economist; The New York Times; The Wall Street Journal; O Globo; El Universal; The Times of India; Gulf News; Today's Zaman; The Australian Financial Review; The Daily News Egypt   | 13 |

Tabla 4. Especificaciones de metadatos aparecidas en los códigos fuente de los periódicos de la muestra

## Anexo 2

The diagram illustrates various semantic markup elements in an HTML document. It shows code snippets with callouts explaining their purpose:

- Elementos propios del formato de marcado semántico Microdatos:** Points to `itemid` and `itemtype` attributes.
- URI que lleva a un espacio d nombre, que se refiere a un modelo de descripción:** Points to the `itemtype` value.
- Prefijo que abrevia un espacio de nombre:** Points to the `og:` prefix in `og:description`.
- Uno de los elementos que compone un esquema de metadatos:** Points to the `og:description` element.
- QName:** Points to the `twitter:` prefix in `twitter:description`.
- Etiqueta meta, en HTML, para el elemento "keywords". Su contenido puede ser libre o basado en un vocabulario:** Points to the `keywords` and `news keywords` meta elements.

Code snippets shown include:

```

<!--[if (g
app-arti
js section-africa format-long tone-news
itemid="http://www.nytimes.com/2015/05/19/world/africa/boko-haram-militants-raped-
hundreds-of-female-captives-in-nigeria.htm" itemtype="http://schema.org/NewsArticle
itemscope xmlns:og="http://opengraphprotocol.org/schema/">
<!--![endif]-->
<head>
<title>Boko Haram Militants Raped Hundreds of Female Captives in Nigeria
NYTimes.com</title>
<meta itemprop="description" name="description" conte
terrorist group, has long targeted women, rounding them up a
the northeast of Nigeria." />
<meta property="og:description" content="Boko Haram an Islamist terrorist group, has
long targeted women, rounding them up as it captures
Nigeria." />
<meta property="twitter:description" content="Boko haram, an islamist terrorist group,
has long targeted women, rounding them up as it captures
of Nigeria." />
<meta name="keywords" content="Women and Gir
Crimes,Refugees and Displaced Persons,Kidnapping and Hostages,Boko Haram,Jonathan
Goodluck,Shekau,Abubakar,Chibok (Nigeria),Damaruru (Nigeria),Abuja (Nigeria),Nigeria" />
<meta name="news keywords" content="Women and Girls;Human
trafficking;Rape;Refugees;Boko Haram;Goodluck Jonathan;Abubakar Shekau;C
Nigeria;Damaruru;Abuja Nigeria;Nigeria" />
(...)
</head>

```

Figura 1. Parte del código fuente de una noticia de The New York Times, de 18 de mayo de 2015, Boko Haram Militants Raped Hundreds of Female Captives En Nigeria

---

# Perspectivas metodológicas para a representação da informação do documento de arquivo

*Methodological Perspectives for the Representation of the archive Document Information*

---

**Gilberto Gomes Cândido (1), João Batista Ernesto de Moraes (2)**

(1) Faculdade de Filosofia e Ciências, UNESP, Av. Higyno Muzzi Filho, 737, Marília, São Paulo, Brasil, CEP: 17525-900, ggcandido@marilia.unesp.br

(2) jota@marilia.unesp.br

## Resumo

O objeto deste estudo é o documento de arquivo em sua fase permanente, que é submetido à Descrição de seu conteúdo para elaboração do ponto de acesso, para proporcionar a recuperação da informação. Observa-se, deste modo, um déficit existente de *procedimentos metodológicos para a representação da informação do documento arquivístico que vão além de normas e convenções*. As hipóteses que aqui são levantadas têm como função buscar solucionar o seguinte problema: *Quais são as metodologias existentes que poderiam vir a contribuir para a representação da informação do documento arquivístico por meio dos pontos de acesso? Tendo como objetivo demonstrar que a Diplomática Contemporânea, a Análise Documental e o Percurso Gerativo de Sentido, como procedimentos metodológicos, podem auxiliar na representação da informação do documento de arquivo*. Posto isto, pode-se apontar que a Diplomática Contemporânea permite extrair elementos intrínsecos e extrínsecos ao texto do documento de arquivo; contudo, sua aplicação dá-se em conjunto com a Análise Documental, o que permite que se faça a leitura do documento de arquivo de forma dinâmica. Portanto, a Diplomática Contemporânea passa a ser um roteiro de análise utilizado pela Análise Documental como modelo de leitura documental. Por sua vez, o Percurso Gerativo de Sentido proporciona compreensão do conteúdo textual do documento de arquivo permitindo identificar a espécie documental. Observa-se assim que, tais metodologias podem ser utilizadas para a representação da informação do documento de arquivo.

**Palavras-chave:** Análise Documental. Percurso Gerativo de Sentido. Diplomática. Pontos de acesso. Descrição.

## 1. Introdução

A Ciência da Informação (CI), como área transdisciplinar, é voltada para a observação de fenômenos da produção, organização e uso da informação, de modo a proporcionar a elaboração de instrumentos e estabelecer metodologias

## Abstract

The object of this study is the archival document in its permanent phase whose content is described for determining its access points to facilitate information retrieval. It is observed, therefore, a lack of methodological procedures for the representation of the information in archival documents that go beyond the standards and conventions. The hypotheses here address the following problem: What are the existing methodologies that could potentially contribute to the representation of the information in archival documents through access points? Aiming to demonstrate that the Contemporary Diplomatics, the Document Analysis, and Generative Meaning Course, as methodological procedures, may assist in the representation of the information in archival documents. Provided this, we can point out that the Contemporary Diplomatic allows to extract intrinsic and extrinsic elements of the archival document; however, its application takes place in conjunction with Document Analysis, which allows to dynamically analyze the archival document. Therefore, Contemporary Diplomatics becomes a guideline for Document Analysis as a model for document reading. On the other hand, the Generative Meaning Course provides understanding of the textual content of the archival document, allowing to identify the document type. This way It is observed that such methods can be used for representing the information of the archival document

**Keywords:** Document Analysis. Generative Trajectory of Meaning. Diplomatics. Access points. Description.

para a disseminação, o acesso e o uso da informação, levando-a a grupos sociais diversificados, exercendo, assim, o seu cunho social.

Desta forma, alguns autores da CI, tais, como Borko (1968); Brookes (1980); Foskett (1973); Le Coadic (1996) e Saracevic (1996) discutem

sobre essas inter-relações que a CI possibilita fazer com outras áreas do conhecimento, firmando assim, a sua transdisciplinaridade ao emprestar métodos de outros Campos de Atuação Profissional (CAP).

Sendo assim, essa pesquisa é calcada na transdisciplinaridade, ao apresentar métodos originados outros CAPs, com propósito de proporcionar discussões teóricas sobre a representação da informação do Documento de Arquivo, tendo como base as seguintes metodologias: Diplomática Contemporânea, Análise Documental e Percurso Gerativo de Sentido.

À vista disso, o objeto desta pesquisa é o documento de arquivo em sua fase permanente, cujo valor primário é o probatório; entretanto, nesta fase também lhe é atribuída outra acepção, o valor histórico, que permite fazer a reconstrução do passado com a finalidade de auxiliar no futuro.

Nesta fase, o documento de arquivo passa pela Descrição de seu conteúdo, de forma a assegurar o acesso e a recuperação da informação. Ainda nesta fase, faz-se uso de normas, tais como: Norma Geral Internacional de Descrição Arquivística (ISAD (G)) e a Norma Internacional de Registro de Autoridade Arquivística para Entidades Coletivas, Pessoas e Famílias (ISAAR (CPF)), bem como, o uso de normas nacionais. Tais normas recomendam a padronização da Descrição, a partir de pontos em comum em seus níveis de Descrição.

As normas de Descrição de documento de arquivo são bem mais usuais em arquivos permanentes, embora essas ressaltem que a Descrição do documento de arquivo pode ser aplicada em quaisquer fases documentais. Os instrumentos de pesquisa, da mesma maneira como a Descrição, de maneira geral são elaborados na fase permanente. Contudo, evidencia-se que as normas não apresentam entradas ou saídas de dados, bem como os instrumentos de pesquisas estão a cargo de cada instituição custodiadora por sua elaboração.

Observa-se, deste modo, um déficit existente de *procedimentos metodológicos para a representação da informação do documento arquivístico por meio dos pontos de acesso que vão além de normas e convenções*. Sendo assim, os métodos, até então apresentados, estão intrinsecamente ligados às normas, e não oferecem outros meios para a representação da informação do documento arquivístico.

As proposições aqui levantadas têm como função buscar resolver o seguinte problema: *Quais*

*são as metodologias existentes que poderiam vir a contribuir para a representação da informação do documento arquivístico, no que tange os pontos de acesso? Assim sendo, se tem como objetivo demonstrar que a Diplomática Contemporânea, a Análise Documental e o Percurso Gerativo de Sentido, como procedimentos metodológicos, podem auxiliar na representação da informação do documento de arquivo, visto que, essas podem apresentar interlocução entre si, por meio de seu objeto de estudo que é o texto.*

## **2. Os Instrumentos de Representação do Documento de Arquivo: e a Descrição de pontos de acesso.**

As normas de Descrição arquivística, tanto a ISAD (G) quanto a ISAAR (CPF) possuem elementos comuns no que tange à Descrição do Documento de Arquivo, sendo que a intenção daquelas é promover o intercâmbio da informação entre as instituições, ressaltando, ainda, que alguns elementos são obrigatórios na Descrição do documento de arquivo.

No entanto, vale realçar que, somente a ISAAR (CPF) usufrui do nível de Descrição para a elaboração de *pontos de acesso*; ambas advertem, além disso, que alguns níveis definidos na norma são obrigatórios para elaboração de instrumentos de pesquisa, embora haja outros, que segundo elas, são opcionais.

Todavia, ressalta-se que, antes das normalizações da Descrição, ou seja, da elaboração de normas internacionais, a Descrição do documento de arquivo tinha respaldo em três elementos básicos, expostos pelo manual dos holandeses (1973), e recuperados por Bonal Zazo (2002, p.155) como:

- a atividade (a resenha de caracteres/elementos internos e externos dos documentos);
- o resultado da atividade (instrumentos de Descrição);
- o objetivo da Descrição (fornecer acesso ao conteúdo dos documentos).

Esses argumentos penduraram até a elaboração de normas de Descrição internacionais e nacionais, as quais trazem um sentido a mais ao conceituar a Descrição no campo da Arquivologia.

Com a elaboração da norma *Règles pour la description des documents d'archives* (RDDA/RDA) a primeira norma de Descrição é introduzida a utilização do contexto (produtor) neste procedimento, que é a descrição do documento de arquivo.

Atenta-se, assim, que o termo Descrição agora passa a conter acoplados quatro sentidos no âmbito da arquivologia que foram identificados, estabelecendo assim, o conceito de Descrição, sendo eles: contexto, atividade, resultado e objetivo.

- Atividade: processo de análise dos caracteres/elementos internos e externos ao documento de arquivo.
- Contexto: onde foi elaborado, quem é o produtor do documento de arquivo.
- Resultado: instrumento de pesquisa que foi elaborado com base no processo de análise realizado, sendo esses produtos: guias, inventários, catálogos.
- Objetivo: promover o acesso ao documento de arquivo, tanto para o usuário quanto para o arquivista.

Sendo assim, tendo-se como fundamento a RDDA/RAD, foi elaborada a Norma geral internacional de Descrição arquivística (ISAD (G)) com o objetivo de auxiliar os arquivistas.

Porém, no momento em que a ISAD (G) se encontrava em elaboração, percebeu-se que não se podia levar em conta tão somente a Descrição do documento de arquivo, mas também quem o produziu.

Como resposta a esta indagação, é elaborada a Norma Internacional de Registro de Autoridade Arquivística para Entidades Coletivas, Pessoas e Famílias ISAAR (CPF), para controlar a elaboração destes *pontos de acesso* relativos ao registro de autoridade; contudo a própria ISAAR (CPF) elucida que os pontos de acesso não estão condicionados apenas ao registro de autoridade.

Sendo assim, entende-se que os pontos de acesso não estão mais somente vinculados e restritos ao registro de autoridade, mas a todos os caracteres/elementos que contribuem para recuperação e localização da informação, por meio da Descrição do documento de arquivo.

Desta forma, pode-se ressaltar que o termo ponto de acesso surge por meio da automação, vindo a substituir os termos que eram até então utilizados, ao se descrever os principais aspectos do documento de arquivo que eram até então denominados de: entrada principal e secundária de título.

Entretanto, tais termos – entrada principal e secundária de título, jamais foram utilizados no âmbito da Arquivologia, como Stibbe (1992, p.110 tradução nossa) enfatiza que, nem na primeira e muito menos na segunda edição do

*Dicionário de Terminologia Arquivística* elaborado pelo ICA continham tais termos, ou sequer no Livro *Glossário Básico para Arquivistas, Curadores de Manuscritos e Gestores de Registros*, elaborado pelos autores Frank B. Evans et al. e publicado na *Journal American Archivist (Revista do Arquivista Americano)* em julho de 1974.

O não aparecimento destes termos nos léxicos arquivísticos ocorre em um primeiro momento, segundo Stibbe (1992), porque esses estão relacionados com os *padrões de Descrição, como regras ou códigos de catalogação*. Esses padrões se encontravam ainda em discussões no âmbito Arquivístico.

Todavia, na Biblioteconomia os *padrões de Descrição* se apresentam bem mais desenvolvidos e discutidos. Durante a década de 1960, ocorrem as primeiras divulgações de normas que vinham a auxiliar na Descrição de modo a proporcionar o acesso, tais como: a Norma *Anglo American Cataloguing Rules – AACR* (Código de Catalogação Anglo-Americano) e a norma *Machine-Readable Cataloging - MARC* (Catalogação Legível por Máquina). Com automação crescente, os termos que eram referenciados como *entrada título* passou a ser denominado de *pontos de acesso*, sendo esse termo oriundo da Ciência da Computação, que foi vinculado ao AACR em um primeiro momento como sinônimo de *entrada principal*, mas que, com o passar do tempo, assumiu características próprias, como elementos necessários para a identificação do registro bibliográfico.

No domínio arquivístico, não ocorreu de forma diferente a busca pela normalização da Descrição, também sendo propulsionada pela automação. Entretanto, isso só ocorreu entre as décadas de 1980 e 1990 por meio da Norma dos arquivistas canadenses, nomeada como *Règles pour la description des documents d'archives*. Essa norma trouxe consigo o conceito de *entrada principal* e os mesmos problemas de sinônimos também ocorreram sobre o termo *ponto de acesso*, como visto no AACR e MARC, no entanto, adaptados para a Descrição Arquivística.

Na Arquivologia, os respectivos termos se apresentavam da seguinte forma, o termo *título* foi substituído pelo termo *acesso*, ficando:

*Acesso Principal*: nome pelo qual um fundo e todas as suas partes podem ser pesquisadas e identificadas (Ponto de acesso primário). *Acesso Secundário*: nome, palavra ou frase em adição à categoria principal de acesso, em que Descrição Arquivística pode ser pesquisada e

identificada (Ponto de acesso secundário) (BU-REAU CANADIEN DES ARCHIVISTES 2008, s/n, tradução nossa)".

O termo *ponto de acesso* na arquivologia só assume um significado mais coeso com as Normas de Descrição Internacionais que são a ISAD (G) e ISAAR (CPF), que reconhecem como ponto de acesso, quaisquer elementos de informação que proporcionam a identificação e a localização da informação.

Compreende-se, assim, que a representação da informação e os instrumentos de pesquisa, no âmbito do arquivo, se realizam por meio da Descrição do documento de arquivo, que segundo as normas ISAD (G) e ISAAR (CPF), apresenta níveis que devem ser seguidos para a elaboração da Descrição do documento de arquivo, sendo alguns deles elementos obrigatórios.

Todavia, apenas a ISAAR (CPF), apresenta nível que proporciona a elaboração e o tratamento de *pontos de acesso*, ao demonstrar que esses podem se basear no uso de qualificadores quando são julgados essenciais, ou não, para recuperação.

A Descrição dos documentos de arquivo se reflete na elaboração dos instrumentos de pesquisa que buscam proporcionar o acesso e a difusão da informação, apresentando caracteres/elementos que são intrínsecos ao conteúdo do documento descrito, permitindo a elaboração de *pontos de acesso* que são apresentados pelos os instrumentos, sejam eles: guia, inventário ou catálogo, etc.

Esses instrumentos de pesquisas são complementados pelos *pontos de acesso* que, podem ser entendido como "Um nome, termo, frase ou código usado como um título em um catálogo, especialmente para agrupar informações relacionadas a este título" (Frank B. Evans, p.50, 1974, tradução nossa).

Sendo assim, os instrumentos de pesquisa são elaborados por meio da Descrição do documento de arquivo e são contemplados com os *pontos de acesso*, pois permitem identificar o fundo, localizar, recuperar e controlar o documento de arquivo, cada um dentro de sua peculiaridade ao servir como instrumento de pesquisa.

Deixa em evidencia também, que segundo a definição apresenta da pelo ISAAR (CPF) os *pontos de acesso* não estão restritos apenas ao controle do registro de autoridade, mas a todos os caracteres/elementos que contribuam para o acesso e recuperação da informação.

Na elaboração dos *pontos de acesso*, não se deve repetir termos e eles devem ser controlados, com intuito de aumentar a eficácia e diminuir a imprecisão no momento da busca da informação.

### 3. Perspectivas Metodológicas

Desta forma, os pontos de acesso, inicialmente foram discutidos tendo como propósito o controle de autoridades, contudo, depois da publicação da ISAAR (CPF) o que era até então restrito passou a abranger outros caracteres/elementos.

Entretanto, mesmo que a Descrição esteja relacionada aos instrumentos de pesquisa, as normas não se apresentam com relação a sua elaboração, e nem com entrada ou saída de dados.

Com isso, percebe-se que a representação da informação do documento de arquivo está limitada às normas, e que estas não oferecem qualquer metodologia que possa auxiliar no trabalho do arquivista na extração de caracteres/elementos para representação da informação do documento de arquivo.

Sendo assim, dando seguimento ao levantamento bibliográfico identificou-se três metodologias que têm como objeto de estudo o texto, sendo elas: *Diplomática, Análise Documental e Percurso Gerativo de Sentido*. Estas buscam identificar alguns traços que se sobressaem no texto, porém, cada uma com a sua singularidade.

Deste modo, neste tópico são apresentadas as três metodologias identificadas, com o intuito de que na sequência se possa fazer a interlocução entre essas.

Posto isso, a Diplomática Contemporânea será abordada por meio da sua estrutura de análise, que busca a auxiliar na identificação do conteúdo do documento de arquivo bem como conhecer o seu contexto.

A segunda metodologia a ser abordada é Análise Documental, mediante as suas duas vertentes: Analítica e Sintética. Na fase analítica, ocorre a leitura técnica, que busca identificar quais são os pontos mais relevantes do texto com objetivo de extrair termos apropriado do texto analisado para representá-lo. Por sua vez, a fase sintética decompõe os termos que foram extraídos em uma linguagem documental, como o vocabulário controlado, tesouros, etc.

E a última metodologia a ser apresentada é o Percurso Gerativo de Sentido que surge dos estudos de Greimas. Seu propósito é o de direcionar o leitor sobre o entendimento geral do

texto que foi analisado, e se divide em três patamares sendo eles: *fundamental*, *narrativo* e *discursivo*.

O *patamar narrativo* é a transformação entre dois estados sucessivos e diferentes que garantem uma narrativa mínima, quando se tem um estado inicial, uma transformação e um estado final. Mas, os textos são bem mais complexos, pois emitem uma série de enunciados de estado (fazer e ser) que são denominados de sequência canônica, sendo esta por: manipulação, competência, performance e sanção.

Já no *patamar discursivo*, as formas abstratas do nível narrativo são revestidas de concretude, ou seja, o discurso que enriquece o texto por meio de temas e figuras.

Sendo assim em síntese, este serão os pontos principais percorridos nos tópicos que se seguiram.

### 3.1. Análise Documental

Análise Documental (AD) na Ciência da Informação, de acordo com alguns autores, como: Gardin et al.1981; Garcia Gutierrez, 1984; Cunha,1987; Pinto Molina, 1989; Coyaud, 1966; Chaumier,1971, 1974; Clauso Garcia, 1993; Ruiz Perez 1992; Guimarães, 2003; Pinto; Galvez,1999; Pinto; Molina 1991. A Análise Documental pode ser entendida, como sendo, um conjunto de etapas que auxiliam na identificação do conteúdo documento, e posterior síntese, tendo como finalidade sua representação por meio de um documento secundário (produtos: catálogo, vocabulário controlado etc.) considerado instrumento provedor de acesso, consulta e disseminação da informação.

Chaumier, J. (1974) entende a AD como sendo uma técnica que visa a, por meio da representação expor o conteúdo de um determinado documento, perante a um formato distinto do original, com intuito de promover posteriormente o seu acesso e consulta.

Segundo García Gutierrez (1984), a AD permite por meio de operações mentais objetivas, identificar e transformar os documentos analisados em produtos como: catálogo, vocabulário controlado etc. De modo, a promover o acesso dos documentos originais e auxiliar as comunidades científicas.

Pinto Molina (1991) compartilhando do entendimento de García Gutierrez expõe que AD se constitui mediante a um conjunto de operações, sendo elas de ordens: Intelectual, Mecânica e Repetitiva.

Ainda na concepção da referida autora, tais operações, influenciam o conteúdo (temática) e as formas (Descrição bibliográfica) dos documentos originais, uma vez que, a esses são atribuídos outra característica considerada secundária que é instrumental, visando a facilitar ao usuário a identificação, recuperação e difusão desses.

Contudo, Pinto Molina (1991) ressalva que, esse é um processo geral de caráter analítico, que busca sintetizar o documento original, isso é, o que permite que a forma final do documento secundário, ou seja, os produtos que auxiliam no acesso e disseminação da informação.

Deste modo, a AD se vale de dois procedimentos sendo eles: fase analítica e fase sintética. A primeira ocorrer por um processo de leitura técnica que busca identificar os conceitos do texto analisado; a segunda fase, sintética, é a seleção dos conceitos, condensação documental e tradução dos conceitos.

Posto isso, a fase analítica como processo de leitura, utiliza tanto de aspectos cognitivo quando de modelos de leituras proposto por alguns autores como: Chaumier (1988); Tálamo (198); Kobashi (1994); Fujita (2003) que permite conhecer a estrutura textual das partes que compõe o texto e realçá-las, de modo que se possa chegar ao seu conteúdo (GUIMARÃES, 1994).

Observa-se assim que, para ocorrer a compreensão da estrutura textual do documento, assim como o entendimento de seus elementos mais relevantes, devem ser levadas em conta as estratégias metacognitivas de leituras e técnicas para a identificação da temática.

As estratégias são adquiridas de forma costumeira no desenvolvimento de certa atividade, de modo a permitir que essa ocorra outra vez. De acordo com Silva; Fujita (2004, p. 148) é essa atividade que possibilita a leitura documentária e dessa forma evitar erros e incompreensões.

Assim sendo, as estratégias cognitivas e metacognitivas estão relacionadas, além disso, com os objetivos que serão refletidos no decorrer da leitura documentária, dos quais seus aspectos são abordados por Pinto Molina (1993, p. 162-163), quando enfatiza que “As estratégias são desenvolvidas durante o período de leitura e não dependem somente do leitor e do texto, mas também dos objetivos documentários”.

Ou seja, os objetivos documentários no momento da leitura são levados em conta e devem ser observados, devido que, tal processo está intrinsecamente relacionado à extração de elementos informativos de um determinado docu-

mento, com intuito de elaborarem-se instrumentos documentários.

Ainda segundo a autora, essas operações são ações mentais que operam de formas organizadas, e podem ser evidenciadas em duas categorias. Sendo a primeira “cognitiva, que compreende a interpretação como comportamento automático e inconsciente; e a segunda o metacognitivo, que supõem atividades desautomatizadas, em que o leitor tem a consciência de como está lendo” (PINTO MOLINA, 1993, p. 162-163).

Para explicar tais operações mentais, utiliza-se a título de exemplo o processo de Descrição de conteúdo usual tanto do bibliotecário quanto do arquivista. A primeira tática de leitura é a cognitiva que se apresenta com sendo automática e inconsciente, podendo ser compreendida pelas as ações rotineiras do cotidiano desenvolvidas por esses profissionais que os leva a agir de forma inconsciente, ou seja, automática não tendo percepção das etapas do desenvolvimento do processo de representação. A segunda tática de leitura as metacognitivas são entendidas como um processo não automático, ou seja, durante o desenvolvimento do método de representação da informação os respectivos profissionais são conscientes das etapas do processo aplicado.

Esses recursos de leituras contêm características conceituais linguísticas e cognitivas, sendo utilizados no processamento da informação por meio das abordagens do *bottom-up* e *top-down*, que são considerados como procedentes da psicologia cognitiva.

Pinto Molina (p.161,1993) nos demonstra que o *bottom-up* “é a leitura do tipo linear, indo das partes ao todo textual, que se baseia exclusivamente na estimulação visual”. Já o *top-down* é método de leitura dedutivo, “em que se caminha de forma inversa, do todo para as partes, aproveitando-se do ‘conhecimento prévio’ ou ‘esquemas’”.

Dessa forma, com base na autora entende-se que, os métodos de leitura são dinâmicos e se sucedem por meio de interpretações, sua compreensão se estabelece pelas ações de significação, de forma a se adequar a análise semântica e análise de conteúdo.

Já na fase denominada de *sintética* se constitui por meio da *seleção de conceitos*: em que os enunciados de assunto são categorizados e ordenados de forma lógica, perante os parâmetros estruturais, funcionais e de uso; a *condensação documentária* corresponde à redução do

documento por meio de síntese; e a *representação documentária*, consiste na tradução do conteúdo temático (GUIMARÃES, 1994). Esses irão ser representados por meio de Linguagem Documentária (LD).

A seleção do conceito acontece no momento em que, os enunciados são categorizados em: principais, secundários e periféricos, e ordenados de forma lógica, em que se têm como parâmetros a estrutura, a função e os usos, (Guimarães, 2003 p.112).

Deste modo, sobre a perspectiva de Cunha (1987), após a seleção dos conceitos (termos), para evitar a ambiguidade e imprecisão, há a necessidade confrontação dos termos com vocabulários técnicos, glossários ou dicionários das áreas de conhecimento envolvidas, cuja finalidade é evitar sinonímia como exemplo: casa, moradia, lar, esses termos passam a ideia do local onde se mora, ou seja, um conceito geral que pode gerar a imprecisão na recuperação da informação, e homonímia, tais como: manga (fruta), manga (camisa), palavras que contêm a mesma grafia mais significados diferentes, isto é conceitos individuais que quando não discernidos corretamente causam imprecisão também.

O controle desses conceitos (termos) se faz necessário em razão de sua importância, devido que os conceitos selecionados e extraídos do documento, poderão ser condensados, ou seja, transformados em resumos ou até mesmo utilizados pelos Sistemas de Recuperações da Informação como pontos de acesso ao documento por meio da Linguagem Documental (LD).

Desta forma, Chaumier (1988, p.65) levanta uma indagação que o documentalista/analista deve se fazer, será que “se um usuário fizer uma pergunta sobre esse conceito, e se esse documento lhe for fornecido, será pertinente aos olhos daquele usuário? Essa questão é a chave de toda a indexação no que concerne à seletividade”.

Além disso, nesta fase também são abordadas algumas questões sobre a profundidade da análise do documento e suas vertentes, tais como: a exaustividade e a especificidade (Guimarães, 2003).

Em síntese pode se apontar que, a exaustividade está relacionada aos números de conceitos utilizados na indexação para representar a informação, e a especificidade, é a exatidão com que esses conceitos representam o conteúdo daquele respectivo documento.

Como consequência disso, ocorre a tradução dos conceitos em termos de Linguagens Documentais (LDs) que são utilizadas como produtos que proporcionam a pesquisa e o acesso a informação.

A LD é a Linguagem Natural (LN) controlada, ou seja, a LN é aquela apresentada pelo conteúdo do documento que quando controlada se transforma em uma LD, podendo ser: vocabulário controlado, tesouro, ontologias, dicionários, sistemas de classificação decimal, dentre outros.

Complementando esta explanação, Lima; Boccato (2009, p.132) demonstram que:

“as linguagens documentárias são construídas de acordo com os princípios e métodos estabelecidos pela terminologia, de modo a propiciar a compatibilidade entre as linguagens do usuário e a utilizada pelo sistema de recuperação”.

Segundo Gardin (apud Cintra et al, 2002, p.35) as LD são um conjunto de termos, utilizados para a representação dos conteúdos de documentos, que têm como propósito a classificação ou a busca da informação.

Observa-se assim que o controle terminológico ocorre por meio da LD, deste modo, além de proporcionar o acesso e recuperação da informação, esta também proporciona o controle do termo na representação da informação, assegurando que haja menos imprecisão e mais eficácia na recuperação.

Desta forma, as LDs atuam como representações documentais, que podem ser “[...] vistas como linguagens de comunicação entre a informação documentária e o usuário que dela necessita” (DODEBEI, 2002, p. 56). Essas são utilizadas também para fim de recuperação da informação, pois são:

“construídas para indexação, armazenamento e recuperação da informação e correspondem a sistemas de símbolos destinados a ‘traduzir’ os conteúdos dos documentos.” (CINTRÁ et al, 2002, p. 33).

Compreende-se assim que, AD por meio de sua forma analítica nos permite por meio da leitura estratégica identificar os pontos principais do texto analisado, logo sua análise sintética nos permite extrair termos de uma LN para a LD tendo como propósito de condensar, representar e recuperar a informação.

### 3.2. Diplomática

A Diplomática Contemporânea surge dos estudos de Luciana Duranti, por meio de seu método de análise do documento contemporâneo

com o intuito de demonstrar os caracteres/elementos intrínsecos e extrínsecos a ele, de modo a prover a identificação de seu contexto.

De acordo com Duranti (1991) os elementos extrínsecos da forma documental são aqueles que o constituem, como o material do documento e sua aparência externa. Segundo a autora alguns elementos extrínsecos que eram utilizados na Diplomática Clássica e Moderna caíram em desuso, não se fazendo mais aplicado sobre o documento contemporâneo.

Ainda na concepção da autora, enquanto alguns elementos caíram no desuso, outros aparecem como as anotações, enfatizados por ela como as partes mais relevantes para o documento contemporâneo, que podem ser apresentadas em três categorias, sendo elas:

- 1) anotações incluídas em um documento após a sua compilação como parte da fase de execução de um procedimento administrativo;
- 2) anotações incluídas em um completo e eficaz documento no curso da execução das etapas subsequentes da operação em que o documento participa; e
- 3) as anotações adicionados a um documento pelos registros e/ou arquivos de serviços que os responsáveis pela sua identificação, como parte de um grupo de documentos (arquivo, série) e para a sua manutenção e recuperação (DURANTI, 1991, p.04).

A primeira remete à anotação como autenticação, podendo ser referida com base em assinaturas, no documento ou cópia dele, ou quando ocorrer a atribuição de um registro por um notário. Estes caracteres/elementos têm como propósito assegurar que são autênticos por quem o elaborou. A segunda anotação é referente às “marcas” feitas no documento de arquivo durante o seu curso, sendo elas colocadas ao lado do texto, podendo ser, pontos de interrogações, setas indicadoras, data ocorrente da leitura, locuções tais como, “Urgente”, “Importante”, e assim por diante. A terceira anotação são as atribuições realizadas pelo arquivo no documento, como, o protocolo, código de localização, ou quando o item documental é o processo, suas páginas são numeradas e rubricadas etc.

Deste modo, Duranti apresenta esses caracteres/elementos como extrínsecos, em suma, de acordo como identificado por ela por meio da análise da forma documental: Meio; Escrita; Linguagem; Sinais especiais; Selos; Anotações: incluída na manipulação de fase ou incluída na fase de gestão; número de registro; (DURANTI, 1991, p.05).

Essa é a categorização realizada por Duranti com o objetivo de evidenciar os caracteres/elementos extrínsecos da forma do documento de arquivo.

Já os elementos intrínsecos da forma documental são componentes de articulação intelectual e os modos existente para a apresentação do documento, sendo eles: *Protocolo e Fórmulas iniciais* têm contexto de ação administrativa, ao fazer indicativo de pessoas envolvidas, apresentando tempo, lugar e assunto. Logo o contexto de ação do *Texto* inclui a apresentação de considerações e o motivo para o qual o documento foi elaborado, advindo de qual é sua função que se reflete em sua atividade, *Escatocolo e Fórmulas finais*, seu contexto de ação esta relacionado à validação e indicação de responsabilidade pelo registro do ato (DURANTI, 1991, p.06).

Sendo assim, de acordo com Duranti (1991, p.07, tradução nossa) os

*“elementos intrínsecos da forma documental que compõe o núcleo do texto de um documento de arquivo, narrando a expressão da vontade do autor e da ação do registro”, ou seja, “o texto é a parte central do documento, onde se encontra a manifestação da vontade do autor, a prova do ato, ou a memória dele”.*

Deste modo, pode se considerar o texto como sendo a parte principal do documento de arquivo, em que nele se encontra a manifestação da vontade do autor, a prova do ato, ou a memória dele. Sendo de um ponto vista administrativo, jurídico ou histórico, o texto se faz como parte de suma importância no documento, pois é nela que se encontra a sua substância, ou seja, a razão para a qual o documento foi elaborado (DURANTI, 1991).

Contudo, para os diplomatas, o texto em um todo não oferece mais material para a crítica do documento, o texto que era até então utilizado na antiguidade de modo amplo, passa a ter os pontos certos para fazer análise crítica.

Com isso, Duranti (1991, p.15) identifica que os caracteres/elementos intrínsecos da forma documental referente ao texto são eles:

**PROTOCOLO**      *Timbre:*  
                           *Título do documento:*  
                           *Data tópica e cronológica:*  
                           *Invocação:*  
                           *Subscrição [cabeçalho]:*  
                           *Inscrição:*  
                           *Saudação:*  
                           *Assunto:*  
                           *Formula Perpetuais:*  
                           *Precação:*

**TEXTO:**            *Preâmbulo:*  
                           *Notificação:*  
                           *Exposição:*  
                           *Dispositivo:*  
                           *Cláusulas finais:*  
**ESCATOCOLO:** *Corroboração:*  
                           *[Datas]*  
                           *[Precação]*  
                           *[Saudação]*  
                           *Cláusula complementar:*  
                           *Atestação:*  
                           *Qualificação de assinatura:*  
                           *Notas secretariais:*

Subentende-se, assim, com Duranti (1991) que as combinações entre os caracteres/elementos tanto internos quanto externos que determinam as formas documentais, ou seja, são eles que permitem a identificação do contexto documental.

Desta forma, segundo Duranti (1991), a crítica Diplomática provém da forma do ato iniciado ou referido pelo documento de arquivo, a análise crítica deste, tem como desígnio compreender o seu contexto jurídico, administrativo e processual para o qual o documento analisado foi elaborado.

Sendo assim, a estrutura da análise Diplomática se dá em um procedimento sistemático, indo do específico para o geral. Isso permite que, quando realizada, possa identificar o contexto do documento em que tais procedimentos estão sendo aplicados. Consequentemente, a crítica Diplomática elaborada por Duranti (1991, p.12) dá os seguintes procedimentos para análise do documento contemporâneo:

*Elementos extrínsecos:* meio/ escrita/ linguagem/ sinais especiais/ selos/ anotações;

*Elementos intrínsecos:* Protocolo (subseções) / texto (subseções) / Escatocolo (subseções);

*Pessoas:* Autor do ato/ autor do documento/ destinatário do ato/ destinatário do documento/ escritor/ pessoa que autentica/ testemunhas;

*Qualificação das assinaturas:* títulos e competências das pessoas envolvidas;

*Tipo do ato:* simples/ contratual/ coletivo/ múltiplo/ continuativo/ complexo ou processual;

*Nome do ato:* venda, autorização, requerimento;

*Relação entre o documento e o procedimento:* especificação da fase do procedimento geral ao qual o documento relaciona-se; e, se o documento resulta de um “ato de procedimento”, a fase do procedimento específico;

*Tipo de documento:* nome (carta, escritura) / natureza (público ou privado) / função (dispositivo, pro-

batório) / estado de transmissão (original, esboço, cópia);

*Descrição Diplomática:* contexto (ano, mês, dia, lugar) / ação (pessoas, ato) / documento (nome da forma, natureza, função, estado, meio quantidade);

*Comentários conclusivos:* qualquer comentário que remete para o documento como um todo e não a um elemento específico de forma documental ou componente de análise Diplomática

Com base na apresentação dos procedimentos esquematizados por Duranti (1991) para a realização da análise Diplomática do documento contemporâneo, Tognoli (2013) ressalta que a análise apresentada por ela se constitui mais abrangente do que aquelas observadas nos manuais modernos de Diplomática. Ainda de acordo com Tognoli (2013, p.124) isso é devido à complexidade estrutural da “forma do documento contemporâneo”, que é um reflexo “da complexidade dos sistemas jurídicos nos quais os documentos são gerados no século XXI”.

De acordo com Duranti (1991, p.12), “o esforço de incluir os elementos reais dos documentos no âmbito da análise Diplomática é um prelúdio necessário para a descoberta e conhecimento”, ou seja, a “Diplomática é uma mentalidade, uma abordagem, uma perspectiva, uma forma sistemática de pensar sobre os documentos de arquivo”.

Após essa apresentação da Diplomática, deixa-se em evidência que este estudo tem seu interesse na Diplomática no que se refere à análise dos caracteres/elementos que ela proporciona identificar por meio de sua sistematização, mais a diante busca fazer a interlocuções desta com as outras metodologias que ainda serão expostas.

Assim sendo, na seção seguinte se apresenta o Percurso Gerativo de Sentido, que dispõem a constituir o conhecimento precedente da estrutura textual, e que será utilizado na estratégia metacognitiva de leitura sobre o documento de arquivo, explanando a compatibilidade entre as metodologias.

### 3.3. Percurso Gerativo de Sentido

O Percurso Gerativo de Sentido surge dos estudos de Greimas (1976), e é constituído por três níveis: *fundamental*, *narrativo* e *discursivo*. Juntos, são distintos em suas formas de análise, contudo, existe uma interdependência entre eles, para que ocorra o pleno entendimento do texto, indo da análise simples à mais complexa.

O Percurso Gerativo de Sentido vincula-se à Semântica Discursiva, sendo que esta, para

Greimas, deveria ser “Gerativa, Sintagmática e Geral”. Em relação a esses aspectos Fiorin (1999) pontua que:

É uma teoria gerativa, porque concebe o processo de produção do texto como um percurso gerativo, que vai do mais simples e abstrato ao mais complexo e concreto, num processo de enriquecimento semântico. Isso significa que vê o texto como um conjunto de níveis de invariância crescente, cada um dos qual suscetível de uma representação metalinguística adequada. [...] É uma teoria sintagmática, porque seu escopo é estudar a produção e a interpretação dos textos. Aqui se produz o primeiro deslocamento produzido pela Semiótica. Sua totalidade não é o plano de conteúdo das línguas naturais, mas o texto [...] *É geral, porque se interessa por qualquer tipo de texto, independentemente de sua manifestação* [...] (FIORIN, 1999, p. 2 grifos do autor).

Pode-se dizer que com a Semântica Discursiva de Greimas a noção de Percurso Gerativo de Sentido passa a ter uma perspectiva mais concreta na medida em que pontua elementos de forma sistemática em uma estrutura que possui níveis de invariância crescente.

Designa-se pela expressão Percurso Gerativo a economia geral de uma teoria semiótica (ou apenas linguística), vale dizer, a disposição de seus componentes uns com relação aos outros, e isso na perspectiva da geração, isto é, postulando que, podendo todo objeto semiótico ser definido segundo o modo da sua produção, os componentes que intervêm nesse processo se articulam uns com os outros de acordo com um “percurso” que vai do mais simples ao mais complexo, do mais abstrato ao mais concreto (GREIMAS; COURTÉS, 2008, p. 232).

De acordo com Fiorin (2011), no nível profundo está o elemento semântico que dá base para a construção do texto. Tais elementos têm como alicerce uma relação de oposição.

Greimas; Courtés (2008) caracterizam a relação de oposição como o quadrado semiótico. Essa oposição é expressa por duas operações que se relacionam: negação e asserção.

Barros (2005, p.11) demonstra que “a semiótica tem como objeto o texto, ou melhor, procura descrever e explicar o que o texto diz e como ele faz para dizer o que diz”.

Segundo Floch (2001, p. 15), o Percurso Gerativo é uma representação dinâmica da produção do sentido:

[...] é a disposição ordenada das etapas sucessivas pelas quais passa a significação para se enriquecer e, de simples e abstrata, tornar-se complexa e concreta. Compreende-se a escolha do termo

“percurso”. Mas por que “gerativo”? Porque todo objeto significante, para a semiótica, pode – e deve – ser definido segundo seu modo de produção, e não segundo a “história” de sua criação: “gerativo” se opõe assim a genético.

Fiorin (1999) enfatiza que o Percurso Gerativo de Sentido se constitui de três níveis:

O percurso gerativo é constituído de três patamares: as estruturas fundamentais, as estruturas narrativas e as estruturas discursivas. Vale lembrar que estamos no domínio do conteúdo. As estruturas discursivas serão manifestadas como texto, quando se unirem a um plano de expressão no nível da manifestação. Cada um dos níveis do percurso tem uma sintaxe e uma semântica (FIORIN 1999, p.17).

O Percurso Gerativo de Sentido vai do mais simples ao de complexidade maior, o fundamental (ou profundo), o narrativo e o discursivo, são níveis composto de uma sintaxe e de uma semântica como apresentado pelo autor.

De acordo com Fiorin (1999, p. 14), a semiótica se atenta para o Percurso Gerativo de Sentido como “[...] um simulacro metodológico, para explicar o processo de entendimento, em que o leitor precisa fazer abstrações, a partir da superfície do texto, para poder entendê-lo”. Deste modo, entende as estruturas profunda ou fundamental (como categorias semânticas), narrativa (como estrutura canônica) e discursiva (com concretização de temas e figuras) (FIORIN, 1999).

No entanto, para a construção dos percursos temático e figurativo, se faz necessária à utilização da *isotopia*, que segundo Greimas e Courtés (2008), “[...] constitui um crivo de leitura que torna homogênea a superfície do texto, uma vez que ela permite elidir ambiguidades”. De acordo com os autores, ao se levantar as isotopias se identificam as continuidades semânticas que tornam o texto num conjunto coerente (GREIMAS; COURTÉS, 2008, p. 278).

Na concepção de Moraes e Guimarães (2006, p.07), na etapa das “*estruturas discursivas*, os resultados obtidos nesta fase devem remeter à etapa inicial, ou seja, às *estruturas fundamentais*”, isto é, no percurso lógico da construção do texto coerente.

Deste modo, é possível afirmar que a contribuição de Greimas para o Percurso Gerativo de Sentido baseou-se na Gramática Gerativa, que de acordo com Moraes (2011) é expressa partindo dos elementos mais abstratos aos mais concretos.

Na sequência será apresentado com mais pontualidade o percurso por meio dos níveis narra-

tivo e discursivo que serão objeto de interlocução entre as metodologias até aqui apresentadas.

### 3.3.1. Nível Narrativo

A sintaxe do nível narrativo é ato de contar um fato que apresenta o seu início e o final, Fiorin (2000, p.21) define a narratividade como sendo a “transformação situada entre dois estados sucessivos e diferentes. Isso significa que ocorre uma narrativa mínima, quando se tem um estado inicial, uma transformação e um estado final”.

Na sintaxe narrativa podem ser encontrados dois tipos de enunciados elementares, o enunciado de estado e o enunciado de fazer, que segundo Fiorin (2000, p.12) são:

a) *enunciados de estado*: são os que estabelecem uma relação de junção (disjunção ou conjunção) entre um sujeito e um objeto (no enunciado “Técnico de informática” é “Contratado”, há uma relação de conjunção, indicada pelo verbo ser, entre um sujeito “Técnico de informática” e um objeto “Contrato de serviço”; em o “Técnico de informática não é contratado”, há uma relação de disjunção, revelada pela negação e pelo verbo ser, entre um sujeito “Técnico de informática” e um objeto “Contrato”);

b) *enunciados de fazer*: são os que mostram as transformações, os que correspondem à passagem de um enunciado de estado a outro (no enunciado “O técnico de Informática foi contratado”, há uma transformação de um estado inicial “não contratado” num estado final “contratado”).

Na narrativa, Fiorin (2000) apresenta que existem duas espécies mínimas que podem ocorrer, sendo caracterizada pela privação e a de liquidação de uma privação. Na privação, o estado conjunto existe e passa a um estado final disjunto, como, por exemplo: homem rico que fica pobre, em um primeiro momento da narrativa o sujeito aparece em conjunção (posse) com a riqueza, no final em disjunção (privação) com ela, observa-se que neste enunciado narrativo há transformação entre os dois estados entre ser rico e ficar pobre. Logo na liquidação da privação há uma inversão entre o estado inicial o disjunto e o final conjunto, como, por exemplo: homem pobre que fica rico, nota-se que, no enunciado narrativo, há transformação entre os dois estados o sujeito esta em disjunção (privação) a riqueza, no final em conjunção (posse) com a riqueza, sendo identifica-se uma transformação entre dois estados ser pobre e ficar rico.

Todavia, segundo Fiorin (2000, p. 22), nas funções dos papéis narrativos, não se deve confundir o termo *sujeito* como sendo pessoas e o

termo *objeto* como sendo coisas, pois esses são papéis narrativos “que podem ser representados por um nível mais superficial por coisa, pessoas ou animais”. Por exemplo, em uma narrativa de perseguição aos leões a serem aprisionados, são os leões o “objeto com que o ser que captura deve entrar em conjunção”. No entanto, “quando se diz o tapete voador pousou no terraço da casa”, temos uma transformação cujo estado final tem como sujeito “tapete voador” e como objeto “terraço da casa” (FIORIN, 2000, p. 22-23) ”.

Entretanto, os textos não são compostos somente de narrativa mínima, mas sim por narrativas complexas, que emitem uma série de enunciados de estado (fazer e ser) que são a organização de forma hierárquica, e que têm como composição a sequência canônica, que é caracterizada por meio de etapas esquematizadas do padrão narrativo, e pode ser compreendida em quatro fases: a manipulação, a competência, a performance e a sanção (FIORIN, 2000; TATIT, 2003).

Deste modo, Fiorin (2000, p. 22-23-24) descreve as quatro sequências canônicas na seguinte conformidade:

*Manipulação:* um sujeito age sobre outro para levá-lo a querer e/ou dever fazer alguma coisa, ou seja, há o manipulador e o manipulado. São vários os tipos de manipulação, aqui foram identificados alguns como: tentação, intimidação, sedução e provocação.

*Competência:* o sujeito que vai realizar a transformação central da narrativa e dotado de um saber e/ou poder fazer.

*Performance:* e a fase em que se dá a transformação (mudança de um estado -a- outro) central da narrativa. Libertar a princesa presa pelo dragão e a performance de muitos contos de fada. Encontrar o pote de ouro no fim do arco-íris, ou seja, passar de um estado de disjunção com a riqueza para um estado de conjunção com ela pode ser uma performance. O sujeito que opera a transformação e o que entra em conjunção ou em disjunção com um objeto pode ser distinto ou idêntico.

*Sanção:* Nela ocorre a constatação de que a *performance* se realizou e, por conseguinte, o reconhecimento do sujeito que operou a transformação. Eventualmente, nessa fase, distribuem-se prêmios e castigos.

A sequência canônica pode ser exemplificada sobre o *Documento de Arquivo*, por meio da interpretação desse, ou ao se fazer a sua representação por meio da Descrição, assim tem-se a *Manipulação* no documento de arquivo expressada pelo ato (ideias) para o qual foi elaborado; já a *Competência*, expõe o ato para o qual

esse foi elaborado, ou seja, a condição que esse tem de informar a todos aqueles que dele se ocupam; para que assim haja a *Performance* que pode ser identificada por uma transformação que gera outros atos ao documento, podendo ser de conjunção ou disjunção sobre esse; logo a *Sanção* é constatação do que houve na *Performance*, isto é, o momento em que se expressa a vontade ou julgamento do ato do documento.

Contudo, algumas destas fases podem ser apenas pressupostas ou uma delas obter mais destaque que a outra, demonstrando assim que, a sequência canônica não se faz por uma estrutura pré-definida, que se apresentará da mesma forma em todos os textos (MORAES; GUIMARÃES; GUARIDO, 2007, p. 97).

Deste modo, se conclui o nível narrativo, passando neste momento para o nível discursivo.

### 3.3.2. Nível discursivo

No nível discursivo, as formas abstratas do nível narrativo “são *revestidas* de termos que lhe dão concretude” por meio de temas e figuras. Assim, “tematização e figurativização são dois níveis de concretização do sentido. Todos os textos tematizam o nível narrativo e depois esse nível temático poderá ou não ser figurativizado (FIORIN, 2002. p.64-65)”.

Ainda para o autor, a figura é o termo que nos remete a algo de um mundo natural: pássaro, lua, brincar, professor, comer, etc. Desta forma, a figura se remete a qualquer língua natural ou algum sistema de representação que se corresponda compreensível no mundo natural. Logo o “tema é um investimento semântico, de natureza puramente conceptual, que não remete ao mundo natural”, ou seja, o tema são categorias que ordenam, organizam e categorizam os elementos do mundo natural, como: elegância, trabalhar, medo, raciocinar, orgulho, etc. (FIORIN, 2000, p.65).

Contudo, quando se fala em figura, não há restrição apenas ao mundo natural em si, esse mundo natural pode também ser construído, como ocorre em textos de ficção, por exemplo, ficção científica, contos de fadas, dentre outros. Como no caso, a título de exemplo, um texto de ficção científica em que surja um ser que de seus olhos é disparado raios, o que seu corpo se regenere, etc. Sendo assim, esse ser é uma figura de um mundo natural só que construído (FIORIN, 2000).

Deste modo, os esquemas figurativos elaboram uma consequência sobre a realidade por meio

de um simulacro que busca representar o mundo. Logo os esquemas temáticos procuram explicar essa realidade, por meio da classificação e organização, constituindo uma relação e dependência. Sendo assim, Fiorin (2000, p.65) apresenta que os “discursos figurativos tem uma função descritiva ou representativa, enquanto os temáticos têm uma função predicativa ou interpretativa”.

Desta forma, “quando tomamos um texto figurativo precisamos descobrir o tema subjacente às figuras, pois para que estas tenham sentido precisam ser a concretização de um tema, que, por sua vez, e o revestimento de um esquema narrativo (FIORIN, 2000. p. 65-66)”. Porém, “como encontrar esse tema? Para que, um conjunto de figuras ganhe um sentido, precisa ser a concretização de um tema, que, por sua vez, e o revestimento de enunciados narrativos (FIORIN, 2000, p.70) ”.

Diante do exposto, o discurso nada mais é, que a narrativa “enriquecida” “por todas essas opções do sujeito da enunciação, que marcam os diferentes modos pelos quais a enunciação se relaciona com o discurso que enuncia (BARROS, 2005, p.53) ”.

Sendo assim, segundo Fiorin (2005) o encadramento entre figura e temas, designa ao nome de percurso temático e percurso figurativo, em que, ambos devem manter a coerência interna do texto. Contudo, quando isso não ocorre e há uma quebra na coerência, esses se tornam contraditório, mais “pode[m] também criar determinados efeitos de sentido (FIORIN, 2005, p.73-74) ”.

De acordo com Moraes (2011, p.50) “essa coerência entre as figuras torna-se essencial, de modo que o leitor, ao ter contato com um texto, possa através desse jogo de figuras ou conexões, entender o contexto em que se insere o mesmo, e a partir daí extrair os temas”.

Consequentemente, os temas estão nos textos cobertos pelas figuras, “a reiteração dos temas e a recorrência das figuras no discurso denominam-se *isotopia*. A isotopia assegura, graças à ideia de recorrência, a linha sintagmática do discurso e sua coerência semântica (BARROS, 2005, p.71) ”.

Fiorin (2000, p.81) enfatiza que, na análise do discurso, a isotopia é a repetição do mesmo traço semântico no decorrer de um texto. “Para o leitor, a isotopia oferece um plano de leitura, determina um modo de ler o texto”. “A recorrência de traços semânticos estabelece a leitura que deve ser feita do texto. Essa leitura não

provém da fantasia do leitor, mas está inscrita no texto (FIORIN, 2000, p. 82) ”.

Barros (2005, p. 71) demonstra dois tipos isotopia, sendo uma delas “a temática decorrente da repetição de unidades semânticas abstratas, em um mesmo percurso temático”, como por exemplo, este estudo, tem uma isotopia temática de *representar*. Tem-se, deste modo, a retomada dos valores como reproduzir ou interpretar, ou seja, o que se representa, onde se representa, ou na ação de representar, deste modo tem-se a isotopia temática, que nos permite conhecer o assunto que se trata o texto, pois, a isotopia temática costura o texto para o seu entendimento em síntese (BARROS, 2005).

Logo a “isotopia figurativa caracteriza-se pela redundância de traços figurativos, pela associação de figuras aparentadas. A recorrência de figuras atribui ao discurso uma imagem organizada e completa da realidade”, sendo assim as figuras concretizam as temáticas, neste caso, a figura desta pesquisa é documento de arquivo, que está ligado ao tema de representar, essa isotopia permite a imagem organizada e completa da realidade (BARROS, 2005, p.71-72).

Demonstra-se assim que, a isotopia é um procedimento de construção do sentido sobre a dimensão sintagmática, que controla os significados das palavras diante de um contexto.

#### **4. Interlocução entre as metodologias por meio do objeto destas o texto.**

A interlocução que ocorre entre as metodologias: Diplomática Contemporânea, Análise Documental e o Percurso Gerativo de Sentido é o objeto de estudo que as três têm em comum, que é o texto. Estas se relacionam entre si, no momento em que uma complementa a outra, ou até mesmo buscando alguns tópicos em comum no texto analisado.

Todavia, por mais que estas metodologias tenham como objeto de estudo o texto, seus procedimentos de análise são peculiares.

A Diplomática Contemporânea utiliza-se para a análise do texto um roteiro estruturado que auxilia na identificação dos caracteres/elementos no texto do documento de arquivo.

Já a Análise Documental, em sua primeira etapa, a analítica, utiliza-se da leitura na análise do texto contido no documento, essa leitura é comumente denominada de leitura técnica, por que faz uso de estratégias metacognitivas ou de Modelos de Leituras para identificar os caracteres/elementos intrínsecos ao texto analisado.

Após a identificação destes caracteres/elementos por meio da leitura, segue-se outra etapa da Análise Documental, a sintética, que categoriza os conceitos e condensa o texto, para a sua representação por meio de uma Linguagem Documentarias (LD).

E por último, o Percurso Gerativo de Sentido apresenta a Estrutura Narrativa que pode ser identificada pelo enunciado (*de estado e de fazer*) sendo esta a narrativa mínima, entretanto, os textos contêm enunciados complexo que se estruturam por meio da sequência canônica, que são revestidas por meio da estrutura discursiva, constituída por temas e figuras, que são complementos de enriquecimento da narrativa.

Sendo assim, pode ser indicado como interlocução ocorrente entre as metodologias, o objeto de estudos que estas têm comum, ou seja, o texto. Contudo se ressalta que os métodos de análise se diferem entre si.

Deste modo se apresenta as interlocuções entre as metodologias por meio de seu objeto de estudo, que estas têm em comum.

Para tanto, veja-se o Quadro - 01 – Objeto de interlocução entres as metodologias o texto, que se encontra no anexo 1.

Observa-se assim, que Análise Documental auxilia a Diplomática Contemporânea e o Percurso Gerativo de Sentido por meio de suas duas etapas de análise sobre o texto, a primeira etapa a *Analítica* ampara ambas as metodologias, mediante seu método de leitura, que auxilia na identificação dos tópicos que cada uma destas traz consigo para a análise do texto. Sua segunda etapa, a *Sintética*, categoriza os conceitos identificados, realiza a condensação do texto e os apresentam por meio de uma LD.

A interlocução que se estabelece, entre a Diplomática Contemporânea e o Percurso Gerativo de Sentido advém de alguns pontos em comum na análise do texto. A Diplomática Contemporânea tem a sua análise voltada no texto buscando identificar os seguintes tópicos: *Protocolo, Texto e Escatocolo* enquanto que, o Percurso Gerativo de sentido busca identificar no texto: *Nível Narrativo e Discursivo*.

Ressalta assim que a interlocução que sobrevém com as metodologias é o objeto de estudo que as três têm em comum, neste caso o texto, do qual buscam identificar alguns tópicos nele contido. A Análise Documental contribui para com a Diplomática Contemporânea e com o Percurso Gerativo de Sentido pela sua etapa Analítica, advindo de seu método de leitura. Entretanto, pode-se observar tal fato de modo

inverso também, em que a Diplomática Contemporânea e o Percurso Gerativo de Sentido fornecem tópicos à Análise Documental que podem ser identificados no texto, podendo ser compreendidos como um Modelo de Leitura.

Pode-se apontar, ainda, tendo como aporte tais metodologias que a Diplomática Contemporânea permite extrair elementos intrínsecos e extrínsecos ao texto do documento de arquivo; contudo, para que ocorra a sua aplicação, essa age em conjunto com a Análise Documental que permite que se faça a leitura do documento de arquivo de forma dinâmica, desta forma, a Diplomática Contemporânea passa a ser um roteiro de análise utilizado pela Análise Documental como modelo de leitura documental. Logo, o Percurso Gerativo de Sentido proporciona a compreensão do conteúdo textual do documento de arquivo permitindo identificar a espécie documental. Isto posto, observa-se que, tais metodologias sinalizam para a uma melhor Descrição do documento de arquivo por meio de pontos de acesso.

## 5. Considerações Finais

Sendo assim, as hipóteses que aqui se levantaram no início dessa pesquisa, tendo como base tais metodologias: Diplomática Contemporânea, Análise Documental (AD) e o Percurso Gerativo de Sentido, tinha como propósito, apresentar métodos de representação da Descrição do documento de arquivo que fossem além de suas normativas, que poderiam, ou não, vir a servir na elaboração dos Pontos de Acesso:

Deste modo, entende-se que, as metodologias apresentadas têm procedimentos para a elaboração dos Pontos de Acesso, uma vez que, buscam identificar no texto fragmentos relevantes no Documento de Arquivo, nos quais poderão, ou não, servir aos fins de acesso dependendo somente *da política adotada na instituição* para isso.

Contudo, vale ressaltar, há não ocorrência de discussões sobre quais procedimentos metodológicos que possam vir a servir para a elaboração dos Pontos de Acesso, no âmbito da Descrição do documento de arquivo.

Relata-se, assim que, os elementos que já estão preestabelecidos nas normativas de âmbito Internacional, só servem como instrumentos para representação do documento de arquivo, de modo, a não possuir qualquer articulação com os produtos de representação e muito menos possuem base procedimental para elaboração desses.

Sendo assim, observa-se que, as metodologias devem sim, ser levadas em contas na Descrição da informação do documento de arquivo, uma vez que, essas se fazem de forma dinâmica em sua aplicação, diferente das normativas que apresenta de forma engessada os elementos que devem ser analisados.

## Referências

- Barros, D. L. P. (2005) *Teoria semiótica do texto*. 5 ed. São Paulo: Ática.
- Bonal Zazo, J. L. (2002). *La descripción archivística normalizada: origen, fundamentos, principios y técnicas*. Gijón: Trea.
- Bureau canadien des archivistes (Org.) (2008). *Règles pour la description des documents d'archives*. Ottawa, rev.
- Chaumier, J. (1974). *Les techniques documentaires*. Presses universitaires de France.
- Cintra, A. M. M. et al (2002). *Para entender as linguagens documentárias*. 2. ed. rev. ampl. São Paulo: Polis.
- Conselho internacional de arquivos (2000). *ISAD(G): Norma geral internacional de Descrição arquivística: segunda edição, adotada pelo Comitê de Normas de Descrição, Estocolmo, Suécia, 19-22 de setembro de 1999, versão final aprovada pelo CIA*. – Rio de Janeiro: Arquivo Nacional, 2000.
- Conselho internacional de arquivos (2004). *ISAAR (CPF): Norma internacional de registro de autoridade arquivística para entidades coletivas, pessoas e famílias*. Trad. Vitor Manoel Marques da Fonseca. 2. ed. Rio de Janeiro: Arquivo Nacional.
- Dodebei, V. L. D. *Tesouro: linguagem de representação da memória documentária*. Niterói: Intertexto, 2002.
- Duranti. L. (1991) Diplomatics: new uses for an old science (Part VI). *Archivaria*, n 33 (Summer), p. 06-24.
- Duranti. L. (1991). Diplomatics: new uses for an old science (Part V). *Archivaria*, n 32 (Summer), p. 06-24.
- Fiorin, J. L. (2000). *Elementos de análise do discurso*. 9. ed. São Paulo: Contexto.
- Fiorin, J. L. (2002). *Introdução à Linguística: objetos teóricos*. Contexto.
- Fiorin, J. L. (1999). *Sendas e veredas da semiótica narrativa e discursiva*. DELTA. fev./jul, vol.15, no.1.
- Floch, J. M. (2001). *Documento de Estudo do Centro de Pesquisa Sociosemióticos*. Periódicos I. São Paulo: Centro de Pesquisas Sociosemióticos.
- Frank B. Evans (1974), A Basic Glossary for Archivists, Manuscript Curators, and Records Managers, *The American Archivist* 37 (July 1974), pp. 415-433.
- Greimas, A. J.; Courtés, J. (2008). *Dicionário de semiótica*. São Paulo.
- Guimarães, J. A. C. (1994). *Análise documentária em jurisprudência: elementos para uma metodologia de indexação de acórdãos trabalhistas brasileiros*. 1994. 250 f. Tese (Doutorado em Ciências da Comunicação) – Escola de Comunicações e Artes, Universidade de São Paulo, São Paulo.
- Guimarães, J. A. C.; Moraes, J. B. E.; Guarido, M. D. M. (2007). Análisis documental de contenido de textos narrativos: bases epistemológicas y perspectivas metodológicas. *Ibersid: revista de sistemas de información y documentación*, 1, 93-99.
- Guimaraes, J.A.C. (2003) Análise documentaria no âmbito do tratamento temático da informação: elementos históricos e conceituais. In: RODRIGUES, G.M.; LOPES, I.L. (Org.). *Organização e representação do conhecimento na perspectiva da Ciência da Informação*. Brasília: The-saurus, p. 100-117.
- Gutiérrez, A. G (1984). *Lingüística documental: Aplicación a la Documentación de la comunicación social*.
- International Council on Archives (1988), *Dictionary of Archival Terminology = Dictionnaire de terminologie archivistique* [English and French with equivalents in Dutch, German, Italian, Russian and Spanish] 2 ed rev. ed. (München, New York, etc., 1988). (ICA Handbooks Series, Volume 7.).
- Lima, V. M. A.; Boccato, V.R.C. (2009). O desempenho terminológico dos descritores em Ciência da Informação do Vocabulário Controlado do SIBi/USP nos processos de indexação manual, automática e semi-automática. *Perspectivas em Ciência da Informação*, v. 14, n. 1, p. 131-151.
- Moraes, J. B. E. de; Guimarães, J. A. C. (2006). Análise documental de conteúdo de textos literários: em busca do diálogo entre as concepções de aboutness/meaning e de percurso temático/percurso figurativo. In: ENCUESTROS INTERNACIONALES SOBRE SISTEMAS DE INFORMACIÓN Y DOCUMENTACIÓN, 11.
- Pearce-Moses, R., & Baty, L. A. (2005). A glossary of archival and records terminology (p. 25). Chicago, IL: Society of American Archivists.
- Pinto Molina, M. (1991). El análisis formal: descripción bibliográfica y catalogación. *Análisis documental: fundamentos y procedimientos: Madrid: EUEDEMA; 1991p.[92], 125*.
- Pinto Molina, M. (1993). Análisis documental: fundamentos y procedimientos. 2.ed. rev. aum. Madrid: Eudema.
- Silva, M. dos R.; Fujita, M. S. L (2004). A prática de indexação: análise da evolução de tendências teóricas e metodológicas. *Transinformação, Campinas*, v.16, n.2, p.133-161, maio/ ago.
- Stibbe, H.L.P. (1992). Implementing the Concept of Fonds: Primary Access Point, Multilevel Description and Authority Control. *Archivaria*, n 34 p.109-137, 1992.
- Tatit, L. (2003). Abordagem do texto. *Introdução à lingüística*, 5, 187-207.
- Tognoli, N. B. (2013). *A construção teórica da diplomática: em busca da sistematização de seus marcos teóricos como subsídio aos estudos arquivísticos*. 2013. 160 f. Tese (Doutorado em ciência da informação) - Universidade Estadual Paulista - Faculdade de Filosofia e Ciências - UNESP Campus de Marília.

## Anexo 1

| OBJETO DE INTERLOCUÇÃO O TEXTO. |  |                                  |   |                                     |   |
|---------------------------------|--|----------------------------------|---|-------------------------------------|---|
| <i>Análise Documental</i>       |  | <i>Diplomática Contemporânea</i> |   | <i>Percurso Gerativo de Sentido</i> |   |
| <i>Fase Analítica</i>           | Leitura Técnica Documental;<br>Identificação dos Conceitos;                  | <i>Protocolo</i>                 | Timbre;<br>Título do documento;<br>Data tópica e Cronológica;<br>Subscrição;<br>[cabeçalho];<br>Inscrição;<br>Saudação;<br>Assunto;<br>Formula Perpetuitais;<br>Precação; | <i>Nível Narrativo</i>              | Manipulação;<br>Competência;<br>Performance |
| <i>Fase Sintética</i>           | Seleção dos Conceitos;<br>Condensação Documental;<br>Tradução dos Conceitos; | <i>Texto</i>                     | Preâmbulo;<br>Notificação; Exposição;<br>Dispositivo;<br>Cláusulas finais;  | <i>Nível Discursivo</i>             | Sanção.<br><br>Temas;<br>Figuras;           |
|                                 |  | <i>Escatocolo</i>                | <i>Corroboração</i> : [Datas]<br>[Precação][Saudação]<br>Cláusula; complementar;<br>Atestação;<br>Qualificação de assinatura;<br>Notas secretariais;                      |                                     |   |

---

# Construção de macroestrutura de categorias de linguagem documentária em economia: proposta de inovação metodológica

*Construction of macrostructure of documentary language categories in economy: methodological innovation proposal*

---

**Heliomar Cavati Sobrinho (1), Mariângela Spotti Lopes Fujita (2)**

(1) Departamento de Ciências da Informação. Universidade Federal do Ceará. Benfica. Avenida da Universidade, nº 2762- Benfica- Fortaleza – Ceará. Brasil. heliomarcavati@yahoo.com.br

(2) Departamento de Ciência da Informação, Universidade Estadual Paulista, Marília. Av. Hygino Muzzi Filho, 737. Cx.P. 181. Campus Universitário. CEP 17525-900. Marília – SP. Brasil fuji-ta@marilia.unesp.br.

## Resumo

As linguagens documentárias, como instrumentos da representação documentária, são construídas para possibilitar a comunicação entre os conteúdos dos documentos e os usuários de um sistema de informação. Neste sentido, um problema a ser resolvido é a necessidade de investigar a representação da informação do domínio da Economia nas LDs, considerando sua especificidade. Por isto, esta pesquisa investiga a construção da macroestrutura do domínio da Economia. Tem como objetivo geral, contribuir com o desenvolvimento da representação da informação, por meio da experimentação científica da aplicação do “Modelo Metodológico Integrado para Construção de Tesouro”, proposto por Cervantes (2009), no domínio da Economia e como objetivos específicos: 1) investigar os aspectos teóricos da representação documentária e da linguagem documentária; 2) desenvolver uma macroestrutura do domínio da Economia; e 3) propor uma categorização deste domínio. A metodologia utilizada é a aplicação do “Modelo Metodológico Integrado para Construção de Tesouro”, proposto por Cervantes (2009), que integra os arcabouços teóricos da Terminologia e da construção de Tesouros. Como plano de trabalho, para realizar esta aplicação, desenvolve os seguintes procedimentos: 1) Realiza uma pesquisa bibliográfica e descritiva, teórico-metodológica da Representação do Conhecimento e da Linguagem Documentária, no contexto da Ciência da Informação; 2) Desenvolve as etapas previstas no modelo, que são: a delimitação do subdomínio; o estabelecimento dos limites da pesquisa terminológica temática e coleta do corpus do trabalho terminológico; o estabelecimento da árvore de domínio; e a coleta e classificação dos termos; 3) Elabora uma macroestrutura documentária do domínio da Economia tendo, como parâmetro, as estruturas das Linguagens Documentárias e os documentos específicos da área, que são o Vocabulário da USP, o Tesouro ISOC de Economia, a tradução da Library of Congress Subject Headings utilizada pela Biblioteca Nacional, o Boletim de Conjuntura do Instituto de Pesquisa Econômica Aplicada; os Ter-

mos da Área Econômica organizados por Gracioso (2002); a Classificação Nacional de Atividades Econômicas, versão 2.0; e o Journal of Economic Literature; e 4) A partir da compilação do corpus de termos econômicos dos boletins, do seu cotejamento e verificação nas linguagens pesquisadas, classifica-os para, em seguida, apresentar uma proposta de categorização deles no subdomínio “Indicadores Econômicos”, inserindo-os no Vocabulário da USP. Por meio desta comprovação a pesquisa conclui que este modelo é um instrumento consolidado e exequível para construção de Linguagens Documentárias.

**Palavras chave:** Organização e Tratamento da Informação. Metodologia para construção de tesouros. Representação Documentária. Linguagem Documentária. Informação do domínio da Economia.

## Abstract

The documentary languages, as tools of documentary representation, are built in order to make possible the communication between the document contents and the users of an information system. In this sense, a problem to be solved is the need of searching into the information representation in the field of Economy in the LDs, taking into account its specificity. Therefore, this research searches into the macrostructure in the field of Economy. Its general aim is to contribute with the development of the information representation by means of the scientific experimentation applying the “ModeloMetodológicoIntegradoparaConstrução de Tesouro”, proposed by Cervantes (2009), in the field of Economy with specific aims: 1) to search into the theoretical aspects of the documentary representation and of the documentary language; 2)to develop a macrostructure in the field of Economy; and 3) to propose a categorization of this field. The methodology consists of applying the “ModeloMetodológicoIntegradoparaConstrução de Tesouro”, proposed by Cervantes (2009) that integrates the theoretical frameworks of the Terminology and construction of Thesaurus. To accomplish this applying, the following procedures are developed: 1) It is made a descriptive

and bibliographic, theoretical-methodological research of the knowledge Representation and of the Documentary Language, in the context of Information Science. 2) It is developed the steps set up in the model, namely: delimitation of the subdomain; establishment of the thematic terminological research limits and the collection of terminology workcorpus; establishment of the domain tree; and the collection and classification of terms; 3) It is made a documentary macrostructure from the Economy field considering as parameter, the Documentary Language Structures and the specific documents of this field, that are: the Vocabulário da USP, the Tesouro ISOC de Economia, the translation of the Library of Congress Subject Headings used by Biblioteca Nacional, the Boletim de Conjuntura do Instituto de Pesquisa Econômica Aplicada; the terms of

the Economic Area organized by Gracioso (2002); the Classificação Nacional de Atividades Econômicas, 2.0 version; and the Journal of Economic Literature; and 4) From the compilation of the corpus of the bulletins' economic terms, of their comparison and verifying in the researched documentary languages, they are classified, and then a proposal of their categorization is presented in the subdomain "Indicadores Econômicos", inserting them in the Vocabulário da USP. It is concluded that this model is a consolidated and feasible tool for constructing Documentary Languages.

**Keywords:** Documentary Language. Documentary Representation. Information in the field of Economy. Information Organization and Treatment. Methodology for constructing thesaurus.

## 1. Introdução

As linguagens documentárias (LDs), como instrumentos da representação documentária, são construídas para possibilitar a comunicação entre os conteúdos dos documentos e os usuários de um sistema de informação.

Neste sentido, um problema a ser resolvido é a necessidade de investigar a representação da informação do domínio da Economia nas LDs, considerando sua especificidade.

Por isto, esta pesquisa investiga a construção da macroestrutura do domínio da Economia, tendo como parâmetro e metodologia a aplicação do "Modelo Metodológico Integrado para Construção de Tesouro", proposto por Cervantes (2009), que integra os arcabouços teóricos da Terminologia e da construção de Tesouros.

## 2. A Representação Documentária no contexto da ORC

A representação da informação e as LDs fazem parte da teoria da classificação, da teoria do conceito e das teorias da comunicação, denominadas, em seu conjunto, teorias da organização do conhecimento. (GHINCHAT; MENO, 1994).

Em pesquisa realizada por Fujita (2008), por meio de estudos apresentados, de 2005 a 2007, no ENANCIB, Encontro da Associação Nacional de Pesquisa e Pós-Graduação em Ciência da Informação (ANCIB), foi realizada uma análise do desenvolvimento da área de estudo da Organização e Representação do Conhecimento (ORC).

O conceito de ORC está sistematizado "em seu próprio nome formado por dois conceitos fundamentais: a Organização do Conhecimento e a Representação do Conhecimento" (FUJITA, 2008, p. 6), cuja atividade gera instrumentos de

Representação do Conhecimento, ou seja, as LDs – sendo tesouro uma tipologia dessas -, assim como as listas de descritores, dentre outros.

A representação do conhecimento, portanto, é realizada por seus processos, sistemas e instrumentos. Os processos são a indexação e a classificação, auxiliados por sistemas e instrumentos de representação.

O processo de classificação é auxiliado por sistemas de classificação bibliográficos, que traduzem os assuntos para notações alfanuméricas, e o processo de indexação é auxiliado por LDs alfabéticas, tais como os tesouros e as listas alfabéticas de assunto. Ou seja, a representação da informação não ocorre só pela indexação, mas também por outros processos e instrumentos elaborados para organizar e representar um dado aspecto da realidade.

Um desses aspectos é a informação do domínio da Economia - que em sua maioria é informação governamental, cujo tratamento, visando a sua organização e disseminação, só recentemente foi reconhecido e incorporado pelas instituições públicas, principalmente com o advento das novas tecnologias como a Internet (LARA, 1999). Tal fato nos coloca em um momento favorável de aprendizado e desenvolvimento de métodos cada vez mais fidedignos de democratização dessa informação, por meio de sua representação.

## 3. Linguagem Documentária

As LDs são instrumentos de estruturação e representação do conhecimento, em que a "representação por conceitos assume função preponderante entre o significado do conteúdo documentário e o termo que o representa. Em continuidade, o termo que representa o conteúdo

documentário é representado por uma linguagem documentária”. (FUJITA, 2013, 49).

As LDs, como instrumentos da representação documentária, são construídas para possibilitar a comunicação entre os conteúdos dos documentos e os usuários de um sistema de informação.

Segundo Lara (1999), a representação documentária

deve ressaltar a representação como algo que se desenvolve no universo da linguagem e como uma construção que modeliza, a seu modo, a significação, mas cuja apropriação está condicionada ao modo como operacionaliza e veicula tal sistema de significação. Gênero de sistema intermediário de comunicação, as representações através das linguagens documentárias configuram um sistema modelizante secundário que imprime aos universos representados um ponto de vista de organização fundado na previsibilidade de associações e princípios de compartilhamento. (LARA, 1999, p. 161).

Pode-se considerar, então, o conceito de representação da informação como “uma estrutura de significação que agencia a produção do conhecimento”, por meio de “atividades documentárias de representação”, nas quais estão inseridas as LDs, que auxiliam a “decodificação do sistema de significação veiculado”, que, nesta pesquisa, é a representação documentária do domínio da Economia. (LARA, 1999, p.154-161).

Segundo Moraes (2013) a representação da informação e do conteúdo do documento, só é possível pela utilização de uma LD que permite a sua “síntese” temática, ou seja,

pode-se dizer que a área de análise documental de conteúdo pode ser definida como um conjunto de procedimentos de natureza analítico/sintética, que envolve os processos de análise do conteúdo temático dos documentos, e sua síntese, por meio da condensação ou da representação em linguagens documentárias. MORAES (2011, p. 27).

A Linguagem Documentária segundo Gardin (1968), citada por Cintra et al (1994, p. 25), é “um conjunto de termos providos ou não de regras sintáticas, utilizado para representar conteúdos de documentos técnico-científicos, com fins de classificação ou busca retrospectiva de informações”.

Na LD, o “sistema de símbolos” constitui-se do vocabulário controlado e da sintaxe, correspondendo o primeiro ao conjunto de termos e às regras próprias de utilização, sendo a sintaxe a combinação lógica entre os termos. (FUJITA, 2005).

As tipologias das linguagens documentárias podem ser classificadas (ou hierárquicas) e alfabéti-

cas (de indexação ou combinatórias) no que se refere à forma de apresentação dos conceitos e, quanto ao princípio da coordenação, elas podem ser pré-coordenadas ou pós-coordenadas (GUIMARÃES, 1990, VAN SLYPE, 1991, LANCASTER, 1995). No âmbito das linguagens de estrutura hierárquica, os termos relacionam-se entre si a partir da subordinação e “co-subordinação”, constituindo-se como uma relação assimétrica entre dois elementos, em que um é superior ao outro por caráter normativo, isto é, pela correspondência exata existente entre eles (VIZCAYA ALONSO, 1997). (CERVANTES, 2009, p. 35-36).

Estas relações são estruturadas lógica e semanticamente em torno das unidades informacionais (classes e descritores), apresentadas em hierarquias – verticais e horizontais – genéricas, específicas e partitivas, ou seja, Todo/Parte ou Gênero/Espécie, sendo coordenadas por superordenação e subordinação, apresentando, também, relações não-hierárquicas associativas. (CINTRA et al, 1994, p. 30-33).

As hierarquias e as estruturas são organizadas no escopo das LDs, apresentando-se por macroestruturas, que são “as entradas organizadas em ordem alfabética dentro dos campos conceituais que têm a função de cabeçalhos para facilitar e agilizar a consulta”. (CERVANTES, 2009, p. 143).

#### 4. Metodologia

Cervantes (2009), ao investigar os procedimentos metodológicos para construção de Tesouros em áreas de especialidades, relaciona os autores da área e sintetiza as etapas de construção de tesouros, após analisar e utilizar as normas NBR 12676 (ABNT, 1992), NBR 13789 (ABNT, 1997a), NBR 13790 (ABNT, 1997b), ANSI/NISO Z39.19 (2005), ISO 1087 (1996), ISO 1087-1 (2000), ISO 2788 (1986), além das Diretrizes da UNESCO (1993) e do IBICT (1984), cujos princípios norteadores se encontram nos seguintes apêndices de sua pesquisa: A (1) (Etapas de Construção de um Tesouro de acordo com as Diretrizes); e B (2) (Métodos de Compilação de termos: terminologias e definições).

Em sua tese de doutorado sobre “a construção de tesouros com a integração de procedimentos terminográficos”, utilizando-se, portanto, do arcabouço teórico e metodológico da Terminologia, a citada autora realizou a análise e a síntese das etapas de construção de tesouro, agrupando-as por categorias temáticas, segundo os principais autores de cada uma, conforme o Quadro 1 (Anexo 1).

Essa análise e síntese serviram de arcabouço teórico para a elaboração do “Modelo Metodológico Integrado para Construção de Tesouro”,

com a sistematização das suas etapas, conforme o Quadro 2 (Anexo 2), elaborado pela autora, como resultado de sua pesquisa.

Dentre os procedimentos, destacamos os que estão relacionados com a tese (3) recém defendida, de onde se originou este trabalho, que consiste na apresentação da metodologia utilizada, que são a escolha do domínio e da língua do tesouro, a delimitação do subdomínio, o estabelecimento dos limites da pesquisa terminológica temática, assim como a coleta e a classificação dos termos, conceituados e descritos no Quadro 3 (Anexo 3).

Esse modelo, a análise e a pesquisa de Cervantes (2009), portanto, demonstra a importância da delimitação do domínio para a compreensão das estruturas de representação documentária da informação.

Após estas considerações iniciais, aplicamos, experimentalmente, o “Modelo Metodológico Integrado para Construção de Tesouro” de Cervantes (2009), no domínio da Economia.

Assim, as etapas do Modelo: trabalho preliminar, método de compilação, registro de termos e verificação de termos (Esta etapa completa será objeto de estudo em outra pesquisa), possibilitaram o desenvolvimento das etapas desta pesquisa, conforme demonstrado no Quadro 4 (Anexo 4).

## 5. Resultados

Utilizando, portanto, o arcabouço teórico e metodológico da Terminologia, Cervantes (2009) realizou a análise e a síntese das etapas de construção de tesouro, agrupando-as por categorias temáticas, cujo modelo utilizamos para obter os resultados por meio da elaboração da estrutura de representação documentária do domínio da Economia, descritas nas etapas seguintes.

### 5.1. Etapa A - Delimitação do Subdomínio

Segundo Cervantes (2009, p. 147), a delimitação do subdomínio é necessária por não ser recomendável uma pesquisa terminológica de um domínio completo, devido à sua complexidade.

Assim, após definido o domínio que nesta pesquisa é o de Economia, buscamos delimitar o subdomínio por meio dos termos pertinentes nas Linguagens Documentárias (LDs) e nos documentos específicos da área.

A delimitação do subdomínio corresponde à revisão de literatura sobre representação documentária e estruturas dessa representação do-

documentária, assim como a identificação e descrição das LDs e das estruturas de representação documentária em documentos específicos do domínio da Economia.

Inicialmente, pesquisamos os referenciais teóricos da ORC, apresentados por Fujita (2008), que esclarece que seu domínio está sistematizado em seu próprio nome, demonstrado por dois conceitos fundamentais: a Organização do Conhecimento e a Representação do Conhecimento, combinados entre as categorias Ação + Objeto. Em seguida, analisa-se a estrutura da área – ORC – dividida em seus três núcleos básicos: Fundamentos, Organização e Representação. (GUIMARÃES, 2001).

Nesse contexto, abordamos as LDs, como instrumentos de estrutura de Representação Documentária, construídas para possibilitar a comunicação entre os conteúdos dos documentos e os usuários de um sistema de informação.

Após um breve relato epistemológico sobre teorias e normas de construção de LDs, destacando a sistematização de etapas da construção de tesouros por meio do “Modelo Metodológico Integrado para Construção de Tesouro”, com a interface da Terminologia, elaborada por Cervantes (2009), constata-se que domínio é o “subconjunto de uma área determinado por um sistema de noções, sendo a área uma parte do saber cujos limites são definidos segundo um ponto de vista particular de uma ciência ou técnica” (ISO 1087, 2000). Aponta-se, então, para a construção de vocabulários controlados, as LDs, que, nesta pesquisa se configuram como estruturas de representação da informação no domínio da Economia.

Apresenta-se, ainda, as principais características de três LDs do domínio da Economia: o Vocabulário Controlado da USP (VC-USP) (VOCABULÁRIO, 2013; KOBASHI et al, 2014) ; o Tesouro ISOC de Economia (IEDCYT, 2013); e a tradução da Library of Congress Subject Headings (LCSH) utilizada pela Biblioteca Nacional (BN) para, em seguida, descrever as estruturas de representação documentária, contextualizadas no estudo sobre o conceito de informação do domínio da Economia.

Conceitua-se a informação do domínio da Economia, assim como a própria Economia como a “Ciência que estuda a atividade produtiva”, ou seja, a Economia como sendo o somatório de toda produção realizada pelo ser humano para fins de industrialização, comércio e serviços, assim como as escolas que a teorizam e agregam valor à sociedade como um todo, relacionando os conceitos de indicador econômico e indicador de conjuntura.

Em seguida, são identificadas e descritas as estruturas de representação documentária dos documentos específicos do domínio da Economia. São elas os termos econômicos e específicos principais do Boletim de Conjuntura do IPEA (BC-IPEA); os Termos da Área Econômica (TAE) organizados por Gracioso (2002); a Classificação Nacional de Atividades Econômicas, versão 2.0 (CNAE 2.0); e as Categorias principais do sistema de classificação do Journal of Economic Literature (JEL).

## 5.2. Etapa B – Estabelecimento dos limites da pesquisa terminológica temática e coleta do corpus do trabalho terminológico

Na etapa B, são estabelecidos os limites da pesquisa terminológica temática e a coleta do corpus do trabalho terminológico, pela execução da fase de identificação e descrição dos Boletins de Conjuntura Econômica (BCEs) (4) do Brasil e das instituições que produzem a informação do domínio da Economia.

Os Boletins de Conjuntura são documentos importantes para o usuário se manter atualizado com relação à evolução recente da economia e suas perspectivas na medida em que veiculam as interpretações de técnicos especializados quanto a um conjunto amplo de indicadores de diferentes fontes e de diferentes dimensões do sistema econômico, com um nível maior ou menor de antecedência/consequência de fatores macroeconômicos nacionais e internacionais, expressos sob forma de números, índices, taxas, variações mensais ou taxas dessazonalizadas.

Os principais produtores da informação do domínio da Economia, no Brasil, são, na esfera pública, o Instituto Brasileiro de Geografia e Estatística (IBGE), o Instituto de Pesquisa Econômica Aplicada (IPEA), os Ministérios da Fazenda, do Desenvolvimento e do Trabalho, o Banco Nacional de Desenvolvimento Econômico e Social (BNDES) e o Banco Central (BC).

Na esfera não pública, tem-se a Fundação Getúlio Vargas (responsável há longo tempo pelas Contas Nacionais e índices oficiais de preços); as Confederações Patronais (como CNI, FIESP, FECOMÉRCIO); Associações e Sindicatos Patronais (como ANFAVEA, Sindipeças, ELETTROS, ABPO); os grupos editoriais especializados (como Gazeta Mercantil, EXAME, Valor); o SEBRAE e empresas de pesquisa, como IBOPE e NIELSEN.

Assim, a pesquisa terminológica para a coleta do corpus terminológico é delimitada, nos termos extraídos do Boletim de Conjuntura do

IPEA (BC-IPEA), utilizando-se o Novíssimo Dicionário de Economia do Sandroni (2004), enquanto documento da área de Economia.

Desse modo, foi necessário o levantamento dos indicadores e dos modelos de recuperação da informação dos BCEs, por fornecer informações para reflexão sobre como essas estão sendo organizadas e representadas, haja vista a relação de causa e efeito entre a indexação e a recuperação da informação.

Em seguida, aplicamos e descrevemos os resultados das etapas “C” e “D”, que são, respectivamente, o “estabelecimento da árvore de domínio” e a “coleta e classificação dos termos”, e as suas fases correspondentes de: Elaboração da estrutura documentária do domínio da Economia tendo, como parâmetro, as estruturas das LDs e as estruturas dos documentos específicos da área; Compilação do corpus de termos econômicos dos BCEs e das instituições que produzem a informação do domínio da Economia; Cotejamento do corpus de termos econômicos nas LDs; Classificação dos termos econômicos do corpus, selecionados nas categorias correspondentes que compõem a estrutura de representação documentária elaborada com base nas estruturas das linguagens e dos documentos específicos.

## 5.3. Etapa C - Estabelecimento da -árvore de domínio

Nesta etapa, desenvolve-se a fase 4, que trata da elaboração da estrutura documentária do domínio da Economia, tendo como parâmetro as estruturas das LDs e os documentos específicos da área.

Segundo Cervantes (2009, p. 148), “a árvore de domínio representa o conjunto nocional que tem a função de situar o campo nocional a ser estudado”.

Inicialmente, foram agregadas as macroestruturas das LDs e dos documentos específicos da área em uma planilha do Excel, o que nos permitiu realizar as inferências das fases seguintes.

No total, foram agregados 169 termos das suas macroestruturas, sendo 37 do VC-USP, 11 do Tesouro ISOC de Economia, 33 da LCSH-BN, 8 do BC-IPEA, 24 dos Termos da Área Econômica (TAE) organizados por Gracioso (2002), 37 da CNAE 2.0, e 19 da JEL.

Essas macroestruturas, enquanto categorias principais, representam 746 termos do VC-USP, 3195 termos do Tesouro ISOC de Economia, 581 termos do BC-IPEA, 24 termos da área econômica da TAE (Gracioso, 2002), 1301 ter-

mos da CNAE 2.0, 1044 termos da JEL, sem contar com a LCSH-BN, com 33 termos, apesar dela não ter disponibilizado a sua estrutura, totalizando 6924 termos do domínio da Economia.

O estabelecimento da árvore de domínio, portanto, ficou composto pelas macroestruturas das LDs e dos documentos específicos”.

#### 5.4. Etapa D - Coleta e classificação dos termos

A coleta e a classificação de termos, segundo Cervantes (2009, p. 149), são efetuadas “a partir do corpus do trabalho terminológico selecionado”, respeitando-se o contexto do enunciado do qual o termo é extraído.

Foi realizada uma compilação dos termos econômicos dos BCEs, tendo, como parâmetro inicial, o BC-IPEA, com 581 termos, que foram agregados para análise dos conteúdos dos BCEs em que foram definidas as categorias principais que constituíram o corpus de termos econômicos para cotejamento nas LDs.

O cotejamento serviu para a fase de “verificação de termos”, que os admite e/ou exclui, de acordo com sua especificidade, seguindo as etapas de verificação, classificação e confirmação de termos e o uso do vocabulário de especialidade para estabelecer as relações entre os descritores e sua organização.

Os 581 termos do BC-IPEA foram confirmados nas LDs utilizadas nesta pesquisa e validados enquanto representações que, em conjunto com os documentos específicos da área, constituem a árvore de domínio da Economia.

Para viabilizar o cotejamento dos “termos identificados” nos boletins e a identificação dos “termos selecionados” nas LDs, foi criado o “Formulário de identificação e seleção de termos” (Quadro 4, no Anexo 4) com o objetivo de realizar a coleta de dados e, conseqüentemente, a sua análise, servindo de subsídio complementar aos resultados desta metodologia por meio da mensuração e confirmação da representatividade dos termos.

O que se observou na coleta de dados para cotejamento do corpus de termos econômicos nas LDs, é que a maioria dos termos estão representados nas LDs, sendo que dos dezenove boletins pesquisados, dezoito tiveram, aproximadamente, 90% dos termos confirmados, conforme exemplo do boletim Política Econômica em Foco, da UNICAMP. (Quadro 5).

A fase de cotejamento possibilitou realização de uma “classificação provisória dos termos e uma exploração sumária das noções” (CERVANTES,

2009, p. 149), pela classificação dos termos econômicos do corpus selecionado nas categorias correspondentes que compõem, a estrutura de representação documental elaborada com base nas estruturas das LDs e dos documentos específicos.

Logo após, fez-se a ordenação de todos os termos e sua classificação por ordem alfabética, tomando-se o cuidado de associar cada termo à sua respectiva origem documental, o que torna possível organizá-los e visualizá-los por aproximação terminológica.

A princípio, a fase de finalização da estrutura de representação documental do domínio da Economia, que contém os termos econômicos do corpus selecionado, seria a totalidade das fases anteriores, que representam a informação do domínio de Economia.

No entanto, indo além das pretensões iniciais, percebemos, de forma inédita, que os termos do BC-IPEA se configuravam em um tesouro específico dentro do domínio de Economia, permitindo apresentar a última etapa do “Modelo Metodológico Integrado de Construção de Tesouros” (CERVANTES, 2009), que é a “Forma de apresentação de um tesouro”, concluindo toda sua aplicação e experimentação.

Ou seja, após o desenvolvimento e análise das fases anteriores, percebemos que os termos agregados dos BCE’s constituem um subdomínio do domínio da Economia, que denominamos de “Indicadores Econômicos”.

Ao constatarmos que no VC-USP existe a categoria CH712.22 LEVANTAMENTO ECONÔMICO, vazia, decidimos criar e inserir a categoria CH712.22.1 INDICADOR ECONÔMICO, incluindo os 71 termos agregados (Quadro 6) segundo os parâmetros lógicos do próprio vocabulário

## 6. Considerações finais

Dessa forma, conclui-se que o “Modelo Metodológico Integrado para Construção de Tesouro”, de Cervantes (2009) é um instrumento consolidado e exequível para construção de LDs.

Esse instrumento destaca-se, principalmente, por apontar a importância das primeiras etapas de delimitação do domínio da área e subárea que se pretende representar, assim como as demais etapas e fases, fundamentadas em critérios teóricos e metodológicos da Terminologia e da Ciência da Informação.

O modelo de Cervantes (2009) permitiu não só a delimitação do domínio da Economia, como, também, descobrir e apontar o subdomínio dos

“Indicadores Econômicos”, além da sua própria aplicação experimental, que não havia sido realizada ainda, caracterizando a originalidade desta pesquisa e a eficiência do modelo.

## Notas

(1) Ver em Cervantes, 2009, p. 200, o Apêndice A - Quadro 6 – Síntese das Etapas de Construção de um Tesouro segundo as Diretrizes IBICT (1984); Diretrizes UNESCO (1993); e Diretrizes ANSI/NISO (2005).

(2) Ver em Cervantes, 2009, p. 207, o Apêndice B - Quadro 7 – Métodos de Compilação de termos: terminologias e definições.

(3) CAVATI SOBRINHO, Heliomar. **A representação documentária do domínio da Economia: análise de estruturas de representação em linguagens documentárias e documentos específicos de economia**. 149 f. Tese (Doutorado) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2014.

(4) Ver Referências dos Boletins de Conjuntura.

## Referencias

- AMERICAN NATIONAL STANDARDS INSTITUTE; NATIONAL INFORMATION STANDARDS ORGANIZATION (2008). *ANSI/NISO Z39.19-2005: guidelines for the construction, format, and management of monolingual controlled vocabularies*. Bethesda, Ma: NISO Press, 2005. 184 p. Recuperado 15-02-2008, de <http://www.niso.org/standards/resources/Z39-19-2005.pdf>
- ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS.(1992). *NBR 12676: Métodos para análise de documentos – Determinação de seus assuntos e seleção de termos de indexação*. Rio de Janeiro: ABNT.
- ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS (1997a) *NBR 13789: Terminologia – Princípios e métodos – Elaboração e apresentação de normas de terminologia*. Rio de Janeiro: ABNT.
- ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS (1997b). *NBR 13790: Terminologia – Princípios e métodos – Harmonização de conceitos e termos*. Rio de Janeiro: ABNT.
- BOLETIM DE CONJUNTURA (2004) Rio de Janeiro: IPEA, n. 65,(Junho, 2004).
- CAVATI, Heliomar (2005) *Informação econômica no Brasil: uma análise dos Boletins de Conjuntura Econômica*. 2005. 141 p. Dissertação (Mestrado em Ciência da Informação) – Pontifícia Universidade Católica de Campinas.
- CERVANTES, Brígida Maria Nogueira (2009). **A construção de tesouros com a integração de procedimentos terminográficos**. 2009. 209f. Tese (Doutorado) – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília.
- CINTRA, A. M. M. et al.(1994) *Para entender as linguagens documentárias*. 2. ed. São Paulo: Polis.
- CLASSIFICAÇÃO nacional de atividades econômicas - CNAE: versão 2.0 (2007). Rio de Janeiro: IBGE. 423 p. Recuperado 01-01-2014, de <http://www.ibge.gov.br/home/estatistica/economia/classificacoes/cnae2.0/cnae2.0.pdf>
- FUJITA, M. S. L. (1992). *Linguagem documentária em Odontologia: uma aplicação do sistema de indexação PRECIS*. 1992. Tese (Doutorado em Ciências da Comunicação) – Escola de Comunicação e Artes, Universidade de São Paulo, São Paulo. 3v.
- FUJITA, M. S. L. (2005). *Linguagens documentárias alfabéticas em análise documentária: aspectos de estrutura e funcionalidade*. Marília. Material didático.
- \_\_\_\_\_. (2008). Organização e representação do conhecimento no Brasil: análise de aspectos conceituais e da produção científica do ENANCIB no período de 2005 a 2007. *Tendências da Pesquisa Brasileira em Ciência da Informação*, v. 1, n. 1, p. 1-32.
- \_\_\_\_\_. (2013). Representação documentária no processo de indexação com o modelo de leitura documentária para textos científicos e livros: uma abordagem cognitiva com protocolo verbal. *PontodeAcesso*, Salvador, v. 7, n.1, (Abril 2013). 42-66.
- GOMES, H. E. (Coord.)(1990). *Manual de elaboração de tesouros monolíngües*. Brasília: MEC/MCT, Programa Nacional de Bibliotecas das Instituições de Ensino Superior.
- GRACIOSO, Luciana de Souza (2002). *Disseminação de informações estatísticas no Brasil: práticas e políticas das agências estaduais de estatísticas*. 2002. 191 p. Dissertação (Mestrado em Ciência da Informação) – Pontifícia Universidade Católica de Campinas, 2002.
- GUIMARÃES, J. A. C. (1990). A recuperação temática da informação. *Revista Brasileira de Biblioteconomia e Documentação*, São Paulo, v.16, n.3/4, (jan./dez. 1990), 112-130.
- GUIMARÃES, J. A. C. (2001). Perspectivas de ensino e pesquisa em organização do conhecimento em cursos de Biblioteconomia: uma reflexão. In: CARRARA, K. (Org.). *Educação, Universidade e Pesquisa*. Marília: Unesp-Marília-Publicações; São Paulo: FAPESP.
- GUINCHAT, Claire; MENOU, Michel. (1994). *Introdução geral às ciências e técnicas da informação e documentação*. Brasília: IBICT.
- IEDCYT (2013). Recuperado 10-12-2013, de <http://www.cindoc.csic.es>
- INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA (1984). *Diretrizes para elaboração de tesouros monolíngües*. Brasília: IBICT.
- INTERNATIONAL ORGANIZATION FOR STANDARDIZATION (1986). *ISO 2788-1986: Documentation - Guidelines to establishment and development of monolingual thesauri*. Geneva : ISO.
- INTERNATIONAL ORGANIZATION FOR STANDARDIZATION (1996). *ISO 1087 – Terminologia – Vocabulário*. Proposta de versão brasileira. [Rio de Janeiro: ABNT].
- INTERNATIONAL ORGANIZATION FOR STANDARDIZATION (2000). *ISO 1087-1 – Terminology work – Vocabulary – pt. 1: Theory and application*. [Geneva] : ISO, 1995].
- KOBASHI, N. Y.; LIMA, V. M. A.; LEME, M. A. T (Orgs)(2006). *Manual de indexação de assuntos com uso do vocabulário controlado USP: versão preliminar*. São Paulo: USP, 2006. Recuperado 1-1-2014, de <http://www.usp.br/sibi/AreaTecnica/manuais/Vocabulario.pdf>
- LARA, M. L. G. de (1999). *Representação e linguagens documentárias: bases teórico-metodológicas*. 1999. Tese (Doutorado em Ciências da Comunicação) – Universidade de São Paulo, Escola de Comunicações e Artes, São Paulo.
- LANCASTER, F. W (1995). **El control del vocabulario en la recuperación de información**. Tradução de Alejan-

- dro de la Cueva Martín. Valência: Universitat de València. (Educació. Materials,12).
- MORAES, João Batista Ernesto de. (2001). *A questão do aboutness no texto narrativo de ficção: perspectivas metodológicas para a Ciência da Informação*. Marília, 2011. 93f. Tese (Livre-Docência em Lingüística e Documentação) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista.
- RONDEAU, G.(1984). *Introduction à la terminologie*. 2.ed. Québec, Canadá: Gaëtan Morin.
- SANDRONI, Paulo(2004). *Novíssimo dicionário de economia*. 14. ed. São Paulo: Best Seller, 2004.
- UNESCO (1993). **Diretrizes para o estabelecimento e desenvolvimento de tesouros monolíngües**. [2ª ed. rev. por Derek Austin e Peter Dale]; Trad. de Bianca Amaro de Melo; rev. de Lígia Maria Café de Miranda. Brasília: IBICT; SENAI.
- VOCABULÁRIO (2001) Controlado SIBi/USP: base de dados em língua portuguesa para indexação e recuperação da informação. São Paulo: USP. Recuperado 15-12-2015, de <http://www.usp.br/sibi/AreaTecnica/index2.html>
- LEVANTAMENTO DE CONJUNTURA (2004). São Paulo: FIESP, (out. 2004). Recuperado 03-01-2005, de [http://www.fiesp.com.br/pesquisas\\_estatisticas/](http://www.fiesp.com.br/pesquisas_estatisticas/)
- NÍVEL DE EMPREGO (2004). São Paulo: FIESP, (out. 2004). Recuperado 03-01-2005, de [http://www.fiesp.com.br/pesquisas\\_estatisticas/](http://www.fiesp.com.br/pesquisas_estatisticas/)
- POLÍTICA ECONÔMICA EM FOCO (2004). Campinas: UNICAMP, n. 4, (maio/out. 2004). Recuperado 02-01-2005, de [http://www.eco.unicamp.br/asp-scripts/boletim\\_cecon/boletim\\_cecon.asp](http://www.eco.unicamp.br/asp-scripts/boletim_cecon/boletim_cecon.asp)
- RELATÓRIO DE INFLAÇÃO (2004). Brasília: BACEN, v. 6, n. 2, (jun. 2004).
- RETAIL HIGHLIGHTS (2005). São Paulo: ACNielsen, (nov. 2004). Recuperado 02-01-2005, de <http://www.acnielsen.com.br/>

## Referencias dos boletins

- ANÁLISE DO MERCADO DE TRABALHO FORMAL (2004). Brasília: MTE, jun 2004. Recuperado 02-06-2004, de <http://www.mtb.gov.br/Empregador/CAGED/Estatistica/Textos/2004/Junho/Analise.asp>
- BOLETIM DE CONJUNTURA (2004). Rio de Janeiro: IPEA, n. 65, (jun. 2004).
- BOLETIM DE INFORMAÇÕES FIPE (2004). São Paulo: FIPE, n. 285, jun. 2004. Recuperado 02-01-2005, de <http://www.fipe.com.br/publicacoes/bif.asp?codigo=1>
- BOLETIM DO BANCO CENTRAL DO BRASIL (2004). Brasília: BACEN, v. 40, n. 6, (jun. 2004).
- CARTA IEDI (2004). São Paulo: IEDI, n. 136, (31 dez. 2004). Recuperado 02-01-2005, de <http://www.iedi.org.br/cgi/cgilua.exe/sys/start.htm?sid=50>
- CONJUNTURA ECONÔMICA (2004). Rio de Janeiro: FGV, v. 58, n. 12, (dez. 2004).
- DÍVIDA PÚBLICA MOBILIÁRIA FEDERAL INTERNA E MERCADO ABERTO. Brasília: Ministério da Fazenda, nov. 2004. Disponível em: [http://www.stn.fazenda.gov.br/hp/downloads/divida\\_publica/relatorio\\_nov04.pdf](http://www.stn.fazenda.gov.br/hp/downloads/divida_publica/relatorio_nov04.pdf)
- ECONOMIA & CONJUNTURA (2004). Rio de Janeiro: IE/UFRJ, v. 4, n. 59, (dez. 2004). Recuperado 04-01-2005, de <http://www.ie.ufrj.br/conjuntura/pdfs/Dez2004.pdf>
- EVOLUÇÃO DA CONJUNTURA ECONÔMICA (2004). Brasília: CNC, v. 3, n. 36, (dez. 2004). Recuperado 03-01-2005, de <http://www.cnc.com.br/pes/publ/Peri/evolconj/evolconj.pdf>
- INDICADORES DIESP (2004). São Paulo: FUNDAP/DIESP, v. 13, n. 98, (jan./fev. 2004).
- INDICADORES INDUSTRIAIS CNI (2004). Brasília: CNI, v. 15, n. 7, (jul. 2004).
- INFORME SEBRAE-SP (2004). São Paulo: Sebrae, (dez. 2004). Recuperado 04-01-2005, de <http://www.sebraesp.com.br/principal/conhecendo%20a%20mpe/indicadores%20de%20conjuntura/default.aspx>
- INFORMAÇÕES ECONÔMICAS. São Paulo: IEA-SP, v. 34, n. 6, jun. 2004. Disponível em: <http://www.iea.sp.gov.br/out/iie.php> Acesso em: 2 nov. 2004.

## Anexo 1. Síntese das etapas de construção de tesouros segundo autores.

| ETAPAS DE CONSTRUÇÃO DE TESAURO SEGUNDO AUTORES   |   |
|---|---|
| CATEGORIAS TEMÁTICAS  | AUTORES   |
| <b>1) Fase do planejamento</b><br>tipo de usuário, suas necessidades; abrangência e nível de especificidade do tesouro; identificação de fontes de procedimentos e de coleta de termos.   | Batty (1989)<br>Gomes (1990)<br>Fujita (1992)<br>Gomes ([2004])   |
| <b>2) Formas/métodos de Compilação de termos</b><br>Dedutivo<br>Indutivo<br>Combinação de Métodos (Dedutivo/Indutivo)   | Aitchison; Gilchrist(1979)<br>Lancaster (1987)<br>Batty (1989)<br>Gomes (1990)<br>Fujita (1992)<br>Gomes ([2004])                   |
| <b>3) Compilação de termos</b><br>a) coleta - registro e seleção dos termos compilados e b) validação – registro do vocabulário básico; coleta e validação de termos.   | Aitchison; Gilchrist (1979)<br>Lancaster (1987)<br>Batty (1989)<br>Gomes (1990)<br>Fujita (1992)<br>Gomes ([2004])                  |
| <b>4) Estabelecimento de relações entre termos/ Categorização</b> estruturação de conceitos com controle terminológico dos termos; ordenação dos termos; estabelecimento de categorias elementares; organização dos termos básicos em categorias (critério de afinidade semântica); definição de subcategorias; estabelecimento de relações entre termos. | Aitchison; Gilchrist (1979)<br>Lancaster (1987)<br>Batty (1989)<br>Gomes (1990)<br>Fujita (1992)<br>Fujita (1998)<br>Gomes ([2004]) |
| <b>5) Especificidade</b><br>Estabelecimento de limites de especificação/ dependendo da complexidade do vocabulário.   | Lancaster (1987)<br>Gomes (1990)<br>Gomes ([2004])  |
| <b>6) Uso de equipamento informático para processamento de dados</b> estruturação automática das partes alfabética e sistemática do tesouro (etapa 7); produção de uma estrutura final (etapa 7)  | Lancaster (1987)<br>Gomes (1990)<br>Fujita (1992)<br>Gomes ([2004])   |
| <b>7) Formas de Apresentação</b> alfabética; sistemática; alfabética/classificada; facetada. estruturação automática das partes alfabética e sistemática do tesouro; produção de uma estrutura final.   | Lancaster (1987)<br>Gomes (1990)<br>Fujita (1992)<br>Gomes ([2004])   |

Fonte: Cervantes, 2009, p. 113.

## Anexo 2. Sistematização de etapas da construção de tesouros

| MODELO METODOLÓGICO INTEGRADO PARA CONSTRUÇÃO DE TESAURO   |   |
|--|---|
| Sistematização de etapas da construção de tesouros (normalização, literatura e tesouros) - Procedimentos terminográficos |   |
| <b>1. Trabalho preliminar</b><br>(Orientações gerais/Uso de equipamento automático de processamento de dados)            | <ul style="list-style-type: none"> <li>- escolha do domínio e da língua do tesouro;</li> <li>- delimitação do subdomínio;</li> <li>- estabelecimento dos limites da pesquisa terminológica temática;</li> <li>- consulta a especialista do domínio/subdomínio.</li> </ul>   |
| <b>2. Método de compilação</b><br>(Abordagem de compilação)  | <ul style="list-style-type: none"> <li>- coleta do <i>corpus</i> do trabalho terminológico;</li> <li>- estabelecimento da árvore de domínio;</li> <li>- expansão da representação do domínio escolhido.</li> </ul>  |
| <b>3. Registro de termos</b>   | <ul style="list-style-type: none"> <li>- coleta e classificação de termos.</li> </ul>   |
| <b>4. Verificação de termos</b><br>(Admissão e exclusão de termos /Especificidade)                                       | <ul style="list-style-type: none"> <li>- verificação, classificação e confirmação de termos;</li> <li>- elaboração de definições;</li> <li>- uso do vocabulário de especialidade para o estabelecimento de relações entre os descritores e de relações entre descritores e não descritores.</li> <li>- organização das relações entre descritores.</li> </ul> |
| <b>5. Forma de apresentação de um tesouro</b>  | <ul style="list-style-type: none"> <li>- trabalhos de apresentação do tesouro.</li> </ul>   |

Fonte: Cervantes, 2009, p. 163.

## Anexo 3. Descrição das etapas de construção de tesouros

| DESCRIÇÃO DAS ETAPAS DE CONSTRUÇÃO DE TESAUROS                        |  |
|---|--|
| <b>Escolha do domínio e da língua do tesouro</b>                      | A escolha do domínio e da língua de trabalho, geralmente, são estabelecidas em função das necessidades dos usuários.   |
| <b>Delimitação do subdomínio</b>                                      | Recomenda-se não desenvolver uma pesquisa terminológica sobre um domínio completo: por um lado, devido à complexidade e amplitude que supõe uma tal tarefa; e, por outro lado, porque em grande parte do tempo, um domínio compreende não somente uma rede nocional que lhe é própria, mas também numerosas redes nocionais conexas.   |
| <b>Estabelecimento dos limites da pesquisa terminológica temática</b> | O limite da extensão da pesquisa terminológica, quanto ao número aproximado de termos, é estabelecido em função dos objetivos propostos, da disponibilidade de tempo e de meios financeiros. Desse modo, pode-se escolher um levantamento básico, compilando-se, em média, 300 termos, ou exaustivo, por volta de 2.500 termos.  |
| <b>Coleta do <i>corpus</i> do trabalho terminológico</b>              | A etapa da coleta do <i>corpus</i> do trabalho terminológico tem a finalidade de reunir os documentos especializados necessários para o desenvolvimento da pesquisa terminológica. Rondeau (1984, p. 50-51) estabelece oito categorias de documentos de conteúdo terminológico, ou seja, nos quais se encontram termos: 1) normas internacionais ou nacionais; 2) manuais, catálogos, guias de utilização de produtos, entre outros; 3) livros e revistas especializados; anais de eventos científicos, relatórios de pesquisa, teses, entre outros 4) vocabulários, <i>thesaurus</i> , glossários, léxicos; 5) dicionários gerais e especializados, de língua ou enciclopédicos, enciclopédias, entre outros; 6) bancos de termos, fichários automatizados ou não; 7) consulta a especialistas da área; 8) bibliografias ou listas relacionadas com o domínio. Os materiais utilizados como fontes que dão origem ao <i>corpus</i> do trabalho terminológico devem respeitar os princípios da atividade terminológica no que tange à confiabilidade e à representatividade. |
| <b>Estabelecimento da árvore de domínio</b>                           | A árvore de domínio representa o conjunto nocional que tem a função de situar o campo nocional a ser estudado. Cabe alertar que, antes de estabelecer a árvore de domínio, o pesquisador deverá consultar os seguintes documentos: sistemas de classificação, glossários, entre outros. Alerta-se, ainda, que, em alguns domínios, esses instrumentos são até abundantes, mas, em outros domínios, podem não existir.  |

|  |   |
|--|---|
| <b>Coleta e classificação dos termos</b> | A coleta de termos efetua-se a partir do <i>corpus</i> do trabalho terminológico selecionado. Consiste, geralmente, em fazer uma leitura do texto, assinalando as unidades terminológicas a extrair. Essa operação requer da parte do pesquisador algum conhecimento metodológico do trabalho terminológico e também algum conhecimento sobre o domínio ou subdomínio. De acordo com a norma ISO 1087-1 (2000), que estabelece os critérios para a coleta dos termos e recorte do contexto de uso, o contexto é o “enunciado no qual figura o termo estudado” ou parte de um texto no qual ocorre o termo. Nesse sentido, o contexto tem um papel fundamental nas operações de coleta dos termos porque permite reduzir os riscos de erros no momento da sua identificação e recorte. |
|--|---|

Fonte: Rondeau (1984) citado por Cervantes, 2009, p. 147-149, adaptado.

#### Anexo 4. Formulário de identificação e seleção de termos

| BOLETIM DE CONJUNTURA 1 |                     |    |      |    |         |    |
|-------------------------|---------------------|----|------|----|---------|----|
| Termos Identificados    | Termos Selecionados |    |      |    |         |    |
|                         | VC-USP              | Qt | ISOC | Qt | LCSH-BN | Qt |
| TOTAL                   |                     |    |      |    |         |    |

Fonte: Elaborado pelo autor.

#### Anexo 5. Cotejamento do boletim Política Econômica em Foco da UNICAMP

| POLÍTICA ECONÔMICA EM FOCO DA UNICAMP |                        |    |                        |    |                        |    |
|---------------------------------------|------------------------|----|------------------------|----|------------------------|----|
| Termos Identificados                  | Termos Selecionados    |    |                        |    |                        |    |
|                                       | VC-USP                 | Qt | ISOC                   | Qt | LCSH-BN                | Qt |
| Economia internacional                | Economia internacional | 1  | Economía Internacional | 1  | Economia internacional | 1  |
| Setor externo                         |                        | 0  | Sector exterior        | 1  |                        | 0  |
| Moedas                                | Moeda (Economia)       | 1  |                        | 0  | Moedas                 | 1  |
| Finanças privadas                     | Finanças privadas      | 1  |                        | 0  |                        | 0  |
| Finanças públicas                     | Finanças públicas      | 1  |                        | 0  | Finanças públicas      | 1  |
| Produção                              | Produção (Economia)    | 1  | Política de producción | 1  | Produção               | 1  |
| Renda                                 | Renda                  | 1  | Inglesos               | 1  | Renda                  | 1  |
| Emprego                               | Emprego                | 1  | Empleo                 | 1  | Emprego                | 1  |
| TOTAL: 8                              |                        | 7  |                        | 5  |                        | 6  |

Fonte: Elaborado pelo autor.

#### Anexo 6. Proposta de estrutura do subdomínio Indicadores Econômicos

| REPRESENTAÇÃO DOCUMENTÁRIA DO SUBDOMÍNIO INDICADORES ECONÔMICOS |  |
|---|--|
| CÓDIGO  | ASSUNTO                                  |
| CH712   | ECONOMIA                                 |
| CH712.22  | LEVANTAMENTOS ECONÔMICOS                 |
| CH712.22.1  | INDICADORES ECONÔMICOS                   |
| CH712.22.1.1  | <b>NÍVEL DE ATIVIDADE</b>                |
| CH712.22.1.1.1  | INDICADORES CONJUNTURAIS DA INDÚSTRIA    |
| CH712.22.1.1.2  | INDICADORES DE PRODUÇÃO INDUSTRIAL       |
| CH712.22.1.1.3  | INDICADORES CONJUNTURAIS DA AGROPECUÁRIA |
| CH712.22.1.1.4  | INDICADORES DE PRODUÇÃO AGROPECUÁRIA     |
| CH712.22.1.1.5  | SERVIÇOS                                 |
| CH712.22.1.1.6  | COMÉRCIO                                 |

Cavati Sobrinho, Heliomar; Fujita, Mariângela Spotti Lopes. *Construção de macroestrutura de categorias de Linguagem Documentária em Economia: proposta de inovação metodológica*. En XII Congreso ISKO España y II Congreso ISKO España-Portugal, 19-20 de noviembre, 2015, Organización del conocimiento para sistemas de información abiertos. Murcia: Universidad de Murcia.

| <b>REPRESENTAÇÃO DOCUMENTÁRIA DO SUBDOMÍNIO INDICADORES ECONÔMICOS</b> |   |
|--|---|
| <b>CÓDIGO</b>  | <b>ASSUNTO</b>  |
| CH712.22.1.1.7   | PRODUTO INTERNO BRUTO – PIB   |
| CH712.22.1.1.8   | FORMAÇÃO BRUTA DE CAPITAL FIXO (FBCF) E TAXA DE INVESTIMENTO (FBCF/PIB)                                 |
| CH712.22.1.2   | <b>EMPREGO E RENDA</b>  |
| CH712.22.1.2.1   | INDICADORES CONJUNTURAIS DA INDÚSTRIA   |
| CH712.22.1.2.2   | INDICADORES DE PRODUÇÃO INDUSTRIAL  |
| CH712.22.1.2.3   | INDICADORES CONJUNTURAIS DA AGROPECUÁRIA  |
| CH712.22.1.2.4   | INDICADORES DE PRODUÇÃO AGROPECUÁRIA  |
| CH712.22.1.2.5   | SERVIÇOS  |
| CH712.22.1.2.6   | COMÉRCIO  |
| CH712.22.1.2.7   | PRODUTO INTERNO BRUTO – PIB   |
| CH712.22.1.2.8   | FORMAÇÃO BRUTA DE CAPITAL FIXO (FBCF) E TAXA DE INVESTIMENTO (FBCF/PIB)                                 |
| CH712.22.1.3   | <b>INFLAÇÃO</b>   |
| CH712.22.1.3.1   | INDICADORES DE PREÇOS, JUROS E CÂMBIO   |
| CH712.22.1.3.2   | IGPs  |
| CH712.22.1.3.3   | IPCA  |
| CH712.22.1.3.4   | IPCA E SEUS GRUPOS  |
| CH712.22.1.3.5   | IPCA-15   |
| CH712.22.1.3.6   | IPC   |
| CH712.22.1.3.7   | IPC E NÚCLEO DO IPC   |
| CH712.22.1.3.8   | INCC  |
| CH712.22.1.4   | <b>SETOR EXTERNO</b>  |
| CH712.22.1.4.1   | BALANÇO DE PAGAMENTOS   |
| CH712.22.1.4.2   | SERVIÇOS E RENDAS   |
| CH712.22.1.4.3   | BALANÇA COMERCIAL BRASILEIRA: EXPORTAÇÕES, IMPORTAÇÕES E SALDO EFETIVOS                                 |
| CH712.22.1.4.4   | EXPORTAÇÕES E IMPORTAÇÕES POR SETORES   |
| CH712.22.1.4.5   | ÍNDICES DA TAXA DE CÂMBIO EFETIVA REAL PARA O TOTAL DAS EXPORTAÇÕES E PARA EXPORTAÇÕES DE MANUFATURADOS |
| CH712.22.1.4.6   | RELAÇÃO CÂMBIO/SALÁRIO  |
| CH712.22.1.4.7   | ÍNDICE DE PREÇO E QUANTUM DE EXPORTAÇÃO   |
| CH712.22.1.5   | <b>ECONOMIA INTERNACIONAL</b>   |
| CH712.22.1.5.1   | COTAÇÕES DAS MOEDAS INTERNACIONAIS  |
| CH712.22.1.5.2   | ESTADOS UNIDOS: TAXA BÁSICA DE JUROS – FED FUNDS  |
| CH712.22.1.5.3   | TAXAS DE JUROS INTERNACIONAIS E ÍNDICE DE AÇÕES   |
| CH712.22.1.5.4   | ESTADOS UNIDOS: ÍNDICE DE PREÇOS  |
| CH712.22.1.6   | <b>POLÍTICA MONETÁRIA</b>   |
| CH712.22.1.6.1   | COEFICIENTE DE MONETIZAÇÃO COMO PROPORÇÃO DO PIB  |
| CH712.22.1.6.2   | SPREAD DE RISCO SOBERANO – BRASIL E BÔNUS GLOBAIS   |
| CH712.22.1.6.3   | SELIC   |
| CH712.22.1.6.4   | TAXA DE JUROS PREFIXADA – MÉDIAS MENSIS   |
| CH712.22.1.6.5   | TAXA DE JUROS   |
| CH712.22.1.6.6   | OPERAÇÕES DE CRÉDITO DO SISTEMA FINANCEIRO  |
| CH712.22.1.6.7   | RISCO BRASIL e dos PAÍSES EMERGENTES  |
| CH712.22.1.7   | <b>POLÍTICA FISCAL</b>  |
| CH712.22.1.7.1   | ARRECADAÇÃO TRIBUTÁRIA  |
| CH712.22.1.7.2   | EVOLUÇÃO DAS DESPESAS FISCAIS DO TESOURO NACIONAL   |
| CH712.22.1.7.3   | NECESSIDADES DE FINANCIAMENTO DO TESOURO NACIONAL   |
| CH712.22.1.7.4   | NECESSIDADES DE FINANCIAMENTO DO SETOR PÚBLICO  |
| CH712.22.1.7.5   | DÍVIDA LÍQUIDA TOTAL DO SETOR PÚBLICO   |
| CH712.22.1.7.6   | DÍVIDA INTERNA E EXTERNA DO SETOR PÚBLICO   |
| CH712.22.1.7.7   | RECEITA DO IMPOSTO SOBRE CIRCULAÇÃO DE MERCADORIAS (ICMS) – BRASIL                                      |

| <b>REPRESENTAÇÃO DOCUMENTÁRIA DO SUBDOMÍNIO INDICADORES ECONÔMICOS</b> |   |
|--|---|
| <b>CÓDIGO</b>  | <b>ASSUNTO</b>                              |
| CH712.22.1.8   | <b>PROJEÇÕES</b>                            |
| CH712.22.1.8.1   | PROJEÇÕES PARA NÍVEL DE ATIVIDADES          |
| CH712.22.1.8.2   | PROJEÇÕES PARA EMPREGO                      |
| CH712.22.1.8.3   | PROJEÇÕES PARA SALÁRIOS                     |
| CH712.22.1.8.4   | PROJEÇÕES PARA UTILIZAÇÃO DE CAPACIDADE     |
| CH712.22.1.8.5   | PROJEÇÕES PARA TAXA DE INVESTIMENTO         |
| CH712.22.1.8.6   | PROJEÇÕES PARA INFLAÇÃO                     |
| CH712.22.1.8.7   | PROJEÇÕES PARA CÂMBIO                       |
| CH712.22.1.8.8   | PROJEÇÕES PARA JUROS                        |
| CH712.22.1.8.9   | PROJEÇÕES PARA USOS E FONTES DE RECURSOS    |
| CH712.22.1.8.10  | PRODUTO INTERNO BRUTO                       |
| CH712.22.1.8.11  | PREVISÕES E ESTIMATIVAS DE SAFRAS AGRÍCOLAS |

Fonte: Elaborado pelo autor.

---

# Política de indexação para bibliotecas universitárias: um estudo com docentes da área de Organização e Representação do Conhecimento

*Indexing policy for university libraries: a study with academical teachers of the Knowledge Organization and Representation*

---

**Paula Regina Dal'Evedove (1), Mariângela Spotti Lopes Fujita (2)**

(1) Universidade Federal de Pernambuco/UFPE. Av. da Arquitetura, s/n, Campus Universitário, CEP 50740-550 – Recife/PE. p.dallevedove@gmail.com.

(2) Universidade Estadual Paulista, Av. Hygino Muz-zi Filho, 737, Campus Universitário, CEP. 17525900, Marília, SP, Brasil, mariangelaslf57@gmail.com

## Resumo

Buscou-se contribuir com novos subsídios teóricos e metodológicos acerca de políticas de indexação para bibliotecas universitárias. Para tanto, foram aplicados questionários com docentes da área de Organização e Representação do Conhecimento que ministram disciplinas relacionadas ao Tratamento Temático da Informação nos cursos de graduação em Biblioteconomia no Brasil, a fim de obter o discurso individual e coletivo dos sujeitos participantes acerca dos aspectos a serem considerados na elaboração de políticas de indexação. De modo geral, os resultados apontam a necessidade da política de indexação prever questões éticas relacionadas à atuação do bibliotecário indexador; o encaminhamento de estudos interdisciplinares com a Gestão do Conhecimento visando avançar nas pesquisas destinadas ao tema e contribuir para a produção de novos conhecimentos; bem como o reconhecimento da comunidade usuária interna e externa como ponto central dos esforços empreendidos na elaboração de diretrizes de política de indexação.

**Palavras chave:** Organização e Representação do Conhecimento. Tratamento Temático da Informação. Política de indexação. Bibliotecas universitárias. Questionários..

## 1. Introdução

A política de indexação é um tema pouco trabalhado pela literatura especializada da área de Organização e Representação do Conhecimento, fato que potencializa novas possibilidades e olhares reflexivos sobre a questão. Reconhecendo a responsabilidade que cerca toda e qualquer discussão sobre política de indexação, não houve nesta pesquisa a pretensão de esgo-

## Abstract

We sought to contribute new theoretical and methodological subsidies about indexing policies for university libraries. Therefore, questionnaires were applied to teachers of the Organization and Knowledge Representation area, that teach disciplines related to the Thematic Treatment Information in undergraduate courses in Library Science in Brazil in order to get collective and individual discourse of the participants about the subject aspects consider in developing indexing policies. Generally, the results indicate the need for guidelines indexing policy predict ethical issues related to the activity of the indexer librarian; forwarding interdisciplinary studies with knowledge management in order to advance research for the subject and contribute to the production of new knowledge; as well as recognition of internal and external user community as the focal point of efforts in developing indexing policy guidelines.

**Keywords:** Knowledge Organization and Representation. Information Subject Treatment. Indexing policy. University libraries. Questionnaires.

tar o assunto, nem seria tal empreendimento possível, pois esta prática contempla o acompanhamento de fenômenos que se revestem de aspectos sociais, cognitivos e culturais – o fazer profissional, o sistema de informação e a comunidade usuária. Todavia, admitiu-se a necessidade de novas alternativas de produção do conhecimento sobre política de indexação que sejam construídas coletivamente.

Nesse sentido, a presente pesquisa visa contribuir com subsídios teóricos e metodológicos para a definição de políticas de indexação direcionadas às bibliotecas universitárias brasileiras a partir da esfera científica. Para tanto, foram aplicados questionários com docentes da área de Organização e Representação do Conhecimento que ministram disciplinas relacionadas ao universo científico do Tratamento Temático da Informação nos vinte e sete cursos de graduação em Biblioteconomia vigentes no Brasil. Portanto, o interesse maior foi verificar a opinião dos docentes sobre os principais elementos, na esfera teórica e metodológica, que devam ser considerados e contemplados nas políticas de indexação, visando contribuir para a melhoria da atividade de tratamento temático da informação realizada em bibliotecas universitárias brasileiras.

## 2. Política de indexação em bibliotecas universitárias

A inter-relação existente entre a universidade e a biblioteca recai para a autonomia intelectual e promoção social que ambas representam. A biblioteca, tal como a universidade, é considerada um espaço de socialização do conhecimento e, por isso, consideradas instituições dependentes. Para Gómez Hernández (2002) pensar em universidade é ao mesmo tempo considerar a biblioteca, pois ambas partilham dos mesmos objetivos quando se reconhece que o conhecimento nestes espaços “do saber” é produzido a partir da informação, fato que permite considerar que a biblioteca é quem faz a universidade; não havendo sentido se consideramos esta relação de outro modo.

A biblioteca, na qualidade de esteio da universidade, possui como funções primeiras atividades relacionadas, em especial: a) armazenagem do conhecimento (destinadas à preservação e conservação da memória da produção intelectual); b) organização do conhecimento (qualidade dos produtos e serviços que favoreça a recuperação da informação); e c) acesso ao conhecimento (possibilitar o acesso simultâneo por todos, posto que a exigência e necessidade por informação transcenda questões relacionadas ao valor, lugar, forma ou necessidade de acesso) (FUJITA, 2005).

Com a inserção das tecnologias de informação e comunicação nas bibliotecas universitárias, suas atribuições com vistas à socialização do conhecimento foram potencializadas. O mundo atual aspira as tecnologias de informação e comunicação na mesma velocidade em que estas são desenvolvidas e apresentadas à sociedade e, naturalmente, no espaço das bibliote-

cas universitárias não poderia ser diferente. Dentre seus benefícios, o acesso à informação ganha destaque, visto ser o conhecimento o produto mais valioso em tempos de contemporaneidade, pois se tem ampliada as possibilidades dos sujeitos, em qualquer espaço geográfico, obter as informações desejadas; ação esta propiciada pelos recursos tecnológicos disponíveis na atualidade.

Ao longo de sua história, as bibliotecas universitárias atravessaram diversas mudanças. Num esforço de acompanhar e refletir sobre as alterações conjunturais inerentes à sociedade do conhecimento, as inovações tecnológicas provocam alterações substanciais em suas atividades, em especial no que tange à organização<sup>1</sup> ou tratamento e oferta da informação. Certamente, a questão das tecnologias de informação e comunicação nas atividades biblioteconômicas é um tema a ser desdobrado em muitas outras facetas, mas, em conformidade com os objetivos da pesquisa, a atenção aqui recai para o uso dos catálogos coletivos que, ao permitirem o acesso *online*, desterritorializam a informação que antes era limitada às paredes das bibliotecas universitárias.

Fujita (2012a, p. 19-20) lembra que, diante das novas tecnologias de informação e comunicação, o catálogo *online* de bibliotecas passa a ser uma base de dados que, além de armazenar representações temáticas e descritivas em formatos bibliográficos, possibilita o “[...] armazenamento de objetos digitais com diversidade de conteúdos, desde textos até audiovisuais imagéticos, que são facilmente acessíveis junto aos registros bibliográficos através da questão de busca”. Em tempos de inovações tecnológicas, permitir o acesso e a recuperação da informação por meio de catálogos coletivos *online* faz com que as bibliotecas universitárias contribuam para que, em meio à multiplicidade de informações produzidas segundo a segunda pela efervescente sociedade do conhecimento, o usuário encontre a informação adequada, sanando suas necessidades informacionais e indo ao encontro com a produção de um novo conhecimento.

Assim, os catálogos coletivos *online* de bibliotecas universitárias, enquanto formas contemporâneas de coleta de dados, sistematização e disseminação dos conteúdos documentais para a recuperação da informação por comunidades usuárias locais e/ou remotas, rejeitando limites geográficos e culturais, precisam refletir os serviços e produtos desenvolvidos e prestados com qualidade e rapidez, bem como serem pertinentes e compatíveis com os interesses informacionais de sua clientela.

Frente a essa nova realidade, Fujita, Rubi e Boccato (2009, p. 21) alertam sobre a importância de processos, condutas e instrumentos mais

Dal'Evedove, Paula Regina; Fujita, Mariângela Spotti Lopes. *Política de indexação para bibliotecas universitárias: um estudo com docentes da área de Organização e Representação do Conhecimento*. En XII Congreso ISKO España y II Congreso ISKO España-Portugal, 19-20 de noviembre, 2015, Organización del conocimiento para sistemas de información abiertos. Murcia: Universidad de Murcia.

adequados para o contexto de bibliotecas universitárias, uma vez que:

O contexto sociocognitivo do catalogador em bibliotecas universitárias apresenta mudanças significativas quanto ao acesso e recuperação por assuntos no catálogo, em parte, influenciadas pelas funções de organização e representação documentárias para a diversidade de documentos e de áreas de conhecimento e, principalmente, pelo comportamento informacional do usuário que exige análise e planejamento de política de tratamento documentário.

Diante disto, a atividade de tratamento temático da informação realizada neste contexto informacional precisa ser amparada por meio de uma política de indexação que propicie coordenadas e diretrizes específicas acerca das ações empregadas pelos bibliotecários no tratamento temático dos diversos tipos de documentos, visando contribuir com melhorias nos produtos e serviços oferecidos, bem como na recuperação da informação por parte da comunidade usuária.

O importante papel da política de indexação para o contexto de bibliotecas universitárias e especializadas é observado em Fujita (2012a, p. 22), para quem “[...] a política de indexação é um conjunto de procedimentos, materiais, normas e técnicas orientadas por decisões que refletem a prática e princípios teóricos da cultura organizacional de um sistema de informação”, os quais ganham significado se aplicados ao contexto de um sistema de armazenagem e recuperação da informação que possua finalidades e objetivos bem definidos, que abrigue condições quanto à natureza da informação produzida e solicitada e, também, conheça as características da comunidade de usuários. De modo geral, a política de indexação compreende “[...] a seleção de tipos de documentos a serem indexados, procedimentos de análise e representação de assuntos, aspectos qualitativos da indexação como precisão, especificidade, exaustividade e revocação, instrumentos de controle de vocabulário [...], além da avaliação da indexação pela consistência e pela recuperação” (FUJITA, 2012a, p. 22).

O ambiente do sistema de informação do bibliotecário indexador é uma variável importante na busca por uma consistência dos resultados da indexação, que serão refletidos na recuperação e uso da informação. Fujita (2012b, p. 187) alerta que o ambiente de trabalho proverá ao bibliotecário indexador conhecimentos específicos acerca da “[...] cultura organizacional, política de indexação expressa em um manual de indexação, linguagem de indexação e o objetivo de atendimento das demandas de sua comunidade usuária”. Sendo assim, a política de indexação

precisa estar bem formulada, esclarecida e em constante avaliação. Contudo, diversos estudos presentes na literatura especializada apontam a inexistência da política de indexação nos diversos sistemas de informação e, ainda, a despreocupação com que estas diretrizes são elaboradas ou utilizadas no cotidiano da prática profissional nas atividades de indexação e catalogação de assunto, tais como Rubi (2004, 2008); Gil Leiva y Fujita (2012); Dal'Evedove (2014), dentre outros.

Na ótica de Rubi (2012, p. 182),

A sistematização da teoria sobre política de indexação forma um conjunto que permite aos bibliotecários a visão da sua realidade, a reflexão sobre sua ação de indexação e o significado que essa reflexão pode trazer ao seu dia-a-dia, representando um impacto direto na mudança de ações e postura frente a novos desafios que lhes são colocados todos os dias.

Com efeito, em se tratando de atividades relacionadas ao tratamento da informação em sistemas de informação especializados em áreas de assunto, tal como a biblioteca universitária, os profissionais precisam direcionar esforços “[...] para a definição de princípios, métodos e orientações quanto às características da indexação e efeitos na recuperação” (FUJITA, 2012a, p. 13). Para tanto, torna-se imprescindível a elaboração e avaliação periódica de uma política de indexação descrita e registrada no manual de indexação, como forma de aprimorar a recuperação da informação ao refletir os objetivos da biblioteca universitária, identificar as condutas teóricas e práticas da equipe de trabalho, bem como para definir um padrão de cultura organizacional que seja coerente com a demanda da comunidade usuária (RUBI; FUJITA, 2010).

Diante da significativa função que a política de indexação exerce nos sistemas de informação, considera-se importante direcionar o olhar para os docentes que ensinam temáticas no eixo do Tratamento Temático da Informação, como forma de se obter novos subsídios a serem considerados na elaboração e atualização de políticas de indexação para o contexto de bibliotecas universitárias.

### 3. Procedimentos Metodológicos

Foram aplicados questionários com docentes que ministram disciplinas relacionadas à área de Organização e Representação do Conhecimento no âmbito do Tratamento Temático da Informação.

Referente à seleção dos docentes, tem-se o ensino de Biblioteconomia que se materializa em diferentes regiões do Brasil. Em sintonia

com este cenário, a amostra foi definida por meio de consultas às grades curriculares oficiais dos vinte e sete cursos de Biblioteconomia oferecidos em Instituições de Ensino Superior (IES) públicas no Brasil (Apêndice A). Para tanto, realizou-se uma análise documental na grade curricular dos cursos identificados para diagnosticar as disciplinas relacionadas ao ensino de temas ligados ao Tratamento Temático da Informação (fundamentos, processos, produtos e instrumentos).

A referida ação resultou na identificação do quadro de disciplinas obrigatórias ofertadas no Brasil nesta temática, viabilizando a identificação de quarenta e um (41)<sup>2</sup> docentes enquanto sujeitos de pesquisa representativos do contexto acadêmico e responsáveis pela produção da massa crítica, pelo arcabouço teórico-metodológico da área e pela formação de futuros profissionais. Do total dos questionários respondidos, obteve-se um grupo social composto por doze (12) docentes. Isto quer dizer que 29,3% do total dos sujeitos identificados na vertente científica participaram da pesquisa.

É proveitoso destacar que na presente pesquisa são trabalhadas e discutidas três questões abertas presentes no questionário eletrônico aplicado por Dal'Evedove (2014), as quais abordam assuntos relacionados ao processo de tratamento temático da informação e política de indexação. Entrementes, neste questionário buscou-se verificar as representações dos temários abordados no discurso coletivo dos docentes ligados aos cursos de graduação em Biblioteconomia do Brasil.

A primeira questão indagou os atores sociais sobre o papel da política de indexação para a atividade de indexação/catalogação de assunto. A análise das respostas obtidas permitiu observar os traços que identificam e definem a política de indexação na perspectiva desses atores sociais. A segunda questão buscou reconhecer alguns dos aspectos relativos à política de indexação que merecem ser discutidos e/ou aprofundados pela área de Organização e Representação do Conhecimento. Por sua vez, a terceira questão foi elaborada como forma de verificar quais são os elementos que devem ser considerados na elaboração de política de indexação para o contexto de bibliotecas universitárias no entendimento dos docentes da área.

Esclarece-se que o questionário figura no grupo de técnicas de pesquisa que permitem a coleta de dados de forma direta extensiva. A vantagem deste instrumento para a coleta de dados decorre pela oportunidade de se observar o discurso individual e coletivo dos docentes sobre os te-

mas abordados e, como resultado, obter a exposição das opiniões, percepções e conhecimentos dos sujeitos investigados. Desse modo, cada sujeito participante recebeu um código de identificação para não prejudicar a análise e compreensão das respostas, mas garantir o anonimato.

Ademais, considerou-se oportuno o emprego deste instrumento em virtude da distância entre as autoras e os respondentes, da quantidade de sujeitos de pesquisa e da qualidade dos recursos tecnológicos disponíveis na atualidade, os quais viabilizaram o envio e retorno dos questionários em tempo hábil.

A linha orientadora que conduziu a interpretação dos dados e a construção do conhecimento foi a abordagem qualitativa, visando preservar os conhecimentos cotidianos dos sujeitos de pesquisa para se chegar a resultados mais substanciados. Conduziu-se a interpretação dos dados por meio da técnica de análise de conteúdo, enquanto conjunto de técnicas de análise de comunicações.

Preconizada por Laurence Bardin (2004, p. 37), a técnica em tela comporta um:

[...] conjunto de técnicas de análise das comunicações visando obter, por procedimentos, sistemáticos e objetivos de descrição do conteúdo das mensagens, indicadores (quantitativos ou não) que permitam a inferência de conhecimentos relativos às condições de produção/recepção (variáveis inferidas) destas mensagens.

A técnica de análise de conteúdo viabiliza a determinação de pontos de inferência que se pretende identificar em uma determinada comunicação, a construção das variáveis analisadas e a verificação de seu sentido no contexto apresentado. Nesta, a contextualização comporta-se como um dos principais requisitos para a obtenção de resultados relevantes; característica que enriquece a análise dos dados em investigações científicas que envolvam percepções e conhecimentos particulares de sujeitos inseridos em contextos sociais, como é o caso desta pesquisa.

Do ponto de vista prático, toma-se o aporte da análise de conteúdo para identificar os núcleos temáticos mais recorrentes, principalmente os que figurarem de forma isolada nas respostas dos questionários. Para tanto, pauta-se nas três fases distintas e complementares que abrangem a técnica, sendo estas: pré-análise, exploração do material e tratamento dos resultados, inferência e interpretação.

Dal'Evedove, Paula Regina; Fujita, Mariângela Spotti Lopes. *Política de indexação para bibliotecas universitárias: um estudo com docentes da área de Organização e Representação do Conhecimento*. En XII Congreso ISKO España y II Congreso ISKO España-Portugal, 19-20 de noviembre, 2015, Organización del conocimiento para sistemas de información abiertos. Murcia: Universidad de Murcia.

#### 4. Apresentação e Discussão dos resultados

Os resultados de natureza exploratória alcançados com a aplicação dos questionários foram dispostos respeitando a natureza dos dados e suas especificações, conforme segue:

1. Qual o papel da política de indexação para a atividade de indexação/catalogação de assunto?

De modo geral, os docentes são conscientes sobre o papel que a política de indexação exerce na atividade de indexação/catalogação de assunto. As respostas obtidas foram consideradas favoráveis ao objetivo pretendido com a elaboração da primeira questão. Na visão dos sujeitos participantes, a política de indexação representa o conjunto de diretrizes capazes de orientar as decisões a serem tomadas no tratamento da informação, que tem como resultado a recuperação da informação com qualidade para o usuário.

No discurso desses atores sociais, o papel da política de indexação é definir as ações, níveis, padrões, mecanismos e ferramentas adotadas pela biblioteca universitária para a realização do processo de tratamento temático da informação. Ao orientar a prática profissional por meio de elementos que embutem um mínimo de regras visando a padronização do processo, desde que apropriadas para a realidade de cada sistema de informação, a política de indexação “[...] permite que, progressivamente, o próprio processo de indexação economize energias, tanto do usuário quanto do bibliotecário, contribuindo para o desenvolvimento da biblioteca” (NUNES, 2004, 60).

*A política de indexação ocupa um papel bastante importante na atividade de indexação e envolve toda a instituição, uma vez que ela sintetiza as diretrizes acordadas por uma instituição e seus colaboradores, no sentido de uniformizar e otimizar os recursos humanos e materiais disponíveis nessa instituição* (D-7).

*Nortear os procedimentos e a avaliação da atividade de indexação/catalogação de assunto* (D-2).

*Instrumento básico e fundamental para nortear as atividades dos indexadores* (D-4).

*A política orienta o trabalho do indexador e promove a padronização dos procedimentos numa mesma biblioteca ou rede de bibliotecas* (D-6).

*Extremamente importante, pois sem ela não há os parâmetros necessários para que a indexação cumpra seu objetivo dentro de determinada instituição e para determinado domínio* (D-12).

Um dos discursos obtidos com as análises dos questionários ganha notoriedade neste momento. Nele, verifica-se a presença de uma concepção bastante difundida na literatura especializada sobre política de indexação, a qual considera que o processo de tratamento temático da informação deva ser norteado e respaldado por uma filosofia que abarque questões relacionadas à parte operacional, mas também contemple o ponto de vista da biblioteca universitária enquanto organização.

*A Política de Indexação dentro de um sistema de informação deve refletir os objetivos da unidade de informação, não se constituindo somente como uma lista de procedimentos, mas sim como [uma] diretriz a ser seguida, uma filosofia de trabalho* (D-7).

Ao explicitar o seu posicionamento dessa forma, o referido docente vai ao encontro com o que é preconizado pelas principais teorias na área de Organização e Representação do Conhecimento no âmbito brasileiro. Destaque seja dado à Rubi (2008, p. 50) ao sinalizar que “[...] a política de indexação não deve ser vista como uma lista de procedimentos a serem seguidos, e sim uma filosofia que reflete os interesses e objetivos da biblioteca”. Este entendimento também é observado em outros estudos sobre a temática, tais como em Carneiro (1985) e Gil Leiva e Fujita (2012).

Para um dos sujeitos participantes, a política de indexação quando bem delimitada e elaborada por uma equipe multidimensional assevera que os produtos e serviços informacionais oriundos do processo de tratamento temático da informação promovam a democracia e o respeito necessário às demandas de sua comunidade usuária. Esta preocupação foi manifestada da seguinte forma:

*Somente com uma política de indexação bem delimitada, elaborada por uma equipe multidimensional e horizontal, permitirá a Instituição indexadora, prover ao seu sistema de indexação, a democracia e o respeito necessário às demandas informacionais de seus usuários* (D-10).

De modo prático, significa imprimir nestes produtos e serviços informacionais uma diferenciação que os torne atraentes aos olhos dos usuários reais ou potenciais de bibliotecas universitárias. Devidamente fundamentada, a política de indexação preserva os interesses e expectativas informacionais da comunidade acadêmica interna e externa. Conseqüentemente, o processo como um todo “[...] deixa de ser um foco de problemas para se transmutar em instrumento indutor de avanços qualitativos dos serviços prestados pela biblioteca” (NUNES, 2004, 60). À medida que a biblioteca universitária admite o

usuário como elemento central do processo e estabelece regras mínimas para o alcance desse propósito, passa-se a ter assegurado que os seus esforços serão dispensados na promoção de produtos e serviços informacionais que reflitam os interesses, expectativas e desejos informacionais de sua comunidade real e potencial.

Ainda, na concepção deste sujeito de pesquisa, o mérito da política de indexação é propiciar novas formas de se pensar as regras existentes na biblioteca universitária diante das particularidades de cada sistema de informação e, sobretudo, avançar na persecução dos objetivos institucionais.

*Mas creio que o principal papel da política de indexação é buscar através dela a superação dos recursos existentes, levando a instituição para onde se quer que esta chegue e não apenas para reproduzir o existente (D-5).*

Como resultado, as características e especificidades de cada biblioteca universitária, além de representadas na política de indexação por meio de diretrizes claras e precisas, passam a ser respeitadas quando da inserção desta na prática cotidiana do processo.

2. Quais questões sobre política de indexação merecem ser discutidas/aprofundadas pela área?

Com base na vivência acadêmica, nesta questão os sujeitos participantes deveriam opinar sobre os pontos que merecem ser discutidos/aprofundados nos estudos encaminhados na área de Organização e Representação do Conhecimento acerca da política de indexação. Dentre os resultados, foram apontados alguns elementos apresentados pela literatura especializada e compilados por Rubi (2008), quais sejam: o contexto, a linguagem documental e os tipos de índices utilizados, a questão da especificidade e exaustividade, os recursos (financeiros, materiais, físicos e humanos), os assuntos cobertos pelos documentos e área de interesse, bem como o uso dos vocabulários controlados. Do total, um grupo formado por três docentes entende que tais elementos ainda carecem de discussões mais verticalizadas, em especial quando se admite as características e particularidades de cada biblioteca universitária, com objetivos e características próprias.

A conscientização dos bibliotecários e instituições sobre a necessidade de definir políticas de indexação específicas para a realidade de cada sistema de informação foi apontada por um dos docentes, para quem a área de Organização e Representação do Conhecimento precisa desenvolver estudos dedicados à operacionaliza-

ção de política de indexação e não somente propor diretrizes, a saber:

*Conscientização de profissionais e instituições sobre a necessidade de definir políticas; relação da política de indexação com as atividades de planejamento e como 'operacionalizar' a política (D-6).*

Relativo a isso, entende-se que para a política de indexação ter condições exequíveis de ser aplicada na prática cotidiana dos bibliotecários e compor efetivamente o quadro de documentos vigentes das bibliotecas universitárias, o desenvolvimento de estudos que cerquem questões inerentes aos atores sociais que figuram no processo de tratamento temático da informação mostra-se uma linha de raciocínio bastante favorável. Isto porque, passa-se a considerar nas diretrizes de políticas de indexação elementos que não estejam apenas centrados nas concepções teóricas da literatura especializada, mas que, especialmente, tragam características sobre a realidade própria do processo realizado no contexto de bibliotecas universitárias mediante o discurso dos diferentes atores sociais envolvidos, quais sejam: docentes, pesquisadores, bibliotecários dirigentes e indexadores, bem como usuários da informação.

Um dos sujeitos participantes indica que algumas das dificuldades existentes em se delimitar com maior precisão e rigor as diretrizes de política de indexação decorrem da inércia com que as características dos usuários são assumidas nos produtos e serviços informacionais gerados pelos sistemas de informação, conforme segue:

*A sintonia e sincronia com as necessidades e características dos usuários, muitas vezes são decisões ou diretrizes formuladas sem levar em consideração o principal que é o usuário (D-1).*

Complementando a discussão, na fala de outro sujeito participante identifica-se uma indagação sobre um ponto relevante a ser assumido nas pesquisas empreendidas no âmbito da temática, qual seja:

*As questões éticas envolvidas no tratamento temático estão fazendo parte da elaboração da política de indexação? (D-2).*

Em um trabalho de análise acerca das diretrizes de política de indexação propostas pela literatura especializada, verifica-se que as questões éticas ainda não foram pontualmente assumidas e/ou mesmo inseridas como elementos constituintes desse conjunto de orientações. Nessa concepção, nota-se uma abertura favorável por indicações desta natureza quando se considera os fenômenos, situações e características que envolvem a atividade de tratamento temático da informação desempenhada por profissionais bibliotecários que possuem hábitos e, por sua

vez, estão inseridos em sistemas de informação com particularidades e objetivos institucionais específicos.

Segundo Almeida (2010, p. 19) o hábito determina as condições da cognição dos sujeitos e, especificamente no caso do bibliotecário indexador, alerta que este profissional “[...] está submetido a forças reais que delimitam as disposições das interpretações futuras”. Isto implica considerar que no processo de interpretação realizado pelo bibliotecário indexador existem variáveis que condicionam o resultado do tratamento temático da informação e, por consequência, o da recuperação nos sistemas de informação. Prever diretrizes operacionais que cerquem tais questões nem sempre é assegurar neutralidade aos produtos e serviços informacionais gerados no processo, tendo-se que traçar orientações de conduta ética pelas quais os profissionais deverão se pautar. Nesta esfera, “sendo um ser ideológico e culturalmente social, o aspecto da neutralidade dificilmente será zelado” (DAL’EVEDOVE, 2010, p. 127). Por consequência, o bom sendo profissional acaba incidindo preconceitos, valores e opiniões pessoais dos bibliotecários indexadores a respeito de determinado assunto, prejudicando a representação do assunto do documento e, posteriormente, a recuperação deste item informacional nos catálogos coletivos *online* das bibliotecas universitárias.

Esta realidade traz a ascensão da política de indexação como forma de orientar o cotidiano da prática profissional, sendo necessária uma mudança de postura por parte do bibliotecário indexador, conscientizando-se sobre a importância da adoção de critérios de qualidade para que o resultado do processo resulte efetivamente na recuperação da informação (RUBI, 2008). Entende-se que uma alternativa à questão é a indicação no corpo da política de indexação de um padrão de conduta ética pelo qual o processo de tratamento temático da informação deva ser conduzido e amparado no contexto de bibliotecas universitárias.

Na concepção de outro sujeito participante, as pesquisas sobre política de indexação precisam considerar as contribuições da área de Gestão da Informação, especificamente no que tange às teorias relativas à gestão do conhecimento como forma de, por meio dos métodos, modelos e técnicas desta literatura, identificar com maior propriedade as necessidades de cada biblioteca universitária e dos seus respectivos profissionais.

*Acredito que, para além do que a literatura clássica sobre indexação nos transmite, o desenvolvimento*

*de uma política de indexação deveria caminhar ao lado da teoria da gestão do conhecimento (GC) nas organizações. Uma vez que a GC nos oferece métodos, modelos e técnicas para identificarmos as reais necessidades de nossas instituições e de nossos colaboradores (D-5).*

Valentim (2008, p. 4) esclarece que a gestão do conhecimento visa trabalhar a cultura e a comunicação organizacional e informacional em contextos organizacionais, como forma de propiciar “[...] um ambiente positivo em relação à criação/geração, aquisição/apreensão, compartilhamento/socialização e uso/utilização de conhecimento [...]”. Autores como Carneiro (1985), Rubi (2008) e Fujita (2012a) destacam a importância da política de indexação ser compreendida como uma decisão administrativa, representada por meio de uma filosofia que reflita os objetivos da instituição. Por contribuir na definição de condutas teóricas e práticas dos profissionais envolvidos no tratamento temático da informação, a política de indexação possibilita que cada biblioteca universitária defina um padrão de cultura organizacional coerente com a demanda de sua comunidade usuária.

Somado a isso, tem-se a capacidade da política de indexação em conferir às bibliotecas universitárias os meios eficazes para se conhecer, com maior propriedade, seus usuários e oportunizar que tais sujeitos informacionais encontrem a informação que desejam.

Complementando a questão, o referido sujeito participante alerta sobre a necessidade desta abertura interdisciplinar, uma vez que a literatura especializada acerca de política de indexação é bastante limitada e, por vezes, concisa no que diz respeito ao desenvolvimento de diretrizes que considerem a realidade organizacional em que o processo de tratamento temático da informação é realizado. Em sua fala, o docente defende que, ao ser admitido nos estudos de política de indexação o desdobramento interdisciplinar com a área de Gestão do Conhecimento, creditar-se-ia uma possibilidade investigativa proveitosa para o avanço dos estudos. Corroborar-se com o presente posicionamento, posto que a política de indexação precisa prever questões relativas ao campo teórico e metodológico do processo e contemplar atividades de gestão, nas quais estão os requisitos para o bom planejamento de um sistema de recuperação da informação, conforme sinalizam os estudos de Carneiro (1985); Rubi (2004, 2008); Gil Leiva y Fujita (2012) e Dal’Evedove (2014).

*Digo isso porque muitas vezes ficamos presos no que a teoria tradicional da indexação nos oferece e ela é bastante lacônica no sentido do desenvolvi-*

mento de políticas baseadas na realidade em que estamos inseridos (D-5).

Na prática, o bom entendimento da biblioteca universitária enquanto contexto social e cultural de atividade, que possui características e objetivos específicos, bem como uma realidade de trabalho própria, sinaliza com maior clareza seus objetivos institucionais. Dessa forma, ao serem identificados com maior precisão quais são os objetivos da biblioteca universitária e, por consequência, estes serem representados por meio de uma filosofia que reflita o padrão de cultura organizacional coerente com a demanda da comunidade acadêmica, facilita o encontro da informação certa e sem maiores esforços físicos e/ou cognitivos por parte de seus usuários.

Políticas servem para repensar estruturas e não para mantê-las. Por isso creio que nossa teoria clássica sobre esta temática está um tanto defasada (D-5).

Em um estudo que alia a gestão e o tratamento da informação, Jambeiro (2000) defende que em razão das conquistas tecnológicas nos setores de informática, comunicações e eletrônica, o profissional da informação precisa adquirir novos atributos, tornando-se um “estrategista” ou um “humano multifacetado” (CARVALHO, 2002), sendo capaz de captar, compreender, analisar e interpretar de maneira crítica a realidade a sua volta. Dentre um conjunto de exigências listadas pelo autor para a atuação do profissional da informação nas diversas organizações, três ganham destaque por darem consistência às reflexões encaminhadas neste ponto da pesquisa sobre os contributos da Gestão do Conhecimento para o tratamento temático da informação realizado em bibliotecas universitárias, a saber:

- Exercitar visão crítica sobre a produção, distribuição e consumo de informação, porque somente com esta visão poderão produzir, selecionar, organizar e disseminar adequada e eficientemente a informação;
- Analisar o conteúdo e dialogar com a fonte ou produtor e o consumidor sobre a qualidade da informação obtida e seu adequado tratamento;
- Valorizar o conhecimento sobre o ambiente em que vivem, buscando identificar possíveis facilidades e dificuldades ao exercício de sua missão (JAMBEIRO, 2000, p. 7).

No primeiro item, nota-se uma preocupação com o exercício de uma visão crítica por parte do bibliotecário indexador de todo o ciclo documental, como um diferencial para a organização, acesso e transferência da informação ade-

quada e eficiente. Para tanto, cabe ao profissional possuir uma visão ampla do processo de tratamento temático da informação, não limitada apenas ao tratamento, mas considerando os meios de distribuição e consumo das informações. No item posterior, tem-se o ápice do processo de tratamento temático da informação realizado em contexto de bibliotecas universitárias, mediante uma política de indexação bem definida na ação de analisar o conteúdo e dialogar com o usuário acerca da qualidade da informação obtida e seu adequado tratamento.

Esta manobra preserva dois importantes elementos presentes nas diretrizes de política de indexação propostas pela literatura especializada da área: a) *identificação dos usuários* no momento da preparação da política de indexação contemplando áreas de interesse, níveis de experiências, atividades que exercem volume e características das questões propostas pela comunidade usuária; e b) *avaliação* do sistema de recuperação da informação feita pelo bibliotecário com o objetivo de determinar o grau de satisfação dos usuários no uso (RUBI, 2008). Adicionalmente, o último item revela o mérito da política de indexação como forma de refletir as características da cultura organizacional de cada biblioteca universitária, identificando suas potencialidades e dificuldades no exercício de sua missão enquanto organização direcionada ao provimento de informações que sanem os desejos e interesses informacionais de seus usuários.

Em um estudo mais verticalizado sobre as contribuições da gestão do conhecimento organizacional para a política de indexação, Lousada et al. (2011) defendem que os elementos de política de indexação podem ser enriquecidos ao passo que a criação e utilização do conhecimento organizacional são considerados. Ao se beneficiar de um modelo de gestão do conhecimento, por exemplo, que oriente a política de indexação para a criação de valor ao usuário, as autoras sinalizam que as instituições, independentemente do seu sistema de informação, “[...] passam a ser orientadas não apenas pelos serviços baseados na infraestrutura física, mas também nas relações humanas” (LOUSADA et al., 2011, p. 2000). Este contributo é de grande valia para o encaminhamento de abordagens centradas no sujeito e suas relações na coletividade a partir de um olhar organizacional, aproximando o fazer profissional às características e particularidades da comunidade usuária; fator indutor para a geração de conhecimento.

A presença da política de indexação nos cursos de formação e capacitação profissional do bibliotecário também foi um dos pontos suscitados

por um dos docentes. Em sua compreensão, a presença de discussões sobre a política de indexação na formação inicial e na formação em serviço contribuiria para capacitar os sujeitos sobre os critérios, os métodos, a construção e uso das linguagens de indexação, considerando-se que a construção e posterior manutenção deste instrumento requer dinamismo.

*De modo geral a política de indexação como um todo deve ser melhor discutida na formação do Bibliotecário, e também via cursos de capacitação para profissionais atuantes. Há uma compreensão equivocada por parte de alguns profissionais do campo, sobre os critérios, os métodos, a construção e uso de Linguagens documentárias (...). Esta competência precisaria ser constantemente reciclada dentre os profissionais, uma vez que a construção e a manutenção das Linguagens requer um dinamismo (D-10).*

Pode-se inferir que tal dinamismo decorra do papel que a linguagem de indexação exerce na representação de conteúdo e na recuperação da informação, em que a permanente atualização é um atributo indispensável por vincular o documento ao usuário. A confluência entre o uso adequado de uma linguagem de indexação na prática de um tratamento temático da informação qualitativo, com fins de recuperação da informação por um público especializado, como é o caso das comunidades usuárias de bibliotecas universitárias, foi evidenciada por Boccato (2009, p. 231) para quem:

O uso adequado da linguagem documentária no processo de representação para indexação está vinculado ao processo de avaliação constante da busca por assunto, proporcionando elementos norteadores para o aperfeiçoamento e a adequação de um sistema de organização do conhecimento que reflita o contexto informacional, cultural e social em que está inserido.

Complementando a questão, o referido docente destaca a necessidade de alunos e profissionais bibliotecários compreenderem o processo de busca da informação realizado nos sistemas de recuperação da informação pelo usuário.

*Dentre os pontos que precisam ser mais explorados, destaco a necessidade de compreensão sobre o funcionamento do processo de busca da informação pela sua comunidade usuária, e isto significa compreender as formas de uso e de significação da linguagem nesta comunidade (D-10).*

Mais do que saber o modo como a comunidade usuária realiza a busca no sistema de informação, alunos e bibliotecários indexadores entenderiam com maior propriedade as formas de uso e significação da linguagem de indexação pela comunidade usuária.

3. Quais os elementos a serem considerados na elaboração de política de indexação para bibliotecas universitárias?

Na terceira questão, buscou-se identificar os principais elementos a serem considerados na elaboração de uma política de indexação própria para o contexto de bibliotecas universitárias. No geral, o conjunto de respostas obtidas foi pouco expressivo, não havendo uma correspondência direta e linear entre o enunciado da pergunta e as opiniões apresentadas pelos docentes. Alguns dos sujeitos participantes não se mostraram a par do universo teórico da política de indexação ou mesmo do propósito destas diretrizes no tratamento temático da informação.

Apesar do pequeno material empírico obtido, uma das falas dos sujeitos participantes indica a caracterização do usuário e os níveis de indexação como elementos a serem considerados.

*(...) a caracterização do usuário, incluindo mapeamento de necessidades, tipo de material utilizado, tipo de linguagem documentária a ser utilizada, tipo de serviço e produto a ser fornecido etc. Estabelecer os níveis de indexação - exaustividade, precisão, quantitativo de descritores a serem utilizados, preferência por descritores mais específicos (D-1).*

Outro sujeito participante apontou a necessidade da política de indexação abarcar a realidade própria de cada sistema de informação, considerando, obrigatoriamente, as necessidades diversas dos distintos usuários. Mais do que isso, o encaminhamento de políticas de indexação que imprimam a importância de se delinear os contornos de cada realidade em que a informação é organizada e disponibilizada para o acesso, com fins de transferência e geração de conhecimento.

*No caso das bibliotecas universitárias, a necessidade de políticas que contemplem as necessidades diversas dos distintos usuários (fator relevante em qualquer tipo de biblioteca, mas, no caso das universitárias, essa característica é muito importante e com perfis bem definidos de necessidades informacionais) (D-6).*

Sobre isto, Smit (2009, p. 59) esclarece que a apropriação da informação é dependente do acesso, mas que a passagem do acesso para a transferência não é uma ação simples ou automática, existe “[...] uma relação de causalidade entre ambas (mas não uma relação de pressuposição recíproca): não há transferência sem acesso à informação, mas pode haver acesso à informação sem que ocorra a respectiva transferência”. Para uma efetiva transferência da informação, isto é, assimilação cognitiva e pessoal da informação pelos usuários no contexto da biblioteca universitária, deve-se admitir que as

ações humanas não são aleatórias, mas correspondem às práticas sociais com um conteúdo que lhes dão fundamento. O mesmo ocorre na busca pela informação; tem que haver significado na informação para que esta satisfaça o usuário. Logo, questões que tenham uma abordagem mais completa do complexo processo de tratamento temático da informação a partir das diferentes realidades institucionais em que a informação se apresenta se mostram bastante significativas.

*Acredito que são todos os elementos que nos rodeiam, sejam eles humanos (comportamentais, sociais, conhecimentos técnicos, teóricos, práticos), materiais (espaço físico, recursos) e organizacionais (estrutura da instituição, tipo de gestão - verticalizada ou mais horizontal, foco da gestão - se centrado no [usuário] ou não, etc.). Digo isso porque geralmente adotamos uma visão mais fechada ao pensarmos na elaboração de políticas de indexação como número de indexadores/quantidade de descritores, recursos financeiros/nº de profissionais contratados, cuidados linguísticos para padronização do vocabulário (...). Claro que todas estas questões são importantes, mas não são questões que modelam o sistema, mas que sim o mantém intacto como está (D-5).*

Complementando o exposto, sinaliza-se que, no âmbito das bibliotecas universitárias, os catálogos coletivos *online* diante das inovações tecnológicas passaram a oferecer produtos e serviços informacionais em âmbito local e global. Essa dimensão assumida pelo catálogo, por consequência, acarreta novas responsabilidades ao fazer profissional, pautadas no “[...] compromisso com a construção de catálogos condizentes com a realidade não somente de sua comunidade usuária local, mas também de uma comunidade usuária potencial virtual, cada vez mais exigente” (RUBI, 2008, p. 145). Em atenção a essa nova realidade, um dos docentes indicou que a política de indexação deve prever o trabalho em rede, uma vez que as bibliotecas universitárias podem apresentar características particulares, de acordo com a realidade institucional de cada uma.

*Prever o trabalho em rede, onde diferentes unidades de informação podem ter características diferentes (D-7).*

As afirmações acima são particularmente interessantes para esta pesquisa. No entendimento de Smit (2009, p. 62) “a organização da informação, para conferir sentido, opera por opções” que são pré-estabelecidas por cada instituição, no caso a biblioteca universitária. Sobre isto, torna-se importante reafirmar que, para conferir sentido à informação, esta deve ser vinculada a um objetivo que a norteie, conforme esclarece a autora:

[...] não faz sentido imaginar uma organização da informação “em si”, desvinculada de objetivos e, portanto, opções. A decisão de preservar determinadas informações e de organizá-las de acordo com objetivos institucionais acarreta a “institucionalização” da informação, ou seja, a atribuição de um status diferenciador àquela informação, pois a mesma agora foi acrescida de um “carimbo” institucional. Este carimbo tem dupla função: afirmar que aquela informação foi considerada digna de guarda e que a mesma encontra seu lugar, ou seu sentido, no contexto maior das opções institucionais (SMIT, 2009, p. 62).

Dito desta maneira, as informações organizadas nas bibliotecas universitárias são dotadas de valor, direcionadas por objetivos institucionais bem definidos. De outro modo, perdem a sua utilidade potencial no futuro. Com efeito, a organização da informação no contexto de bibliotecas universitárias “[...] não constitui somente uma imperiosa necessidade para que o acesso à mesma possa ser ativado, mas é a condição *sine qua non* para que o sistema de informação ‘faça sentido’, ou seja, que o mesmo cumpra seu papel social” (SMIT, 2009, p. 61).

Portanto, é pertinente prever nas diretrizes de política de indexação a definição clara das características e objetivos institucionais, conforme sinalizado por um dos sujeitos de pesquisa:

*Ela deve funcionar [como] um guia a ser seguido na tomada de decisões. Fundamentalmente ela deve considerar as características e objetivos da instituição. Identificar os usuários da unidade de informação para poder atender as suas necessidades de informação. Prever os recursos humanos, materiais e financeiros, que delimitam o funcionamento de um sistema de recuperação de informações. Prever treinamento para os profissionais que trabalham com a indexação. Estabelecer previamente e com conhecimento os documentos fontes que forem necessários para a indexação (D-7).*

Diante do exposto, tem-se o pensamento de que a política de indexação propicia novas formas de se pensar e aplicar as regras vigentes para avançar na persecução dos objetivos estabelecidos pela biblioteca universitária.

## 5. Considerações Finais

A política de indexação deve ser pensada como eixo integrador entre o ‘saber’ e o ‘fazer’ profissional, isto é, como um corpo de diretrizes que aliam literatura especializada e especificidades da prática cotidiana da profissão.

Nesse cenário, verificou-se o conhecimento dos docentes da área de Organização e Representação do Conhecimento que ministram disciplinas relacionadas ao Tratamento Temático da

Informação acerca da política de indexação por meio do uso de questionário. Esta linha investigativa mostrou-se bastante oportuna, ao passo que permitiu sinalizar algumas das opiniões, impressões, concepções e expectativas dos sujeitos participantes sobre política de indexação para bibliotecas universitárias.

De modo geral, verificou-se que os docentes brasileiros são conscientes sobre o posicionamento estratégico que a política de indexação assume no contexto de bibliotecas universitárias ao contribuir para a promoção da democracia e respeito às demandas de suas comunidades usuárias.

A partir da reflexão empreendida, verifica-se uma interface vantajosa entre a área de Organização e Representação do Conhecimento e a área de Gestão do Conhecimento para o encaminhamento de investigações sobre política de indexação, em especial quando tais diretrizes são direcionadas para sistemas de informação multifacetados, como é o caso das bibliotecas universitárias. Outro aspecto relevante a ser considerado é a reflexão realizada pelos docentes acerca do usuário ser o ponto central dos esforços empreendidos pelas pesquisas destinadas à política de indexação. Tal postura viabiliza que os produtos e serviços informacionais gerados não fiquem alheios aos valores, exigências e necessidades dos usuários de informação, os quais devem ser capazes de construir sentido a partir das informações que recebem no momento da recuperação em sistemas de informação.

Com base nos resultados apresentados, conclui-se que a política de indexação precisa ser pensada e elaborada de modo a contribuir para que o bibliotecário indexador tenha uma postura ética e compromissada no momento da representação de assunto, reconhecendo, para tanto, a necessidade desta atividade no contexto de bibliotecas universitárias. Para tanto, recomenda-se o encaminhamento de estudos que trabalhem com as questões de natureza ética relacionadas à atuação do bibliotecário indexador e que possam ser descritas na política de indexação.

## Notas

(1) Autores como Fujita (2003), Guimarães (2003) e Dias e Naves (2007) tratam como sinônimos a organização e o tratamento da informação. Na concepção destes autores, a organização da informação compreende as atividades e operações do tratamento descritivo e do tratamento temático da informação.

(2) Este número não representa o montante total de docentes que ministram disciplinas relacionadas ao eixo Tratamento Temático da Informação no Brasil, uma vez que em algumas Escolas de Biblioteconomia os planos de ensino estavam indisponíveis ou mesmo desatualizados.

## Referencias

- Almeida, C. C. (2010). Pragmatismo e Semiótica de Peirce na organização da informação. // *Scire*, Zaragoza, 16, 15-20.
- Bardin, L. (2004). *Análise de conteúdo*. Lisboa: Edições 70, 2004.
- Boccatto, V. R. C. (2009). *Avaliação do uso de linguagem documentária em catálogos coletivos de bibliotecas universitárias: um estudo sociocognitivo com protocolo verbal*. Tese (Doutorado em Ciência da Informação) – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 299f.
- Carneiro, M. V. (1985). Diretrizes para uma política de indexação. // *Revista da Escola de Biblioteconomia da UFMG*, Belo Horizonte. 14:2 (set. 1985) 221-241.
- Carvalho, K. (2002). O profissional da informação: o humano multifacetado. *DataGramaZero - Revista de Ciência da Informação*, Rio de Janeiro, 3:5 (out. 2002).
- Dal'Evedove, P. R. (2014). *O tratamento temático da informação em abordagem sociocultural: diretrizes de política de indexação em bibliotecas universitárias*. Tese (Doutorado em Ciência da Informação) – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília. 266f.
- Fujita, M. S. L. (2005). Aspectos evolutivos das bibliotecas universitárias em ambiente digital na perspectiva da Rede de Bibliotecas da UNESP. *Informação & Sociedade: estudos*, João Pessoa, 15:2. Recuperado 24-07-2013 de <http://www.informacoesociedade.ufpb.br/pdf/IS1520504.pdf>.
- Fujita, M. S. L. (2012a). A política de indexação para representação e recuperação da informação. // Isidoro Gil Leiva; Mariângela Spotti Lopes Fujita (Org.). *Política de indexação*. São Paulo; Marília: Cultura Acadêmica; Oficina Universitária, 2012. 17-30.
- Fujita, M. S. L. (2012b). Elaboração e avaliação da política de indexação na formação inicial do indexador. // Isidoro Gil Leiva; Mariângela Spotti Lopes Fujita (Org.). *Política de indexação*. São Paulo; Marília: Cultura Acadêmica; Oficina Universitária, 2012. 187-216.
- Fujita, M. S. L.; Rubi, M. P.; Boccatto, V. R. C. (2009). O contexto sociocognitivo do catalogador em bibliotecas universitárias: perspectivas para uma política de tratamento da informação documentária. // *DataGramaZero - Revista de Ciência da Informação*, Rio de Janeiro, 10:2 (Abril 2009).
- Gil Leiva, I.; Fujita, M. S. L. (2012). Política de indexação latino-americana. // Isidoro Gil Leiva; Mariângela Spotti Lopes Fujita (Org.). *Política de indexação*. São Paulo; Marília: Cultura Acadêmica; Oficina Universitária. 2012.
- Guimarães, J. A. C. (2004). As políticas de indexação como elementos para a gestão do conhecimento nas organizações. In: VIDOTTI, S. A. B. G. (Org.). *Tecnologia e conteúdos informacionais*. São Paulo: Polis, 2004, 43-52.
- Jambeiro, O. (2000). Gestão e tratamento da informação na sociedade tecnológica. // Nídia Lubisco e Lídia Brandão. (Org.). *Informação & Informática*. Salvador: EDUFBA, 2000, 207-232.

- Lousada, M. et. al. Políticas de indexação no âmbito da gestão do conhecimento organizacional. // *Informação & sociedade* (UFPB. *Online*), v. 21, 191-202, 2011.
- Nunes, C. O. (2004). Algumas considerações acerca da ausência de políticas de indexação em bibliotecas brasileiras. // *Biblos*. Rio Grande do Sul. 16, 55-61.
- Rubi, M. P.; Fujita, M. S. L. (2010). Política de indexação na catalogação de assunto em bibliotecas universitárias: a visão sociocognitiva da atuação profissional com protocolo verbal. // *Revista Digital de Biblioteconomia e Ciência da Informação*. 7 (2010) 168-200.
- Rubi, M. P. (2004). *A política de indexação na perspectiva do conhecimento organizacional*. Dissertação (Mestrado em Ciência da Informação) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília. 135 f.
- Rubi, M. P. (2008). *Política de indexação para construção de catálogos coletivos em bibliotecas universitárias*. Tese (Doutorado em Ciência da Informação) – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília. 169f.
- Rubi, M. P. (2012). Proposta para implantação de política de indexação em bibliotecas. // Isidoro Gil Leiva; Mariângela Spotti Lopes Fujita (Org.). *Política de indexação*. São Paulo: Cultura Acadêmica, 2012. 171-183.
- Smit, J. W. (2009). Novas abordagens na organização no acesso e na transferência da informação. // Helen de Castro Silva e Maria Helena T. C. de Barros. (Org.). *Ciência da Informação: múltiplos diálogos*. Marília: Oficina Universitária Unesp, 2009, 57-66.
- Valentim, M. L. P. (2008). Informação e conhecimento em organizações complexas. // Marta Lígia Pomin Valentim (Org.). *Gestão da informação e do conhecimento no âmbito da Ciência da Informação*. São Paulo: Polis: Cultura Acadêmica, 2008. 11-25.

## Anexo 1

## Cursos de formação em Biblioteconomia no Brasil

| REGIÃO       | ESTADO              | IES  | VÍNCULO INSTITUCIONAL  | DENOMINAÇÃO DO CURSO                                  | ANO DE CRIAÇÃO |
|--------------|---------------------|--|--|---|----------------|
| Norte        | Amazonas            | Universidade Federal do Amazonas (UFAM)                    | Instituto de Ciências Humanas e Letras   | Biblioteconomia                                       | 1966           |
| Norte        | Pará                | Universidade Federal do Pará (UFPA)                        | Instituto de Ciências Sociais Aplicadas - Faculdade de Biblioteconomia (FABIB) | Biblioteconomia                                       | 1963           |
| Nordeste     | Alagoas             | Universidade Federal de Alagoas (UFAL)                     | Instituto de Ciências Humanas, Comunicação e Arte (ICHCA)                      | Biblioteconomia                                       | 1998           |
| Nordeste     | Bahia               | Universidade Federal da Bahia (UFBA)                       | Instituto de Ciência da Informação   | Biblioteconomia e Documentação                        | 1942           |
| Nordeste     | Ceará               | Universidade Federal do Ceará (UFC)                        | Campus da UFC no Cariri  | Biblioteconomia                                       | 2006           |
| Nordeste     | Ceará               | Universidade Federal do Ceará (UFC)                        | Centro de Humanidades  | Biblioteconomia                                       | 1964           |
| Nordeste     | Maranhão            | Universidade Federal do Maranhão (UFMA)                    | Centro de Ciências Sociais   | Biblioteconomia                                       | 1969           |
| Nordeste     | Paraíba             | Universidade Federal da Paraíba (UFPB)                     | Centro de Ciências Sociais Aplicadas   | Biblioteconomia                                       | 1969           |
| Nordeste     | Pernambuco          | Universidade Federal de Pernambuco (UFPE)                  | Centro de Artes e Comunicação  | Biblioteconomia                                       | 1950           |
| Nordeste     | Rio Grande do Norte | Universidade Federal do Rio Grande do Norte (UFRN)         | Centro de Ciências Sociais Aplicadas   | Biblioteconomia                                       | 1996           |
| Nordeste     | Piauí               | Universidade Estadual do Piauí (UESPI)                     | Centro de Ciências Sociais e Aplicadas   | Biblioteconomia                                       | 2003           |
| Centro Oeste | Distrito Federal    | Universidade de Brasília (UnB)                             | Faculdade de Ciência da Informação   | Biblioteconomia                                       | 1962           |
| Centro Oeste | Goiás               | Universidade Federal de Goiás (UFG)                        | Faculdade de Comunicação e Biblioteconomia (FACOMB)                            | Biblioteconomia                                       | 1980           |
| Centro Oeste | Mato Grosso         | Universidade Federal de Mato Grosso (UFMT)                 | Instituto de Ciências Humanas e Sociais (IHS)                                  | Biblioteconomia                                       | 2000           |
| Sudeste      | Espírito Santo      | Universidade Federal do Espírito Santo (UFES)              | Centro de Ciências Jurídicas e Econômicas (CCJE)                               | Biblioteconomia                                       | 1974           |
| Sudeste      | Minas Gerais        | Universidade Federal de Minas Gerais (UFMG)                | Escola de Ciência da Informação  | Biblioteconomia                                       | 1950           |
| Sudeste      | Rio de Janeiro      | Universidade Federal do Estado do Rio de Janeiro (UNIRIO)* | Escola de Biblioteconomia  | Biblioteconomia                                       | 1911           |
| Sudeste      | Rio de Janeiro      | Universidade Federal Fluminense (UFF)                      | Instituto de Arte e Comunicação Social   | Biblioteconomia e Documentação                        | 1963           |
| Sudeste      | Rio de Janeiro      | Universidade Federal do Rio de Janeiro (UFRJ)              | Faculdade de Administração e Ciências Contábeis (CCJE)                         | Biblioteconomia e Gestão de Unidade de Informação     | 2005           |
| Sudeste      | São Paulo           | Universidade de São Paulo (USP)                            | Escola de Comunicação e Artes (ECA)  | Biblioteconomia                                       | 1967           |
| Sudeste      | São Paulo           | Universidade Estadual Paulista (UNESP)**                   | Faculdade de Filosofia e Ciências (FFC/Campus de Marília)                      | Biblioteconomia                                       | 1977           |
| Sudeste      | São Paulo           | Universidade Federal de São Carlos (UFSCar)                | Centro de Educação e Ciências Humanas (CECH/Campus São Carlos)                 | Biblioteconomia e Ciência da informação               | 1996           |
| Sul          | Paraná              | Universidade Estadual de Londrina (UEL)                    | Centro de Educação, Comunicação e Artes  | Biblioteconomia                                       | 1972           |
| Sul          | Rio Grande do Sul   | Universidade Federal do Rio Grande (FURG)                  | Instituto de Ciências Humanas e da Informação (ICHI)                           | Biblioteconomia                                       | 1978           |
| Sul          | Rio Grande do Sul   | Universidade Federal do Rio Grande do Sul (UFRGS)          | Faculdade de Biblioteconomia e Comunicação (FABIBO)                            | Biblioteconomia                                       | 1947           |
| Sul          | Santa Catarina      | Universidade do Estado de Santa Catarina (UDESC)           | Centro de Ciências Humanas e da Educação                                       | Biblioteconomia – Habilitação em Gestão da Informação | 1973           |
| Sul          | Santa Catarina      | Universidade Federal de Santa Catarina (UFSC)              | Ciências da Educação (CED)   | Biblioteconomia                                       | 1973           |

Fonte: Pesquisa empírica 2012/2013.

\* Corresponde ao primeiro curso Biblioteconomia criado no Brasil, sob o comando da Biblioteca Nacional do Rio de Janeiro.

\*\* Consta no respectivo como forma de informação, não sendo selecionada para a coleta de dados com os docentes desta IES em virtude do vínculo existente com a pesquisa.

---

# Compatibilidade entre linguagens documentais para construção, atualização e adequação de vocabulário de bibliotecas universitárias

*Compatibility among documentary languages for building, updating and suitability of University Libraries' vocabulary*

---

Mariângela Spotti Lopes Fujita (1), Milena Posinelli Rubi (2)

(1) Universidade Estadual Paulista "Júlio de Mesquita Filho (UNESP), Campus de Marília, Av. Hygino Muzzi Filho, 737, Bairro Mirante, CEP: 17.525-000 - Marília, São Paulo, Brasil, fujita@marilia.unesp.br

(2) Universidade Federal de São Carlos (UFSCar), Campus Sorocaba, Rodovia João Leme dos Santos (SP-264), Km 110, Bairro Itinga, CEP: 18052-780 – Sorocaba, São Paulo, Brasil, milenarubi@ufscar.br

## Resumo

As bibliotecas universitárias contêm acervos de várias áreas do conhecimento. Dessa forma, a linguagem documental utilizada deve conter estrutura e vocabulário que contemple todas as áreas de conhecimento em profundidade, especificidade e exaustividade. Objetivou-se apresentar uma metodologia de compatibilização entre linguagens documentais para a construção de sistemas de organização do conhecimento de áreas científicas especializadas de catálogos online, frente aos fundamentos teóricos e metodológicos da área de Organização e Representação do Conhecimento (ORC) com enfoque na recuperação da informação e perspectiva de bibliotecas universitárias. A proposta metodológica de construção, adequação e atualização de linguagens documentais para bibliotecas universitárias com uso da compatibilização entre linguagens de mesma matriz de estrutura e vocabulário consiste na combinação dos registros de autoridade em formato MARC21 de assuntos tópicos e geográficos de linguagens documentais de mesma origem – "Terminologia de assuntos da Biblioteca Nacional (Brasil)", *Library of Congress Subject Headings* (LCSH), *Descritores em Ciências da Saúde* (DeCS) e *Medical Subject Headings* (MeSH) – para descritores da linguagem UNESP. Concluiu-se que a compatibilização é um recurso viável para a construção de linguagens documentais no que se refere a otimização de tempo, racionalização de recursos financeiro e humano provenientes das instituições e das próprias bibliotecas universitárias envolvidas.

**Palavras-Chave:** Compatibilização entre linguagens documentais. Catálogos online. Indexação. Bibliotecas universitárias.

## 1. Introdução

A área de Organização e Representação do Conhecimento (ORC) e suas dimensões interdisciplinares dispõem de teorias e de metodologias que nos configuram diversos caminhos para a construção de sistemas de organização do conhecimento, identificando-se como possi-

## Abstract

University libraries have collections from several areas of knowledge. So, the used documentary language should contain structure and vocabulary covering all areas of knowledge in depth indexing and specificity. The aim of this paper is to present a methodology of compatibility among documentary languages to build systems of knowledge organization of specialized scientific areas of online catalogs taking into account theoretical and methodological foundations from the area of Knowledge Organization and Representation focusing the information retrieval and perspective of university libraries. The methodological proposal of building, suitability and updating of documentary languages for university libraries using the compatibility among documentary languages of same original structure and vocabulary consists of the combination of the authority records in MARC21 format of topic and geographical subjects of documentary languages – "Terminologia de assuntos da Biblioteca Nacional (Brasil), Library of Congress Subject Headings (LCSH), "Descritores em Ciências da Saúde" (DeCS and Medical Subject Headings (MeSH) – for UNESP language descriptors. It is concluded that the compatibility is a feasible feature for the building of documentary languages with respect to the time optimization and rationalization of financial and human resources from the institutions and their libraries.

**Keywords:** Compatibility among documentary languages; Online catalogs; Indexing; University libraries.

bilidade teórica e sistêmica a realização de compatibilização entre linguagens documentais de mesma matriz de estrutura e vocabulário.

Dada a necessidade de contínua atualização de vocabulários para atender a especificidade e exaustividade do desenvolvimento científico e tecnológico de áreas do conhecimento em acervos documentais de bibliotecas universitárias,

torna-se oportuna a mesclagem de linguagens documentais especializadas para o aproveitamento de vocabulários que já tenham comprovada utilização e façam parte do conhecimento de usuários e catalogadores especializados.

Considerando-se que as bibliotecas universitárias, em sua maioria, contêm acervos de várias áreas do conhecimento, salvo aquelas que possuem acervo documental em uma única área especializada, a linguagem documental utilizada deve conter estrutura e vocabulário que contemple todas as áreas de conhecimento em profundidade, especificidade e exaustividade.

De modo geral, existem várias linguagens documentais construídas por bibliotecas para uso próprio e aquelas cuja construção e atualização são realizadas por instituições que se encarregam de serviços cooperativos.

A Library Congress of Subject Headings (LCSH), o Medical Subject Headings (MeSH), a Lista de Cabeçalhos de Assuntos da Rede BIBLIODATA (LCARB) e a Terminologia de Assuntos da Biblioteca Nacional do Brasil são exemplos de linguagens documentais construídas por instituições cooperativas com destaque para o fato de que possuem a mesma matriz estrutural e de vocabulário gerados pela primeira delas.

O Vocabulário Controlado do SIBI/USP (VocaUSP) é o exemplo de linguagem construída pelo Sistema de Bibliotecas da Universidade de São Paulo para uso próprio, porém, disponível para uso de outras bibliotecas.

As linguagens de mesma matriz estrutural e de vocabulário podem ser consideradas compatíveis, entretanto, para utilizá-las na construção, adequação e atualização de outra linguagem documental é necessário ter uma metodologia de compatibilização para uso colaborativo em um sistema de bibliotecas.

Conforme estudos realizados de avaliação da linguagem documental, realizado por Boccatto, Fujita, Gil Leiva (2011), verificou-se que a linguagem anteriormente adotada no catálogo do Sistema de Bibliotecas da Universidade Estadual Paulista (UNESP) não alcançava a especificidade e exaustividade demandada pelo usuário e catalogador para representação dos assuntos. Dessa forma, propôs-se estudo de compatibilização de vocabulários para construção, adequação e atualização de linguagem documental para a estrutura de áreas de conhecimento especializada de bibliotecas universitárias com uso das linguagens de mesma matriz de estrutura e vocabulário.

Para tanto, tem-se por objetivo apresentar uma metodologia de compatibilização entre linguagens documentais para a construção de sistemas de organização do conhecimento de áreas científicas especializadas de catálogos online, frente aos fundamentos teóricos e metodológicos da área de Organização e Representação do Conhecimento (ORC) com enfoque na recuperação da informação e perspectiva de bibliotecas universitárias.

## 2. Linguagem documental em bibliotecas universitárias

O uso de uma linguagem documental (1) para escolha do descritor ou cabeçalho de assunto reduz a diversidade e a ambiguidade do vocabulário e estabelece uma uniformidade de representação dos termos selecionados pelo indexador para descrever o assunto dos documentos, já que vários autores podem utilizar diferentes palavras para expressar uma mesma ideia, assim como os usuários podem apresentar diversidade de vocabulário quando da expressão de uma estratégia de busca.

A linguagem documental faz-se necessária uma vez que a qualidade dos serviços de disseminação da informação depende da linguagem utilizada pelo sistema para a indexação e recuperação dos documentos. Isso se dá quando a linguagem do sistema permite que se traduza a linguagem natural do autor sem que se perca a ideia principal e quando permite que se traduza a linguagem do usuário de modo que satisfaça suas necessidades de informação. Esse é o foco da representação da informação dentro do processo de indexação que depende de vários componentes para completar o conjunto de procedimentos e atingir seus objetivos de recuperação, quais sejam o indexador, a linguagem documental, elementos e variáveis da política de indexação da biblioteca ou serviço de informação.

A linguagem documental cumpre a função de controle de vocabulário para o catalogador e deveria cumprir a função de mediadora na recuperação por assuntos pelo usuário como se observa na Figura 1 (Gil Leiva, 1999, p. 22, tradução nossa):

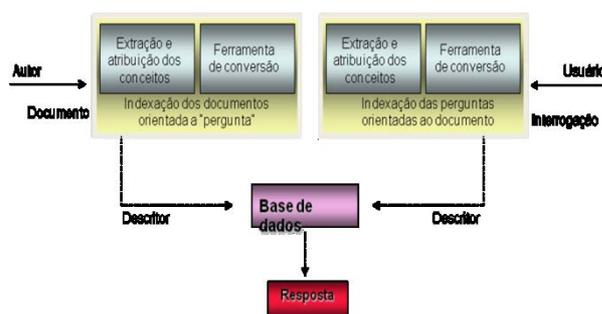


Figura 1. Processo integral da indexação

Na Figura 1 fica claro que a linguagem documental, como “ferramenta de conversão”, deve ser utilizada, tanto pelo indexador após a atribuição de conceitos extraídos do documento, quanto pelo usuário antes da atribuição de conceitos extraídos da necessidade de informação que orientará a estratégia de busca no catálogo.

A linguagem documental é o instrumento de conversão da linguagem natural do documento durante o processo de indexação e, por isso, exerce importante influência na recuperação da informação cujos efeitos são controlados e avaliados pela política de indexação adotada pela biblioteca ou serviço de informação (Fujita, 2012).

O uso de uma linguagem documental pelo indexador e pelo usuário do sistema de recuperação da informação, destina-se, de forma pragmática, à escolha do termo correto para descritor a fim de reduzir a diversidade e a ambigüidade de vocabulário. Desse modo, realiza uma mediação entre a linguagem do texto e a linguagem do usuário e estabelece uma representação precisa por meio de termos que correspondem aos termos utilizados pelo autor para expressar o assunto do texto.

A linguagem documental é formada de vocabulário e sintaxe. O vocabulário se refere à relação dos descritores usados para a identificação do conteúdo de um documento e a sintaxe se refere às regras utilizadas para a combinação dos descritores usados para a identificação do conteúdo de um documento.

Analisados seus aspectos de estrutura e funcionalidade, a linguagem documental como ferramenta de controle de vocabulário apresenta-se por meio de tipologias diferenciadas pela forma e objetivos diversos. O critério de divisão adotado por Vizcaya Alonso ([1997]) para agrupar as diferenças tipológicas de linguagens documentais baseia-se no:

- nível de coordenação: pós-coordenada e pré-coordenada

- estrutura: hierárquicas, alfabéticas e associadas
- expressão gráfica: numéricas, alfabéticas e alfanuméricas
- O critério de divisão defendido por Gil Urdiciain (2004) consiste nas características de:
- controle: livres e controladas;
- coordenação: pós-coordenadas e pré-coordenadas
- estrutura: hierárquica, combinatória e sintática

Considerando-se o aspecto de estrutura, é bem típico referir-se às linguagens documentais alfabéticas e às linguagens documentais hierárquicas, apesar dos autores citados considerarem três tipos de linguagens no aspecto de estrutura.

Guimarães (1990) considera que as linguagens documentais, de acordo com a forma de apresentação dos conceitos, são Classificadas ou hierárquicas e Alfabéticas e conforme a coordenação dos conceitos, pós-coordenada e pré-coordenadas.

Nas linguagens de *estrutura hierárquica*, os termos relacionam-se entre si a partir da subordinação e “co-subordinação”, constituindo-se como uma relação assimétrica entre dois elementos, onde um é superior ao outro por caráter normativo (Viscaya Alonso, [1997]). São exemplos de estrutura hierárquica os sistemas de classificação, como por exemplo: Classificação Decimal de Dewey – CDD, Classificação Decimal Universal - CDU e a Classificação Bibliográfica de Bliss.

Nas linguagens de *estrutura alfabética*, tal como o tesouro e as listas de cabeçalhos de assuntos, os termos que a integram são ordenados alfabeticamente.

O critério de *coordenação* é estabelecido pelo momento da combinação dos termos: se realizada durante a descrição, a linguagem é pré-coordenada e se realizada durante a estratégia de busca, a linguagem é pós-coordenada.

- Pré-coordenadas: Listas de cabeçalhos de assunto e Sistemas tradicionais de Classificação (Classificação Decimal de Dewey, Classificação Decimal Universal, Classificação da Library Congress, Classificações Facetadas); e
- Pós-coordenadas: Tesouros e UNITERMO.

Nos sistemas pré-coordenados os termos são previamente combinados de acordo com regras sintáticas estabelecidas pela linguagem. Por outro lado, nos pós-coordenados os termos serão combinados de acordo com a sintaxe da lógica de busca. Gil Urdiciain (2004, p. 24) con-

clui que “[...] nas linguagens pré-coordenadas a relação entre os termos é gramatical e que na linguagem pós-coordenada a relação é lógica.” As linguagens de estrutura combinatória, como os tesouros, permitem a combinação de seus termos a partir da necessidade de indexação.

Quanto aos Tesouros constituem-se de descritores que podem ser combinados ou não entre si na indexação, relacionados entre si semanticamente e genericamente. Representam primeiramente os conceitos para depois especificar suas relações e são desenvolvidos para sistemas pós-coordenados. Constituem-se como Linguagem documental alfabética, controlada, pós-coordenada e combinatória.

Em relação às Listas de cabeçalhos de Assunto são Linguagens pré-coordenadas, alfabéticas e controladas. A origem remota das listas de cabeçalhos de assunto pode ser identificada nos índices alfabéticos de assunto que suplementaram, em muitos casos, os primitivos catálogos de autores. A seleção dos termos desses índices não obedecia a nenhum esquema. Durante muito tempo, para indicar o assunto de um livro, era considerado suficiente o emprego das próprias palavras do título, na mesma forma com que apareciam na folha de rosto, ou na forma invertida para destacar a palavra-chave. Somente quando se verificou, na prática, a insuficiência desse método é que se procurou normalizar os critérios para utilização dos verdadeiros cabeçalhos. Surgiram, então, as listas, como uma decorrência lógica da experimentação prática, servindo como exemplo e sugestão ao trabalho da catalogação de assunto.

Apresentam sinônimos, quase sinônimos e homógrafos. Prevê regras específicas tais como forma de entrada do descritor, uso de singular/plural, uso de abreviaturas e de termos traduzidos. Geralmente são listas gerais que arrolam termos de todas as áreas do conhecimento, tendo em vista o contexto de bibliotecas que incluem várias áreas do conhecimento. As principais listas de cabeçalhos de assunto, atualmente existentes e atualizadas, são:

- Library Congress of Subject Headings - LCSH

É a mais completa lista de cabeçalhos publicada atualmente, adotada por bibliotecas do mundo inteiro. São editados suplementos mensais, acumulados periodicamente, onde se registram os novos cabeçalhos e as alterações ocorridas entre os já usados. A lista de cabeçalhos de assunto da LC é enumerativa, portanto, torna-se necessário consultar cada entrada (e suas subdivisões) para saber se é possível usar determi-

nada combinação de cabeçalhos de assunto e sua subdivisão.

- Sears List of Subject Headings

A lista de Sears, desde 1923 até a atualidade, foi desenvolvida para interpretar as necessidades das bibliotecas de médio e pequeno porte mantendo, ao mesmo tempo, as possibilidades de expansão. Com edições sucessivas, cada vez mais ampliadas, é destinada, também, para grandes bibliotecas. Uma adaptação da lista de Sears para a língua portuguesa, de autoria de Wanda Ferraz, publicada pelo Instituto Nacional do Livro em 1944 e atualizada em 1972 e 1977, foi uma contribuição às pequenas bibliotecas brasileiras.

- Medical Subject Headings - MeSH

Conforme *Bachrach e Charen (1978)* citado por Nelson, Johnston, Humphreys (2001, p.171) “o MeSH thesaurus é o vocabulário controlado da National Library of Medicine (NLM) para indexação e busca por assunto de artigos de periódicos na MEDLINE, e livros, títulos de periódicos e materiais não impressos no catálogo da NLM”. Existem vocabulários que se intitulam listas de cabeçalhos de assunto e são utilizadas em sistemas pós-coordenados onde sua condição de lista de cabeçalhos de assunto é negada, sendo tratado como tesouro, ou seja, oferecem possibilidades de síntese que nas listas mais tradicionais é representada pelos sub-cabeçalhos. Possuem paralelamente ao arranjo alfabético (usam termos da própria linguagem natural e por isso determinam uma ordenação alfabética), um arranjo sistemático ou facetado.

- Lista de cabeçalhos de Assuntos da rede BIBLIODATA- LCARB

A uniformização e padronização das entradas de assunto são garantidas pela manutenção de uma Lista de Cabeçalhos de Assunto (LCARB) para uso de bibliotecas brasileiras participantes da Rede. Esta Lista, representada como a Bases de Autoridades Assuntos, é fundamentada na *Library of Congress Subject Headings (LCSH)*, mantendo a estrutura de linguagem pré-coordenada, respeitando as características da língua portuguesa (Fundação Getúlio Vargas, 1995, p. 11-12). Desde 2013, a Rede Bibliodata está funcionando sob a responsabilidade do Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT), cujo site não apresenta informações sobre a atual situação da LCARB.

- Terminologia de assuntos da Biblioteca Nacional do Brasil

A Biblioteca Nacional do Brasil ([www.bn.br](http://www.bn.br)) disponibiliza online a Terminologia de Assuntos como linguagem documental atualizada pela equipe de bibliotecários em português e inglês com base na *Library of Congress Subject Headings* e coordena um consórcio formado por bibliotecas públicas, universitárias e governamentais.

- Vocabulário Controlado do SIBI/USP - VO-CAUSP

Com esta mesma condição, outras linguagens surgiram com base no vocabulário de listas de assuntos compiladas pelas bibliotecas mais especializadas na elaboração dos catálogos de assuntos que tem uma mistura de princípios de cabeçalhos de assuntos e de tesouros tal como o Vocabulário Controlado do SIBI/ USP – VO-CAUSP que é “é uma linguagem documentária construída, a partir de procedimentos terminológicos e documentários, pelos bibliotecários do SIBI/USP, com a participação de especialistas de todas as áreas do conhecimento abrangidas pelos seus descritores.” (Lima; Boccatto, 2009).

- Descritores em Ciências da Saúde - DeCS

Assim como o vocabulário estruturado trilingue “Descritores em Ciências da Saúde” (DeCS), utilizado para a indexação e recuperação da informação em fontes de informação disponíveis na Biblioteca Virtual em Saúde (BVS), foi desenvolvido a partir do Medical Subject Headings (MeSH) pela Biblioteca Regional de Saúde – BIREME.

A descrição dessas linguagens documentais indica que a Lista de Cabeçalhos de Assuntos da rede BIBLIODATA (LCARB) e a Terminologia de Assuntos da Biblioteca Nacional são linguagens documentais construídas e atualizadas com base na *Library of Congress Subject Headings* e a indicam como fonte primária nas fichas e registros de autoridade (Figuras 3 e 4), portanto, são similares em estrutura e vocabulário. O DeCS é fundamentado e atualizado com base no MeSH, também apresentam estrutura e vocabulário similares na área de saúde e medicina.

O que caracteriza de modo mais marcante essas linguagens documentais são os cabeçalhos de assunto. Os cabeçalhos de assunto apresentam uma gramática bem estruturada, com rígido controle de sinônimos, quase sinônimos, homógrafos, mas, uma grande limitação é que faziam referência (ver também) apenas do geral para o específico e de alguns assuntos correlatos, por isso, os cabeçalhos de assunto eram sistemas em que o princípio de hierarquia não

era constante. Eram linguagens não-hierárquicas basicamente enumerativas.

Como linguagem controlada, prevê regras fixas à respeito de: - forma de entrada direta ou invertida; - formação de síntese através de cabeçalho; - uso de singular e plural; - uso de palavras abreviadas ou termos por extenso; - uso de termos traduzidos e transliterados; - uso de sinais de pontuação.

De acordo com Torres (2001, p. 2) a sintaxe dos cabeçalhos de assuntos se manifesta no universo dos cabeçalhos compostos de que são exemplos típicos os cabeçalhos-frase ou as expressões compostas. Os cabeçalhos de assuntos servem para expressar assuntos ou o conteúdo e/ou forma da informação registrada e entende-se, como sintaxe de cabeçalhos de assunto, a combinação dos elementos ou palavras que formam os cabeçalhos compostos.

A escolha dos cabeçalhos de assunto, antes dos atuais catálogos online, era condicionada pela classificação, servindo mesmo de complemento à classificação dos livros nas estantes. O que a classificação separava, em seus diferentes aspectos, o cabeçalho tornava a unificar num só conjunto. Esta era a diferença fundamental entre os princípios da classificação e os da escolha dos cabeçalhos. Como característica dos cabeçalhos de assunto, destaca-se o seu aspecto de linearidade, o que o torna aplicável apenas a pesquisas unidimensionais.

A coordenação prévia de termos é denominada pré-coordenação e a coordenação de termos na busca, pós-coordenação. Os sistemas pré-coordenados estabelecem a coordenação dos vários tópicos referentes a um assunto composto no momento de indexação, e com grande subjetividade determinam uma prioridade na citação desses elementos (Fundação Getúlio Vargas, 1995, p. 66).

Um sistema pós-coordenado, segundo Pinto (1985, p. 183), “[...] consiste de entradas que normalmente são conceitos únicos (termos indexadores) e de uma saída que permite a comparação das entradas de assunto, com o objetivo de determinara coincidências que revelem documentos pertinentes ao assunto composto pesquisado.”

Em um sistema pré-coordenado a coordenação de termos envolve uma ordem de citação dos termos, gerando uma entrada única. Considerando que os termos representam conceitos, a coordenação dos termos em uma entrada poderá ser priorizada a partir de uma ordem de citação dos conceitos na qual os termos ocupam funções específicas:

Entrada 1: Banco de dados - ecologia  
 Entrada 2: Ecologia - banco de dados = cabeçalho de assunto

No sistema pós-coordenado a entrada é feita por cada um dos termos sem a necessidade de coordenação e prioridade na ordem de citação:

Entradas: Banco de dados  
 Ecologia

Pela ordem de citação todos os termos da entrada, com exceção do primeiro, ficam ocultos pelo sistema pré-coordenado. Antes da recuperação online pelos sistemas de busca, programas de computador foram construídos para gerar índices coordenados que faziam uma rotação de todos os termos para a entrada principal, o que era bastante oneroso para os sistemas de busca manual.

Com o sistema pós-coordenado, o usuário tem a possibilidade de manipular os termos livremente independentemente do número de pontos de acesso, estabelecendo qualquer combinação com todos os termos da busca, de forma que cada uma das combinações terá a mesma relevância. A principal diferença entre os sistemas pré e pós-coordenados está no momento de combinação dos termos pelo usuário durante a estratégia de busca em sistemas de recuperação da informação.

Lancaster (1995, p. 28) expressa essa diferença por meio de representação gráfica, com o exemplo de um documento que foi indexado com quatro termos (FIGURA 2)

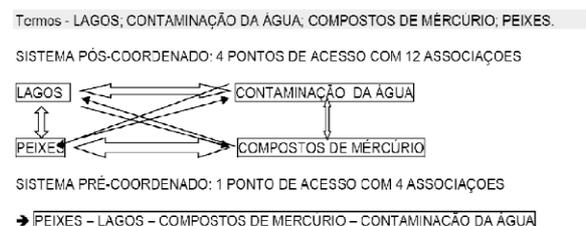


Figura 2. Comparação entre os sistemas pré-coordenados e pós-coordenados (LANCASTER, 1995)

Segundo Lancaster (1995, p. 26) não é preciso uma ordem de termos em um sistema pós-coordenado, pois todos têm o mesmo peso e permitem a recuperação do documento por qualquer um dos termos ou combinação entre eles, enquanto no sistema pré-coordenado o usuário não poderá realizar combinações na ordem que desejar e sim aceita-las na sequência em que foram construídas durante a indexação.

Atualmente, em sistemas de busca online a pré-coordenação perde o efeito de sequência e de

prioridade porque a ferramenta de busca “pega” as palavras onde elas estiverem. A ordem de citação pela prioridade poderia funcionar muito bem para organizar a estratégia de busca em bases de dados e metabuscadores em função da indicação de relevância pelo usuário.

A lista de cabeçalhos de assuntos é a forma mais antiga de linguagem documental baseado no sistema alfabético de assunto. Sua prática desenvolvida na Library of Congress (LC) a partir da elaboração dos catálogos sedimentou o uso do cabeçalho de assunto. São linguagens documentais alfabéticas baseadas em pré-coordenação.

Entretanto, sua atualização para uso nos catálogos online transformou sua configuração e visualização com uma aparência de linguagem documental pós-coordenada tal como a dos tesouros que possuem estrutura lógica e hierarquizada entre os termos.

Houve uma sensível adequação que melhora visivelmente a interface de busca das listas de cabeçalhos de assunto tanto pelo indexador quanto para o usuário como demonstrado nas Figuras 3 e 4, mediante comparação dos registros de autoridade das linguagens LCARB e da Biblioteca Nacional.



Figura 3. Ficha da autoridade – Termo tópico “Tecidos (Anatomia e fisiologia)” da Lista de Cabeçalhos de Assuntos da Rede Bibliodata – LCARB

|                              |   |
|------------------------------|---|
| Descrição                    | Tecidos (Anatomia e fisiologia)   |
| Library of Congress          | ts<br>su<br>es  |
| Remissiva Ver Também (TR)    | TE: Doação de órgãos, tecidos, etc.<br>TE: Epitélio<br>TE: Músculos<br>TE: Periodonto<br>TE: Tecido conjuntivo<br>TE: Venda de órgãos, tecidos, etc.<br>TR: Histologia<br>TR: Órgãos (Anatomia) |
| Fonte positiva dos dados (R) | LC5H  |
| Nota geral pública           | Para anatomia microscópica dos tecidos animais usar Histologia. Para tecidos ou órgãos específicos entrar pelo nome do tecido ou órgão. Ex. Ossos; Cartilagem; Rins; etc.                       |

Figura 4. Ficha da autoridade – Termo tópico “Tecidos (Anatomia e fisiologia)” da Terminologia de Assuntos da Biblioteca Nacional

É possível verificar que na ficha da autoridade do termo tópico “Tecidos (Anatomia e fisiologia)” da LCARB não se demonstra a estrutura hierarquizada de termos genéricos e específicos e tampouco termos relacionados ou associados como ocorre na ficha da Biblioteca Nacional que já avançou com a atualização atualmente praticada pela LCSH.

Partindo da idéia de Jewett de catalogação cooperativa e uso de recursos tecnológicos a Library of Congress (LC) criou a maior Central de Catalogação cooperativa. A introdução da tecnologia computacional na década de 1960 e o desenvolvimento do formato MARC para padronização de registros bibliográficos e formatos de intercâmbio reflete-se até os dias atuais nas bibliotecas e centros de informação.

O desenvolvimento de tecnologias de software e hardware juntamente com a Internet, ferramenta tecnológica de informação e comunicação, proporcionaram às centrais de catalogação cooperativa disponibilizarem seus catálogos em meio magnético, seja online ou em CD-ROM, acelerando a troca de informações.

Para propiciar o intercâmbio de informações, foram criados, junto ao formato MARC, os protocolos Z39.50 e ISO-2709 que orientam sobre a estrutura de um registro e a transferência de dados bibliográficos, o que “[...] permite que um registro criado para uma máquina por um determinado software seja lido por qualquer outro software em qualquer tipo de máquina” (Marasco; Mattes, 1995, p. 42).

Quando um registro bibliográfico é criado, o catalogador faz a descrição bibliográfica do item documentário nos campos devidos e quando se depara com os campos de assunto, recorre à Base de Autoridades Assuntos que mantém os denominados Registros de autoridade que formam, atualmente, a linguagem documental em formato MARC21.

Os registros de autoridade destinam-se a conter informação autorizada para padronizar nomes e assuntos a serem usados como pontos de acesso em registros bibliográficos, bem como permitir os relacionamentos entre eles. Compõem uma base de dados que é consultada, tanto pelo catalogador ao atribuir os termos que representam assuntos, quanto automaticamente pelo próprio sistema para correção de termos e inclusão de remissivas dentro dos registros bibliográficos. Isso significa que os registros de autoridade têm o objetivo de assegurar a constante padronização da linguagem documental de forma automática e podem ser transferidos de uma base de dados para outra na formação dos catálogos de bibliotecas e rede de bibliote-

cas. Abaixo se demonstra um exemplo de registro de autoridade de termo tópico da Biblioteca Nacional na versão de ficha completa para o usuário e no formato MARC21:

|                              |   |
|------------------------------|---|
| Descrição                    | Contratos de afretamento (subdividido geograficamente)  |
| Library of Congress          | h<br>ar<br>te<br>r-<br>p<br>ar<br>ti<br>es  |
| Remissiva Ver (U.S./UF)      | Affreightment<br>Afretamento - Contratos<br>Chartering of ships<br>Contratos de frete marítimo<br>Fretes marítimos - Contratos<br>Ship chartering |
| Remissiva Ver Também (TR)    | TG: Direito marítimo<br>TG: Fretes  |
| Fonte positiva dos dados (R) | LCSH  |

Figura 5. Ficha da autoridade no formato usuário – Termo tópico “Contratos de afretamento (subdividido geograficamente)” da Terminologia de Assuntos da Biblioteca Nacional

```
000 00563cz a2200217o 4500
001 000085744
003 Br
005 20070227114454.6
008 070208d| anznnbabn a ana d
035 _ |a 2007020811064084med
040 _ |a Br |c Br |f Br
150 _ |a Contratos de afretamento
450 1_ |a Affreightment
450 _ |a Afretamento |x Contratos
450 _ |a Chartering of ships
450 _ |a Contratos de frete marítimo
450 _ |a Fretes marítimos |x Contratos
450 _ |a Ship chartering
550 1_ |w g |a Direito marítimo
550 _ |w g |a Fretes
670 1_ |a LCSH
750 _0 |a Charter-parties
```

Figura 6. Ficha de autoridade no formato MARC21 - Termo tópico “Contratos de afretamento (subdividido geograficamente)” da Terminologia de Assuntos da Biblioteca Nacional

Obedecendo ao formato e à Linguagem documental adotada, o catalogador pode formar o cabeçalho, caso não esteja pronto, extraindo os termos autorizados da Base de Autoridades de Assuntos e os distribuindo nos campos e subcampos destinados ao assunto do registro bibliográfico.

Atualmente, os cabeçalhos de assunto são intensamente utilizados por muitas bibliotecas que possuem catálogos manuais e eletrônicos e principalmente por aquelas que disponibilizaram seus catálogos online. Na recuperação da informação, a pré-coordenação de assuntos oferece especificidade e precisão. Na indexação de assuntos, apesar do aspecto prescritivo e enumerativo, a tradução dos termos identificados com a lista de cabeçalhos de assunto pode

ter exaustividade, além de especificidade, porque o formato eletrônico não necessita ser reproduzido na quantidade de termos presentes na entrada de assuntos, além do que, o indexador pode adotar os mesmos conceitos presentes nos cabeçalhos para identificar e selecionar conceitos.

Ressalte-se que no LCSH, a linguagem documental alfabética mais completa e específica da atualidade, evoluiu em suas últimas edições com a introdução e uso da nomenclatura de relações semânticas adotadas por tesouro, incluindo uma estrutura lógica de relações hierárquicas anteriormente atrofiada na linguagem. Essa evolução da LCSH, os formatos de intercâmbio de registros bibliográficos e os catálogos online demonstram que o sistema pré-coordenado e as linguagens alfabéticas pré-coordenadas são importantes para a indexação e recuperação da informação.

Ocorre que, a linguagem documental, no caso das bibliotecas universitárias, nem sempre está disponível na web junto à ferramenta de busca de catálogos online de bibliotecas universitária. Os usuários de catálogos de bibliotecas universitárias não a conhecem e nunca a utilizaram. Pesquisas sobre estratégias de busca da informação em catálogos revelam resultados preponderantes para a recuperação por autor e título e não por assunto. Portanto, a mediação da linguagem documental durante a estratégia de busca em catálogos é inexistente e, em realidade, não precisaria ser utilizada pelos catalogadores, pois a linguagem natural é a mais utilizada pelos usuários.

Para modificar essa situação é necessário que a linguagem seja utilizada no processo de tratamento temático e continuamente atualizada para prever a inserção de termos novos, a adequação e manutenção de vocabulário específico como exigência da especificidade e abrangência de áreas especializadas que caracterizam o acervo bibliográfico de bibliotecas universitárias.

Para tanto, ações para compatibilizar linguagens documentais devem ser implementadas.

### **3. A compatibilização entre linguagens documentais para elaboração, adequação e atualização da linguagem UNESP**

Segundo Gil Leiva (2012) surge nos primeiros anos da década de 1960 os projetos de compatibilidade e integração de vocabulários controlados por meio de tabelas de equivalência para ampliar as buscas em várias bases de dados. Esses projetos foram verificados por Boccato e

Torquatti (2012) mediante estudos de Gardin (1969) e Coates (1970) que introduziram o conceito de léxico intermediário como possibilidade de compatibilização entre linguagens. O léxico intermediário, de acordo com Boccato e Torquatti (2012, p. 83), “opera como referente conceitual, isto é, como uma tabela de correspondência de termos”.

Com base no princípio de construção de léxicos intermediários, Neville (1970) avança no sentido de que a compatibilização além de sintática deve ser semântica ao considerar as significações dos termos descritores expressas em definições/notas de escopo. Conforme explica Boccato e Torquatti (2012, p. 83) o método de Neville (1970) “[...] prevê a elaboração de uma linguagem intermediária, baseada na codificação numérica de conceitos [...] em que se torna possível o estabelecimento da equivalência conceitual entre termos descritores de diferentes linguagens [...]”.

A partir desses estudos, outros foram realizados de modo que a literatura registra a existência de termos convertibilidade e interoperabilidade que são conceitualmente equivalentes.

Atualmente as linguagens documentais são consideradas Sistemas de Organização do Conhecimento na perspectiva das normas internacionais ANSI/NISOZ39:19-2005 e BS 8723-4:2007. Essas normas definem a interoperabilidade como “a capacidade que dois ou mais sistemas tem para trocarem informação e utilizar essas informações” (Tradução nossa).

Segundo Boccato (2011) a interoperabilidade entre vocabulários controlados, focalizada pela norma ANSI/NISO Z39:19-2005 é um ponto de destaque visto como um recurso de construção de sistemas de organização do conhecimento. Segundo a referida autora, aspectos sobre necessidade e os fatores que afetam a interoperabilidade, a construção de vocabulários controlados multilíngues, a fusão de vocabulários controlados, o armazenamento e manutenção das relações entre os termos em vários vocabulários controlados, os métodos de realização da interoperabilidade entre vocabulários controlados, entre outros, são tratados em profundidade, com exemplos e ilustrações que facilitam a compreensão da temática.

Para que a interoperabilidade ocorra entre linguagens documentais, Gil Leiva (2008) considera que dois fatores devem ser considerados: 1) ser fundamental a similaridade temática das diferentes linguagens; 2) compatibilizar linguagens de indexação sobre o mesmo assunto e no mesmo idioma pode economizar tempo e dinheiro.

Decorrente da necessidade de melhorar os resultados de recuperação por assuntos de seu catálogo online, o sistema de bibliotecas universitárias da UNESP decidiu construir a Linguagem UNESP a partir da compatibilização de linguagens com similaridade temática e que possuem os mesmos assuntos com a vantagem de serem procedentes de uma mesma linguagem originalmente construída por e para bibliotecas.

O estudo de avaliação da linguagem documental utilizada pelo sistema de bibliotecas universitárias da UNESP, realizado por Boccato, Fujita, Gil Leiva (2011), verificou o desempenho e o índice de precisão e revocação (LANCASTER, 2004; GIL LEIVA, 2008) na recuperação de assuntos especializadas em 3 áreas de conhecimento no catálogo online ATHENA do sistema de bibliotecas da UNESP. Os resultados obtidos demonstraram que a linguagem natural é a mais utilizada pelos usuários. A explicação para esse problema decorre dos seguintes fatores:

- a Lista de Cabeçalhos de Assuntos da Rede BIBLIODATA (5) (LCARB) utilizada como linguagem documental da UNESP não era disponibilizada na web junto à ferramenta de busca do catálogo ATHENA e, por isso, os usuários da UNESP não a conheceram e nunca a utilizaram;
- além do campo 650 do MARC21, autorizado para os cabeçalhos de assuntos da LCARB, os catalogadores utilizavam o campo 690 para incluir descritores de outras linguagens mais especializadas e termos de linguagem livre;
- inadequação e desatualização da LCARB frente às necessidades de representação específica e precisa da informação de catalogadores e usuários.

Os fatores levantados foram decisivos para a discussão de proposta da política de indexação para as bibliotecas universitárias da UNESP que, após resultados positivos de avaliação de sua aplicabilidade obtidos no período de dois anos, tornou-se o "Manual da Política de Indexação das bibliotecas universitárias da UNESP" (UNESP, 2014).

Com a elaboração de sua política de Indexação a rede de bibliotecas universitárias da UNESP decidiu pelo uso de uma só linguagem que seria construída a partir da compatibilização das linguagens "Terminologia de assuntos da Biblioteca Nacional", *Library of Congress Subject Headings* (LCSH), *Descritores em Ciências da Saúde* (DeCS) e *Medical Subject Headings* (MeSH).

A decisão de usar uma só linguagem foi tomada em função da compatibilização de vocabulário

de uma mesma linguagem a ser utilizada na indexação e na estratégia de busca para obter resultados de recuperação precisa e específica. Dessa forma, a linguagem UNESP poderá ser disponibilizada, futuramente, ao usuário para consulta apresentando contínua avaliação, atualização e adequação com léxico e estrutura lógico-semântica mais especializada e exaustiva.

O processo de compatibilização das linguagens "Terminologia de assuntos da Biblioteca Nacional", *Library of Congress Subject Headings* (LCSH), *Descritores em Ciências da Saúde* (DeCS) e *Medical Subject Headings* (MeSH) resulta na construção da Linguagem UNESP, em desenvolvimento pelo Grupo de Linguagem da UNESP formado por catalogadores e pesquisadores.

De acordo com o "Manual de Política de Indexação das bibliotecas universitárias da UNESP" (UNESP, 2014), a construção da Linguagem UNESP se dá por meio:

- a) da compatibilização do lote de registros de autoridades inicialmente importados da LCARB;
- b) da criação de registros novos para termos importados das linguagens BN, LCSH e MeSH.
- c) da criação de registros novos para termos não encontrados na Bse da UNESP e nas linguagens BN, LCSH e MeSH;

Os registros de autoridade em formato MARC21 destinam-se a conter informação autorizada para padronizar nomes e assuntos a serem usados como pontos de acesso em registros bibliográficos, bem como permitir os relacionamentos entre eles.

Os registros de autoridade de assuntos tópicos e geográficos utilizados para a construção da Linguagem UNESP devem ser elaborados com base em sequência de consulta que sempre se inicia pela Terminologia de assuntos da BN e quando não encontrado segue a consulta na LCSH ou MeSH. Caso o termo não seja encontrado nessas linguagens, um novo registro de autoridade deve ser construído com base em processo terminográfico. Isso significa que os registros importados da LCARB, como descrito no item a) servem de base para o processo de compatibilização. Nos demais itens b) e c) são considerados, respectivamente, registros novos e termos novos sem que haja o processo de compatibilização.

Considerando-se, portanto, os registros inicialmente importados da LCARB, conforme item a),

o processo de compatibilização é realizado com base nos registros de autoridade da Biblioteca Nacional, da LCSH e do MeSH, pois são linguagens que já possuem o registro com estrutura hierárquica e de termos relacionados mais completa e atualizada em relação à LCARB. O processo de compatibilização deve assegurar a verificação de todos os campos a serem completados ou incluídos. Nesse caso, o catalogador solicita ao Grupo de Linguagem, mediante formulário próprio, o processo de compatibilização do registro de autoridade referente ao termo que ele necessita. O catalogador do Grupo de Linguagem, ao abrir o registro da LCARB para compatibilização, verifica todos os campos, inclui a hierarquia dos termos relacionados segundo as linguagens da BN, LC ou MeSH e indica a fonte usada para a correção num campo de notas específico.

Abaixo, verifica-se a sequência da compatibilização com a demonstração de exemplo utilizando o formato:

Ex: Termo solicitado para compatibilização:  
**Plantas alucinógenas**

**Formato:** OPAC Ficha Reduzido Nomarc CamposMARC

|                     |   |
|---------------------|---|
| Número de sistema   | 000223585   |
| Termo tópico        | Plantas alucinógenas  |
| VER - Termo topico  | Hallucinogenic plants DLC<br>Hallucinogens DLC<br>Plants Hallucinogenic DLC<br>Plants Psychedelic DLC<br>Psychedelic plants DLC |
| VER TBM-Termo topic | Alucinógenos<br>Cogumelos alucinógenos  |
| Nota geral interna  | Podem ser subdividido geograficamente<br>Para o Brasil usar a forma direta  |

Figura 7: Registro de autoridade na base de Linguagem da UNESPO (importado da LCARB sem compatibilização)

|   |   |
|---|---|
| Descrição                               | Plantas alucinógenas (subdividido geograficamente)  |
| Library of Congress                     | Hallucinogenic plants   |
| Remissiva Complexa Ver Também - Assunto | Também nomes de plantas alucinógenas. Ex.: Daturas; Maconha; Cogumelos alucinógenos         |
| Remissiva Ver (US/UF)                   | Hallucinogens<br>Plants<br>Hallucinogenic<br>Plants<br>Psychedelic<br>Psychedelic<br>plants |
| Remissiva Ver Também (TR)               | TE: Cogumelos alucinógenos<br>TR: Alucinógenos  |
| Fonte positiva dos dados (R)            | LCSH  |

Figura 8: Registro na Biblioteca Nacional (formato ficha)

**Formato:** OPAC Ficha Reduzido Nomarc CamposMARC

|                     |   |
|---------------------|---|
| Número de sistema   | 000223585   |
| Termo tópico        | Plantas alucinógenas  |
| Ref. VER TBM Assun. | Também nomes de plantas alucinógenas. Ex.: Daturas; Maconha; Cogumelos alucinógenos |
| VER - Termo topico  | Hallucinogens<br>Plants Hallucinogenic<br>Plants Psychedelic<br>Psychedelic plants  |
| VER TBM-Termo topic | Alucinógenos<br>Cogumelos alucinógenos  |
| Nota geral interna  | Podem ser subdividido geograficamente   |
| Fonte encontrada    | BN online   |
| Ligação             | Hallucinogenic plants   |
| Termo específico    | Cogumelos alucinógenos  |

Figura 9: Registro de autoridade da Linguagem UNESP com compatibilização segundo a Terminologia de Assuntos da BN

A proposta metodológica de construção, adequação e atualização de linguagens documentais para bibliotecas universitárias com uso da compatibilização entre linguagens de mesma matriz de estrutura e vocabulário consiste na combinação dos registros de autoridade em formato MARC21 de assuntos tópicos e geográficos de linguagens documentais de mesma origem para descritores da linguagem UNESP. Devem ser realizados com a constante intervenção de um grupo de catalogadores dedicados à esse objetivo que acompanhem a evolução da linguagem com uso de regras específicas de construção de linguagens documentais.

#### 4. Considerações finais

Destacamos que a aplicação da metodologia de compatibilização à elaboração de registros de autoridade em formato MARC21 para a construção de linguagem documental de catálogo online de bibliotecas universitárias distinguiu essa proposta metodológica de outras já desenvolvidas sobre compatibilização para a interoperabilidade entre linguagens documentais. Além disso, ela pode ser vista por uma abordagem mais ampla, isto é, aplicável a outros tipos de bibliotecas (especializadas, nacionais, públicas) e em demais unidades de informação como os arquivos e centros de informação.

Conclui-se que a compatibilização é um recurso viável para a construção de linguagens documentais no que se refere a otimização de tempo, racionalização de recursos financeiro e humano provenientes das instituições e das próprias bibliotecas universitárias envolvidas, entre outros aspectos.

#### Notas

(1) Linguagem documentária é denominação utilizada por Lancaster (2004) e Gil Leiva (2008), autores representativos da linha "Indexing" e Linguagem documental ou Documental é utilizada por autores como Guimarães (2008), Gil Urdiciain (2004), autores representativos da linha de "Análise Documentária"

(2) Medical

(3) *Pós-coordenação*: coordenação de termos na busca - Passa a idéia de uma linguagem em que o instrumento de controle de vocabulário exerce uma função sugestiva) como é o caso do *Medical Subject Heading* (MESH, usado com o sistema MEDLARS

(4) [http://areatecnica.sibi.usp.br/wiki/index.php/Vocabul%C3%A1rio\\_Controlado](http://areatecnica.sibi.usp.br/wiki/index.php/Vocabul%C3%A1rio_Controlado)

(5) A Lista de Cabeçalhos de Assuntos da Rede BIBLIODATA (LCARB) foi utilizada pelo Sistema de Bibliotecas da UNESP até 2013.

## Referencias

Bachrach, C. A.; Charen, T. (1978) Selection of MEDLINE contents, the development of its thesaurus, and the indexing process. *Medical Informatics*. (3) (1978) 237-254.

Boccatto, V. R. C. (2011). Os sistemas de organização do conhecimento nas perspectivas atuais das normas internacionais de construção. *InCID: R. Ci. Inf. e Doc.* 2:1 (janeiro/junho 2011) 165-192.

Boccatto, V. R. C.; Torquetti, M. C. (2012). Interoperabilidade entre linguagens de indexação como recurso de construção de instrumento de representação temática de clippings de coordenadorias de comunicação social em ambientes universitários: uma proposta metodológica. *Informação & Informação*. 17:3 (setembro/dezembro 2012) 76-101.

Boccatto, V. R. C., Fujita, M. S. L., Gil Leiva, I. (2011). Avaliação comparada do uso de linguagens de indexação em catálogos de bibliotecas universitárias para recuperação por assunto. *Scire*. 17 (2011) 55-64.

Batista, G. H. R. (1986). Compatibilidade e convertibilidade entre linguagens de indexação: um estudo de caso. Rio de Janeiro, 1986. 176 f. Dissertação (Mestrado em Ciência da Informação) - Instituto Brasileiro de Informação em Ciência e Tecnologia; Universidade Federal do Rio de Janeiro.

Campos, M. L. de A. (2006). Integração de ontologias: o domínio da bioinformática e a problemática da compatibilização terminológica. En: Encontro Nacional de Pesquisa em Ciência da Informação, 7., 2006, Marília. *Anais...* [S.l.: s.n.]. Recuperado 20-04-2015 de <http://goo.gl/7qB3Bh>

Campos, M. L. A.; Gomes, H. E.; Campos, L. M. (2011). Integração e compatibilização em ontologias. En.: SILVA, F. C. C. da; SALES, R. de (orgs.). (2011). *Cenários da organização do conhecimento: linguagens documentárias em cena*. Brasília: Thesaurus, 169-200.

Cervantes, B. M. N. (2009). *A construção de tesouros com a integração de procedimentos terminográficos*. 2009. 209 f. Tese (Doutorado em Ciência da Informação) – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília.

Fujita, M. S. L. (2010) *Política de indexação para bibliotecas*. Marília: Faculdade de Filosofia e Ciências, UNESP. (Projeto de Produtividade em Pesquisa do CNPq).

Fujita, M. S. L. (2012). A política de indexação para representação e recuperação da informação. En: Gil Lieva, I.; Fujita, M. S. L. (eds.). (2012). *Política de indexação*. São Paulo: Cultura Acadêmica, 17-28. Recuperado 15-04-2015 de [http://www.culturaacademica.com.br/\\_img/arquivos/Livro%20Politica-de-indexacao\\_ebook.pdf](http://www.culturaacademica.com.br/_img/arquivos/Livro%20Politica-de-indexacao_ebook.pdf)

Fujita, M. S. L.; Rubi, M. P.; Boccatto, V. R. C. (2009). O contexto sociocognitivo do catalogador em bibliotecas universitárias: perspectivas para uma política de tratamento da informação documentária. *Datagramazero* -

*Revista de Ciência da Informação*. 10. Recuperado 10-04-2015 de <http://www.dgz.org.br/>

Fujita, M. S. L.; Boccatto, V. R. C.; Rubi, M. P. (2012). A política de indexação elaborada por catalogadores: um estudo de observação participante em cursos de capacitação. En: Pulgarín Guerrero, A.; Vivas Moreno, A. (org.). (2012). *Limites, fronteras y espacios comunes: encuentros y desencuentros en las ciencias de la información*. Badajoz: Universidad de Extremadura, 90-101.

Fujita, M. S. L. (org.). (2009). *A indexação durante a catalogação de livros na percepção de catalogadores e usuários de bibliotecas universitárias*: um estudo de observação do contexto sociocognitivo com protocolos verbais. São Paulo: Editora da UNESP.

Fujita, M. S. L. et al. (2012) *Proposta de política de indexação para as bibliotecas universitárias da UNESP*. Marília: Faculdade de Filosofia e Ciências, UNESP; São Paulo: Coordenadoria Geral de Bibliotecas, Reitoria da UNESP.

FUNDAÇÃO GETÚLIO VARGAS. Rede BIBLIODATA/CALCO. (1995) *Manual de cabeçalhos de assunto: normas e procedimentos: versão 1.0*. Rio de Janeiro.

Gil Leiva, I. (1999) *La automatización de la indización de documentos*. Gijón: Ediciones Trea.

Gil Leiva, I. (2008) *Manual de indización: teoría y práctica*. Gijón: Ediciones Trea.

Gil Leiva, I. Fujita, M. S. L. (orgs.). (2012). *Política de indexação*. São Paulo: Cultura Acadêmica; Oficina Universitária

Gil Urdiciain, B. (2004). *Manual de lenguajes documentales*. Gijón: Ediciones Trea. 280p.

Guimarães, J. A. C. (1990). Recuperação temática da informação. *Rev. Bras. Bibliotecon. e Doc.* 16:3/4 (janeiro/dezembro 1990) 112-30.

GUIMARAES, J. A. C. (2008). A dimensão teórica do tratamento temático da informação e suas interlocuções com o universo científico da International Society for Knowledge Organization (ISKO). *Revista Ibero-Americana de Ciência da Informação*. 1 (2008) 77-99.

Lancaster, F. W. (2004). *Indexação e resumos: teoria e prática*. 2. ed. Brasília: Briquet de Lemos.

Lancaster, F. W. (1995) *El control del vocabulario en la recuperación de información*. 2. ed. Valencia: Universitat de València.

Lima, V. M. A.; Boccatto, V. R. C. (2009) O desempenho terminológico dos descritores em Ciência da Informação do Vocabulário Controlado do SIBI/USP nos processos de indexação manual, automática e semi-automática. *Perspect. ciênc. Inf.* 14:1 (April 2009) 2009 Recuperado 12-03-2015 de <http://goo.gl/xSzCKL>.

Marasco, L. C.; Mattes, R. N. (1995) Avaliação e seleção de software para automação de centros de documentação e bibliotecas. *Cadernos da FFC*. 4(1), (1995) 40-49.

Nelson, S. J.; Johnston, W. D.; Humphreys, B. L. (2001). Relationships in medical subject headings (MeSH). In: Bean, C. A.; Green, R. (eds.). *Relationships in the organization of knowledge*. Boston, MA: Kluwer Academic Publishers, 171-184

Pinto, M. C. M. F. (1985) Análise e representação de assuntos em sistemas de recuperação da informação: linguagens de indexação. *R. Esc Bibliotecon.UFGM*. 14:2 (setembro 1985) 169-186.

Torres, L. M. C. (2001) *Sistematização da sintaxe de cabeçalho de assunto*. [Rio de Janeiro]: Conexão Rio. Recuperado 14-05-2014 de <http://www.conexaorio.com/bit/lecy/lecy.htm>. Acesso em: 14 jun. 2014.

UNESP. Coordenadoria Geral de Bibliotecas. (2014) *Manual da política de indexação para as bibliotecas universitárias*

*rias da UNESP*. São Paulo: Coordenadoria Geral de Bibliotecas.

VISCAYA ALONSO, D. ([1997]) *Lenguages documentarios*.  
Corrientes: Nuevo Parhadigma.

---

# Organização e Representação do Conhecimento sobre Seca no Brasil: o caso do Núcleo Temático da Seca e do Semiárido da UFRN

---

*Organization and Representation of Drought Knowledge in Brazil: the case of Núcleo Temático da Seca at UFRN*

---

**Monica Marques Carvalho Gallotti (1), Jacqueline Aparecida de Souza (2)**

(1) Faculdade de Letras da Universidade do Porto, Porto, monica\_mcg@hotmail.com

(2) jackebci@gmail.com

## Resumo

A seca, por ser um fenômeno natural que assola vários países e atinge grandes proporções, demanda e reforça a necessidade de criar estratégias que visem a organização, representação e difusão da informação e do conhecimento sobre o assunto. Assim, o Núcleo Temático da Seca e do Semiárido (NUTSECA) é um centro de documentação especializado em seca que se dedica a coletar, organizar, representar e difundir informação nesse âmbito. O presente trabalho objetiva descrever ações e projetos empreendidos no referido núcleo em torno da gestão da informação sobre seca com vistas à sua representação e difusão em larga escala em meio digital. Neste artigo, são relatados os projetos realizados e resultados alcançados em parceria com importantes instituições de fomento à pesquisa. Para o apoio à dimensão teórica do trabalho, recorreu-se à metodologia da pesquisa bibliográfica na área de organização e representação do conhecimento. Ao final, é possível inferir que a socialização da informação sobre seca e o relato de experiências são medidas necessárias na atual conjuntura, pois favorecem a criação de novos conhecimentos que potencialmente podem gerar uma inovação desejada na área para tentar superar esse problema secular.

**Palavras-chave:** Seca – Socialização da Informação. Seca – Organização do conhecimento. Núcleo temático da Seca e Semiárido – Projetos.

## 1. Introdução

Contextos informacionais organizados possibilitam a construção de novos conhecimentos, assim como geram inovações no campo social, político ou econômico de uma sociedade. A seca é um fenômeno natural que tem assolado vários países e se tornado um problema mundial complexo. Organizar, representar e difundir informação e conhecimento sobre a seca é de fundamental importância para que se consiga a

## Abstract

Drought is considered a natural phenomenon that has reached global major proportion. Thus there is the need to develop strategies that aim to organize, represent and disseminate information and knowledge regarding this subject. The Núcleo Temático da Seca e do Semiárido (NUTSECA) is a documentation Center that is specializes in collecting, organizing, representing and in the diffusion on the drought information. This paper aims to describe projects realized in NUTSECA that deal with representation and dissemination of information on drought in digital platforms. Results are pointed out and were achieved in partnership with leading development research institutions. To support the theoretical dimension of the work there was a literary review regarding information and knowledge representation. In the end, it is possible to infer that information socialization on drought case study results are made necessary since they favor the creation of new knowledge that can potentially generate desired innovation in the area to try to overcome this secular problem.

**Keywords:** Drought - Information.Socialization Drought - Knowledge Organization - Projects

superação desse problema secular. Este trabalho visa descrever projetos e resultados em preendidos no Núcleo Temático da Seca e do Semiárido (NUTSECA), um centro de documentação especializado em seca da UFRN no Brasil. Para subsidiar a dimensão teórica do trabalho, recorreremos à metodologia da pesquisa bibliográfica, norteando-nos a partir das premissas conceituais presentes na área da organização do conhecimento (OC). O cotejo entre a dimensão teórica e prática do trabalho fornecerá

elementos necessários para compreender como a informação sobre a seca pode ser tratada, visando a sua socialização.

## 2. Organização do conhecimento: conceitos, princípios e premissas

A OC, enquanto disciplina, estuda as leis, os princípios e os procedimentos pelos quais se estrutura o conhecimento especializado. Visa representar tematicamente e recuperar a informação contida no documento de qualquer natureza, por meios eficientes que forneçam resposta rápida às necessidades dos utilizadores (Barité, 2008). Para tanto, nutre-se de aportes teóricos recebidos da Informática, da Linguística e da Terminologia. Enquanto disciplina científica, social, de caráter aplicado, objetiva otimizar a circulação e o fluxo do conhecimento. Seu objeto de estudo é o conhecimento socializado ou registrado, sendo no escopo da Ciência da Informação que se desenvolvem as teorias e práticas para a construção, a gestão, o uso e a avaliação de classificações, taxonomias, nomenclaturas, ontologias e linguagens de indexação. Abarca, além disso um conjunto de conhecimentos vinculados à análise da informação, considerando os aspectos semânticos, cognitivos e informáticos, abrangendo diversos aspectos da OC, como seus atributos, função, objetivos, instrumentos e processos envolvidos. Conforme Dahlberg (2006) advoga, a OC é

the science of structuring and systematically arranging of knowledge units (concepts) according to their inherent knowledge elements (characteristics) and the application of concepts and classes of concepts ordered by this way for the assignment of the worth knowing contents of referents (objects/subjects) of all kinds.

Assim, a construção e a definição de um sistema de conceitos e relações compreendem processos intelectuais e metodológicos, os quais conferem a representação do conhecimento de um domínio de especialidade Hjørland (2008) complementa que a OC é um campo de estudo interessado na natureza e na qualidade dos processos implicados na própria OC: a descrição de documentos, a indexação e a classificação.

No tocante aos problemas de busca e recuperação da informação, o autor destaca nove princípios da OC (Lima & Alvares, 2012):

1. A percepção realístico-ingênua de estruturas do conhecimento não é possível em ciências complexas;
2. Categorizações e classificações devem reunir assuntos relacionados e separar assuntos distintos;

3. Para fins práticos, o conhecimento pode ser organizado de diferentes formas, para diferentes objetivos;
4. Qualquer categorização deve refletir o próprio objetivo;
5. Categorizações e classificações sempre podem ser questionadas;
6. Observar sempre o conceito de polirepresentação;
7. Diferentes áreas do conhecimento podem ser organizadas de diferentes formas para o mesmo fenômeno;
8. A natureza das áreas é variável;
9. A qualidade da produção do conhecimento, em muitas áreas e em alguns momentos, pode ficar vulnerável.

Tais princípios auxiliam e norteiam a construção de sistemas de organização do conhecimento (SOC)(1). Sob as mesmas orientações, Barité (2001 cit. por Fogaça & Cervantes, 2013) expõe dez premissas básicas para a OC:

- 1º. O conhecimento é um produto social, uma necessidade social e um dinamismo social;
- 2º. O conhecimento realiza-se a partir da informação e, ao socializar-se, transforma-se em informação;
- 3º. A estrutura e a comunicação do conhecimento formam um sistema aberto;
- 4º. O conhecimento deve ser organizado para seu melhor aproveitamento individual e social;
- 5º. Existem 'n' formas possíveis de organizar o conhecimento;
- 6º. Toda organização do conhecimento é artificial, provisória e determinista;
- 7º. O conhecimento é sempre registrado em documentos, como conjunto organizado de dados disponíveis, e admite o uso indeterminado;
- 8º. O conhecimento expressa-se em conceitos e é organizado por sistemas de conceitos;
- 9º. Os sistemas de conceitos são organizados para fins científicos, funcionais ou de documentação;
- 10º. As leis que regem a organização dos sistemas de conceitos são uniformes e previsíveis e aplicam-se por igual a qualquer área disciplinar.

É oportuno destacar o primeiro, o segundo e o sétimo princípios elencados por Barité (2011), pois enfatizam a dimensão social do conhecimento, a importância da socialização, o conhecimento registrado e a disponibilização dos dados disponíveis. Assim, ressalta-se que as ações empreendidas estão correlacionadas a esses princípios.

Entende-se por SOC as representações do conhecimento, ou seja, modelos de abstração do mundo real, construídos para determinada finalidade. São sistemas de classificação dicionários, ontologias, taxonomias, tesouros, redes semânticas, entre outros (Schiesl & Shintaku, 2012). Estes explicitam o conhecimento de um domínio; possibilitam a organização e recuperação de informações; auxiliam a construção de mapas de conhecimento, a navegação para busca de informação e a criação de um novo conhecimento a partir do existente. Vickery (2008 cit. por Schiesl & Shintaku, 2012) afirma que os SOC cumprem diferentes funções e cronologicamente podem ser assim classificados:

- Era da pré-coordenação (2): estruturas estáticas que atendem às necessidades dos sistemas manuais de organização e recuperação da informação, como índices e catálogos;
- Era da pós-coordenação (3): mais dinâmicos, possibilitam que cada um de seus elementos seja manipulado de forma independente para representar assuntos de cada documento, como vocabulários controlados e tesouros;
- Era da internet: classificações hierárquicas que orientam o usuário na escolha do termo que melhor expressa sua questão de busca, como taxonomias;
- Era da *web* semântica: são projetados por agentes inteligentes, como ontologias.

Barité (2011) aponta as características essenciais do SOC: referem-se ao conhecimento especializado, que se estabelece de diversas formas (classificação científica, garantia literária, opinião de especialistas etc.); apresentam estrutura lógica, que se constrói de acordo com um método e uma teoria de OC e prevê o controle do vocabulário, que contribui para a seleção, depuração, formalização e normalização da terminologia. Esta inclui o estabelecimento de relações recíprocas entre os termos, considerando critérios semânticos, linguísticos e disciplinares.

Entendemos que a OC deve ser aplicada a contextos informacionais específicos com vistas a produzir impactos pontuais sobre determinada realidade. Neste cenário, as ações empreendidas no NUTSECA expressam essa perspectiva.

### 3. Núcleo Temático da Seca e do Semiárido

O NUTSECA é um Centro de Documentação especializado em seca vinculado ao CCSA da UFRN que iniciou suas atividades em meados de 1980, a partir da compreensão da seca enquanto um fenômeno amplo e, sobretudo, interdisciplinar. Representa um esforço no sentido

de repensar a seca e seus impactos. Nesse sentido, visa a organização, representação, gestão e difusão da informação sobre seca em meio convencional e digital, bem como subsidiar a criação de políticas sociais inovadoras.

Seu acervo reúne diferentes abordagens e pontos de vista sobre o fenômeno e fornece ao pesquisador uma visão epistemológica interdisciplinar. Está dividido em quatro grandes coleções: A Universidade e a Questão Nordestina, Seca e Semiárido, Carnaúba e a do Vale do Assú ou Projeto Baixo-Assú. As informações contidas nessas coleções são socializadas a partir de projetos.

#### 3 Projetos em torno da socialização de informação sobre seca

Destacaremos alguns projetos pontuais que corroboram as premissas conceituais estabelecidas anteriormente.

##### a) Projeto Finep-MCT

“Tecnologia a serviço da Socialização da Memória das Tradições do Rio Grande do Norte: caso do Núcleo Temático da Seca”. Com a participação de uma equipe interdisciplinar, procedeu-se à organização do acervo do Nutseca, com vistas à digitalização de sua coleção. Para subsidiar as ações, foi criado o Laboratório de Tecnologia da Informação (LIBER). Os pressupostos teóricos basearam-se na área da Gestão Eletrônica de Documentos. As informações e resultados alcançados foram socializados em meio digital (4).

##### b) Digitalização da Coleção da Carnaúba e difusão de informação sobre seca

Devido à quase inexistência de informações sobre a Carnaúba no estado do RN e do Piauí, desenvolveu-se o projeto Catálogo Virtual da Carnaúba, de autoria da Professora Isa Maria Freire e equipe, com o apoio do BNB. Visou o levantamento de fontes de informação a respeito do assunto, bem como a sua digitalização e difusão. A digitalização deu-se no âmbito do LIBER tendo como resultado, a disponibilização do catálogo online (5) que norteou o desenvolvimento de outros projetos.

##### c) Vocabulário Controlado da Carnaúba (2011 - 2013)

O Vocabulário Controlado da Carnaúba, foi construído com vistas a auxiliar a indexação das fontes digitalizadas, implicando um tratamento temático do acervo. Foram aplicados os pressupostos teórico-metodológicos da Linguística de Corpus, que se ocupa da coleta e exploração de um conjunto de dados linguísticos

textuais selecionados criteriosamente, com o propósito de servir à pesquisa de uma língua ou variedade linguística, dedicando-se à exploração da linguagem por meio de evidências empíricas, extraídas por computadores (Berber Sardinha, 2001, p. 3). Sob essa orientação, utilizou-se como apoio computacional o E-Termos, um ambiente computacional colaborativo *web* de acesso livre e gratuito dedicado à gestão terminológica, cujo principal objetivo é viabilizar a criação de produtos terminológicos, sejam para fins de pesquisa acadêmica, sejam para divulgação, por meio da semiautomatização das etapas do trabalho terminológico. Ao final, destaca-se que o maior desafio imposto está associado à compilação do *corpus*, especificamente na busca por diferentes gêneros textuais sobre o assunto em tela. Nessa perspectiva, foram obtidos 500 termos representativos do domínio da carnaúba.

#### 4. Conclusão

Conforme a proposta apresentada, relataram-se algumas das ações que, articuladas, proporcionaram a socialização da informação via representação e organização do conhecimento. Todas essas atividades estão vinculadas, correlacionadas e buscaram atender às diferentes funções dos SOC, tanto na era pós-coordenada como na era da internet. Convém observar que os empreendimentos enquadraram-se no escopo das premissas para a OC, com destaque ao primeiro, segundo e sétimo, ao considerar o conhecimento como produto social, realizado a partir da informação e que ao socializar-se, transforma-se em informação, e ao fato de o conhecimento ser sempre registrado em documentos, como conjunto organizado de dados disponíveis. Entendemos que os contextos informacionais devem ser organizados sob a visão de propostas concretas de OC para que seja possível atingir os propósitos básicos de socializar a informação/conhecimento, promovendo ações infocomunicacionais.

O NUTSECA pode ser percebido um lócus privilegiado que cumpre seu papel como agente no intuito de superar problemas advindos da seca e outras questões relacionadas. Embasados nas características peculiares da cibercultura (Lévy, 2009), e da sociedade em rede (Castells, 2000), outros projetos ligados à OC poderão ser desenvolvidos, tais como especificamente, a indexação colaborativa, representação e tratamento de imagens e encontrabilidade, com vistas a promover a acessibilidade e o uso das coleções que compõem o acervo e ampliar e contribuir com as pesquisas sobre o fenômeno da seca.

#### Notas

- (1) Contemplam esquemas que apoiam a organização da informação e do conhecimento, tais como: linguagens de indexação, vocabulário controlado, tesouros, taxonomias, redes semânticas.
- (2) Utilizada em sistemas tradicionais, os termos são combinados no momento da indexação, durante a inserção do material no sistema, como, por exemplo, as classificações bibliográficas.
- (3) Quando os termos são combinados no momento da busca.
- (4) Disponível em: <http://ccsa.ufrn.br/nutseca/NUT-Seca/Nut%20Seca.html>
- (5) Disponível em: <http://ccsa.ufrn.br/nutseca/NUT-Seca/D321B381-DDD9-40B4-A0AE-D03FC542BA32.html>-contem maiores detalhes.

#### Referências

- Alvares, L. (2012). *Organização da Informação e do conhecimento: conceitos, subsídios interdisciplinares e aplicações*. São Paulo: B4Editores.
- Barité, M. (2008). *KODictionary: Dicionário da Organização do Conhecimento*.
- Barité, M. (2011). Sistemas de Organización del Conocimiento: una tipología actualizada; Sistemas de Organización do Conhecimento: uma tipologia atualizada. *Informação & Informação*, 16(2), 122-139.
- Sardinha, A. P. B. (2001). Uso de corpora na formação de tradutores. *Delta*, 19, 43-70.
- Campos, M. L.A., & Gomes, H. E. (2006). Princípios de organização e representação do conhecimento na construção de hiperdocumentos. *Pesquisa Brasileira em Ciência da Informação e Biblioteconomia*, 1(1).
- Carlan, E., & Medeiros, M. B. B. (2011). Sistemas de Organização do Conhecimento na visão da Ciência da Informação. *Revista Ibero-Americana de Ciência da Informação*, 4(2).
- Castells, M. (2000). "A sociedade em rede". 4. ed. São Paulo: Paz e Terra.
- Carvalho, L. M., Carvalho, M. M., Freire, I. M., & Aranha, T. D. Q. (2008). Ampliando o acesso livre a informação: a digitalização do acervo do Núcleo Temático da Seca.
- Carvalho, M. M., Carvalho, L. M., & Freire, I. M. (2011). A digitalização do acervo do Núcleo Temático da Seca da UFRN como possibilidade de socialização da informação. *Interface*, 1(1).
- Carvalho, R. P. F. de. (1999). Núcleo temático da seca/UFRN: uso do acervo informacional. *Informação & Sociedade: Estudos*, 9(2).
- Dahlberg, I. (2006). *Knowledge organization*. Consultado em: junho 01, 2015, Recuperado a partir de: [http://www.iva.dk/bh/lifeboat\\_ko/CONCEPTS/knowledge\\_organization\\_Dahlberg.htm](http://www.iva.dk/bh/lifeboat_ko/CONCEPTS/knowledge_organization_Dahlberg.htm)
- Fogaça, M. G., & Cervantes, B. M. N. (2013). Estudo de terminologias da área jurídica em sistemas de organização do conhecimento à luz do código de processo civil brasileiro: com ênfase na classificação decimal universal e classificação decimal de direito. In *Anais do Congresso Brasileiro de Biblioteconomia, Documentação e Ciência da Informação-FEBAB* (Vol. 25, pp. 5320-5340).
- Freire, I. (2004). A rede de projetos do núcleo temático da seca da UFRN como possibilidade de socialização da informação. *Informação & Sociedade: Estudos*, 14(2).

- Hjørland, B. (2008). What is knowledge organization (KO)? *Knowledge organization. International journal devoted to concept theory, classification, indexing and knowledge representation*.
- Lévy, P. (2009). *Cibercultura*. São Paulo: Editora 34.
- Lima, J.L.O. & Alvares, L. (2012). Organização e representação da informação e do conhecimento. In L. Alvares (org), *A organização da informação e do conhecimento: conceitos, subsídios interdisciplinares e aplicações* (21-47). São Paulo: B4.
- Schiessl, M., & Shintaku, M. (2012). Sistemas de organização do conhecimento. In L. Alvares (org), *A organização da informação e do conhecimento: conceitos, subsídios interdisciplinares e aplicações* (49-118). São Paulo: B4.
- Silva, D. L. D., Souza, R. R., & Ramos, J. A. D. A. (2010). *Representação da informação e do conhecimento em sistemas de informações: contribuições de campos interdisciplinares*.

---

# Organização da Informação em Contexto Dinâmico e Sensível: Tesouro de Inteligência Policial

*Organization of Information in Sensitive and Dynamic Context: Thesaurus of Police Intelligence.*

---

**Rodrigo Benjamin Gesteira (1), Maria Aparecida Moura (2)**

(1) Universidade Federal de Minas Gerais, rodrigo.gesteira@gmail.com

(2) Universidade Federal de Minas Gerais, mamoura@ufmg.br

## Resumo

Esse estudo, desenvolvido no âmbito da Organização da Informação, tomou como referência o caráter informacional das operações de inteligência policial e partiu da hipótese de que seria possível desenvolver uma metodologia que orientasse a elaboração de um tesouro capaz de auxiliar a atividade de inteligência policial na organização das informações sensíveis coletadas e geradas no contexto da atividade e no combate a organizações criminosas. Para tanto, observou-se as peculiaridades do contexto, no qual a informação está inserida em ambiente de diversificadas fontes de dados, revestido do sigilo e de extrema dinamicidade. Como resultado, elaborou-se o Tesouro Eletrônico de Operações de Inteligência Policial (TEOIP) que incluiu como inovação a agregação de termos oriundos da análise efetuada no linguajar da organização criminosa observada. Dessa tarefa, inferiu-se que seria possível fazer do TEOIP, além de uma ferramenta para organizar a informação, oriunda das operações de inteligência policial, um sistema de alerta do cometimento ou da preparação de determinada ação pela Organização Criminosa. Além disso, o estudo realizado permitiu evidenciar que as práticas adotadas pela inteligência policial no combate a organizações criminosas reivindicam a composição de uma base teórica interdisciplinar e que as contribuições consolidadas no campo da Ciência da Informação, notadamente no âmbito da organização e uso da informação e do conhecimento. Dessa perspectiva, a Ciência da Informação ofereceu aporte teórico-metodológico fundamental para o entendimento da ação criminosa como um fenômeno informacional dinâmico e agravado pela sensibilidade das ações que compõem a sua caracterização, monitoramento e combate.

**Palabras clave:** Inteligência policial, tesouros - metodologia, informação sensível.

## 1. Introdução

As operações de inteligência policial (OIP) se constituem essencialmente em complexos informacionais. Durante o desenvolvimento de

## Abstract

This study, developed in the scope of Information Organization, has taken as reference the informational side of the police intelligence's operations and started with the hypothesis that it would be possible to develop a methodology that would guide the creation of a thesaurus, capable of helping the police intelligence's activity, in fighting criminal organizations. In order to do so, the singularities of the context were analyzed, in which the information is inserted in an environment of multiple databases, covered by secrecy and extremely dynamic. As a result, the Electronic Thesaurus of Police Intelligence's Operations (TEOIP) was developed, that includes as innovation the aggregation of terms coming from the analysis made in the language of the criminal organization observed. From this task, it was inferred that it would be possible to make the TEOIP become, besides a tool to organize the information that came from the police intelligence's operations, a system of alert of committal or the preparation of certain action by the criminal organization. Besides that, the study has allowed to evidence that the practices adopted by the police intelligence in the combat of criminal organizations claim for the composition of an interdisciplinary theoretical base and that the contributions consolidated in the field of Information Science were very important to the investigations, notably in the scope of organization and use of information and knowledge. From this perspective, the IS can offer theoretical and methodological contributions, keys to the comprehension of criminal action as a dynamic informational phenomenon and aggravated by the sensibility of the actions that make up its characterization, monitoring and combat.

**Keywords:** Police Intelligence, thesauri - methodology, sensible information.

uma OIP, articulam-se inúmeros desafios ligados à informação, advindos principalmente do crescimento exponencial dos dados disponíveis, que turvam o ambiente e dificultam, ou até

mesmo impossibilitam, a conclusão dos trabalhos de combate a organizações criminosas.

O contexto em que ocorre a atividade de inteligência policial é extremamente dinâmico, envolto em desinformação e, ainda, cercado pelo sigilo, fatores que aumentam o grau de dificuldade para se alcançar o objetivo pretendido.

Esse ambiente dinâmico e sensível requer metodologias para tratar as informações geradas no decorrer das operações integrando documentos tradicionais, documentos eletrônicos e fluxos informacionais produzidos no curso de ações criminosas.

Essa última, o grande desafio, pois exige a implementação de instrumentos documentários capazes de orientar a indexação (1) em tempo real de toda e qualquer ação de determinada organização criminosa (Orcrim), de maneira a que a partir dessa intermediação seja possível recuperar os assuntos de interesse. Para tanto, é necessário indexar os fluxos informacionais ocorridos no desenrolar da operação de inteligência policial.

A ciência da Informação (CI), mais especificamente nos aspectos referentes à organização e uso da informação, permeia as ações desenvolvidas no contexto acompanhado, entretanto estas ações precisam receber também uma reflexão específica da atividade de inteligência policial. Para a Organização da Informação, a garantia literária, a garantia de uso e a garantia estrutural são bases para a construção de um tesouro (MOREIRA; MOURA, 2006). Todavia, no caso de sua aplicação na atividade de inteligência policial, foi preciso incorporar novos elementos terminológicos oriundos do contexto que revelam a dinâmica do campo, mas que não estavam ainda repertoriados na literatura da CI.

Assim, a meta estabelecida foi a de criar um instrumento útil para organizar e recuperar informações registradas e aquelas ainda difusas nos contextos de atuação das operações policiais.

Diante das opções de instrumentos intelectuais para organização e uso da informação, concluiu-se que o tesouro apresentava características que atenderiam às demandas colocadas, já que sua origem está relacionada com a necessidade de manipulação de grande quantidade de documentos especializados de determinadas áreas que exigem maior sistematização para a recuperação da informação (MOREIRA; MOURA, 2006).

Estruturalmente, o tesouro “é um vocabulário controlado e dinâmico de termos relacionados semântica e genericamente cobrindo um domí-

nio específico do conhecimento” (UNESCO, 1976, p. 5). Sua principal característica é a presença de relações associativas e de equivalência entre os termos e sua finalidade é promover a coerência na indexação de conteúdos de objetos, especialmente para sistemas de recuperação e armazenamento da informação pós-coordenados, podendo ser usado também na recuperação de conteúdos de objetos em pesquisa de texto livre (ANSI/NISO Z 39.19: 2005).

Diante desse cenário, e tendo como corpus a Operação de Inteligência Policial Carga Suja (2), idealizou-se o Tesouro Eletrônico de Operações de Inteligência Policial (TEOIP) capaz de abarcar o contexto apresentado.

Durante o desenvolvimento do TEOIP, através de uma perspectiva multidisciplinar, inferiu-se que o tesouro, além de se prestar à organização das informações, também poderia servir como um sistema de alerta para a atividade policial.

Assim, alicerçada na Ciência da Informação, mas utilizando de conhecimentos advindos da Filosofia, Linguagem, Terminologia e Análise de Conteúdo, foi realizada uma análise do linguajar da organização criminosa, cujo resultado se prestou a fornecer o contexto terminológico que caracteriza e segmenta as ações criminosas. Tais termos identificados foram inseridos no tesouro.

Esses termos, quando de sua entrada no Sistema de Recuperação da Informação, teriam a capacidade de emitir proativamente um alerta sobre a preparação ou cometimento de determinada ação de interesse para a operação policial.

Com isso, além de ser uma ferramenta que apoia o trabalho policial ao tratar a informação no contexto dinâmico e sensível, o tesouro proposto ainda tem o potencial de agregar em seu interior a capacidade de alertar a equipe policial para o acompanhamento de possível ação delituosa a partir do monitoramento da linguagem.

O estudo e suas conclusões foram abordados detidamente em pesquisa desenvolvida junto ao PPGCI/UFMG (3). Desse trabalho, optou-se por destacar três de seus tópicos: a metodologia de construção do tesouro voltado a operações de inteligência policial; a indexação proposta, tendo em vista ser uma maneira de indexação que orienta práticas, sujeita ao dinamismo das atividades inerentes às operações policiais; e o sistema de alertas como uma inovação funcional à ferramenta tesouro.

Antes, no entanto, apresentam-se sinteticamente alguns elementos que permeiam a discussão e que ajudarão o leitor a compreender melhor o

ambiente em que foi pensada a pesquisa. Para isso, será abordado o contexto dinâmico e sensível das operações de inteligência policial com suas múltiplas fontes de dados, evidenciando a importância da organização e uso da informação nessas operações.

## 2. Contexto dinâmico e sensível das operações de inteligência

As operações de inteligência policial se desenvolvem em cenários bastante dinâmicos e peculiares, envoltos em uma miríade de variáveis que afetam a forma como ocorre a coleta e sistematização da informação. Nele, a atenção do profissional incumbido de tratar os dados coletados está voltada para um fluxo que corresponde a fatos sociais praticados pela organização criminosa. A indexação, em parte, é efetuada em tempo real, acompanhando o passo a passo da Orcrim.

Diante do contexto, passa-se a tratar a realidade e as etapas das operações policiais como se documentos fossem, indexando as ações. Essas ações ficariam sujeitas a reindexações em função do curso da atividade.

De modo sintético, apresenta-se a seguir algumas funções essenciais no trabalho das equipes de inteligência, que são: coletor, analista e equipe de campo. Esses sujeitos atuam como os “profissionais da informação” no âmbito das operações policiais.

O coletor, conforme o próprio nome já indica, é o policial incumbido de coletar os dados. Sua atuação é normalmente interna, nos escritórios de inteligência. Dessa maneira, vai operar, principalmente, nas estações de Sigint (4) (Inteligência de sinais).

Tomar-se-á como referência o exemplo de Sigint mais difundido pela imprensa que é a interceptação telefônica. Assim, o coletor é aquele policial responsável por ouvir e classificar os áudios decorrentes da interceptação telefônica. Todavia, na atividade policial, suas tarefas não ficam restritas à coleta, pois, devido ao fato de receber em tempo real as ligações telefônicas de um alvo da operação, ele fica responsável também por acionar e orientar a equipe de campo. Dessa maneira, fica a critério desse profissional o momento do acionamento da equipe de campo quando os diálogos indicam a eminência de ocorrer alguma Situação de interesse.

No curso dessas atividades, cabe ao coletor a responsabilidade de tornar acessível os dados importantes gerados durante o acompanhamento da referida Situação, uma vez que deverão

estar disponíveis para serem recuperados por outro membro da equipe policial que porventura venha a ser incumbido de efetuar a análise e a consolidação das informações armazenadas.

Importante salientar que uma Situação é qualquer evento, criminoso ou não, que a equipe de inteligência identifique como relevante. Por isso, sua identificação, indexação e posterior recuperação se consubstanciam como uma das atividades mais importantes em uma operação de inteligência policial.

Vista a função do coletor, passamos à equipe de campo, também conhecida como equipe de rua, que opera em ambiente externo e desenvolve-se também como uma atividade de coleta.

Com atuação bem diversificada, uma de suas funções é o acompanhamento in loco das atividades da Orcrim, de modo a efetuar o registro em imagem e/ou som de determinada Situação. Ao final, sistematiza os dados coletados em um documento que descreverá o ocorrido. Normalmente é redigido um texto remetendo a fotografias, vídeos e demais recursos coletados e utilizados. Esse documento receberá um número e uma indexação que o vincule aos demais documentos gerados na operação, de maneira a que, em um segundo momento, o analista possa recuperá-los em conjunto com outros que também possuam identidade com o evento.

Também é a equipe de campo que efetua levantamentos referentes à localização de pessoas físicas e jurídicas, bem como a busca por imóveis registrados em cartórios, ação importante para identificação e recuperação dos ativos frutos das práticas ilícitas perpetradas pela Orcrim. Normalmente também é o responsável pela coleta de informações com fontes humanas, conhecidas no jargão policial como informantes.

Por último, a função do analista. Esta se distancia um pouco da dinâmica em que estão mergulhados coletores e equipe de campo. Diferente das demais funções que trabalham premidas pelo tempo, uma vez que a coleta para dar um retorno adequado deve ser feita o mais próximo possível do tempo real, a análise dispõe de um período maior para sua execução.

Desenvolvendo o seu trabalho, o analista se assemelha a um pesquisador. Ele utiliza os dados coletados e sistematizados pelos coletores e equipes de campo com a missão de produzir o conhecimento.

O ethos profissional da atividade de análise em inteligência e suas regras de produção do conhecimento são as mesmas que governam qualquer outra atividade de pesquisa (Cepik, 2001, p. 61).

Para aprofundar o entendimento de como se realiza o trabalho do analista, imaginemos que ele se depara com um único documento classificado pelo coletor como relevante, e que o conteúdo indexado remete a uma conversa que poderia indicar um acordo entre empresários e políticos no sentido de direcionamento de uma licitação pública.

De posse dessa informação, o analista vai efetuar pesquisas em bancos de dados que estejam à sua disposição no intuito de encontrar novos conteúdos que possam complementar aquela documentação inicial. Uma estratégia de busca possível é recuperar e escutar as conversas telefônicas que ocorreram em período aproximado da data do áudio classificado como relevante. Essa tarefa não é simples. Dependendo da operação, por vezes coletam-se centenas de áudios em um único dia, que se acumulam por semanas, meses e anos, compondo uma massa informacional de grandes proporções que se torna humanamente impossível conhecer seu conteúdo.

Nesse contexto, a indexação de conteúdo ao disponibilizar descritores específicos sobre os assuntos presentes em cada um desses áudios, possibilita uma busca por assuntos de interesse.

Assim, o analista solicitaria uma busca no sistema, por exemplo, pelo termo “licitação”. Com isso, poderia recuperar os áudios presentes naquele universo que versassem sobre o referido assunto.

De posse de um universo de áudios menor, torna-se possível escutá-los e, a partir deles, obter novos dados que somados aos anteriormente obtidos possibilitam ao analista tecer inferências a respeito da Situação. A partir daí, efetua-se nova coleta em bancos de dados disponíveis, em fontes abertas, como o Diário Oficial que porventura tenha publicado o edital da licitação, e nos demais meios que entenda pertinentes. Além disso, faz requerimentos para que a equipe de campo e os coletores empreendam novos esforços de coleta que venham a suprir os vazios informacionais em sua análise.

Feito esse apanhado geral, passa-se a um exemplo fictício que congrega o trabalho dos coletores, analistas e equipes de campo.

O policial coletor está em sua estação de trabalho acompanhando os áudios oriundos de uma interceptação telefônica. A cada áudio que escuta, necessariamente deverá indexar os assuntos de seu conteúdo, ainda que aquela conversa não aparente relevância. A indexação é obrigatória, pois no trabalho posterior do analista, este pode entender, devido a conhecimento advindo

de múltiplas fontes, que esse áudio passou a ter relevância.

Caso o coletor identifique de imediato a relevância do áudio, indexará o conteúdo e indicará a necessidade de sua transcrição para uso pelo analista.

Por vezes, o áudio é importante para a consolidação do Relatório de Inteligência (RI) (5) e precisa apenas ser indexado para propiciar sua recuperação pelo analista no futuro. No entanto, há áudios relevantes nos quais o coletor precisa fazer as devidas indexações e, por fazer referência a algum evento que demanda acompanhamento pela equipe de campo, acioná-la para registrar o evento in loco.

Nesse momento, a atividade do coletor aumenta em complexidade e responsabilidade, pois ao mesmo tempo em que escuta e indexa o áudio, deverá acionar e manter informada a equipe de campo. Além disso, o acompanhamento demanda novas tarefas a serem atendidas pelo policial coletor que se encontra na base de operações, que, em paralelo, efetuará diversos levantamentos demandados pela equipe de campo, como o levantamento de endereços, placas de carros e consultas a mapas e operadoras de telefonia, por exemplo.

O sucesso desse acompanhamento em contexto dependerá de fatores inerentes tanto ao trabalho do coletor quanto da própria equipe de campo. O primeiro precisa captar e repassar em sua completude as informações relevantes. E mais, tem que fazer com que essas informações cheguem a tempo de serem úteis à equipe de rua. Quanto à equipe de rua, deve possuir um nível de coordenação entre seus componentes de maneira que suas ações converjam para a consecução do resultado. A sensibilidade desses policiais em intuir os limites de suas atuações de maneira a que os alvos acompanhados não percebam a presença deles é primordial. Às vezes, o que está em jogo são meses de trabalho que podem ser perdidos devido a um pequeno erro da equipe de campo, que pode propiciar aos integrantes da Orccrim descobrirem que estão sendo acompanhados. Fato que pode comprometer a continuidade da operação de inteligência.

Longe dessa turbulência, o analista pesquisa no banco de dados da operação os documentos gerados em momentos anteriores pelos coletores e equipes de campo, para transformar aquela massa de dados em conhecimento potencial para compor o Relatório de Inteligência.

Durante uma operação de inteligência, os fatos narrados acima se repetem com frequência,

acarretando num número enorme de Situações acompanhadas. Daí a necessidade de se organizar os dados coletados em cada uma dessas Situações de maneira a que o policial encarregado de efetuar a análise consiga identificar aqueles documentos que guardam relação com a mesma Situação. Esse é o maior desafio enfrentado pelo analista, haja vista o volume de dados que compõem o acervo de uma operação de inteligência.

Entendida a dinâmica presente nas operações de inteligência policial, na próxima seção é apresentada a adequação metodológica realizada na construção do Tesouro Eletrônico de Operações de Inteligência Policial (TEOIP), iniciando pelo campo de abrangência que o TEOIP cobre tematicamente.

### 3. Metodologia de construção do TEOIP

O problema de pesquisa direcionou os estudos para o universo relativo à atividade de inteligência policial. O corpus utilizado foi a operação contra organização criminosa especializada em roubo de caminhões e cargas denominada Carga Suja. Essa operação foi sintetizada em um Relatório de Inteligência (RI), que se configurou no principal produto analítico gerado pelo escritório de inteligência policial. Todo fluxo informacional tem por finalidade a construção desse relatório. O modelo de RI adotado na Operação Carga Suja foi composto por 04 camadas analíticas: Análise Geral; Análise de Situação; Análise de Alvo; Análise Financeira e Patrimonial.

Esses parâmetros delimitaram o campo específico de conhecimento que foi coberto pelo TEOIP: partindo do pressuposto que os três grandes assuntos de interesse de uma operação de inteligência contra Orccrim são os crimes que praticam; a forma como operam e se organizam; e a sua atividade financeira e patrimonial, idealizou-se três dimensões que abrangeram os referidos assuntos. Elas estão demarcadas da seguinte forma:

Uma dimensão adjetiva foi nomeada como Dimensão Aspecto da Orccrim, abrangendo os conceitos que qualificam determinado grupo de indivíduos como uma organização criminosa, bem como aqueles que agravam essa condição. Os conceitos compreendidos nessa dimensão foram aglutinados na *Faceta Aspectos da Orccrim*.

O fluxo material da ação criminal foi nomeado como Dimensão Financeira e Patrimonial e abrangeu os conceitos referentes ao patrimônio e ao fluxo financeiro da Orccrim e de seus integrantes. Os conceitos que identificam o patri-

mônio obtido por meio da atividade ilícita foram aglutinados na *Faceta Financeira e Patrimonial*.

A dimensão verbal nomeada Eventos abrangem os conceitos que representam os fatos sociais praticados pela Orccrim. Nesse sentido, os conceitos relativos a cada um dos fatos sociais relevantes foram aglutinados na *Faceta Eventos*.

Dessa forma, ficou definido o facetamento do TEOIP, conforme figura abaixo.

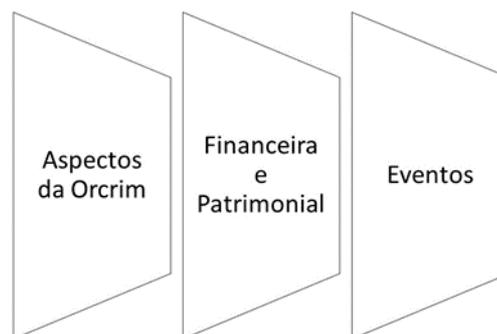


Figura 01. Facetas do TEOIP

Chamamos a atenção para o fato de que não foi idealizada uma dimensão correspondente à Análise de Alvo, que também compõe o Relatório de Inteligência. O motivo é explicado pelo fato de que não identificamos uma dimensão específica que contivesse os assuntos pertinentes à Análise de Alvo. Na verdade, a Análise de Alvo vai buscar nas três dimensões propostas, quais sejam, Dimensão Aspectos da Orccrim, Dimensão Financeira e Patrimonial e Dimensão Eventos, o subsídio para sua construção.

Definidas as facetas, passa-se a descrever o procedimento utilizado para conceber o conteúdo de cada uma delas, de maneira a destacar o caráter metodológico dos procedimentos adotados. Nesse percurso, destaca-se o método dedutivo, através do qual foram extraídos os conceitos mais genéricos, para em seguida relacionar outros termos a essa categoria, partindo do geral para o particular. Para isso, buscou-se identificar em outros tesouros que, porventura, contivessem os termos eleitos, além de dicionários e glossários especializados de modo a compreender o alcance contextual e abrangente de cada termo incorporado ao tesouro (CAMPOS; GOMES, *on line*).

#### 3.1. Faceta Aspectos da Orccrim

O primeiro passo para construção dessa faceta foi a consulta a fontes que versam sobre organizações criminosas. Com base na revisão de

literatura, foram extraídos conceitos da Lei 12.850/2013, que dispõe acerca das Organizações Criminosas, bem como aqueles presentes na doutrina da área. Além disso, foram utilizados os elementos obtidos a partir da experiência dos autores junto a policiais que atuaram nas diversas operações de inteligência. Da análise desse material e dos elementos obtidos na prática policial, foi possível extrair os termos que representam os principais conceitos contidos nessa faceta.

A título ilustrativo, apresentamos três dos termos escolhidos, apontando respectivamente, o conceito representado e a justificativa da escolha.

**USO DE ARMA DE FOGO** - Demonstra que na atuação da Orccrim houve emprego de arma de fogo. Justificativa: O uso de arma de fogo por membros da Orccrim majora as penas aplicadas.

**ASSOCIAÇÃO DE PESSOAS** - Demonstra a associação de 4 (quatro) ou mais pessoas estruturalmente ordenadas. Justificativa: A associação de pessoas é característica necessária para a comprovação da existência de uma Orccrim.

**CONEXÃO COM POLICIAL** - Demonstra a participação de policial nas ações da Orccrim. Justificativa: A existência de conexão com servidor público policial majora as penas aplicadas.

Foi possível constatar que a Faceta Aspectos da Orccrim não trouxe grande número de termos preferenciais, somando 18 ao todo. Seu escopo se restringe basicamente às características de uma organização criminosa.

Também foi possível concluir que ela não apresenta variação à medida que o tesouro seja voltado a outra operação de inteligência policial em tema diverso do roubo a caminhão. Explicase: independente de uma Orccrim ser especializada em fraudar licitações ou traficar drogas, as características relevantes ao domínio que se pretende cobrir com essa faceta irão se manter, pois o que caracteriza a Orccrim não tem a ver com o tipo de crime que pratica, mas sim elementos ligados à forma como operaram e se organizam. Estes continuarão a existir ainda que em nichos diferentes.

A partir das observações feitas, afirma-se que a faceta Aspectos da Orccrim é composta por elementos gerais, no sentido de que perpassa e poderá ser utilizada em toda e qualquer operação de inteligência contra o crime organizado, independente do ramo de atividade em que a Orccrim atue.

Diante dessa estabilidade e do número reduzido de termos descritores, e levando em consideração a dinâmica do trabalho do policial encarregado da tarefa de coleta, agregou-se à Faceta Aspectos da Orccrim dezoito acrônimos notacionais formados por apenas três letras que representarão os termos eleitos, diminuindo o tempo de digitação dos indexadores.

Essa decisão teve por objetivo dar mais agilidade ao trabalho de indexação efetuada pelo policial coletor. Sua função exige rapidez na execução das tarefas, pois há a necessidade de se acompanhar em tempo real as ações da Orccrim, indexando os áudios e ao mesmo tempo coordenando a equipe de campo.

O padrão adotado para os acrônimos notacionais é composto por três letras, sendo que a última sempre será o "X". O motivo de tal escolha se deveu a pouca incidência no vocabulário da língua portuguesa de palavras terminadas em "X", o que aumenta a precisão na recuperação da informação. Segue alguns exemplos dos acrônimos e os termos para os quais serão usados:

| <b>Acrônimo</b> | <b>Relação</b> | <b>Termo</b>           |
|-----------------|----------------|------------------------|
| <b>ARX</b>      | UF             | Uso de arma de fogo    |
| <b>ASX</b>      | UF             | Associação de pessoas  |
| <b>ATX</b>      | UF             | Atuação transnacional  |
| <b>CPX</b>      | UF             | Conexão com policial   |
| <b>CRX</b>      | UF             | Corrupção              |
| <b>DIX</b>      | UF             | Dimensão interestadual |
| <b>DVX</b>      | UF             | Divisão de valores     |
| <b>FOX</b>      | UF             | Forma de operar        |
| <b>LDX</b>      | UF             | Conotação de liderança |
| <b>ORX</b>      | UF             | Organização            |
| <b>OVX</b>      | UF             | Obtenção de vantagem   |
| <b>VIX</b>      | UF             | Violência              |

Tabela I. *Acrônimos e respectivos termos*

### 3.2. Faceta Financeira e Patrimonial

Seu escopo compreende elementos que caracterizem aspectos fiscais, financeiros e patrimoniais. Engloba termos que representam conceitos conexos com patrimônio, fluxo financeiro, demonstração da situação contábil das pessoas e empresas, mensuração da quantidade de dinheiro transacionada por um indivíduo ou empresa, identificação dos bens móveis e imóveis, e declarações de Imposto de Renda, por exemplo.

A título ilustrativo, apresentam-se cinco de seus termos, apontando, respectivamente, o conceito representado e a justificativa da escolha:

**CONTA BANCÁRIA** - Demonstra a existência de conta bancária movimentada pela Orccrim ou por pessoas relacionadas a ela. Justificativa: Relevante ao levantamento financeiro, fiscal e patrimonial.

**LAVAGEM DE DINHEIRO** - Demonstra a existência de ocultação de patrimônio ou rendimentos obtidos de forma ilícita e também a posterior inserção de tais bens no mercado como se lícito fossem: Justificativa: Relevante ao levantamento financeiro, fiscal e patrimonial.

**PESSOA INTERPOSTA** - Demonstra o emprego por parte do alvo de pessoa (s) interposta (s). Pessoa interposta é aquela popularmente conhecida como “laranja”. Justificativa: Relevante ao levantamento financeiro, fiscal e patrimonial.

**PATRIMÔNIO** - Demonstra a posse, pelo alvo ou seus aliados, de bens móveis e/ou imóveis. Justificativa: Relevante ao levantamento financeiro, fiscal e patrimonial.

**SONEGAÇÃO FISCAL** - Demonstra a ocorrência de irregularidades fiscais ou tributárias praticadas tanto por pessoas físicas, quanto jurídicas, envolvidas na investigação. Justificativa: Relevante ao levantamento financeiro, fiscal e patrimonial.

Na sequência, repetiu-se o procedimento desenvolvido na faceta Aspectos da Orccrim, com buscas em tesouros, dicionários especializados e outras fontes. Construiu-se assim o *lôcus* e as relações de cada termo.

Com a consolidação dessa parte do TEOIP, foi possível identificar também nessa faceta um caráter geral, no sentido de que não apresentou aspectos particulares que trariam diferenças na sua construção caso a Orccrim empreendesse sua atividade para a prática de crimes diferentes do roubo a carga e caminhões. Ainda que a Orccrim seja voltada para outros ilícitos, as características relevantes ao domínio que pretendemos cobrir se mantêm: patrimônio, sonegação fiscal, uso de pessoa interposta, conta bancária e lavagem de dinheiro.

Pelos motivos já expostos na Faceta Aspectos da Orccrim, e tendo em vista o número pequeno de termos, foram eleitos acrônimos notacionais em substituição aos termos preferidos, os quais estão apresentados na tabela abaixo.

| <b>Acrônimo</b> | <b>Relação</b> | <b>Termo</b>        |
|-----------------|----------------|---------------------|
| CBX             | UF             | Conta bancária      |
| LDX             | UF             | Lavagem de dinheiro |
| PIX             | UF             | Pessoa interposta   |
| PTX             | UF             | Patrimônio          |
| SNX             | UF             | Sonegação fiscal    |

Tabela II. *Acrônimos e respectivos termos*

A partir dos termos descritores, completamos a faceta com outros termos que compõem o universo pesquisado, trazendo sua relação com os descritores.

### 3.3. Faceta Eventos

Conforme determinado na escolha do campo de conhecimento do TEOIP, ele possuirá dentre suas facetas, uma relativa à dimensão verbal nomeada Dimensão Eventos, abrangendo os conceitos que representam os fatos sociais praticados pela Orccrim que denotem interesse à operação de inteligência policial.

Os termos da Faceta Eventos serão destinados a indexar os documentos que se refiram às Situações que ocorrem em determinada operação de inteligência policial. Por exemplo, o conjunto de documentos relacionados à movimentação de uma Orccrim no sentido de fraudar uma licitação.

O TEOIP está focado para servir ao corpus dessa pesquisa, que é uma operação de inteligência sobre roubo de carga e caminhões. Assim, essa faceta tem por missão congregar o universo de termos que possa representar os fatos sociais envolvidos nesse nicho que sejam relevantes.

O primeiro passo para a extração dos termos foi a leitura dos documentos denominados como Análise de Situação afetos à Operação de Inteligência Policial Carga Suja. Dessa leitura, elegemos os termos ‘roubo’, ‘furto’, ‘extorsão’, ‘receptação’, ‘sequestro’ e ‘falsificação de documentos’.

De posse desses termos, passou-se à consulta de tesouros já existentes, identificando no Microtesouro de Ciências Criminais (SALDANHA, 2005) e no Tesouro do Superior Tribunal de Justiça os termos ‘roubo’, ‘furto’, ‘extorsão’, ‘receptação’, ‘sequestro’ e ‘falsificação de documentos’ com seu *lôcus* e relações convencionadas de acordo com a política dos referidos tesouros.

Foram lançados os termos e suas relações e iniciada a modelagem da Faceta Eventos de

acordo com os propósitos e convenções de interesse ao TEOIP.

Com as pesquisas e o início da consolidação dos termos da Faceta Eventos, foi possível perceber que ela não guarda semelhanças com as características das Facetas Aspectos da Orçim e Financeira e Patrimonial.

Não se pode classificá-la como geral, pois caso a Orçim atue em atividade criminosa diferente do roubo de cargas e caminhões, os delitos envolvidos seriam diversos dos que compõem essa faceta do TEOIP, forçando a identificação e eleição de outros termos ainda não repertoriados. Esse fator também orienta a opção de não fazer uso de acrônimos notacionais, pois a variedade de termos descritores tende a ser grande. O uso nas demais facetas foi decidido segundo o argumento de dar agilidade ao trabalho do coletor no momento da indexação. Concluiu-se que nessa faceta essa medida não seria funcional. A quantidade de acrônimos criados seria grande. Isso impossibilitaria uma pesquisa rápida em lista que apresentasse os acrônimos e os termos aos quais se refere, demandando mais esforço do coletor, o que somente atrapalharia.

A título ilustrativo, apresenta-se abaixo o termo Roubo de caminhão e suas relações na Faceta Eventos:

#### **Roubo**

BT1: Crime contra o Patrimônio

BT2: Crime

BT3: Delito penal

BT4: Delito

BT5: Situação (Fato social)

NT1: Roubo Circunstanciado

NT1: Roubo de Carga

NT1: Roubo de Veículo Automotor

NT2: Roubo de caminhão

NT2: Roubo de Carro

NT1: Roubo Impróprio

NT1: Roubo Próprio

Figura 02. *Termo Roubo e suas relações*

Concluída a descrição da metodologia que orientou a concepção das facetas do TEOIP, passa-se a descrever a funcionalidade agregada ao tesouro no intuito de que servisse também para alertar acerca da prática de atividade ilícita pela organização criminosa.

#### **4. Nova funcionalidade do Tesouro: o sistema de alerta**

O TEOIP incorporou um pacto de significação da linguagem dos envolvidos na operação Carga Suja. A comunicação desses indivíduos foi tomada como uma dimensão importante, pois, durante os trabalhos em uma operação policial, é possível observar comportamentos, flexões na voz ou uso de determinados termos pelos alvos da operação que indicam o cometimento de ilícitos. Algumas expressões utilizadas pelos indivíduos servem como códigos que indicam a prática de ilícitos ou se relacionavam com objetos frutos dessa atividade.

Dessa constatação, veio o desafio de tentar representar um fluxo de informação que 'não queria aparecer'. Ou melhor, aparecia sob o manto de uma codificação, como numa espécie de criptografia. Para enfrentar esse obstáculo, buscou-se aporte teórico na Análise de Conteúdo, que forneceu subsídios, dentre outros, para a escolha do campo de incidência da análise (BARDIN, 1997).

A partir da apreciação da estrutura dos documentos presentes no Relatório de Inteligência da Operação Carga Suja, conclui-se que a extração deveria incidir sobre as análises e transcrições dos diálogos efetuadas pelos policiais nas Análises de Situação, Análise de Alvos e do Relatório Geral, documentos que apresentaram o insumo necessário aos objetivos da pesquisa, uma vez que estarão presentes os diálogos ocorridos entre os alvos, bem como aquilo que foi apontado em suas conclusões pelos policiais responsáveis pela operação.

Preparado o material, partiu-se para a etapa de extração dos termos e seus significados.

Identificou-se no corpus aquele fluxo de informação escamoteada. A estratégia de identificação foi analisar os diálogos a partir do contexto em que estariam inseridos. O desafio aqui foi considerável, pois o sentido dado ao termo está repleto de nuances e subjetividades.

Nesse aspecto, tomaram-se por referência as reflexões de Wittgenstein acerca dos jogos de linguagem, nos quais se observa a existência de regras no uso da linguagem que transformam as palavras, vistas como unidades de significação, de acordo com o contexto em que são empregadas, pois, através desse jogo, os indivíduos envolvidos em um diálogo vão construindo de maneira dinâmica os significados. O jogo de linguagem é um conjunto de regras que afetam os sistemas de comunicação devido ao fato de essas regras estarem relacionadas ao cotidiano dos indivíduos e as suas práticas diárias, resul-

tando que os indivíduos aprendam o significado das palavras com o exercício da prática, utilizando-as em determinado contexto. Sendo assim, os referenciais utilizados na linguagem estarão associados ao contexto de seu uso. Caso esse contexto sofra mudanças, isso incidirá diretamente na significação (WITTEGENSEIN, 1999).

Na prática da Operação Carga Suja, os jogos de linguagem estavam presentes, o que dificultou o entendimento do que foi falado pelos integrantes da organização criminosa. Foi preciso entendê-los a partir do nicho social e cultural ao qual pertenciam.

Exemplo emblemático da codificação existente nos diálogos foi identificado na expressão “Lombardi”. Na linguagem da Ocrim, Lombardi é uma referência a caminhão que não pode circular em rodovias fiscalizadas, pois não possui documentação, ficando restrito para uso em pátio interno de determinada empresa ou propriedade particular. O “Lombardi” é uma comparação com o locutor de um tradicional programa da televisão brasileira, o programa Silvio Santos. A peculiaridade do uso do termo “Lombardi” pelos criminosos se deve ao fato de que o verdadeiro Lombardi da televisão era apenas uma voz, pois o indivíduo nunca aparecia na tela da televisão. Dessa forma, a expressão “Lombardi”, citada na Carga Suja, se referia ao conceito “caminhão roubado e que não recebeu uma documentação falsa em substituição à original”. Esse tipo de caminhão é negociado no mercado clandestino para ser usado em propriedade particular, ficando livre de fiscalização policial. Ou seja, é um caminhão que não pode ser visto pelas autoridades.

Outro exemplo que serve de ilustração foi o uso da expressão “tomar uma”, que no contexto dos indivíduos acompanhados significava um chamamento ou um aviso de que iriam se reunir para praticar roubo de caminhão. Se observada sob a ótica da população brasileira em geral, a expressão “tomar uma” estaria relacionado com um convite para beber uma cerveja ou algo similar.

Para encontrar os termos referenciados acima e outros que seguiam a mesma lógica, foi realizada uma leitura intensiva nas transcrições e análises, bem como a audição dos diversos áudios que compunham o corpus da pesquisa, de maneira a identificar as expressões que se encaixavam naquilo que o referencial teórico assinalava.

Desse esforço, foi possível identificar sessenta e duas (62) expressões, das quais se inferiu, de acordo com a leitura do ambiente em que foram

travadas, o seu significado. Apresentar-se-á, abaixo, quadro demonstrativo com um extrato de algumas dessas expressões.

| <i>Expressão</i>                  | <i>Possível significado</i>   |
|-----------------------------------|---|
| Lombardi.                         | Caminhão que não pode circular em vias com fiscalização.                                |
| Estar a pé.                       | Receptador reclamando por não ter recebido o veículo ainda.                             |
| Hoje vai ter.                     | Assaltante confirmando que vai roubar veículo.  |
| Nós vai quebrar o vidro (SIC)     | Indicação de que estão prestes a proceder ao roubo do veículo.                          |
| Carro de bandido.                 | Caminhão roubado.   |
| Desembolar.                       | Resolver, desenrolar.   |
| Você vai hoje?                    | Convocação para participar da equipe que vai praticar o roubo de caminhão naquela data. |
| Caíram.                           | Foram presos.   |
| Caminhão esquentado.              | Caminhão roubado que recebe documentação e fica com aparência de legalizado.            |
| Fez só o papel.                   | Fez apenas a documentação falsa do veículo.   |
| Os meninos estão na pista.        | A quadrilha está agindo no roubo a caminhão.  |
| Passei uma folha.                 | Entregou dinheiro.  |
| Vai sair hoje pra trabalhar?      | Questionamento ao chefe se irão cometer roubo de caminhão naquele dia.                  |
| Caixa de negócio.                 | Caixa de ferramentas apropriadas a desativar o aparelho GPS de rastreamento do veículo. |
| Os Homens.                        | A polícia.  |
| Vai me dar.                       | Vai me entregar para a polícia.   |
| Nós vai beber cerveja hoje. (SIC) | Comparsas combinando a prática do roubo.  |
| Lombrou.                          | Deu errado. A polícia apareceu.   |

Tabela III. *Expressões utilizadas pela Ocrim*

Extraídas as expressões, a tarefa seguinte foi efetuar a inclusão das que guardariam características comuns aos termos pertencentes a uma das três facetas já constituídas do tesouro, de maneira a que pudessem ser incorporadas como novos termos do TEOIP. Nem todas as expressões extraídas do corpus foram utilizadas, pois não possuíam características em comum com os termos já presentes no TEOIP que per-

mitissem sua inclusão em uma das três facetas idealizadas.

Abaixo, é apresentada a inserção como termos do TEOIP das expressões “Estou Trabalhando”; “Hoje vai ter”; “Hoje vai, né?”; “Negócio”; “Nós vai beber cerveja hoje”; “Nós vai quebrar o vidro”. Essas expressões, na conotação do falar dos envolvidos na Orccrim, faziam referência ao roubo de caminhões. Elas foram inseridas na Faceta Eventos. Assim, no exemplo, tem-se o termo mais geral da faceta Eventos, “Situação (Fato social)”, e, na sequência, são apresentadas as relações até chegar ao termo “Roubo de caminhão”. Abaixo dele, expressões precedidas de “UF”, indicativo de que são termos não preferidos, para os quais o TEOIP indica que sejam substituídos por “Roubo de caminhão”.

### **TERMO: Roubo de caminhão**

Situação (Fato social)

Delito

Delito penal

Crime

Crime contra o Patrimônio

Roubo

Roubo de Veículo Automotor

#### **Roubo de caminhão**

UF: Estou Trabalhando

UF: Hoje vai ter

UF: Hoje vai, né?

UF: Negócio

UF: Nós vai beber cerveja hoje

UF: Nós vai quebrar o vidro

UF: Os meninos estão na pista.

UF: Pegou?

UF: Tomar uma

UF: Trabalhar

UF: Vai sair hoje pra trabalhar?

UF: Vamos embora?

UF: Tô passando aí pra te buscar

UF: Vai hoje não?

Com essa construção, objetivou-se adicionar mais uma funcionalidade ao TEOIP. A ideia foi de que ao inserir essas expressões como termos do tesouro, além de possibilitar uma correta indexação por parte do coletor, esse pudesse ser de alguma maneira alertado pelo sistema de coleta, por exemplo, o sistema de interceptação telefônica, caso as referidas expressões surjam nos diálogos de operações versando sobre roubo de cargas e caminhões, o que serviria de um indicativo ao policial coletor de que a ação da

Orccrim relativa ao cometimento do referido delito estivesse ocorrendo ou na eminência de ocorrer.

## **5. Conclusões**

Esse trabalho partiu da hipótese de que seria possível desenvolver uma metodologia que orientasse a construção de tesouro capaz de auxiliar a atividade de inteligência policial no combate a organizações criminosas.

Foi delineado e entendido o contexto em que se desenvolvem as citadas operações, bem como sua abrangência conceitual, o que possibilitou determinar o campo coberto pelo tesouro e suas facetas.

Levou-se em conta peculiaridades desse contexto, no qual a informação está inserida em ambiente de diversificadas fontes de dados, revestido do sigilo e de extrema dinamicidade. Buscou-se consolidar uma metodologia para organizar e acompanhar informações presentes não só em documentos, mas em relação a diálogos que ocorrem em tempo real, com os efeitos que a dinâmica de uma Operação de Inteligência Policial imputa aos policiais designados para a coleta e análise de dados, pois esses profissionais personificam os usuários da informação do universo pesquisado.

O método foi descrito e experimentado através da materialização do TEOIP. Buscou-se inovação, agregando ao tesouro termos oriundos da análise efetuada no linguajar da Orccrim.

Dessa tarefa, inferiu-se possível fazer do TEOIP, além de uma ferramenta para organizar a informação, um sistema de alerta do cometimento ou da preparação de determinada ação pela Orccrim. Todavia, tem-se a consciência de que essa tarefa demanda a aglutinação sistemática de conhecimento que só poderá ser realizada com o acúmulo de pesquisas que tomem como referência diferentes operações de inteligência policial.

O conhecimento adquirido no percurso da pesquisa foi relevante ao comprovar o caráter informacional das operações de inteligência policial. Além disso, permitiu evidenciar que as práticas adotadas pela inteligência policial no combate a organizações criminosas reivindicam a composição de uma base teórica interdisciplinar e que as contribuições consolidadas no campo da Ciência da Informação, notadamente no âmbito da organização e uso da informação e do conhecimento, podem oferecer aporte teórico-metodológico fundamental para o entendimento da ação criminosa como um fenômeno informacional dinâmico e agravado pela sensibilidade

das ações que compõem a sua caracterização, monitoramento e combate.

É sabido que múltiplos fatores têm influência na questão da segurança, todavia o trabalho desenvolvido no âmbito desse estudo evidencia que mudanças estruturais no entendimento e condução dos processos nas polícias brasileiras são requeridas. Assim, espera-se que essa pesquisa contribua com a atividade de inteligência policial, no sentido de potencializar o combate a organizações criminosas. Em decorrência, que possa contribuir para um constante aperfeiçoamento na segurança pública do País.

## Agradecimentos

Agradecimentos são devidos à **Associação Internacional para Estudos de Segurança e Inteligência** – **INASIS** (<https://www.inasis.org/>).

## Notas

- (1) A indexação é a operação que consiste em descrever e caracterizar um documento com o auxílio de representações dos conceitos contidos nesses documentos. [...] A indexação permite uma pesquisa eficaz das informações contidas no acervo documental (OLIVER, 2010, p.150).
- (2) A Carga Suja prendeu, no início de setembro de 2011, 14 pessoas componentes de uma organização criminosa especializada no roubo e receptação de caminhões e cargas na região metropolitana de Belo Horizonte. Os autos das ações criminais decorrentes da operação foram colocados à disposição dos pesquisadores da Universidade Federal de Minas Gerais pela Justiça de 1ª Instância, Comarca de Pedro Leopoldo/MG, para fins de pesquisa, estudos e elaboração de dissertação de mestrado, compondo assim o *corpus* da pesquisa.
- (3) Dissertação intitulada Contribuições à organização de informações dinâmicas e de aplicações sensíveis: a construção do Tesouro Eletrônico de Operações de Inteligência Policial (TEOIP), defendida em setembro de 2013 na Escola de Ciência da Informação da UFMG. Disponível em: <http://www.bibliotecadigital.ufmg.br/>
- (4) *Sigint*, ou inteligência de sinais, refere-se a transmissões eletrônicas que podem ser recolhidas por navios, aviões, locais em solo, ou satélites (FBI).
- (5) O Relatório de Inteligência consiste no produto final de uma Operação de Inteligência Policial. Nele vão estar contidas todas as informações relevantes geradas pela equipe de inteligência policial.

## Referencias

- ANSI/NISO Z 39.19: 2005 – *Guidelines for the construction, format and management of monolingual controlled vocabularies*. Bethesda, Md, NISO, 2005.
- Bardin, L. *Análise de Conteúdo*. Lisboa: Edições 70, 1977.
- Brasil. Lei nº 12.694, de 24 de julho de 2012.
- Campos, M. L. A; Gomes, H. E. *Taxonomia e classificação: o princípio de categorização*. DataGramaZero, v. 9, n. 4, ago. 2008. Disponível em: <[http://www.dgz.org.br/ago08/Art\\_01.htm](http://www.dgz.org.br/ago08/Art_01.htm)>. Acesso em: 06 nov. 2011.
- Campos, Maria Luiza de Almeida; Gomes, Hagar Espanha; Motta, Dilza Fonseca. *Elaboração de tesouro documental - Tutorial*. Disponível em: < <http://www.conexao.org/bit/tesouro/index.htm>> Acesso em 11 ago. 2013.
- Cepik, Marco. *Serviços de Inteligência: Agilidade e Transparência como Dilemas de Institucionalização*. Tese de doutorado em ciências humanas: ciência política. Rio de Janeiro, IUPERJ, 2001.
- Cintra, A. M. M.; Tálamo, M.F.G.M.; Lara, M.L.G.; Kobashi, N.Y. *Para entender as linguagens documentárias*. 2. ed. São Paulo: Polis, 2002.
- Moreira, Alexandra; Alvarenga, Lídia; Oliveira, Alcione de Paiva. *O nível do conhecimento e os instrumentos de representação: tesouros e ontologias*. DataGramaZero: revista de Ciência da Informação, v. 5, n. 6, dez. 2004. Disponível em <[http://www.dgz.org.br/dez04/Art\\_01.htm](http://www.dgz.org.br/dez04/Art_01.htm)>. Acesso em 28 abr. 2013.
- Moreira, Manoel Palhares; Moura, Maria Aparecida. *Construindo tesouros a partir de tesouros existentes: a experiência do TCI – Tesouro em Ciência da Informação*. DataGramaZero: revista de Ciência da Informação, v. 7, n. 4, 2006.
- Oliver, Paulo Roberto da Costa. *Projetos de ECM/BPM: Os Segredos da Construção*, Volume 2. 1ª ed. São Paulo: Editora: Biblioteca 24 Horas, 2010. 292p.
- Ranganathan, S. R. *Prolegomena to library classification*. 3. ed. London: 1967.
- Saldanha, Gustavo Silva. *Micro Tesouro de Ciências Criminais*. Projeto de conclusão de curso de Biblioteconomia na Escola da Ciência da Informação – UFMG. Belo Horizonte, Dez. 2005.
- UNESCO. Unisist: *Guidelines for the establishment and development of monolingual thesauri*. Paris. Dez de 1976.
- Wittgenstein, Ludwig - *Investigaciones filosóficas*, Barcelona: Altaya, 1999.

---

# Terminología del huipil triqui

*Terminology of the "triqui huipil"*

---

Patricia Alejandra Méndez Zapata (1), Catalina Naumis Peña (2)

- (1) Posgrado en Bibliotecología y Estudios de la Información, alejandra.mendez@san-pablo.mx.  
(2) Instituto de Investigaciones Bibliotecológicas y de la Información, naumis@unam.mx

## Resumen

Se muestra la metodología utilizada para investigar la terminología relacionada con el huipil tradicional triqui de San Andrés Chicahuaxtla, Oaxaca, con la finalidad de representarlo a través de un modelo terminológico de dominio, el cual será diseñado a partir de las categorías de Aristóteles y Ranganathan. El resultado es útil no sólo en la organización del conocimiento sobre el tema sino en la propuesta de un marco normativo para representar objetos culturales relacionados con la indumentaria indígena.

La terminología se obtuvo de una serie de entrevistas a tejedoras triquis de una región de Oaxaca, las cuales fueron transcritas y posteriormente sometidas a un proceso de análisis.

La dificultad de obtener la terminología es la equivalencia idiomática que supone adaptar al español los términos usados en triqui y como el triqui además consta de cuatro variedades fue necesario restringir el estudio a una sola de ellas para hacerlo posible.

A partir de la terminología obtenida, se diseñó un modelo de dominio sobre el huipil como objeto cultural. El huipil es una vestimenta femenina común en los pueblos originarios, que en cada comunidad adquiere características diferentes. Para organizar los diferentes huipiles es necesario poder designar las características de cada una de ellos. A través de la indumentaria las mujeres triquis en este caso, preservan su historia, cultura y visión del mundo.

**Palabras clave:** Objetos culturales, Huipil triqui, Terminología del huipil triqui, Modelo de dominio, Organización del huipil triqui.

## 1. Introducción

El objetivo general de este trabajo es "proponer una metodología para desarrollar un modelo terminológico del huipil triqui de San Andrés Chicahuaxtla, Oaxaca."

En este objetivo se destacan tres particulares: investigar las designaciones terminológicas de los equivalentes en español de una lengua indígena; definir la metodología para organizar la indumentaria femenina como objeto cultural y obtener elementos de representación sobre el huipil triqui para los sistemas de información.

## Abstract

The methodology used to investigate the related terminology in traditional huipil Triqui of San Andrés Chicahuaxtla, Oaxaca, in order to represent it through a terminological domain model, which will be designed from the categories of Aristotle shown and Ranganathan. The result is useful not only in the organization of knowledge on the subject but on a proposed regulatory framework for representing cultural objects related to indigenous indumentary.

The terminology was drawn from a series of interviews with knitters Triqui of one region of Oaxaca, which were transcribed and then subjected to a process of analysis.

The difficulty of obtaining the terminology is the linguistic equivalents that means adapting the terminology used in the Triqui Language to the Spanish. The triqui has four varieties and that is why it was necessary to restrict the study to one of them to make it happen.

Since the terminology obtained it was designed the domain model of the huipil as a cultural object. The huipil is a common garment in women indigenous people, which each community has different characteristics. To organize the different huipiles is necessary to designate the characteristics of each of them. Through clothing Triqui women in this case, preserving their history, culture and world view.

**Keywords:** Cultural objects, Triqui huipil, Terminology Triqui huipil, Domain Model, Organization Triqui huipil.

Las lenguas triquis pertenecen al grupo mixtecano y ésta a la familia otomangue. El Instituto Nacional de Lenguas Indígenas (INAI) reporta que existen cuatro variantes de la lengua triqui.

Las variedades del triqui junto con todas las lenguas indígenas de México y el español son lenguas oficiales según la reforma constitucional del año 2003.

En el presente estudio se trabaja únicamente con la variedad hablada por la comunidad triqui de Chicahuaxtla.

El modelo terminológico es diseñado a partir de las categorías propuestas por Aristóteles y Ran-

ganathan para organizar el conocimiento. La terminología será desprendida de entrevistas a tejedoras de la comunidad e identificada a partir de un marco normativo para representar artefactos culturales.

Para ello se reconoce al huipil como un “documento”, a partir de esto se trata de desprender la información a través de la cual puede ser conceptualizado y representado.

El huipil es una especie de túnica o vestido utilizada principalmente por mujeres de los diversos pueblos indígenas desde tiempos prehispánicos. Está formado por dos o más lienzos o tiras de tela rectangulares, tejidas en telar y unidas entre sí, con una abertura en la parte alta para el cuello y a los lados para los brazos.

Los diseños, tamaños, formas, colores, materiales, técnicas de elaboración e iconografías, varían considerablemente dependiendo del contexto en que se elaboran, es una de las prendas que más ha perdurado a través del tiempo y en la que las mujeres han logrado preservar y representar su historia, cultura, costumbres y tradiciones.

En el huipil y alrededor de él, se puede identificar información que muestra la cosmovisión de un pueblo, lo que lo convierte en un documento de importante valor histórico y cultural susceptible de análisis y a las tejedoras en especialistas en el tema. Con ambos elementos es posible generar un modelo de dominio a través del cual se encuentre representado el objeto antes mencionado.

Un modelo de dominio se puede conceptualizar como la representación de una pequeña parte del mundo. “Para Bunge (2001, p. 8), la representación es “una traducción conceptual, visual, auditiva o por medio de artefactos de un objeto (material o idea)”.

La representación de los dominios de especialidad se hace a través de esquemas en donde haciendo uso del lenguaje, se conceptualiza el conocimiento. En este caso se concluirá con un mapa conceptual desde donde se presenta el esquema que comprende los términos relacionados con el huipil triqui.

## **2. Modelos terminológicos de dominio**

Un modelo de dominio sirve para recopilar, organizar y recuperar el conocimiento de una especialidad. Para su diseño se hace uso de la terminología, en su vertiente sobre la designación de los conceptos. La representante más destacada de la especialidad es María Teresa Cabré.

“La terminología, es la materia de intersección que se ocupa de la designación de los conceptos de las lenguas de especialidad. Su objetivo es la denominación de los conceptos [...] es una disciplina con unos principios teóricos de base (la teoría terminológica) y unas finalidades aplicativas (la elaboración de vocabularios, léxicos y diccionarios, y la normalización de las denominaciones). Las bases conceptuales de la teoría no son genuinas, sino que proceden –como en cualquier materia interdisciplinaria- de las disciplinas que participan en la misma: la lingüística, la lógica, la ontología y la teoría de la información.”. (Cabré, 1993, p. 82)

Es por tanto, la disciplina que ha sido utilizada para diseñar modelos de dominio de diversas áreas del conocimiento, con la finalidad de facilitar principalmente la comunicación entre especialistas a través del rescate de las denominaciones en el área de estudio.

Schwarz es más específica con respecto al modelo de dominio:

“los elementos que conforman un modelo de dominio son los conceptos, las relaciones entre los conceptos y las propiedades de los conceptos y sus relaciones”. (2005, p. 5),

Para diseñar una herramienta lingüística como la propuesta, se parte del principio que no existe una única metodología. Esto dependerá del esquema que se quiera utilizar para representarlo, por ello será importante tener en cuenta los siguientes aspectos:

“especificidad, selección de términos, tipos de relaciones, propiedades, tipos de jerarquías (monojerárquico o polijerárquico) y el nivel de coordinación (pre o post coordinado)”. (ibidem, p. 6)

En las ciencias de la información y la documentación, los esquemas de modelos de dominio más utilizados son los tesauros, las taxonomías y las ontologías. Cada uno de ellos está estructurado de diferente manera pero en su generalidad buscan compartir el conocimiento que representan. Los términos son la base con la que se diseñan estos modelos de dominio presentados en forma de mapas conceptuales:

“se utilizan herramientas terminológicas basadas sobre todo en los conceptos contenidos en los documentos, y que se expresan por medio de designaciones o términos”. (Cabré 1993, p.101)

Cada uno de estos esquemas tiene su origen en la necesidad de organizar el conocimiento que existe y que se sigue generando, con el principal objetivo de representarlo. Para ello se han buscado formas de hacerlo, describiendo y agrupando las cosas u objetos, de acuerdo a características similares, de ese modo se originan los llamados sistemas de clasificación y organización del conocimiento.

### 3. Organización del conocimiento

El proceso de conocimiento se produce al interior de los seres humanos pero se transmite mediante objetos culturales y éstos a su vez son organizados con otros objetos similares para una difusión más amplia:

“el problema de la clasificación del saber y el conocimiento científico surgió por la necesidad del hombre de sistematizar todos los conocimientos sobre el mundo exterior y sobre el proceso de conocimiento”. (San Segundo 1996, p. 25)

Diversos filósofos y científicos se han ocupado de encontrar una forma de organizar el conocimiento, diseñando para ello formas de representarlo a partir de entender el origen y las relaciones entre los diversos objetos de estudio, generando para ello sistemas diseñados a partir de una categorización general y una subcategorización que muestra de manera más detallada las características y relaciones de cada área del conocimiento.

#### 3.1. Categorías Aristotélicas

Dentro de las diferentes teorías sobre organización del conocimiento destaca la propuesta por Aristóteles. Él, por medio de una categorización, designa atributos o predicados a las cosas de tal modo que puedan ser descritas, esta clasificación la da a partir de creer que “los conceptos que utilizamos pueden reducirse a un cierto número limitado de grupos típicos” (Aristóteles 1962, p. 16).

Aristóteles explica las categorías como la expresión de las ideas que tenemos de los objetos de manera generalizada sin perder de vista el objeto particular, con predicados que describen al sujeto.

“un estudio de la expresión humana, considerada aún al nivel del concepto, de la idea simple y la palabra sola, antes de unirse a otra idea o palabra, para formar un juicio o una proposición” (ibidem, p.10)

Esta conceptualización la expresa a partir de diez categorías, haciendo una división entre la primera categoría y las nueve siguientes, identificando a la primera como el individuo en sí y las siguientes como aquello que determina al individuo: Sustancia (el qué), Cantidad (magnitud), Relación (con qué se relaciona), Cualidad (qué clase de cosa es), Lugar (dónde está), Tiempo (cuando), Posición (en qué actitud está), Hábito (cuáles son sus circunstancias), Acción (actividad), Pasión (pasividad).

#### 3.2. Clasificación facetada de Ranganathan

La categorización hecha por Aristóteles fue retomada por Ranganathan para diseñar un sistema de clasificación para bibliotecas llamado facetado o analítico sintético, el cual permite por medio del análisis, identificar las partes del sujeto y por medio de la síntesis, establecer las relaciones entre los componentes.

Para clasificar, establece tres planos de actuación, interconectados entre sí, estos son: el plano de las ideas (mente), el plano verbal (lenguaje) y el plano notacional (representación de la idea en lenguaje controlado); es decir, identificar la clase principal y posteriormente ubicarla dentro de las “categorías fundamentales”. “Las categorías son “las clases mayores o clases más generales que pueden ser formadas (Ranganathan 1967, p. 399)”, y pertenecen únicamente al contexto de la clasificación como disciplina y no al plano de la metafísica o la física. Estas categorías son: Tiempo, espacio, energía, materia y personalidad.

Rodríguez Bravo (2011, p. 159), menciona que

“estas categorías responden a los elementos principales de una proposición gramatical: verbo (energía), sujeto u objeto (personalidad), complemento determinativo (materia) y complementos de tiempo y lugar (espacio y tiempo),”

lo que lo hace ser muy similar a la categorización hecha por Aristóteles, quien también las ubicaba como predicativos del sujeto; es decir, complementos del objeto.

#### 4. Triquis de San Andrés Chicahuaxtla

El “triqui” es uno de los 16 pueblos originarios del estado de Oaxaca. Está localizado en la región de la mixteca alta, ocupando porciones de los distritos de Tlaxiaco, Putla y Juchitán, pertenecientes a la zona occidental del estado.

Según el sistema de indicadores sobre población indígena de México, hay aproximadamente 29,01874 hablantes de lengua triqui asentados en los pueblos de San Juan Copala, San Andrés Chicahuaxtla y San Martín Itunyoso.

San Andrés Chicahuaxtla es una agencia municipal de Putla Villa de Guerrero, en el poniente del estado de Oaxaca, cuenta con una población de aproximadamente 1015 personas, en su mayoría bilingües, hablantes del español y triqui. De acuerdo con Fidel Hernández, “el nombre viene de la lengua náhuatl y significa lugar fuerte y refleja el dominio y la influencia que los aztecas tuvieron en ella”, (2009, p. 50)

El mismo Fidel Hernández sostiene que:

“el nombre no es funcional ya que a la comunidad se le conoce como “Yuma’ niko”, que en lengua triqui significa “pueblo grande” con una noción cuantitativa..., es la principal comunidad triqui, posiblemente la cuna de la cultura.” (ibidem, p. 51)

#### 4.1. Lengua triqui

“Los triquis de Chicahuaxtla se han propuesto el nombre Nanj nín’in que significa lengua completa para referirse exclusivamente a su lengua (Hernández, 2009, p. 60)”

Solo a través de la lengua los triquis pueden transmitir sus saberes, por ello la nombran lengua completa, ya que para ellos no existe otro modo de dar a conocer su cultura. Esto se pudo observar claramente en las entrevistas con las tejedoras, cuando al preguntar el nombre en español de algunas partes del huipil, ellas respondían que no conocían como se les decía en español, solo en triqui, aunque podían describir el objeto o el uso del mismo.

Así mismo sucede con los procesos de enseñanza; a través de la tradición oral, los triquis (aunque no es exclusivo), transmiten su historia y conocimiento de generación en generación buscando que esta perdure, es así también como las abuelitas enseñan a tejer a sus hijas, y sus hijas a sus hijas, y así sucesivamente, buscando siempre la trascendencia.

### 5. Los textiles tradicionales triquis de Chicahuaxtla

Desde su región, cada uno de los pueblos indígenas de Oaxaca, posee características que lo hacen diferente de las demás; entre estas encontramos la alimentación, la vestimenta, la forma de organización, la forma de trabajar, las fiestas, las ceremonias, la religión, la música, y la lengua.

“Vestirse parece ser; antes que nada, una necesidad para combatir los efectos del entorno. Sin embargo, la ropa es también producto de un hábito social. Por lo mismo, contribuye a definir parcialmente los criterios de una sociedad a cuyas exigencias y reglas debe plegarse (Stresser-Péan 2012, p.35)”.

“El traje es un signo distintivo de cada grupo étnico e inclusive entre distintos pueblos pertenecientes a una misma etnia (Lechuga, 1996, p. 12)”, y consisten en la mayoría de los casos para una mujer, de un huipil, un enredo y un ceñidor o faja, y en el caso de los hombres consta de un calzón, una camisa y un ceñidor o faja. Los diseños, materiales, herramientas, colores e iconografías, están relacionados con la zona de donde proceden, la vegetación que

los rodea, el clima, los materiales a que tienen acceso y en algunos casos con su cosmovisión.

La vestimenta es un símbolo de identidad, y aunque con el paso del tiempo, tanto los triquis, como otros pueblos indígenas, han modificado su forma de vestir, estudiosos de los pueblos indígenas de México y en especial de Oaxaca (Lechuga, Basauri, Huerta, Mapelli y Cordero), coinciden al mencionar que hay un traje tradicional, el cual está conformado por enredo, huipil, cobija, faja, y diversos adornos, en el caso de la mujer; para el hombre, calzón, faja, sombrero, sarape y huaraches. Es importante mencionar que las características de estos, pueden variar.

En la zona triqui de Chicahuaxtla, el traje, al igual que la lengua, está conformado por elementos particulares a ellos, mismos que se presentan a continuación.

#### 5.1. El traje de mujer

El traje tradicional consta de los siguientes elementos:

**Rollo o enredo:** falda elaborada de un lienzo de algodón tejido en telar de cintura en color azul, decorado con franjas en color rojo y blanco, o de color rojo decorado con franjas en color blanco y azul. En la parte superior cuenta con una franja roja más ancha.

**Huipil:** vestido de tres lienzos de algodón o estambre en color blanco con franjas horizontales de color rojo decoradas con figuras en colores muy llamativos. Lenzos unidos con bordado de hilos de colores. Cuenta con listones de colores que cuelgan por la espalda.

**Cobija:** rebozo de dos lienzos de lana o hilo vela, color blanco y negro, decorado con franjas rojas, blancas y negras, unidas con bordado de colores y barbas color negro a los lados.

**Trenzas:** tejidas de estambre en color negro o azul. Se adhieren al cabello natural y se adornan con listones de colores y una peineta.

**Tenate con mecapal:** cesto con asa larga, elaborada de palma o plástico, tejida a mano. Va colgada del pecho hacia atrás o de los hombros en forma de bolso.

**Ceñidor o faja:** Cinturón de algodón tejido en telar de cintura, en colores blanco y rojo, con decoración geométrica, complementado con soyate de palma natural tejida a mano.

**Collar:** adorno para el cuello, formado de cuentas de plástico de diversos colores, sobre hilo de algodón.

Huaraches: sandalia de cuero o plástico, en forma cruzada o de pata de gallo.

Antiguamente, cuando salían al campo llevaban en la cabeza una jícara (fruto del jícaro, de aspectos similar a una calabaza pequeña, pulposo y con la corteza leñosa y dura, que se usa para confeccionar vasijas) que les servía para cubrirse del sol y como recipiente. (Figura 1) (Comisión Nacional para el Desarrollo de los Pueblos Indígenas, 2009)

Actualmente es más común el uso de sombreros de palma, gorra de tela, rebozo o manto sobre la cabeza para cubrirse del sol.



Figura 1. Mujeres triquis

## 5.2. El traje de hombre

Camisa: Para cubrir el torso, elaborada de algodón y de colores fuertes, bordada o no, en ocasiones puede ser de tela de artisela.

Calzón: Pantaloncillo hecho con algodón o manta, puede ser cosido a mano o a máquina y llega a la altura de la rodilla. Actualmente se observa a pocos hombres portar esta prenda. (Figura 2) (Archivo Fotográfico México Indígena. Universidad Nacional Autónoma de México, 2014)

Faja o ceñidor: Cinturón de algodón tejido en telar de cintura, en colores blanco y rojo, con decoración geométrica, complementado con soyate de palma natural tejida a mano.

Sombrero: De ala ancha y copa baja, hecho de palma o de lana tejida a mano.

Sarape o cobija: dos lienzos de lana o hilo vela, color blanco y negro, decorado con franjas rojas, blancas y negras, unidas con bordado de colores y barbas color negro a los lados. Es muy similar al de la mujer.

Huaraches: sandalia de cuero, en forma cruzada o de pata de gallo.



Figura 2. Hombre triqui

Es importante mencionar que en el cotidiano de los triquis de Chicahuaxtla, se puede observar como la vestimenta suele variar, como se advierte en las mujeres en la protección de su cabeza. Para el caso de los hombres es común observarlos usar camisa y pantalón comercial de mezclilla, botas de diferentes modelos o zapatos deportivos, gorra de tela o sombrero de palma.

Sin embargo, el traje tradicional aún es utilizado en fechas o eventos importantes, aunque es más común en el caso de las mujeres que en el de los hombres.

## 6. El huipil triqui

El huipil es una prenda de vestir, como se explicó líneas arriba, utilizada desde épocas prehispánicas por las mujeres. Es una especie de vestido, de forma rectangular o cuadrada, elaborado con dos o tres lienzos unidos. Los materiales utilizados para su elaboración, así como los adornos, colores y el tamaño, varían dependiendo del lugar en donde son hechos y de los gustos de la tejedora.

“The huipil is an untailed woman's garment which is tunic-like in appearance. The principle of construction is the same as that which has been standard practice in Mesoamerica from earliest times. It may be constructed of one, two, or three rectangular webs of cloth, sewn together lengthwise (Weitlaner Johnson 1957, p. 189)”.

### 6.1. Descripción

Para los triquis se llama huipil o “ri'nin”. Los tres lienzos están unidos por medio de una randa de trenza o de tres puntos. Es de forma cuadrada, llega hasta los tobillos y va abierto por los lados desde las mangas hasta la altura de las pantorrillas.

Está decorado con franjas de color blanco a las que llaman gusanos u orugas “yiluu” y de color rojo, decoradas con imágenes a las que llaman mariposas “yawí”, mismas que representan animales, plantas, objetos, personas, figuras geométricas, entre otras, (pájaros, pinos, canastas, jarras, conejos, estrellas, rombos, mariposas, mujeres, soldados, manzanas, etc.), diseñados de forma real o estilizados de modo que en ocasiones solo las tejedoras pueden identificar de qué se trata.

La parte central se diferencia de las otras debido a que a la altura del pecho tanto en la parte de atrás como en la de adelante lleva una franja gruesa de color rojo a la que llaman “mamá de las mariposas” decorada con figuras más grandes que las demás, con una franja en color azul, roja o negra en cada lado.

El cuello está decorado con triángulos de colores a los que relacionan con los rayos del sol, (y para proteger el corte del cuello), la cabeza de la mujer representa el sol, el centro del universo.

Por la espalda le cuelgan listones de colores a los que llaman “arcoíris” los cuales parten del cuello y llegan hasta la parte baja del huipil. En general los colores que adornan el huipil se ponen de acuerdo al gusto de la tejedora o de quien vaya a ser la portadora del textil, aunque la tejedora elige siempre los colores más fuertes debido a que serán más duraderos (Figura 3).

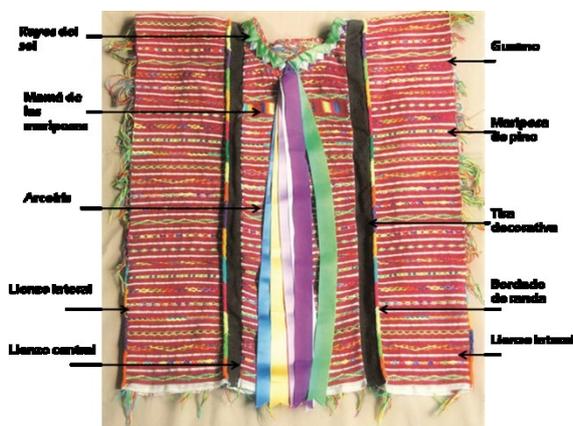


Figura 3. Huipil triqui

Como se puede observar en la figura 3, el diseño del huipil triqui está rodeado de simbolismos, esto se debe principalmente a que se considera que la vida de las mujeres triquis es como la vida de las mariposas y esto es lo que se intenta representar en el huipil: el proceso de metamorfosis de la mariposa o lo que es lo mismo, “la vida de una mujer triqui”.

En una investigación llevada a cabo hace unos años se explica un aspecto de la relación de la mujer triqui con las mariposas:

“La vida de la mujer triqui es comparada con la de una mariposa, es la depositaria del conocimiento y de la reproducción, por ello, una de las primeras actividades o aprendizajes principales que debe de conocer una mujer triqui es el tejido del huipil”. (Hernández, 2009, p. 71)

Razón por la que las mujeres aprenden a tejer desde muy pequeñas.

“Para los triquis, la mariposa es más que un insecto, es el símbolo sorprendente de lo vivo y lo bello en continua metamorfosis, concepto que sitúan más allá de su existencia cotidiana y que expresan con especial énfasis en sus huipiles.” (Ramírez López, 1990, p. 1).

Existen antecedentes en donde se describe y muestra el huipil diferente a lo que actualmente se conoce. Era de color blanco con algunas franjas delgadas en color rojo y con relativamente poca decoración de diversos colores.

## 6.2. Materiales y técnicas de elaboración

El huipil triqui está hecho con hilo de algodón de color natural y/o teñido con tintes naturales o de estambre industrial; tejido en telar, y decorado con brocados de hilo industrial, con aplicaciones de listón de artíselas cosidos a mano con hilo industrial.

En general el proceso de elaboración del textil consta de 4 pasos: urdir, tejer, unir y decorar.

Anteriormente también se hilaba; es decir, las tejedoras extraían manualmente el hilo de algodón o de lana, y también teñían sus hilos con elementos de la naturaleza. En la actualidad, aún se hace pero en menor medida, debido principalmente a los costos y tiempo que es necesario invertir para hilar o teñir los materiales.

El proceso de urdir consiste en “arreglar los hilos en la posición exacta que tendrán en el telar, calculando longitud, anchura y tipo de diseño que va a elaborarse (Velasco, 1995, p. 138)”.

Para tejer los lienzos del huipil, se hace uso del telar de cintura o de otates. Este nombre lo recibe debido a que “la tejedora lo amarra al talle; se conoce como telar de otate, porque se construye con palos de esta vara (Lechuga 1990, p. 25)”, el cual está integrado por diversos elementos con los cuales la tejedora se apoya para realizar el tejido. (Figura 4).

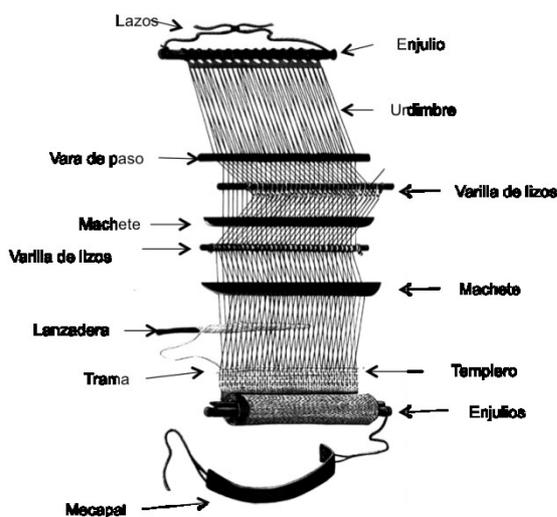


Figura 4. Telar de cintura

Las mujeres triquis utilizan principalmente dos tipos de tejido: cerrado y abierto, y el brocado y el bordado como técnicas de decoración.

El tejido cerrado o de tafetán, “es cuando un hilo de trama se entrecruza siguiendo el orden de pasar una vez por arriba y otra vez por debajo de un filamento de urdimbre” (Velasco, 1995, p. 130).

El tejido abierto o de gasa, “se forma por el cruce de los hilos impares de la urdimbre sobre los hilos pares, o viceversa, antes de que se introduzca la pasada de trama” (ibidem, p.31)

El brocado “es una técnica decorativa que se realiza simultáneamente a la operación del tejido” (ibidem, p. 30). Consiste en insertar hebras de hilo de estambre de colores a lo largo del proceso del tejido, formando las mariposas.

El bordado se hace “cuando el tejido está completamente terminado y fuera del telar” (ibidem). Para ello se utiliza una aguja metálica e hilos de estambre o algodón.

Para unir los lienzos, utilizan una técnica de tejido que se llama de randa en forma de trenza de múltiples colores.

Y por último, se coloca la decoración del cuello y los listones que formarán el arcoíris.

## 7. Metodología

Una vez que se definieron las categorías como se explica en los párrafos siguientes, se definió la estrategia para investigar y completar la terminología del huipil triqui de San Andrés Chicahuaxtla.

Para identificar la terminología, se decidió partir de la tradición oral, debido a que es de esta manera en la que las mujeres triquis han transmitido sus conocimientos de generación en generación.

No todos los elementos del huipil están claramente designados y las versiones entre antropólogos y estudiosos del tema muestran discordancias, por lo que es necesario seguir confrontando la terminología y agregando elementos que no muestran ni siquiera una designación.

Otro objetivo era además identificar no solo el proceso de elaboración del huipil o los materiales, técnicas e iconografías, sino también, conocer el contexto histórico y social existente en torno a este aspecto, lo que da mayor significado a las palabras de los hablantes.

Al contar con los términos y el conocimiento del lugar que ocupan en las prioridades del grupo indígena, se puede proseguir con el modelo de dominio.

### 7.1. Recopilación de información

Con la finalidad de obtener la mayor cantidad posible de información sobre el huipil, se diseñó una entrevista, la cual se estructuró con base en las categorías de Aristóteles y Ranganathan y de acuerdo a tres objetivos principales: Identificar los detalles de la obra (descripción detallada), Identificar técnicas de elaboración (procedimientos, materiales y herramientas) y Conocer usos y simbologías, esto a partir de un marco normativo para describir artefactos culturales.

Dicho marco fue definido a partir de normas internacionales que en su estructura cuentan con un área especial para la descripción de textiles. Estas fueron: “Reglas de catalogación Angloamericanas”, las cuales son normas para la descripción de materiales, el “Cataloging Cultural Objects”, el cual contiene normatividad para catalogar y clasificar objetos culturales, y la “Guía para la Clasificación de los Datos Culturales o Guía Murdock”, que es una herramienta utilizada para la clasificación de los datos culturales.

Teniendo en cuenta lo anterior, la guía de entrevista quedó organizada en 10 categorías (Tabla 1), con un total de 37 preguntas.

La entrevista fue piloteada con una tejedora triqui de San Andrés Chicahuaxtla y después de hacer los ajustes necesarios, se aplicaron. Esto se llevó a cabo en las casas de las tejedoras, obteniendo de ellas, una demostración de la forma cómo hacen su huipil, una entrevista de aproximadamente una hora en cada caso y la

posibilidad de entrar en sus casas y su comunidad.

Para participar en el estudio se eligieron tres tejedoras que como características tenían: tejer actualmente, ser triquis de San Andrés Chichauxtla, pertenecer a la misma familia, hablar español (lo suficiente para poder conversar con ellas), contar con experiencia en el uso del telar y ser mayores de edad.

El corpus obtenido de las charlas con las tejedoras fue grabado con una micro grabadora sony, modelo ICDUX512F 2010 de 2 GB, y fue transcrito en un documento de extensión "txt" para posteriormente poder trabajar con el.

Cabe resaltar que la transcripción de las entrevistas en algunos momentos resultó compleja debido al español que utilizan las tejedoras; sin embargo, haciendo las acotaciones necesarias fue posible concluirlo satisfactoriamente, respetando en lo posible los diálogos originales.

| Categoría                        | Objetivo  |
|----------------------------------|---|
| Datos personales                 | Conocer de manera general el perfil de la persona entrevistada                  |
| Denominación                     | Identificar la manera como nombran el huipil                                    |
| Uso                              | Conocer con qué fines hacen el huipil   |
| Descripción física               | Reconocer las características físicas del huipil                                |
| Relación con                     | Identificar con qué otras prendas lo usan y cómo                                |
| Materiales                       | Saber de qué está hecho el huipil   |
| Lugar                            | Reconocer las similitudes y diferencias con el huipil de otras regiones triquis |
| Tiempo                           | Listar las etapas, tiempo de elaboración y temporalidad o fechas de elaboración |
| Técnica y proceso de elaboración | Saber el procedimiento para elaborar el huipil                                  |
| Iconografía                      | Identificar los simbolismos existentes en el huipil                             |

Tabla 1. Marco normativo para estructurar la guía de entrevista.

## 7.2. Análisis de la información (corpus)

El corpus lingüístico desprendido de las entrevistas fue ingresado en el analizador de corpus "wordsmith".

WordSmith Tools fue desarrollado por Mike Scott de la Universidad de Liverpool, es un software para PC publicado por Lexical Analysis Software y Oxford University Press. El programa cuenta con una serie de herramientas para ana-

lizar los textos, entre las que encontramos: Concord, WordList y KeyWords.

Una vez ingresado el corpus, se siguieron diversos procedimientos: El primero consistió en aplicar la herramienta "lista de palabras" para identificar las palabras del texto ordenadas de acuerdo a la frecuencia de aparición. El resultado fue de 1342 entradas.

A este resultado se le aplicó un "stopword" o lista de palabras vacías, con el objetivo de eliminar algunas conjunciones, preposiciones, artículos, etc. Posterior al filtrado, el total de palabras en lista fue de 1189, las cuales se ordenan de acuerdo a la frecuencia de aparición.

Las palabras resultantes fueron sometidas a un proceso de lematización. Con la lematización se busca "la reducción de todas las posibles variantes flexivas de una palabra a su respectivo lema representado por el lema (Álvarez, p. 188)".

Finalmente se contabilizaron un total de 621 palabras lemas.

Las palabras obtenidas de la lematización fueron sometidas a un proceso de análisis para identificar las palabras clave o keywords. Estas se extrajeron de la comparación de la lista de palabras final con un corpus de referencia mayor, esto permitió al programa identificar las palabras que tienen una frecuencia media y extraerlas como palabras clave.

La comparación se realizó con un documento en formato "txt" del libro "Don Quijote de la Mancha" en español.

El texto se introdujo a WordSmith y se analizó con la herramienta "stoplist", proceso del cual se extrajo una lista de 24, 648 palabras.

El resultado fue de 147 "palabras clave", organizadas en una tabla en donde se muestra la frecuencia de aparición en el texto origen, posteriormente la frecuencia de aparición en el corpus de referencia y en último lugar se muestra el valor de "keyness", que corresponde a la distancia que existe entre la palabra principal "tema" o la de mayor frecuencia y las demás palabras clave.

Posteriormente, las palabras fueron analizadas por medio de la herramienta "concordancias" de WordSmith, para identificar las relaciones entre estas y poder así comenzar a determinar la terminología y las relaciones. Es importante mencionar que para el análisis no se tomaron en cuenta las palabras que tienen una frecuencia de aparición de 3 a 1.

A partir de las categorías previamente establecidas y del listado de palabras clave, se identificaron las palabras que podrían formar parte de la terminología; sin embargo, fue necesario hacer una revisión al listado de palabras final y de las concordancias de algunas palabras de este listado, debido a que se identificaron palabras candidatas a término que no habían sido consideradas.

Con las palabras resultantes se conformó una tabla en donde se identificó el número que ocupa la palabra en la lista, la palabra, la frecuencia de aparición, el porcentaje que tiene la palabra en el texto, el texto de donde es tomada, el porcentaje del texto que fue analizado y las lematizaciones que se hicieron de la misma, principalmente datos estadísticos que arroja el programa.

De esta manera se obtuvo una primera visualización de la terminología del huipil triqui.

## 8. Resultados

A partir de las palabras organizadas por categorías de la guía de entrevista, se desarrolló una tabla (Tabla 2), en la que de acuerdo con las categorías de Aristóteles y Ranganathan se empezaron a desprender descriptores y términos equivalentes y se identificaron nuevas categorías, las cuales fueron comparadas con categorías establecidas en el Tesouro sobre pueblos indígenas de México y América Latina y en el tesouro de Arte y arquitectura, de tal modo que las categorías resultantes fueron:

|   |                                 |
|---|---------------------------------|
| 1 | Traje indígena                  |
| 2 | Traje indígena y traje femenino |
| 3 | Materiales                      |
| 4 | Herramientas                    |
| 5 | Procesos y técnicas             |

Tabla 2. *Categorías principales del mapa conceptual sobre el huipil triqui*

Con los resultados obtenidos fue posible construir un modelo de dominio ahora sí, con un formato aproximado a un tesouro del huipil triqui (faltó su inclusión en un software para tesouros), el cual está conformado por 98 términos ordenados alfabéticamente, precedidos por una notación y de acuerdo a la siguiente estructura:

|   |                        |
|---|------------------------|
| 1 | Descriptor principal   |
| 2 | UP Término equivalente |
| 3 | NA Nota de alcance     |
| 4 | TG Término general     |
| 5 | TE Término específico  |

Tabla 3. Relaciones terminológicas de un tesouro como modelo de dominio elegido

A partir del tesouro, se desarrolló un mapa conceptual, tratando de seguir las categorías de organización del tesouro y buscando reflejar los descriptores del tesouro y las relaciones que existen entre estos. (Anexo 1)

Para poder identificar las categorías, el mapa fue organizado en diferentes colores quedando estructurado de la siguiente forma:

| Color      | Categoría                       |
|------------|---------------------------------|
| Azul       | Traje indígena y traje femenino |
| Rosa       | Formas                          |
| Verde      | Iconografía                     |
| Morado     | Materiales, procesos y técnicas |
| Anaranjado | Aplicaciones                    |

Tabla 4. *Colores asignados a las categorías en el mapa conceptual del anexo 1.*

## 9. Conclusiones y trabajos futuros

Por medio de esta investigación, se logró proponer la metodología adecuada para desarrollar un modelo de dominio diseñado desde la interdisciplinariedad, combinando elementos básicos de lingüística, antropología, bibliotecología y ciencias de la información.

En el desarrollo de la investigación se observó que no existe un solo término en español, equivalente al triqui y tampoco los términos triquis son muy consistentes en su traducción al español. Los significados también exhiben variaciones, que se observan desde el significado de la palabra triqui.

Sin embargo, se logró establecer un cuerpo terminológico que cumple con las funciones para utilizar en sistemas de información e indizar contenidos coherentes y eficientes para recuperar información.

El resultado permitió generar una herramienta lingüística en formato de tesouro, a partir de la cual es posible organizar el conocimiento de un área de especialidad. (Tabla 3)

En este caso, los cuestionarios aplicados y las relaciones establecidas de acuerdo a la utilización de elementos extraídos de la clasificación de Aristóteles y Ranganathan permitieron obtener un cuerpo terminológico y un modelo de dominio que representa el conocimiento sobre el objeto cultural "el huipil triqui".

La utilización de los resultados obtenidos no sólo incide en el conocimiento de los elementos relacionados con el huipil triqui, sino que colabora en la promoción de la tradición textil de la comunidad estudiada y de otras y preserva tradiciones culturales que son parte de la historia y del presente.

El siguiente paso es inclinar la investigación hacia algo más amplio “diseñar un modelo ontológico del textil tradicional indígena de Oaxaca” para colecciones digitales, porque prestaría un servicio funcional y valioso para las búsquedas de información sobre el tema.

La recuperación puede incluir todas las equivalencias del español en las lenguas originarias, que solo pueden ser limitadas en un modelo de dominio como el desarrollado en la presente investigación. También se puede obtener información sobre las prendas resultantes de los procesos textiles, los materiales utilizados y las relaciones entre los elementos usados en una comunidad y en otras que también produzcan elementos textiles tradicionales.

Desde luego la utilización de estas investigaciones son fundamentales tanto en las bibliotecas que contienen documentación sobre el tema como en la elaboración de catálogos sobre las colecciones en museos que contengan muestras del textil oaxaqueño y en cualquier sistema en el que se incluya la temática.

Como se puede observar, el utilizar análisis de dominio implica un trabajo multidisciplinario, en donde será necesario contar con el apoyo de expertos en sistemas o de antropólogos y/o estudiosos del dominio; sin embargo, es también relevante hacer la propuesta desde el área de la bibliotecología para crear mejores sistemas de indización y recuperación.

Las tendencias actuales imponen la necesidad de interactuar en espacios nuevos como la web y disponer de software para sustentar cualquier herramienta lingüística que interactúe en los sistemas de información.

La única manera de organizar sistemas de información efectivos es trabajar con otras disciplinas para disponer de herramientas que correspondan a los intereses de los destinatarios de la información.

La propuesta ahora es reutilizar la información ya obtenida para elaborar el nuevo modelo de dominio más amplio que se explicó con párrafos anteriores.

La presente investigación facilita cualquier otra que se realice en relación a textiles indígenas porque es un antecedente en el que se cuenta con una parte de la terminología identificada. El resultado es útil porque se dispone de un cuerpo de términos aceptados y validados, así como con el conocimiento de la metodología para desprender nuevo vocabulario.

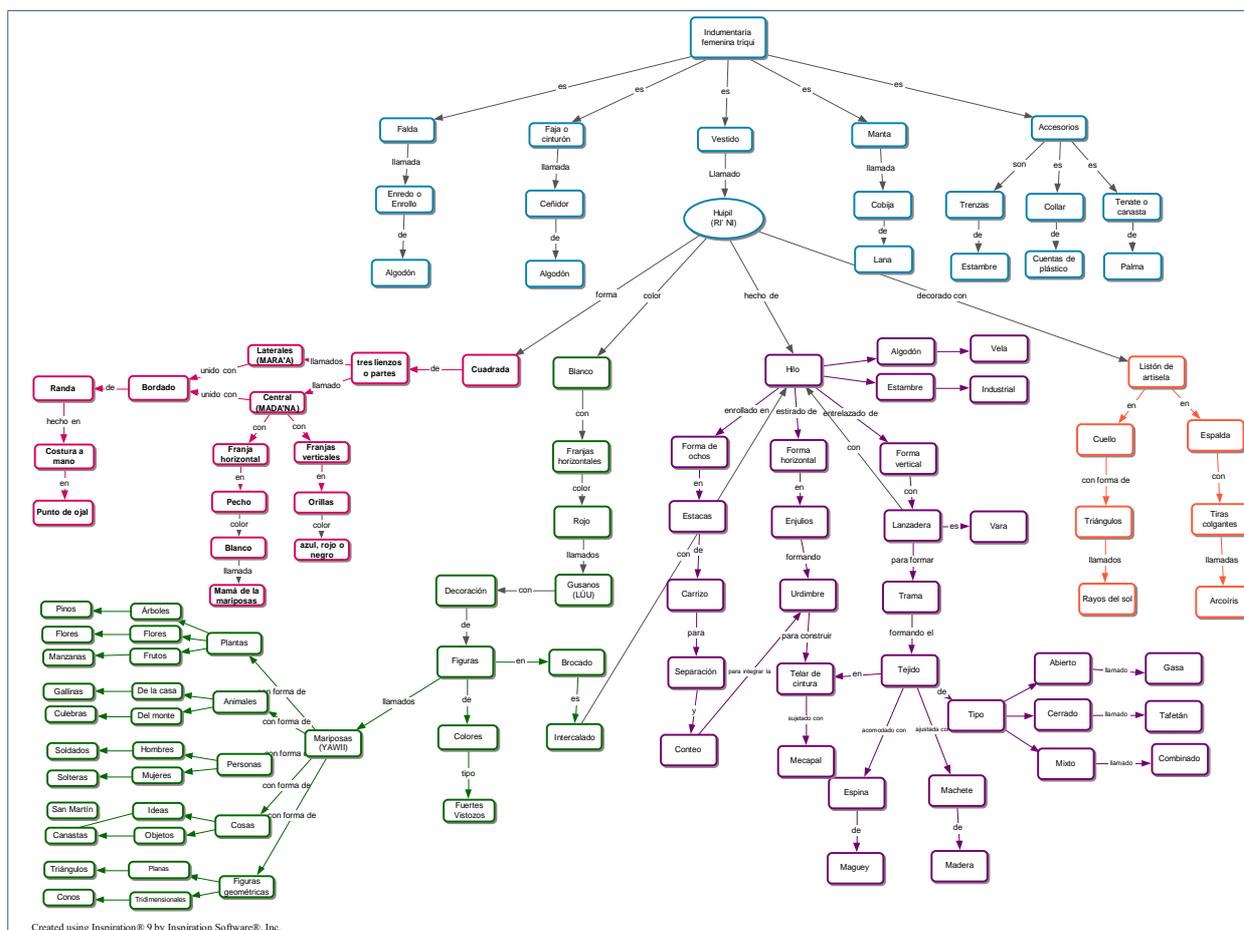
En la página final se presenta el mapa conceptual que condensa las categorías, clases y sub-

clases que conforman la terminología del huipil triqui y sus relaciones principales de generalidad y especificidad.

## Referencias

- Álvarez, José. *Problemas de lematización verbal en una lengua aglutinante con infinitivos múltiples: el caso del guarijio o wayuunaiki*. Venezuela: Universidad de Zulia. (De, 6 de julio, 2015: <http://bit.ly/V8tonl>)
- Aristóteles (1962) *Categorías*. Traducción del griego y prólogo de Francisco de P. Samaranch. México: Aguilar.
- Basauri, Carlos (1940). *La población indígena de México*. México: SEP.
- Bunge, Mario (2001). *Diccionario de filosofía*. México: Siglo xxi editores.
- Cabré, M. Teresa (1993). *La terminología: teoría, metodología, aplicaciones*. Barcelona: Empúries.
- Cardero García, Ana María (2004). *Lingüística y terminología*. México: UNAM.
- Comisión Nacional para el Desarrollo de los Pueblos Indígenas (2009). *Triquis de Oaxaca (el traje tradicional indígena)*. Recuperado el 2/07/2015 de [http://www.cdi.gob.mx/index.php?option=com\\_content&task=view&id=693&Itemid=63](http://www.cdi.gob.mx/index.php?option=com_content&task=view&id=693&Itemid=63)
- Hernández Mendoza, Fidel (2009). *Daj nadure' sinugun: los usos de la lengua Nanj nin'in y del español en las comunidades triquis de Chichauaxtla*. Tesis para obtener el título de Magister en educación intercultural bilingüe. Bolivia: Universidad Mayor de San Simón.
- Huerta Ríos, César (1994). *Triquis*. Pueblos indígenas de México. México: INI.
- Lechuga, Ruth D. (1996). *Mirando los textiles oaxaqueños*. En Artes de México. Textiles de Oaxaca México: Artes de México, no. 35.
- Lechuga, Ruth D. (1990). *El traje indígena de México: su evolución, desde la época prehispánica hasta la actualidad*. México: Panorama editorial.
- Mapelli Mozzi, Carlotta (1965). *El traje indígena de México*. Dibujos de Teresa Castelló Yturvide. México: INAH.
- Ramírez López, Heladio (1990). *Triquis: la región de las mariposas*. México: Artesanías e Industrias Populares del Estado de Oaxaca.
- Ranganathan, S. R. (1967). *Prolegomena to library classification*. 3a ed. Bombay: Asia Publishing House.
- Rodríguez Bravo, Blanca (2011). *Apuntes sobre representación y organización de la información*. España: Trea.
- San Segundo Manuel, Rosa (1996). *Sistemas de organización del conocimiento: la organización del conocimiento en las bibliotecas españolas*. Monografías. Madrid: Universidad Carlos III: Boletín Oficial del Estado.
- Schwarz, Katharina (2005). *Domain model enhanced search: a comparison of taxonomy, thesaurus and ontology*. Thesis Master of content and knowledge engineering, Utrecht: University of Utrecht.
- Stresser-Péan, Claude (2012). *De la vestimenta y los hombres: una perspectiva histórica de la indumentaria indígena en México*. México: Fondo de Cultura Económica.
- Velasco Rodríguez, Griselle J. (1995). *Origen del textil en Mesoamérica*. México: Instituto Politécnico Nacional.
- Weitlaner Johnson, Irmgard (1957). *Survival of feather ornamented huipiles in Chiapas, Mexico*. En: *Journal de la Société des Américanistes*. Tomo 46. París: Au siège de la Société.

## Anexo 1. Mapa conceptual del huipil triqui



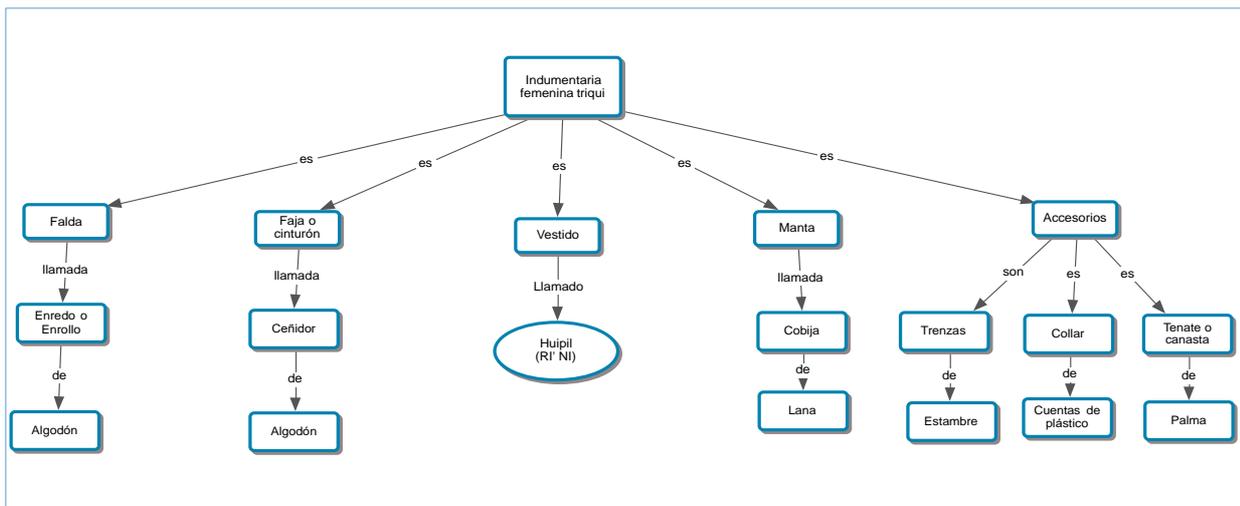
### Color

- Azul
- Rosa
- Verde
- Morado
- Anaranjado

### Categoría

- Traje indígena y traje indígena femenino
- Forma
- Iconografía
- Materiales, procesos, técnicas
- Aplicaciones

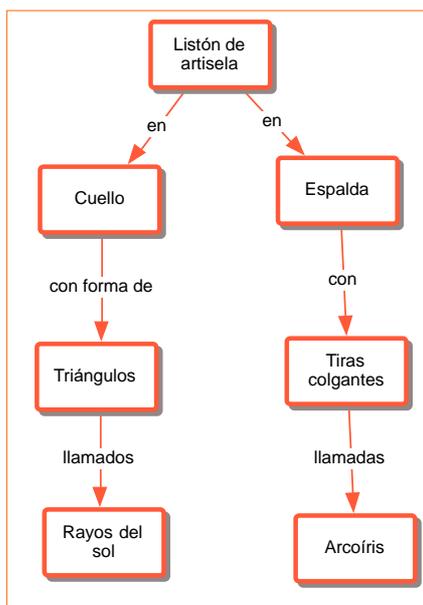
Anexo 2. Estratos del mapa conceptual



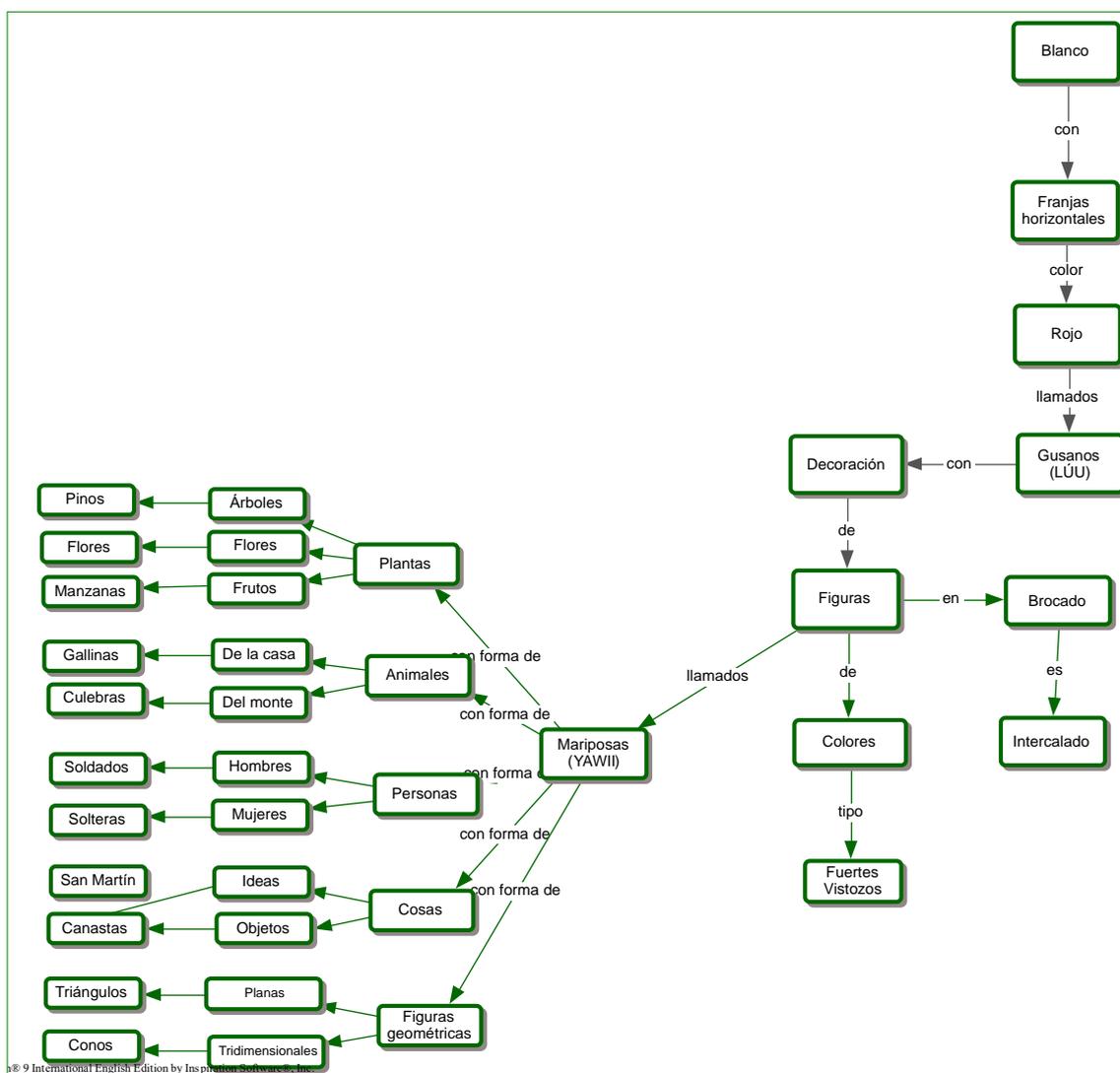
PARTE 1. TRAJE INDÍGENA Y TRAJE INDÍGENA FEMENINO



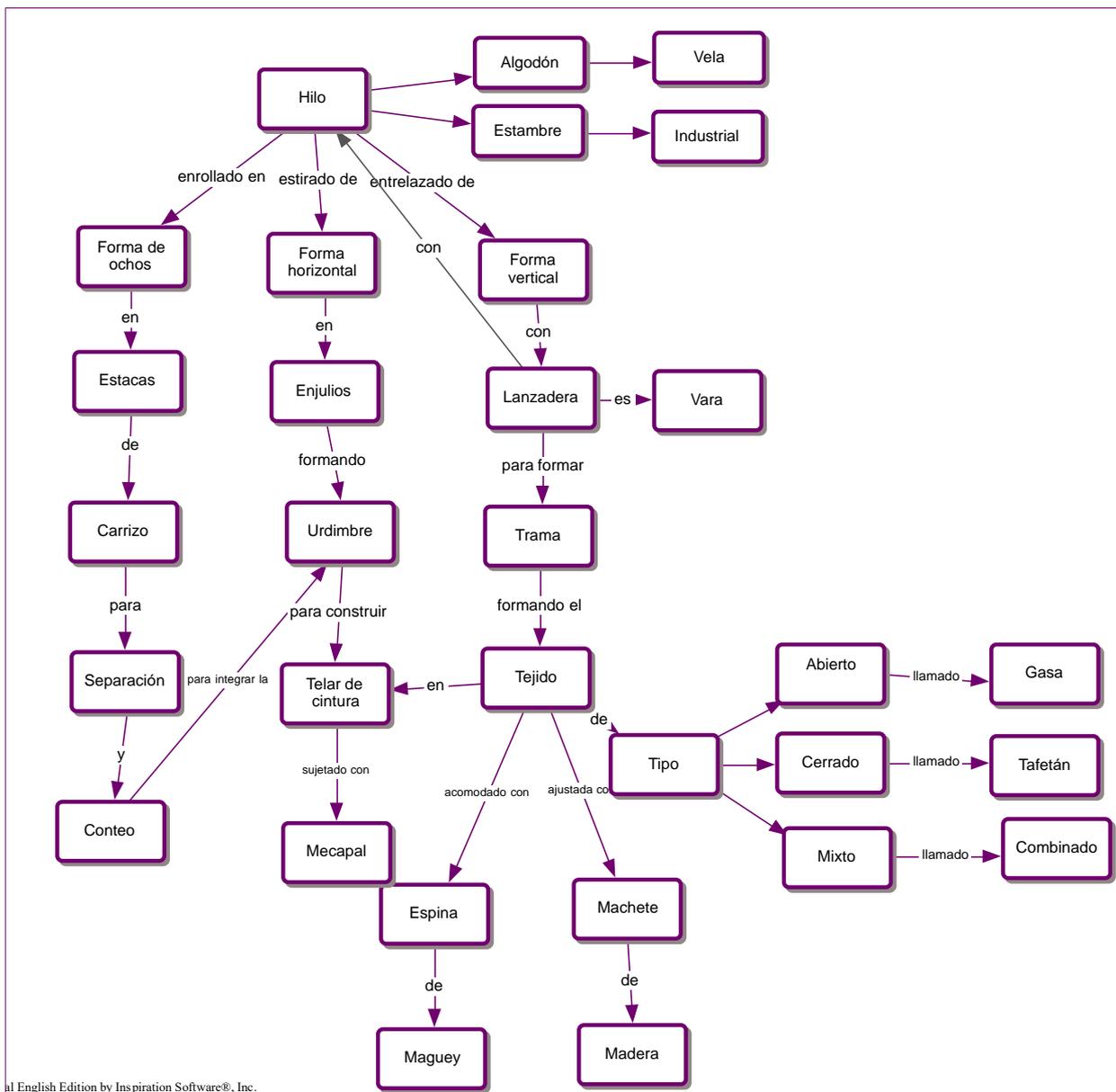
PARTE 2. FORMA/DESCRIPCIÓN



PARTE 3. APLICACIONES



PARTE 4. ICONOGRAFÍA



al English Edition by Inspiration Software®, Inc.

PARTE 5. MATERIALES, PROCESOS Y TÉCNICAS

---

# Representação de Contexto em Acervos Digitais

*Context representation in digital repositories*

---

**Denysson Mota (1), Nair Yumiko Kobashi (2)**

(1) Universidade Federal do Cariri. Av. Tenente Raimundo Rocha S/N - Bairro Cidade Universitária Juazeiro do Norte – Ceará, 63048-080; Doutorando do Programa de Pós-graduação em Ciência da Informação – USP, Email: denysson@usp.br.

(2) Escola de Comunicações e Artes – USP, Av. Prof. Lúcio Martins Rodrigues, 443, 05508-900 São Paulo - SP - Brasil, Email: nykobash@usp.br

## Resumem

Abordagem crítica sobre a representação e recuperação de informação em acervos digitais, indicando a necessidade de aprimorá-la com a introdução de dados de contexto. A proposta, integrada às discussões sobre a Web Semântica e a Web Pragmática, foi construída na interlocução entre a Ciência da Informação e as Ciências da Computação. O estudo apoia-se teoricamente, de um lado, na perspectiva pragmática de Peirce (1977), Morris (1985), Eco (1987) e Armengaud (2006) e, de outro, na abordagem sociocognitiva de Ingwersen e Järvelin (2005). A proposta teve como campo de experimentação a Base de Dados Referencial de Artigos de Periódicos em Ciência da Informação (BRAPCI). Neste trabalho, após a discussão de diferentes abordagens sobre o conceito de contexto, são operacionalizados dois contextos: a definição das palavras-chave e as citações dos trabalhos de artigos científicos. Verificou-se que há duas possibilidades de representar contextos: tanto por meio do RDF quanto com o uso do XTM. Os pontos fortes e fracos dessas aplicações são discutidos neste trabalho. Os resultados mostram que o uso do XTM é mais vantajoso do que o RDF para representar contextos em acervos científicos digitais.

**Palabras clave:** Acervos Digitais. Representação da Informação. Representação de contextos informacionais. Web Pragmática. Web Semântica.

## 1. Introdução

Este texto discute a representação e a recuperação de informações em acervos digitais. Nesse sentido retoma-se, inicialmente, a proposta da Web Semântica (WS), de descrever diferentes recursos por meio de metadados. A Web semântica foi concebida com o fim de tornar os computadores mais inteligentes para processar informação semântica. As metodologias criadas para esse fim têm promovido avanços consideráveis nos processos de representação e recuperação de informação. Elas não cobrem po-

## Abstract

It is presented, in this paper, a critical approach to information representation and retrieval in digital collections and the need to improve it by introducing context data. The proposal, contextualized in the discussions on the Semantic and Pragmatic Web, is an interdisciplinary approach that looks for a cross fertilization between Information science and Computer science. This proposal is based in one hand on the pragmatic perspective of Peirce (1977), Morris (1985), Eco (1987) and Armengaud (2006) and, on the other hand, in Ingwersen and Järvelin's (2005) socio-cognitive approach. The tests of this proposal had been made using BRAPCI, a Brazilian database of journal articles on Information Science. In this paper it is first discussed the different concepts on context. Then, two contexts had been operationalized: the definition of key words and quotations present in scientific journal articles. It has been found that it is possible to use both the RDF as XTM patterns to represent contexts. The strengths and weaknesses of these two patterns are discussed. The results showed that the use of XTM is more advantageous than the RDF to represent contexts in digital scientific collections.

**Keywords:** Digital repositories. Information representation. Informational context representation. Pragmatic Web. Semantic Web.

rém, aspectos importantes da significação, necessários para os usuários humanos das informações disponíveis nas redes eletrônicas. O passo seguinte proposto é adicionar contexto aos recursos, para caminhar em direção a uma Web Pragmática (WP).

Em face das diversas concepções do que pode ser considerado contexto, e diante das diferentes possibilidades de representá-lo, limitamos, neste trabalho, a abordar dois tipos de dados de contextos: os intratextuais, ou de contexto, como definido por Eco (1987), e os extra-

textuais, que se referem, neste caso, à intertextualidade (CORTINA, 2000).

## 2. Marco teórico do estudo

Os conceitos de sintaxe, semântica e pragmática, presentes nas teorias semiótico-pragmáticas de Peirce (1977) e Morris (1985), inspiraram a criação da Web Pragmática (WP), considerada como proposta que pretende ir além da Web Semântica. Estas perspectivas vêm sendo discutida por autores como Signh (2002) e de Moor, Keeler e Richmond (2002), assim como por Sowa (2000), que apontam os problemas teóricos e tecnológicos da aplicação de ferramentas semânticas para tratar e recuperar informação em ambientes digitais.

Antes de discutir as características da Web Semântica e da Web Pragmática, é necessário distinguir Semântica de Pragmática. Grosso modo, pode-se afirmar que o objeto da Semântica é a significação dos termos em um sistema, enquanto o estudo das expressões linguísticas em contextos de uso, de acordo com as variações situacionais, é objeto da Pragmática. O contexto é portanto, o cerne da Pragmática, um objeto de pesquisa intrigante para diferentes áreas do conhecimento. A definição do termo 'contexto' varia historicamente, adquirindo, além disso, um significado específico em cada área do conhecimento que o utiliza (VALLET et al, 2006). Não há, portanto, um único significado universalmente aceito do termo (EDMONDS, 1999).

No campo da Linguística, o contexto é concebido como algo composto por vários elementos (sociais, situacionais e/ou pessoais), que interferem no comportamento linguístico dos indivíduos. Armengaud (2006) define contexto como as situações comunicativas concretas que ocorrem em situações determinadas por lugar, tempo e identidade dos falantes. Por outro lado, Ingwersen e Järvelin (2005), pesquisadores da Ciência da informação, tratam o contexto como um conjunto de fatores, uma série de elementos determinados por itens como "a tarefa de uma pessoa, sua fase e situação" (INGWERSEN; JÄRVELIN, 2005, p. 1, tradução livre).

Armengaud (2006, p. 82) afirma que o "conceito de contexto adquire vigor e consistência ao se aproximar da noção de mundo possível", restringindo-se, assim, as possibilidades contextuais a um conjunto de situações, denominado de conjunto-contexto, que são as mais prováveis de ocorrer. O conceito de mundos possíveis é compartilhado por Eco (2000), que exemplifica-o com a palavra baleia. Para Eco, existem dois mundos possíveis: um onde 'baleia' é interpreta-

tado como um peixe, e outro onde é interpretado como mamífero. Ambas as interpretações podem estar corretas porque pertencem a diferentes mundos possíveis.

De forma geral, tudo o que é necessário para avaliar o significado dos enunciados ou mensagens pode ser considerado contexto. O sentido das palavras deve ser apreendido a partir do sentido das frases onde elas figuram, das frases dentro de um parágrafo, e assim sucessivamente. Para compreender, um enunciatário há que situar as palavras em seu contexto de uso na frase. Esta percepção vai ao encontro da definição de Eco (1987), para quem o significado de um termo está relacionado à aparição deste junto de outros termos pertencentes ao mesmo sistema de significação, e que, juntos, apontam para o significado adequado ou preciso. Quando isto ocorre, tem-se o co-texto, um contexto delimitado pelo próprio texto. Para o autor, as "seleções contextuais preveem possíveis contextos: quando estes se realizam, se realizam em um co-texto" (ECO, 1987, p. 29, tradução nossa).

Um outro contexto importante para a correta interpretação de uma mensagem é a citação. Segundo a norma ABNT 10520 (2002, p. 1), uma citação é uma "menção de uma informação extraída de outra fonte", podendo ser direta, quando é uma transcrição textual do original, ou indireta, quando cria-se um texto baseado na obra consultada. As citações são elementos textuais utilizados para reafirmar, negar ou referenciar informações dentro do texto, e "no universo do discurso científico, as influências teóricas e metodológicas se manifestam, via de regra, por meio de citações" (FRANCELIN, 2010, p. 120).

Com a possibilidade de identificar as diferentes influências exercidas sobre um trabalho, a adequada representação das citações desempenha importante papel na Web Pragmática. O contexto das citações é um elemento de difícil identificação mediante análises puramente sintáticas e/ou semânticas. A possibilidade que se apresenta é realizar inferências mais complexas com a adição das citações presentes à representação de um dado texto.

## 3. Metodologia

Os contextos identificados na literatura como relevantes para a interpretação de textos, como os previamente abordados, foram analisados e modelados em Diagrama Entidade-Relacionamento (DER) complementando o modelo da BRAPCI (BUFREM et al., 2010), e posteriormente representados em Resource Des-

cription Framework (RDF) e Topic Maps (TM), na sua forma expressa em XML, chamada de XTM (ISO, 2013).

Ao abordar duas tecnologias como o RDF e os TM, surgem naturalmente questionamentos sobre as semelhanças e diferenças entre elas porque ambas são ferramentas para representar conceitos e suas propriedades e relacionamentos. As abordagens são, porém, fundamentalmente diferentes.

Enquanto o RDF se baseia na lógica formal e tem na simplicidade triádica sua principal força e flexibilidade, os TM utilizam como inspiração os índices remissivos de livros, assim como os glossários e tesouros. Estes últimos têm estrutura consideravelmente mais complexa. É possível afirmar, ainda, que o RDF foi criado sob a perspectiva da programação para que os computadores entendessem as ontologias, enquanto os TM foram pensados sob o ponto de vista humano (PEPPER, 2002).

Porém, a principal diferença, em termos práticos, é a existência de qualificadores, principalmente os *Scopes* presentes nos TM. Como dito anteriormente, esse elemento permite definir, dentro de uma única representação, quais relacionamentos e valores são adequados, ou que afirmações são consideradas verdadeiras, em diferentes situações (Figura 1), o que permite que um conceito se relacione com diferentes recursos ou conceitos, de acordo com o contexto escolhido (PEPPER, 2000).

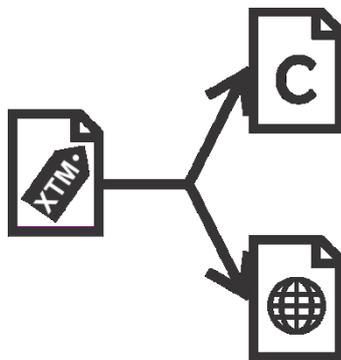


Figura 1 – Exemplo gráfico de relacionamento em XTM

Enquanto é possível representar desta forma em TM, no RDF essa característica não foi desenvolvida. Para realizar a indicação de contexto é necessário um retrabalho, um processo denominado de reificação, que envolve a criação de novas estruturas que aumentam a complexidade da representação e de programação (Figura 2). Com isso, compromete-se a simplicidade da sua concepção, além de realizar as

indicações e restrições de contexto de forma limitada (GARSHOL, 2003).

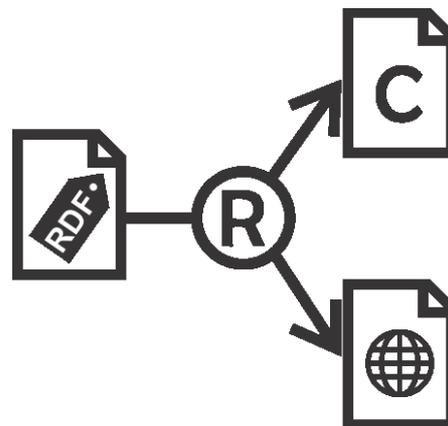


Figura 2 – Exemplo gráfico de relacionamento reificado em RDF

A possibilidade de definição formal de contextos torna viável a Web Pragmática. A capacidade de definir, dentro dos recursos informacionais, não apenas o conceito, mas também o contexto de aplicação do mesmo, permite a seleção e recuperação de informação adequada às diferentes necessidades e perfis de usuários.

O tamanho do código necessário também é diferente nas três situações. Em RDF simples, sem contexto, o código necessário para referenciar uma palavra chave para um determinado documento é curto e simples (Figura 3 em Anexo 1).

Esse mesmo relacionamento, entre obra, palavra-chave e autor, para ser expresso em RDF, é reificado, sendo necessário, além de maior trabalho para criar as novas entidades e estruturas, ser feito o controle das entidades reificadas (Figura 4 em Anexo 2).

Em XTM, mesmo sendo criada uma estrutura, que requer quase tantas linhas de código quanto a reificação em RDF, a relação não precisa ser um objeto separado com um identificador único, que necessita de controle e gerenciamento, como em RDF, sendo também mais simples a estrutura necessária para identificar a relação entre os diferentes objetos (Figura 5 em Anexo 3).

Por meio dos Topic Maps é possível representar as relações contextualizadas de forma mais elegante e com gerenciamento mais simples das entidades e relacionamentos. Esta possibilidade mostra ser interessante ao se considerar as formas de representação de informações com adição de contexto na Web Pragmática.

#### 4. Resultados preliminares

A análise de dados via DER de Peter Chen (1976) permite a visualização simples e direta da estrutura de dados de um repositório. Dentro da BRAPCI (BUFREM et al., 2010), as estruturas propostas dão suporte a diversas funcionalidades já implantadas por sua equipe de desenvolvimento. No entanto, para a adequada representação contextual, novas estruturas precisam ser criadas. Dentre as representações contextuais sugeridas por Mota (2015), tem-se a representação contextual da citação.

As citações podem ser expressas mediante relacionamentos entre artigos. Esta entidade, [Artigo], já está representada no DER da BRAPCI (Figura 6 em Anexo 4) e não serão necessárias modificações para representá-las. Para adequadamente representar a relação de citação em um DER, uma das possibilidades é criar uma entidade [Citação] que receberá os códigos dos dois artigos: o <citante> e o <citado> (Figura 7).



Figura 7 – Relacionamento de Citação entre Artigos

Ou, no caso de obras que não façam parte da base, é possível criar uma outra entidade chamada [Obra], que manterá os registros de obras externas e elaborar as relações com esta entidade (Figura 8). De qualquer forma, a [Citação] será uma entidade intermediária que sinaliza a relação entre um objeto <citante> e um <citado>.



Figura 8 – Relacionamento de Citação entre Artigo e Obra

Esta flexibilidade permitirá relacionar o artigo com qualquer obra que exista, esteja ela armazenada no próprio repositório, de forma interna (Figura 7), ou com alguma obra externa que tenha uma representação na base (Figura 8).

Outros elementos podem ser incluídos nesta representação: o tipo de citação (direta ou indireta), o trecho extraído (no caso de citação direta), o uso da citação (como fonte, concordando, discordando etc), entre outros. Estes elementos são, também, parte do contexto da citação, que direcionam a interpretação dos textos e podem ser representados na [Citação] por meio de

propriedades, valores que qualificam a relação entre o [Artigo] e a [Obra].

Dentro do escopo da Web Semântica, já existe uma ontologia que disponibiliza uma notação para representar as citações e seus diferentes usos dentro do texto. A CiTO (SHOTTON; PERONI, 2014) contém em sua definição diversas formas de representar um ato de citação. As distintas formas de citação são representadas por meio de propriedades do ato de citação, na ontologia, e podem assumir as formas ilustradas na Figura 9 (em Anexo 5).

Os elementos da Figura 9 podem ser reaproveitados para a Web Pragmática, pois foram criados segundo a estrutura tripla proposta para a Web Semântica, e não necessitam da criação de um relacionamento n-ário entre as partes, pois é representada mediante OWL2, que não permite representação de relacionamentos n-ários. No entanto, a forma atual não permite melhor descrever a citação além do já expresso nas propriedades. A possibilidade de representar, também, o trecho citado ou o tipo de citação (direta/indireta) não é permitida caso se utilize esta ontologia, ou as tecnologias atuais da Web Semântica.

Como mostrado na Figura 10 (em Anexo 6), a ontologia permite fazer a asserção de quem é citado pelo documento e de que forma isso foi realizado. A limitação desta ontologia é no que tange à qualificação das palavras-chave. Mesmo sendo possível determinar quem foi citado, e de que forma, não é possível identificar a citação com palavras-chave, podendo ocorrer conflitos entre autores semelhantes que definem conceitos de formas distintas.

Esta representação visa facilitar o processo de recuperação de informações, possibilitando a identificação pontual da definição dos termos, relacionando-os aos autores que os definem.

A outra possibilidade é utilizar a reificação, a forma mais próxima dos padrões e tecnologias definidas e utilizadas para a Web Semântica. Nesta situação, cria-se uma classe intermediária, com RDFS, chamada neste exemplo de [Keyword\_Context], cujos objetos realizam a relação entre a palavra-chave e o contexto de citação. O problema, neste modo, é que para cada palavra-chave ou autor utilizado é necessário criar um novo objeto da classe [Keyword\_Context].

No exemplo ilustrado na Figura 11 (em Anexo 7), há apenas uma palavra-chave e autor, denominado "Keyword\_Context\_1", mas caso exista um número maior de palavras-chave ou de autores para cada uma delas, novos objetos,

com diferentes nomes (Keyword\_Context\_2, Keyword\_Context\_3, Keyword\_Context\_4, etc.), precisariam ser criados.

Cada um desses objetos se relaciona com o documento, criando-se uma lista de relacionamentos entre o documento e diversos objetos [Keyword\_Context]. Um exemplo de registros RDF reificados pode ser visto na Figura 12 (em Anexo 8), que contém uma lista de palavras-chave adequadamente contextualizadas para a descrição do documento.

Em Topic Maps, as associações entre os documentos e as palavras-chave seriam descritas de forma semelhante ao apresentado na Figura 13 (em Anexo 9). À medida que novas palavras-chave forem incluídas, novas relações serão expressas desta forma.

Esta forma de representação pode ser escolhida para evitar a criação e manutenção de diversos objetos de relacionamento entre palavra-chave e contexto, o que evita que, acidentalmente, se referencie no documento algum objeto [Keyword-Context] inexistente ou existam objetos que não sejam referenciados por nenhum documento no repositório.

## 5. Considerações finais

O objetivo desta pesquisa foi avaliar a possibilidade de representação contextual em um ambiente informacional digital. Como ambiente de estudos, foi escolhida a Base de Dados Referencial de Artigos de Periódicos em Ciência da Informação (BRAPCI), por reunir, em um só dispositivo, os artigos publicados em periódicos brasileiros de Ciência da Informação.

Como abordado no início do texto, os diferentes contextos, intra ou extratextuais, são um conjunto de elementos que são considerados pelos sujeitos no processo de interpretação de textos. Porém o computador não tem como compreender ou interpretar esses elementos sem a adequada representação e programação.

A Web Semântica, como diz o próprio nome, representa informação semântica, carecendo portanto de informação contextualizada. A Web Pragmática procura percorrer novos caminhos para que a compreensão se dê para além do nível semântico, de modo a incorporar os aspectos pragmáticos indispensáveis à interpretação.

Os resultados da pesquisa indicaram que os Topic Maps são um recurso elegante e efetivo para representar informação contextual (intra-textual e intertextual) em dispositivos digitais de

informação científica na perspectiva da Web Pragmática.

## Referências

- Armengaud, F. (2006) *A Pragmática*. São Paulo: Parábola Editorial. 159 p. ISBN 85-88456-50-8.
- Associação Brasileira de Normas Técnicas (ABNT). (2002). *NBR 10520: Informação e documentação – Citações em documentos – Apresentação*. Rio de Janeiro: ABNT, 2002.
- Bufrem, L. S.; Costa, F. D. de O.; Gabriel Junior, R. F.; Pinto, J. S. de P. (2010). Modelizando práticas para a socialização de informações- a construção de saberes no ensino superior. *Perspectivas em Ciência da Informação*. 15(2), 22-4.
- Chen, P. (1976). The Entity-Relationship Model-Toward a Unified View of Data. *ACM Transactions on Database Systems*. 1(1), 9-36.
- Cortina, A. (2000). *O Príncipe de Maquiavel e seus leitores: uma investigação sobre o processo de leitura*. São Paulo: Editora UNESP. ISBN
- De Moor, A.; Keeler, M.; Richmond, G. (2002). Towards a pragmatic web, En: Uta, Priss et al. (2002). *Conceptual Structures: Integration and Interfaces. Lecture Notes in Computer Science*. 235-249. ISBN 978-3-540-45483-0.
- Eco, U. (1987). *Lector in Fabula: la cooperación interpretativa en el texto narrativo*. Barcelona: Lumen. ISBN 84-264-1122-3.
- Eco, U. Os limites da interpretação. São Paulo: Pioneira, 2000. ISBN 85-27301-78-4.
- Edmonds, B. (1999). The Pragmatic Roots of Context. En: Bouquetm P.; Benerecetti, L.; Brézillon, P.; Castellani, F. (1999) *Modeling and Using Context*. Berlin; Heidelberg; New York: Springer. 119-132. ISBN 978-3-540-48315-1.
- Francelin, M. M. (2010). *Ordem dos conceitos na organização da informação e do conhecimento*. São Paulo: Tese (Doutorado em Ciência da Informação) PPGCI. 220 p.
- Garshol, L. M. (2003). Living with topic maps and RDF. 2003. Recuperado 08-04-2015, de <http://www.ontopia.net/topicmaps/materials/tmrd.html>
- Ingwersen, P.; Järvelin, K. (2005). *The Turn: integration of information seeking and retrieval in context*. Dordrecht, The Netherlands: Springer. ISBN 978-1-4020-3850-1.
- International Organization for Standardization (ISO). (2003). *ISO/IEC 13250: Information technology -- SGML applications -- Topic maps*. Genebra: ISO.
- International Organization for Standardization (ISO). (2013). *ISO/IEC 13250-3: Information technology -- Topic Maps -- Part 3: XML syntax*. Genebra: ISO.
- Morris, C. (1985). *Fundamentos de la teoría de los signos*. Barcelona: Paidós. 118 p. ISBN 84-7509-331-0.
- Mota, D. (2015). *Representação e Recuperação de Informação em Acervos Digitais nos Contextos da Web Semântica e Web Pragmática: um estudo crítico*. São Paulo: Qualificação (Doutorado em Ciência da Informação) PPGCI. 82 p.
- Mota, D.; Kobashi, N. Y. (2014). *Inclusão de contexto em artigos científicos: em busca da web pragmática*. Recuperado 14-12-2014, de [http://biredial.ucr.ac.cr/index.php/Biredial-ISTEC\\_2014/2014/paper/view/133](http://biredial.ucr.ac.cr/index.php/Biredial-ISTEC_2014/2014/paper/view/133)
- Peirce, C. S. (1977) *Semiótica*. São Paulo: Perspectiva. 337 p. ISBN 85-273-0194-6.
- Pepper, S. (2000) *The TAO of Topic Maps*. Recuperado 08-04-2015, de <http://www.ontopia.net/topicmaps/materials/tao.html>

- Shotton, D.; Peroni, S. (2014). *CiTO, the Citation Typing Ontology*. Recuperado 08-07-2015, de <http://www.essepuntato.it/lode/http://purl.org/spar/cito>
- Singh, M. P. (2002). *The Pragmatic Web: Preliminary thoughts*. Recuperado 14-12-2014, de <http://lsdis.cs.uga.edu/SemNSF/Singh-Position.pdf>
- Sowa, J. F. (2000) *Knowledge Representation: Logical, Philosophical, and Computational Foundations*. Pacific Grove: Brooks/Cole. 608 p. ISBN 978-0-534-94965-5.
- Vallet, D.; Fernández M.; Castells, P.; Mylonas, P.; Avrithis, Y. (2006) *Personalized Information Retrieval in Context*. Recuperado 15-04-2015, de <http://ir.ii.uam.es/~acemedia/publications/mrc06.pdf>
- World Wide Web Consortium (W3C). (2006). *Defining N-ary Relations on the Semantic Web*. Recuperado 20-02-2015 de <http://www.w3.org/TR/swbp-n-aryRelations>

## Anexo 1

```

1 <rdf:Description rdf:about="http://www.csc.ncsu.edu/faculty/mpsi-
  ngh/papers/positions/nsf-02.pdf">
2 < dct:title>The Pragmatic Web: Preliminary Thoughts</dct:title>
3 < dct:language rdf:resource="http://lexvo.org/id/iso639-3/eng"/>
4 < dct:subject rdf:resource="http://dbpedia.org/page/Pragmatics"/>
5 < dct:subject rdf:resource="http://dbpedia.org/page/Char-
  les W. Morris"/>
6 < dct:subject rdf:resource="http://bnb.data.bl.uk/id/con-
  cept/ddc/e21/940.548673092"/>
7 </rdf:Description>

```

Figura 3 – Exemplo de definição de palavras-chave em RDF

## Anexo 2

```

1 <rdf:RDF>
2 < rdfs:Class rdf:ID="Resource"/>
3 < rdfs:Class rdf:ID="Keyword_Context"/>
4 < rdf:Property rdf:ID="context">
5 < rdfs:domain rdf:resource="#Keyword_Context"/>
6 < rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema-
  chema:string"/>
7 </rdf:Property>
8 < rdf:Property rdf:ID="keyword">
9 < rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema-
  chema:string"/>
10 < rdfs:domain rdf:resource="#Keyword_Context"/>
11 </rdf:Property>
12 < rdf:Property rdf:ID="has_Keyword">
13 < rdfs:range rdf:resource="#Keyword_Context"/>
14 < rdfs:domain rdf:resource="#Resource"/>
15 </rdf:Property>
16 < Keyword_Context rdf:nodeID="Keyword_Context_1">
17 < keyword rdf:datatype="http://www.w3.org/2001/XMLSchema-
  chema:string"> http://dbpedia.org/page/Pragmatics </keyword>
18 < context rdf:datatype="http://www.w3.org/2001/XMLSchema-
  chema:string"> http://dbpedia.org/page/Charles W. Morris
  </context>
19 </Keyword_Context>
20 < Resource rdf:ID="http://www.csc.ncsu.edu/faculty/mpsingh/pa-
  pers/positions/nsf-02.pdf">
21 < has_Keyword rdf:resource="#Keyword_Context_1"/>
22 </Resource>
23 </rdf:RDF>

```

Figura 4 – Exemplo de reificação em RDF

## Anexo 3

```

1 <association>
2   <type>
3     <subjectIdentifierRef
4       href="http://xmlns.com/foaf/spec/#term_topic"/>
5   </type>
6   <member>
7     <type>
8       <subjectIdentifierRef
9         href="http://xmlns.com/foaf/spec/#term Document"/>
10    </type>
11    <subjectIdentifierRef xlink:href="http://www.csc.ncsu.edu/faculty/mpsingh/papers/positions/nsf-02.pdf"/>
12  </member>
13  <member>
14    <type>
15      <subjectIdentifierRef href="http://purl.org/dc/terms/subject"/>
16    </type>
17    <subjectIdentifierRef xlink:href="http://dbpedia.org/page/Pragmatics"/>
18  </member>
19  <scope>
20    <subjectIdentifierRef xlink:href="http://dbpedia.org/page/Charles W. Morris"/>
21  </scope>
22 </association>

```

Figura 5 – Exemplo de associação em XTM

## Anexo 4

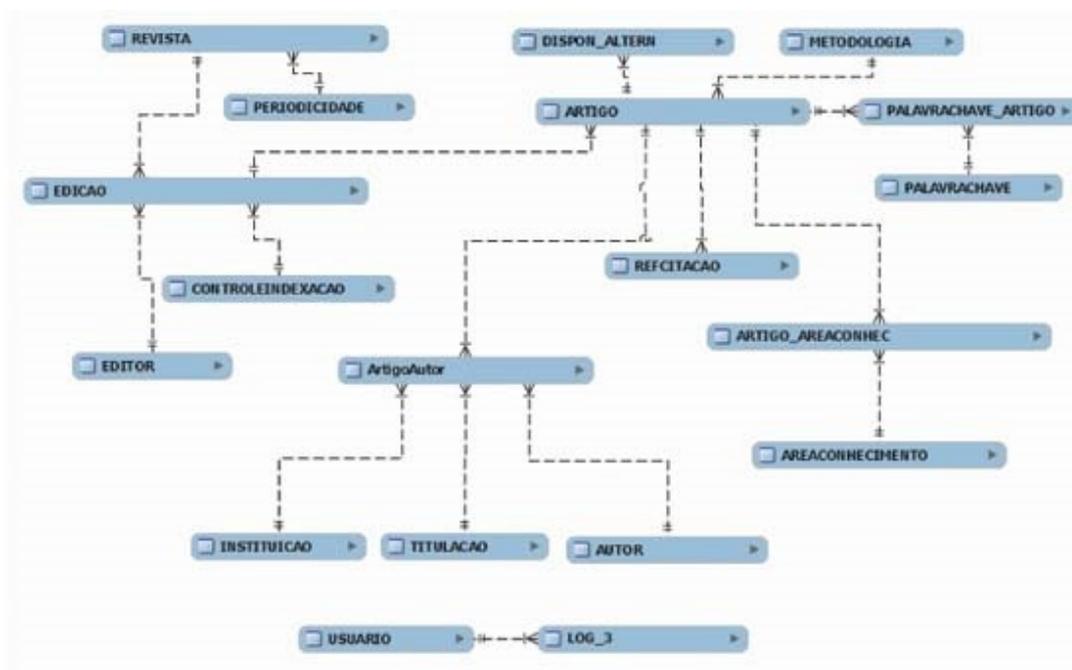


Figura 6 – Diagrama Entidade-Relacionamento da BRAPCI (Base de Dados Referenciais de Artigos de Periódicos em Ciência da Informação). Fonte: BUFREM; COSTA; GABRIEL JUNIOR, 2005.

## Anexo 5



Figura 9 – Tipos de Citação da Ontologia CiTO. Fonte: Elaborado pelo autor, 2015. Adaptado de Shotton e Peroni (2014).

## Anexo 6

```

1 <rdf:Description rdf:about="http://biredial.ucr.ac.cr/in-
2 dex.php/Biredial-ISTEC_2014/2014/paper/view/133">
3 < dct: title>Inclusão de contexto em artigos científicos: em busca
4 da web pragmática</dct: title>
5 < dct: language rdf: resource="http://lexvo.org/id/iso639-3/por"/>
6 < dct: subject rdf: resource="http://dbpedia.org/page/Pragmatics"/>
7 < cito: cites_as_authority rdf: resou-
8 rce="http://bnb.data.bl.uk/id/resource/002554543"/>
9 </rdf: Description>

```

Figura 10 – Exemplo de citação em RDF. Fonte: MOTA, 2015.

## Anexo 7

```

1 <rdf:RDF>
2   <rdfs:Class rdf:ID="Resource"/>
3   <rdfs:Class rdf:ID="Keyword_Context"/>
4   <rdf:Property rdf:ID="context">
5     <rdfs:domain rdf:resource="#Keyword_Context"/>
6     <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#string"/>
7   </rdf:Property>
8   <rdf:Property rdf:ID="keyword">
9     <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#string"/>
10    <rdfs:domain rdf:resource="#Keyword_Context"/>
11  </rdf:Property>
12  <rdf:Property rdf:ID="has_Keyword">
13    <rdfs:range rdf:resource="#Keyword_Context"/>
14    <rdfs:domain rdf:resource="#Resource"/>
15  </rdf:Property>
16  <Keyword_Context rdf:nodeID="Keyword_Context_1">
17    <keyword rdf:datatype="http://www.w3.org/2001/XMLSchema#string"> http://dbpedia.org/page/Pragmatics </keyword>
18    <context rdf:datatype="http://www.w3.org/2001/XMLSchema#string"> http://lccn.loc.gov/a41002443 </context>
19  </Keyword_Context>
20  <Resource rdf:ID="http://biredial.ucr.ac.cr/index.php/Biredial-ISTEC_2014/2014/paper/view/133">
21    <has_Keyword rdf:resource="#Keyword_Context_1"/>
22  </Resource>
23 </rdf:RDF>

```

Figura 11 – Exemplo de citação com reificação em RDF. Fonte: Mota, 2015.

## Anexo 8

```

1 <Keyword_Context rdf:nodeID="Keyword_Context_1">
2   <keyword rdf:datatype="http://www.w3.org/2001/XMLSchema#string">
3     http://dbpedia.org/page/Pragmatics </keyword>
4   <context rdf:datatype="http://www.w3.org/2001/XMLSchema#string">
5     http://lccn.loc.gov/a41002443 </context>
6 </Keyword_Context>
7 <Keyword_Context rdf:nodeID="Keyword_Context_2">
8   <keyword rdf:datatype="http://www.w3.org/2001/XMLSchema#string">
9     http://dbpedia.org/page/RDF </keyword>
10  <context rdf:datatype="http://www.w3.org/2001/XMLSchema#string">
11    http://w3.org/2001/sw/wiki/Special:URIResolver/RDF</context>
12 </Keyword_Context>
13 <Keyword_Context rdf:nodeID="Keyword_Context_3">
14  <keyword rdf:datatype="http://www.w3.org/2001/XMLSchema#string">
15    http://dbpedia.org/page/Knowledge Representation </keyword>
16  <context rdf:datatype="http://www.w3.org/2001/XMLSchema#string">
17    http://bnb.data.bl.uk/doc/resource/009418545 </context>
18 </Keyword_Context>
19 <Resource rdf:ID="http://biredial.ucr.ac.cr/index.php/Biredial-IS-
20   TEC\_2014/2014/paper/view/133">
21   <has_Keyword rdf:resource="#Keyword_Context_1"/>
22   <has_Keyword rdf:resource="#Keyword_Context_2"/>
23   <has_Keyword rdf:resource="#Keyword_Context_3"/>
24 </Resource>

```

Figura 12 – Exemplo de várias citações com reificação em RDF. Fonte: Mota, 2015.

## Anexo 9

```

1 <association>
2   <type> <subjectIdentifierRef
3     href="http://xmlns.com/foaf/spec/#term_topic"/> </type>
4   <member>
5     <type> <subjectIdentifierRef
6       href="http://xmlns.com/foaf/spec/#term_Document"/> </type>
7     <subjectIdentifierRef xlink:href="http://www.csc.ncsu.edu/faculty/mpsingh/papers/positions/nsf-02.pdf "/>
8   </member>
9   <member>
10    <type> <subjectIdentifierRef
11      href="http://purl.org/dc/terms/subject"/> </type>
12    <subjectIdentifierRef xlink:href="http://dbpedia.org/page/Pragmatics"/>
13  </member>
14  <scope>
15    <subjectIdentifierRef href="http://lccn.loc.gov/a41002443"/>
16  </scope>
17 </association>

```

Figura 13 – Exemplo de citação em XTM. Fonte: Mota, 2015.

---

# Aplicação da Teoria da Análise Facetada e Mapas Conceituais para navegação facetada em uma ontologia de domínio: resultados preliminares

*An application of Theory Faceted Analysis and Concept Maps for faceted search in a domain ontology: preliminary studies*

---

**Cristiane Mendes Netto (1), Gercina Ângela Borém de Oliveira Lima (2), Ivo Pierozzi Junior (3)**

(1) Universidade Vale do Rio Doce, cristiane@univale.br

(2) Universidade Federal de Minas Gerais, glima@eci.ufmg.br

(3) Embrapa Informática Agropecuária, ivo.pierozzi@embrapa.br

## Resumo

Este artigo apresenta os resultados parciais de um projeto de pesquisa em desenvolvimento que tem como objetivo estudar a Teoria da Análise Facetada e Mapas Conceituais para navegação em uma ontologia de domínio. A pesquisa consiste em apresentar uma solução que permita aos usuários níveis de abstração para a navegação facetada na área de domínio representado em uma ontologia. O problema, analisado no contexto real de uma empresa brasileira de pesquisa em agropecuária, constitui-se um desafio, pois a estrutura computacional formal de descrição semântica de uma ontologia de domínio não se mostra viável do ponto de vista da cognição de seus usuários. Considerando isto, apresenta-se neste trabalho os resultados de estudos preliminares de Teoria da Análise Facetada e Mapas Conceituais com a utilização do Facetlog, um protótipo de ferramenta web para navegação facetada aplicado em uma amostra da ontologia OntoAgroHidro. Esta ontologia foi criada por uma empresa de pesquisa brasileira visando a organização do conhecimento do domínio referente ao impacto da agricultura e das mudanças climáticas nos recursos hídricos. Os resultados obtidos demonstram inicialmente a viabilidade de navegação e recuperação de informação na ontologia através do protótipo web facetado. Acredita-se através deste estudo que seja possível desenvolver uma solução computacional que possa facilitar a criação do modelo conceitual para navegação em facetadas e mapas conceituais em todo o domínio representado pela OntoAgroHidro, permitindo assim avaliar o aprendizado humano sobre o domínio, bem como a recuperação e compartilhamento de informação em grupos de usuários.

**Palavras-chave:** Organização do conhecimento. Teoria da análise facetada. Mapas conceituais. Ontologias de domínio..

## Abstract

This paper presents partial results of a research project still in development that aims to study the theories of faceted analysis and concept maps for faceted search in a domain ontologies. The research presents a solution that enables abstraction levels for users in order to retrieve information in the domain area represented in ontology. The problem, which is analyzed in a real context of Embrapa (Brazilian Agricultural Research Corporation), is considered a challenge because the formal computational structure of semantic description of domain ontology does not present itself as feasible from the point of view of its users' cognition. Therefore, this paper presents studies of Faceted Analysis Theory and Concept Maps and the results using a web tool prototype for faceted navigation in a sample of *OntoAgroHidro* ontology that was created by Embrapa and which is aimed at the organization of the domain knowledge regarding the impact of agriculture and climatic changes on water resources. The results show the feasibility of navigation and retrieval of information in the ontology using the Web faceted prototype. It is believed that through this study a computational solution can be developed that can facilitate the creation of the conceptual model for faceted and concept map navigation around the area represented by *OntoAgroHidro*. This way human learning on the domain can be assessed, as well as the recovery and sharing of information on user groups.

**Keywords:** Organization of knowledge. Theory of faceted analysis. Concept maps. Domain ontologies.

## Agradecimientos

Agradecimentos ao apoio das Agências de fomento FAPEMIG e CNPq/Brasil.

## 1 Introdução

Um dos objetivos da organização do conhecimento é criar uma visão compartilhada sobre um domínio para determinada comunidade de usuários. Assim, no contexto de um grupo de pessoas trabalhando em conjunto dentro de um mesmo domínio, tem-se que a organização de conceitos e informação em Sistemas de Organização do Conhecimento é essencial para o desenvolvimento de um trabalho eficaz.

Considerando essa necessidade e a possibilidade de uso das tecnologias para apoiar a organização do conhecimento, as ontologias de domínio têm-se apresentado como uma alternativa para explicitar de maneira formal conceitos e relações compartilhadas dentro de uma determinada área de assunto.

Conforme apresentado por Sales *et al* (2008), um importante elemento das ontologias é a representação do conhecimento em um contexto semântico, podendo prover aos computadores maior inteligência de processamento devido às representações de relações estarem explícitas em uma linguagem formal. Desta forma, melhores resultados de recuperação de informação podem ser apresentados para os usuários.

No âmbito computacional, a criação de ontologias tem sido beneficiada pelo desenvolvimento de linguagens formais como a OWL (*Web Ontology Language*) (1) e de ferramentas sofisticadas como o *Protégé* (2) que oferecem os recursos necessários para que se possa fazer a recuperação de informação pelos computadores. No entanto, do ponto de vista navegacional e da cognição humana, a estrutura é extremamente complexa e carece de uma interface que possibilite níveis de abstrações para recuperação de informação que favoreça aos usuários, sem domínio dessas tecnologias, facilidade para acesso, recuperação e aprendizagem dos diversos conteúdos do domínio, representados pela ontologia.

Esse problema identificado, no contexto real da Empresa Brasileira de Pesquisa Agropecuária, constitui-se um desafio e carece de uma solução eficiente. Assim, este trabalho apresenta os resultados parciais de um projeto de pesquisa em desenvolvimento, que tem como objetivo estudar a Teoria da Análise Facetada e Mapas Conceituais para representação de informação em ontologias de domínio. A ontologia de domínio analisada, neste estudo de caso, é a *Onto-*

*AgroHidro*, que foi criada, visando a organização do conhecimento do domínio referente ao impacto da agricultura e das mudanças climáticas nos recursos hídricos. A sua construção foi feita por especialistas e a estrutura contém mais de 8500 conceitos do domínio descritos em OWL 2, usando a ferramenta *Protégé*.

A justificativa para este estudo da Teoria da Análise Facetada foi baseada nos resultados dos trabalhos desenvolvidos por Pontes e Lima (2012) e Maculan (2011), onde se mostra que um modelo conceitual, no contexto digital baseado nesta teoria, proporcionou uma representação e recuperação mais eficiente. Além disso, têm-se os resultados positivos do trabalho realizado por Lima (2004), em que a Teoria da Análise Facetada foi utilizada para modelagem conceitual de um documento hipertextual, sendo também feito o uso da teoria de Mapas Conceituais para criar uma estrutura de navegação.

No percurso metodológico de desenvolvimento deste trabalho, a partir dos estudos do referencial teórico e de revisão de literatura, encontrou-se no protótipo *Facetlog* desenvolvido por Silva (2013), uma ferramenta para representar a ontologia *OntoAgroHidro* em uma interface web de navegação facetada. Para esta representação, foi escolhida de forma arbitrária uma amostra da ontologia e criado um modelo de mapeamento desta para a estrutura do protótipo.

Os resultados obtidos demonstram inicialmente a viabilidade de navegação e recuperação de informação na ontologia através do protótipo web facetado. Acredita-se que é possível desenvolver uma solução computacional que possa facilitar a criação do modelo conceitual para navegação em facetadas e mapas conceituais em todo o domínio representado pela *OntoAgroHidro*, permitindo, assim, avaliar o aprendizado humano sobre o domínio, bem como a recuperação e compartilhamento de informação entre usuários.

A organização do artigo apresenta-se da seguinte forma. A seção 2 relata a metodologia de desenvolvimento do trabalho. Em seguida, são descritos, nas respectivas seções, os fundamentos teóricos, a revisão de literatura, a ontologia *OntoAgroHidro*, os resultados do trabalho e as considerações finais.

## 2 Metodologia

No que se refere ao objetivo proposto, este trabalho se enquadra como uma pesquisa exploratória, descritiva e de estudo de caso. Segundo Gil (1991), a primeira tem como objetivo proporcionar maior familiaridade com o problema, com

vistas a torná-lo mais explícito, envolvendo o aprimoramento de ideias ou a descoberta de intuições. A pesquisa descritiva visa a descrição das características de determinada população ou fenômenos ou, então, o estabelecimento de relações entre variáveis. O estudo de caso é em razão do uso da *OntoAgroHidro*, dentro do seu contexto real. A natureza da pesquisa é de caráter aplicado, que, segundo Marconi e Lakatos (2003), caracteriza-se por seu interesse prático, isto é, que os resultados sejam aplicados, imediatamente, na solução de problemas que ocorrem na realidade.

As etapas a seguir descrevem o percurso metodológico utilizado para desenvolvimento deste trabalho:

- *Pesquisa Bibliográfica.* Nesta etapa foi realizada uma pesquisa de referencial teórico para estudo sobre os conceitos de Teoria da Análise Facetada, Mapas Conceituais e Ontologias, bem como um estudo de trabalhos relacionados, onde se buscou analisar o uso desses conceitos em aplicações de organização da informação. A pesquisa foi realizada nas bases de dados do Portal Capes, Google Acadêmico, Revistas Científicas nacionais e internacionais, bem como em bibliotecas de teses e dissertações de diversas instituições.
- *Estudo da Ontologia OntoAgroHidro.* Visando compreender o domínio a ser modelado, estudou-se o processo de construção da ontologia, bem como a sua estrutura, relações e instâncias representadas na ferramenta *Protégé*. Durante esta etapa houve contato contínuo com profissionais especialistas do domínio que trabalharam na criação da ontologia e no estudo de ferramentas de visualização para compreensão das estruturas da ontologia.
- *Estudo e escolha da Interface.* Para avaliar como a representação da *OntoAgroHidro* poderia ser feita, em facetadas e mapas conceituais, estudaram-se algumas ferramentas e escolheu-se aprofundar no estudo do protótipo *Facetlog* (3), desenvolvido por Silva (2013), onde se apresenta uma interface web para navegação facetada.
- *Modelagem Conceitual.* Dentro desta etapa, buscou-se identificar como as estruturas da ontologia poderiam ser abstraídas e mapeadas para a ferramenta de navegação web facetada *Facetlog*.
- *Testes e Análises de Resultados.* A partir do modelo representado, foram realizados testes e avaliações de navegação e recuperação de informação.

### 3 Fundamentos Teóricos

Nesta seção é apresentada a fundamentação teórica dos assuntos em que se baseia a proposta de desenvolvimento deste trabalho.

#### 3.1 Teoria da Análise Facetada

A Teoria da Análise Facetada tem sua origem no trabalho de um matemático indiano que se tornou bibliotecário, Shialy Rammarita Ranganathan (1892-1972), publicado na década de 1930. Ele é considerado um dos estudiosos que mais contribuiu para a teoria da biblioteconomia no século XX, especialmente na área da classificação de assunto (Lima, 2004). A teoria consiste em um tipo de classificação capaz de identificar características comuns a diversas categorias de um assunto, organizando-o em partes denominadas facetadas; envolve dois processos diferenciados, porém complementares: a análise de assunto em facetadas e a síntese dos elementos constituintes do mesmo, sendo portanto, aplicável a qualquer que seja a área do conhecimento (Duarte e Cerqueira, 2007). Na definição de Barbosa (1972) tem-se que:

Classificação facetada é o sistema que agrupa termos estruturados, na base da análise de um assunto, para identificação de suas facetadas, isto é, dos diferentes aspectos nele contidos. A análise em facetadas coordena conceitos, significando que um assunto, por mais complexo que seja, pode ser representado pela síntese de mais de uma facetada, cada uma indicando conceitos diferentes [...].

Conforme apresentado por Lima (2004), para elaborar uma classificação facetada, examina-se a literatura de um assunto com a finalidade de identificar seus conceitos e termos e então se estabelecem as suas características e facetadas. Após levantar e definir a terminologia própria do assunto, os termos são analisados e distribuídos em facetadas. Neste processo, um assunto é fragmentado em facetadas e representado através da seleção de elementos de cinco categorias fundamentais: Personalidade (*Personality*), Matéria (*Matter*), Energia (*Energy*), Espaço (*Space*) e Tempo (*Time*), conhecida pela sigla PMEST. A categoria Personalidade é a característica que distingue o assunto. Na categoria Matéria, localiza-se o material físico do qual um assunto pode ser composto. A categoria Energia pode ser entendida como uma ação que ocorre com respeito ao assunto, enquanto a categoria Espaço é o componente geográfico da localização de um assunto e a categoria Tempo representa o período associado com um assunto.

As categorias fundamentais de Ranganathan foram estudadas de forma dedicada, a partir da

década de 50, por um grupo de pesquisadores de Londres que formou o *Classification Research Group* (CRG). A partir das ideias propostas por Ranganathan, esse grupo propôs modificações em alguns aspectos dessa teoria, por considerá-la muito restritiva. Dentre as modificações, o CRG ampliou as cinco categorias do PMEST para dez: Tipos de produto final, Partes, Materiais, Propriedades, Processos, Operações, Agentes, Espaço, Tempo e Forma de Apresentação. A Tabela 1, conforme apresentada por Maculan (2011), demonstra que as categorias indicadas pelo CRG parecem se acomodar dentro das categorias indicadas por Ranganathan, sendo apenas mais segmentadas e subdivididas.

| <i>Ranganathan</i> | CRG                    |
|--------------------|------------------------|
| Personalidade      | Tipos de produto final |
| Matéria            | Partes                 |
|                    | Materiais              |
|                    | Propriedades           |
|                    | Formas de Apresentação |
| Energia            | Processos              |
|                    | Operações              |
|                    | Agentes                |
| Espaço             | Espaço                 |
| Tempo              | Tempo                  |

Tabela 1 - *Comparação entre as categorias de Ranganathan e CRG*

Conforme explica Maculan (2011), tanto os fundamentos de Ranganathan quanto do CRG propõem segmentar um assunto dentro de um domínio em particular. Assim, tem-se que todo domínio possui, na visão de Ranganathan, as categorias: “Personalidade”, que é o assunto ou objeto; “Matéria”, que é a decomposição desse assunto ou objeto em diferentes atributos ou características (materiais, propriedades, qualidade); “Energia”, que são as ações, atividades ou processos em relação ao assunto ou objeto investigado; “Espaço”, que é a ambientação do assunto ou objeto, tal como seu lugar geográfico; e “Tempo”, que geralmente é a data ou período histórico-contextual do objeto estudado, ou mesmo o ano de produção de tal conhecimento. As categorias “Espaço” e “Tempo” são iguais e, por isso, não causam polêmica, porém, as outras merecem melhor esclarecimento. Traçando um paralelo conforme as demais categorias indicadas por Ranganathan e o CRG, consideram-se as categorias: “Personalidade” como “Tipos de Produto Final”, pois é o tema sobre o qual o autor irá investigar; “Matéria”, como “Partes, Materiais, Propriedade”, que é a segmentação do tema pesquisado e “Energia”,

como “Processos, Operações, Agentes”, uma vez que indica as formas de execução e as variáveis analisadas na ação. Desse modo, nota-se que, mesmo ampliando o número de categorias, as mesmas podem ser consideradas especificidades das categorias representadas pelo PMEST, idealizado por Ranganathan.

Dentre as vantagens da Teoria da Análise Facetada, Costa (2012) apresenta a possibilidade de um conjunto de informação ser organizado de diversas formas. A organização de uma revista, por exemplo, pode ser realizada por assuntos: notícias, esportes, cultura, arte, saúde, beleza, dentre outros. Esse conjunto pode ser denominado de classificação facetada, pois cada assunto é uma dimensão na estrutura que organiza a informação. Esse tipo de classificação permite ao usuário buscar informações a partir de diversos termos, podendo relacioná-los, o que se caracteriza como uma das vantagens da classificação facetada, pois acomoda diferentes estratégias de busca e modelos conceituais.

No desenvolvimento deste trabalho, buscou-se o entendimento dos princípios da organização do conhecimento pela análise facetada, representação de conceitos descritos em uma ontologia de domínio.

### 3.2 Mapas Conceituais

A teoria referente a Mapas Conceituais foi desenvolvida, na década de 70, pelo pesquisador Joseph Novak, a partir da Teoria da Aprendizagem Significativa de David Ausubel. Novak e Kañas (2008) apresentam o mapa conceitual como uma ferramenta gráfica para organização e representação do conhecimento. Na definição apresentada por Moreira (2010) tem-se que os mapas conceituais podem ser entendidos como diagramas bidimensionais, que mostram relações hierárquicas entre conceitos de um corpo de conhecimento e que derivam sua existência da própria estrutura conceitual desse corpo.

Os mapas conceituais permitem representar os conceitos, geralmente dentro de círculos, e as relações entre os conceitos, que são indicadas por linhas que os interligam. Os conjuntos conceito-relação-conceito formam proposições, que são afirmações significativas sobre os objetos ou eventos envolvidos. As proposições definem as unidades semânticas, ou unidades de conhecimento do mapa. Na estrutura do mapa conceitual pode-se representar também ligações cruzadas, identificando relacionamentos entre conceitos em duas diferentes áreas do mapa. A Figura 1 apresenta um mapa conceitual, conforme Souza (2006), representando esses conceitos. Uma característica importante do

mapa conceitual mostrado nesta ilustração é que os conceitos são representados de forma hierárquica com os mais inclusivos, ou genéricos, localizados na parte superior do mapa, e os conceitos mais específicos localizados na parte inferior. Dessa forma, o eixo vertical do mapa define uma estrutura hierárquica para os conceitos.

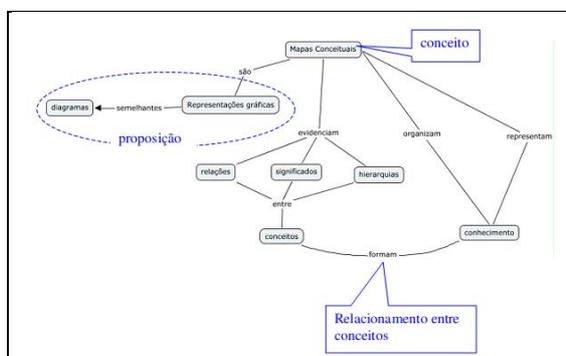


Figura 1: Um mapa conceitual representando o conceito de mapa conceitual

Conforme apresentado por Lima (2004), os mapas conceituais podem apresentar as seguintes vantagens: 1) definição de uma ideia central, através do posicionamento do assunto no centro da página; 2) uma clara indicação da importância relativa de cada ideia; 3) a facilidade para encontrar os links entre as ideias-chave. 4) a visão geral de toda informação básica em uma página; 5) em decorrência, uma revocação e uma revisão mais eficientes; 6) a inserção de novas informações, sem atrapalhar a estrutura informacional; 7) a facilidade para acessar a informação em diferentes formatos e diferentes pontos de vista; 8) a facilidade da compreensão da complexidade de relações entre as ideias; 9) a facilidade de verificar contradições, paradoxos e falhas no material organizado.

Conforme Tavares (2007), embora os mapas conceituais possam transmitir informações factuais tão bem quanto os textos, esses organizadores gráficos são mais efetivos que os textos para ajudar os leitores a construir inferências complexas e integrar as informações que eles fornecem. Eles também têm o potencial de melhorar a acessibilidade e usabilidade materiais durante uma pesquisa na medida em que apresentam marcas visuais-espaciais que podem guiar uma seleção ou categorização. Existe a comprovação empírica sobre a eficiência de buscas, onde se comprova que os interessados localizam mais informações quando elas são apresentadas em formas de mapas ao invés de textos.

A estrutura de um mapa conceitual pode ser representada de diferentes maneiras, conforme a necessidade e interesse de demonstrar os conceitos. Dentre as representações mais comuns estão os mapas do tipo teia, em que o tema central é colocado no centro do mapa, os mapas hierárquicos e os em *flowchart*. Quanto à forma de apresentação, os mapas podem estar em formato 3D, hiperbólico, etc.

Existem diversas ferramentas que permitem o desenho de mapas conceituais, sendo o software *CmapTools*, desenvolvido pelo IHMC (*Institute for Human and Machine Cognition*), uma solução com potencial para auxiliar na exploração de documentos, dando suporte à visualização e ao processo de aquisição, organização, representação, localização e utilização do conhecimento documentado, bem como da informação e do conhecimento presente nos ambientes digitais.

Neste trabalho a compreensão do conceito de Mapas Conceituais foi necessária para estudar a representação dos conceitos descritos no domínio de conhecimento.

### 3.3 Ontologias

Conforme apresentado por Almeida (2013), o termo ontologia tem sido utilizado em diversos campos científicos, apresentando diferentes visões sobre o seu significado. No contexto da Ciência da Informação as ontologias podem ser consideradas um mecanismo de representação do conhecimento, como um meio de observação do conhecimento de um dado domínio.

Guizzard *et al* (2008) apresenta que, no contexto de modelagem conceitual, o termo ontologia é usado de acordo com a definição dada em filosofia, que é a referência a um sistema formal de categorias que pode ser usado para articular conceituações e modelos em domínios específicos de conhecimento.

Desta forma, de acordo com Campos (2010), pode-se entender que uma conceituação é uma abstração, uma visão simplificada do mundo, que se representa com o propósito de auxiliar na compreensão, compartilhamento ou consenso sobre uma área de conhecimento. Embora, na Ciência da Informação, existam outras maneiras de representação da informação em bases compartilhadas, como tesouros, as ontologias oferecem um recurso além e extremamente poderoso que é a possibilidade de vir a oferecer inferência para dar suporte à manipulação do conhecimento explicitado em uma ontologia.

De acordo com Almeida e Bax (2003), as ontologias nem sempre apresentam a mesma estru-

tura, mas existem características e componentes básicos, que são presentes em grande parte delas. Assim, são apresentados como componentes básicos: Classes (organizadas em uma taxonomia), Relações (representam o tipo de interação entre os conceitos de um domínio), Axiomas (usados para modelar sentenças verdadeiras) e Instâncias (utilizadas para representar elementos específicos – os próprios dados).

Quanto aos benefícios relacionados ao uso de ontologias na especificação de um domínio, Guizzardi (2000) apresenta:

- **Comunicação.** As ontologias são úteis para auxiliar a comunicação entre pessoas acerca de determinado conhecimento, pois auxiliam o raciocínio e o entendimento sobre esse domínio. Essa relação auxilia na obtenção de consenso, interação e elicitação de conhecimentos sobre termos técnicos em uma comunidade profissional.
- **Formalização.** A formalização está relacionada à especificação da ontologia, que permite eliminar contradições e inconsistências na representação de conhecimento, além de não ser ambígua. Além disso, essa especificação pode ser testada, validada e verificada.
- **Representação de Conhecimento e Reutilização.** As ontologias formam um vocabulário de consenso que permite representar conhecimento de um domínio em seu nível mais alto de abstração, possuindo, dessa forma, potencial de reutilização.

Existem várias classificações quanto aos tipos de ontologias. Guarino (1998) propõe uma classificação baseada em seus conteúdos. Nesta classificação têm-se as ontologias de alto nível (*Top Level Ontologies*) que descrevem conceitos gerais como espaço, tempo, evento e outros, que são independentes de um domínio em particular. Já as ontologias de domínio (*Domain Ontologies*) e as ontologias de tarefas (*Task Ontologies*) descrevem respectivamente o vocabulário de um domínio e de tarefas ou processos. As ontologias de aplicação (*Application Ontologies*) descrevem conceitos que são dependentes tanto de um domínio quanto de uma tarefa.

Neste trabalho a ontologia *OntoAgroHidro* estudada classifica-se como uma ontologia de domínio.

#### 4 Revisão de Literatura

Nos últimos anos diversos estudos baseados na Teoria da Análise Facetada têm sido realizados, principalmente no contexto de conteúdos e am-

bientes digitais, apontando ser esta a abordagem adequada para se utilizar na representação e organização da informação neste meio.

No trabalho de pesquisa de doutorado de Campos (2001) foi estudada a modelagem conceitual em hiperdocumentos, sendo proposto o modelo conceitual, não somente como etapa da modelagem de dados para a implementação de hiperdocumentos mas como uma prerrogativa da própria essência do objeto hiperdocumento. No trabalho de Campos destaca-se a atenção às necessidades e abordagens da Ciência da Informação no desenvolvimento da modelagem semântica de hiperdocumentos, mostrando preocupação com a preservação do conteúdo, os problemas da fragmentação e com a organização a partir da autoria dos mesmos. Dentre as teorias utilizadas no trabalho de Campos, a Teoria da Análise Facetada, com sua base classificatória, foi utilizada para compor um sistema de conceitos para construção de um domínio de conhecimento. Como resultados desse trabalho foram apresentados um conjunto de requisitos para auxiliar o desenvolvimento de metodologias, para a construção de modelos conceituais, mais apropriadas para hiperdocumentos. Constatou-se que a modelagem conceitual é uma etapa importante em todos os processos ligados a tecnologias de informação que envolvem a construção de modelos de representação.

Lima (2004) apresenta como objetivo de seu estudo de doutorado a construção de um modelo hipertextual para teses e dissertações, visando a organização desses documentos acadêmicos, tanto na instância de produção primária como no momento da organização de sistemas de hipertextos ou bibliotecas digitais. Os resultados desse trabalho foram um modelo conceitual baseado na conjugação de uma estrutura facetada e de um elemento da estrutura formal do documento; bem como demonstrou-se a aplicabilidade da Teoria da Análise Facetada na modelagem conceitual. A autora destaca ainda a importância e necessidade de pesquisas sobre a organização semântica de textos, envolvendo uma estrutura conceitual, a partir de uma análise cognitiva, seguida da correspondente aplicação tecnológica.

No trabalho realizado por Tristão *et al* (2004) constatou-se, através de uma revisão de estudos que a Teoria da Análise Facetada tem sido bastante utilizada como recurso para a organização do conhecimento, estando apta para atender à demanda em todas as áreas do conhecimento. Constatou-se que os sistemas de classificação facetado, além de criarem termos gerais e específicos, distribuem-nos em classes ou categorias gerais e específicas, usando face-

tas ou subclasses, criando-se um sistema correlacionado, que possibilita facilitar a organização e a recuperação.

A Teoria da Análise Facetada é apresentada no trabalho de Broughton (2006) como uma abordagem conceitual importante para métodos de recuperação de informação; essa teoria também é usada para a organização dos conteúdos em sites comerciais e também para a construção de vocabulários e modelos conceituais. Portanto, verificou-se que a utilização de sistemas facetados se tornou comum dentre os métodos de recuperação de informação, sendo útil para navegação e recuperação, podendo também ser explorada para indexação semiautomática e expansão de consultas em sistemas de recuperação de informação.

Hearst (2008) apresenta, em seu trabalho, uma análise sobre a interface de sistemas de navegação facetada construídos, baseados nos princípios da Teoria da Análise Facetada. É relatado que existe um crescente número de serviços oferecendo as possibilidades dessa técnica para navegação e exploração de conteúdos em ambientes web. No entanto, apresenta a necessidade das interfaces apresentarem melhores estratégias para as escolhas das suas categorias de classificação, bem como terem uma integração com as busca por palavra-chave.

Maculan (2011) em seu trabalho de mestrado utilizou os princípios da teoria da análise facetada e elaborou uma taxonomia facetada navegacional para uma biblioteca de teses e dissertações. Os resultados do trabalho de Maculan demonstraram que o mecanismo de navegação facetada proposto facilitou a tarefa de exploração, busca e recuperação, por conteúdos, dos documentos de teses e dissertações de uma biblioteca digital, fornecendo acesso a dados como teorias, métodos e instrumentos de coleta de dados, auxiliando o usuário a explorar os resultados das pesquisas concluídas, uma vez que esses passam a ter maior visibilidade do conteúdo disponível no ambiente, sem sobrecarregá-lo de informações.

No trabalho de Silva e Neves (2011) é apresentada a aplicação da Teoria da Análise Facetada como possibilidade de estruturação do conhecimento, através da organização de conceitos e da criação de relacionamentos em ambiente de banco de dados. O modelo desenvolvido destina-se à organização de materiais pedagógicos de uma instituição da pesquisa. Concluiu-se, neste trabalho, que é possível adotar a teoria em ambientes digitais por meio da modelagem conceitual, sendo possível apresentar uma es-

trutura de classificação facetada ilimitada, flexível e multidimensional.

No trabalho de Pontes e Lima (2012) a Teoria da Análise facetada foi utilizada para representação e organização do acervo de uma biblioteca digital de teses e dissertações, em conjunto com Taxonomias Dinâmicas, para compor o mecanismo de busca e navegação. Ponte considera que uma busca dinâmica facetada fornece suporte a um novo paradigma de acesso, permitindo a exploração guiada do acervo e fazendo uma ponte entre os processos de busca e a navegação. Além disso, o uso dessa abordagem parece ser intuitivo e facilmente utilizado pelos usuários. As taxonomias dinâmicas, projetadas de forma facetada, requerem um esquema conceitual simples e são capazes de se adaptar a relacionamentos novos e não previstos, permitindo aos usuários descobrir esses relacionamentos nas operações de navegação e recuperação.

O trabalho de Silva (2013) relata o uso das taxonomias facetadas aplicadas em ambientes digitais de recuperação de informação e web e apresenta um protótipo de uma ferramenta web de navegação facetada e indexação colaborativa pelo usuário. Na avaliação do protótipo desenvolvido, verificou-se que tanto a navegação facetada quanto a taxonomia facetada são estratégias válidas para complementar a recuperação da informação e a indexação, sendo utilizadas, em geral, em um terço das ações dos usuários nos testes realizados.

Referente aos estudos sobre a utilização de Mapas Conceituais, tem-se, no trabalho apresentado por Amoretti (2000), que a representação do conhecimento sob a forma de mapas conceituais, com conceitos organizados de forma relacional e modular, é uma maneira alternativa de estruturar a informação. A fundamentação teórica dos mapas conceituais decorre da teoria das redes semânticas, que é basicamente uma representação visual do conhecimento.

No trabalho de Lima (2004), os mapas conceituais foram utilizados como uma ferramenta de representação do conhecimento de forma gráfica, construindo redes de conhecimentos constituídas de nodos e links, nas quais os nodos representam os conceitos e os links representam as relações entre os conceitos. Nesse trabalho, os mapas conceituais representaram modelos conceituais criados a partir da Teoria da Análise Facetada. Segundo a autora, a representação concisa e gráfica do conhecimento por meio do mapa conceitual, com ampla possibilidade de relações, resulta em um ambiente ideal para criar uma estrutura navegacional,

onde os usuários possam encontrar a informação com conteúdo semântico, principalmente se esta representação for feita por um especialista na área a ser estruturada.

Berti Júnior *et al* (2011) apresenta em seu trabalho uma reflexão sobre a aplicabilidade da técnica de mapas conceituais na organização da informação e do conhecimento em ambientes digitais e conclui, por meio de experimento, que os mapas conceituais, enquanto método e técnica, podem ser incorporados à Ciência da Informação, sendo necessária a realização de pesquisas que demonstrem a aplicação dos mapas conceituais para esse propósito.

A partir desta revisão de literatura percebe-se que essas teorias já vêm sendo aplicadas para melhorar a organização e a recuperação da informação, o que justifica a sua investigação como solução para o domínio deste trabalho.

## 5 OntoAgroHidro: uma ontologia de domínio

A *OntoAgroHidro* é uma ontologia de domínio criada a partir do interesse de um grupo de pesquisadores especialistas na área de recursos hídricos, mudanças climáticas e de uso da terra dentro da Embrapa. O estudo referente à necessidade de se conciliar o desenvolvimento da produção agrícola com o uso sustentável da água e das terras pode ser considerado de relevância mundial, e a necessidade de se ter uma solução para compartilhar e integrar informação entre as diversas instituições e pesquisadores desencadeou os esforços para criação de uma ontologia de domínio.

O processo de construção da *OntoAgroHidro* é relatado por Bonacin *et al* (2013), onde se apresentam as estratégias de reuso e integração de outras representações existentes para refinar os conceitos da ontologia. O resultado gerado deste processo foi a ontologia descrita em OWL 2.0, usando a versão 4 da ferramenta Protégé, com cerca de 8500 conceitos (incluindo classes e instâncias) do domínio.

Os principais conceitos apresentados na *OntoAgroHidro* são mostrados na Figura 2, gerada pela ferramenta *OntoGraf* no Protégé. Esta Figura contém as sete classes principais destacadas, a partir das quais se têm as descrições de outras subclasses, relações, propriedades e instâncias na *OntoAgroHidro*.

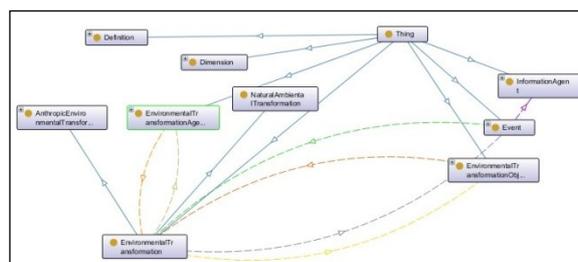


Figura 2. Representação das principais classes da *OntoAgroHidro* na ferramenta Protégé

A classe *Environmental Transformation* (mostrada na parte inferior da Figura 2) é a classe principal da ontologia e representa os fenômenos referentes às mudanças (físicas, químicas, biológicas, sociais ou econômicas) no meio-ambiente. Conforme apresentado na Figura 2 esta classe possui uma relação com as classes *Event*, *Environmental Transformation Object* e *Environmental Transformation Agent*. A classe *Event* representa os eventos que causam mudanças ambientais. A classe *Environmental Transformation Object* representa os objetos que são afetados pelas transformações. A classe *Environmental Transformation Agent* representa os agentes das transformações ambientais. A classe *Dimension* descreve os vários aspectos de um evento. A classe *Definition* representa os conceitos necessários para definir os elementos-chave do domínio e de áreas afins. A classe *Information Agent* é a última classe do primeiro nível hierárquico da *OntoAgroHidro* e representa os agentes que produzem ou contêm informações sobre as transformações ambientais.

No trabalho de Bonacin *et al* (2015) avalia-se os aspectos de interoperabilidade semântica e de recuperação de informação na *OntoAgroHidro*, mostrando, através de um estudo, as potencialidades e limitações desta ontologia, bem como sugestões de trabalhos futuros para uso da ontologia.

## 6 Resultados

Seguindo as etapas que foram descritas na metodologia, apresenta-se a seguir os resultados alcançados com o desenvolvimento deste trabalho.

Como resultado dos estudos teóricos, apresenta-se a Tabela 2, baseada em Silva (2013) e Maculan (2011), que mostra relação entre as propostas de categorização de Ranganathan e CRG com as representações em Ontologias e Mapas Conceituais.

| Ranganathan | CRG | Ontologias | Mapas |
|-------------|-----|------------|-------|
|-------------|-----|------------|-------|

|               |   |                           | Conceituais             |
|---------------|---|---------------------------|-------------------------|
| Personalidade | Tipos de produto final  | Classe e subclasse        | Conceito                |
| Matéria       | Partes<br>Materiais<br>Propriedades<br>Formas de Apresentação | Todo e Parte<br>Atributos | Hierarquia<br>Relações  |
| Energia       | Processos<br>Operações<br>Agentes                             | Relações                  | Proposições/inferências |
| Espaço        | Espaço  | Atributo<br>Origem        | Objeto<br>Origem        |
| Tempo         | Tempo   | Atributo<br>Data          | Objeto<br>Data          |

Tabela 2. Quadro comparativo da proposta da representação de um domínio

O objetivo desta relação apresentada na Tabela 2 não é estabelecer uma equivalência, mas sim mostrar que os fundamentos podem ser usados em conjunto conforme os propósitos desejados. Trata-se de considerar que abstrações podem ser feitas para melhorar as compreensões de domínios, sem perder, no entanto, as propriedades fundamentais.

A Figura 3 apresenta a interface do protótipo da ferramenta desenvolvida por Silva (2014), utilizada para representação da ontologia *OntoAgroHidro*, em cujo ambiente é possível criar catálogos que são acessados através de uma interface web, onde se pode realizar buscas por palavras-chave e navegar por uma estrutura facetada. Esta solução possui ainda funções de colaboração para que outros usuários possam manter a criação e a indexação colaborativa dos conteúdos.

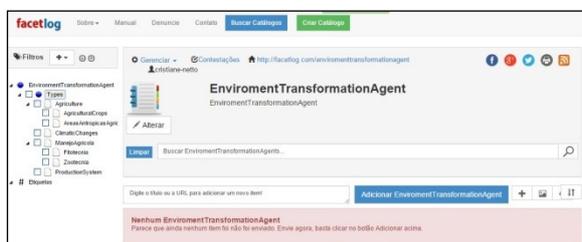


Figura 3. Interface da ferramenta Facetlog

O resultado da modelagem conceitual da ontologia pode ser sintetizado conforme a Figura 4 onde se tem a estrutura de mapeamento da *OntoAgroHidro* para o modelo de catálogos da *Facetlog*.

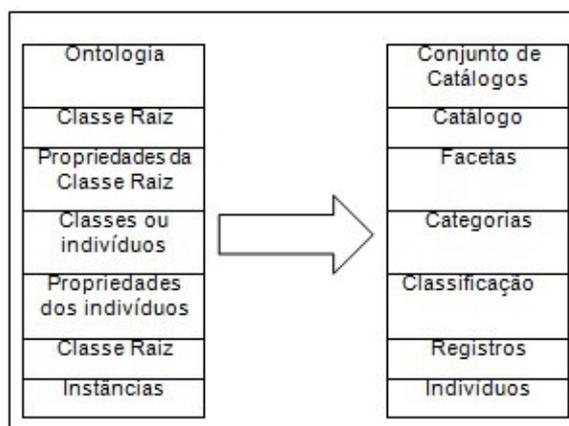


Figura 4. Mapeamento da Ontologia para a estrutura de Catálogos

Este mapeamento permitiu criar manualmente a representação de uma amostra da ontologia na ferramenta. A Figura 5 apresenta a interface gerada com o catálogo *Environment Transformation*. A classe escolhida da *OntoAgroHidro* compreendeu os primeiros níveis das classes *Environment Transformation* e suas relações com as classes *Environment Transformation Agent*, *Environment Transformation Object* e *Event*.

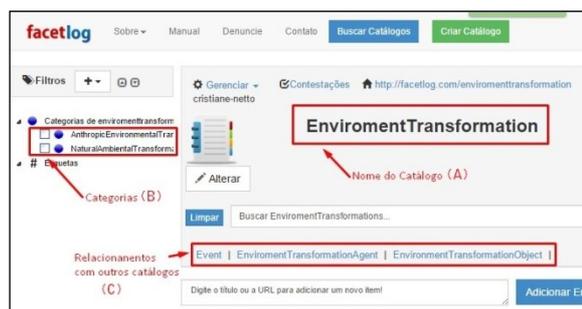


Figura 5. Interface de um catálogo na ferramenta Facetlog

Os destaques nessa interface referem-se à identificação do catálogo *Environment Transformation* (A); à identificação das categorias (B) e aos relacionamentos (C) deste catálogo com outros que também foram criados na ferramenta.

Em uma versão de testes da ferramenta Facetlog avaliou-se a capacidade de representação dos catálogos em mapas em representação de quatro tipos: *Hypertree*, *Spacetree*, *Treemap* e *RGraph*. Esta opção de navegação é mostrada na Figura 6, onde se tem um *menu* para escolha da navegação no catálogo por uma interface visual.

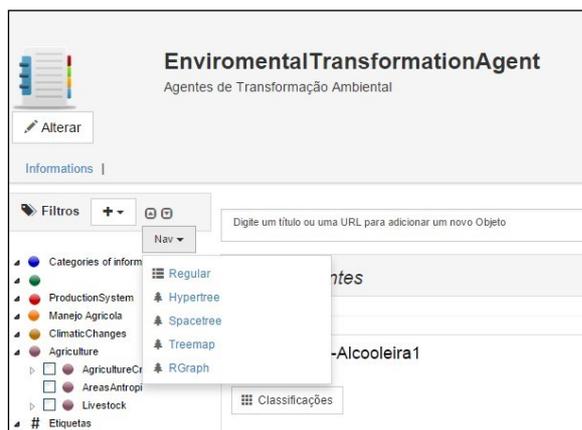


Figura 6. Opções de navegação pelo catálogo

A figura 7 mostra o catálogo *Enviromental Transformation Agent*, representado em árvores hiperbólicas (*hypertree*). Este tipo de mapa conceitual oferece uma interface, onde os componentes diminuem e aumentam de tamanho em função da sua distância ao centro, assim como oferece ao usuário uma visão global da informação e uma navegação por foco e contexto.

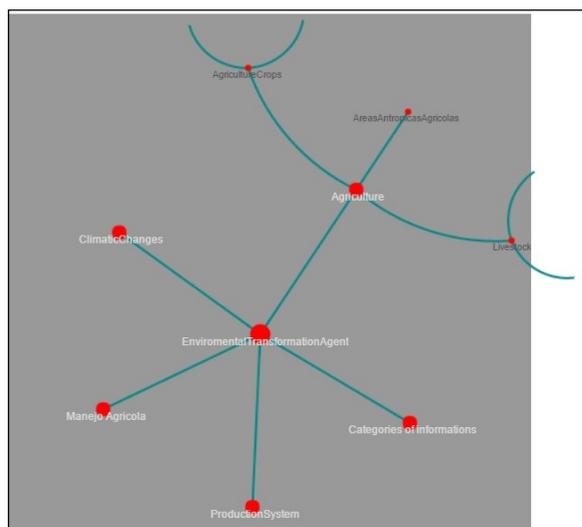


Figura 7. Interface de visualização do catálogo em árvores hiperbólicas (*Hypertree*)

Na figura 8 exibe-se o catálogo *Enviromental Transformation Agent* representado em árvores hierárquicas (*spacetree*). Nesta interface os usuários podem ter uma visão analítica das informações do catálogo e das relações entre os conteúdos.

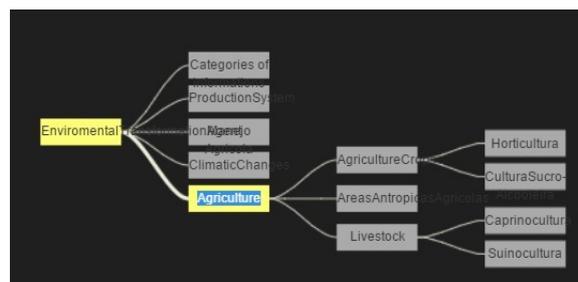


Figura 8. Interface de visualização do catálogo em árvore (*Spacetree*)

Uma representação em retângulos aninhados (*treemap*) é mostrada na Figura 9. Nesta interface as informações hierárquicas são apresentadas com cores e tamanhos variados, podendo auxiliar os usuários na compreensão da organização da informação.

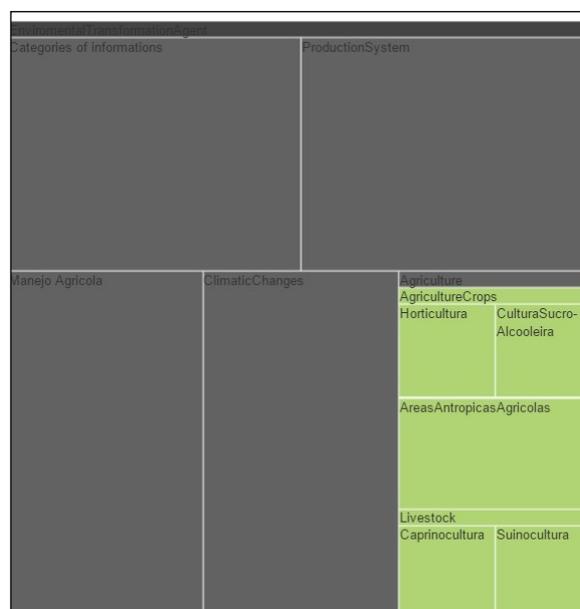


Figura 9. Interface de visualização do catálogo em retângulos (*Treemap*)

Na figura 10 a representação do catálogo é mostrada em grafo, que mostra as categorias de um catálogo, bem como as suas subcategorias.

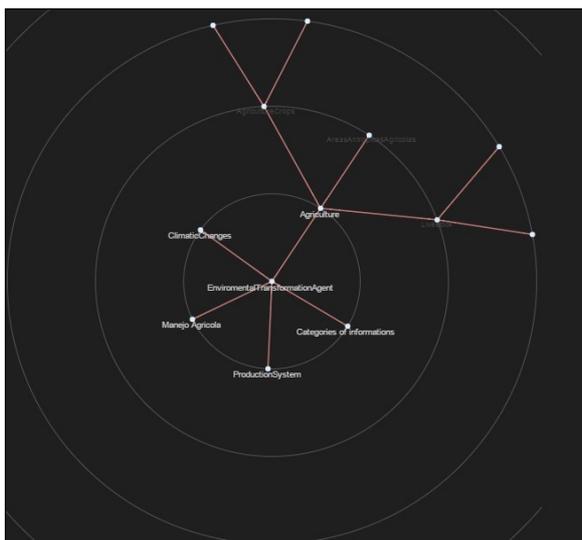


Figura 10. Interface de visualização do catálogo em Grafo (RGraph)

Os resultados dos testes realizados com a ferramenta *Facetlog* demonstraram o potencial dos seus recursos para serem desenvolvidos como uma solução com usuários que desejam navegar e realizar buscas em ontologias. Além disso, os recursos colaborativos podem ser utilizados para construção colaborativa e instanciação de ontologias.

A partir desses primeiros testes realizados visualiza-se a possibilidade de implementar um protótipo baseado no *Facetlog* para usuários do domínio agropecuário brasileiro. Acredita-se que, a partir do conhecimento representado na *OntoAgroHidro*, novos recursos possam ser implementados para se obter uma busca e navegação semântica pela informação. Além disso, considera-se relevante a pesquisa da interface com usuários, avaliando-se os recursos e também os aspectos de auxílio à cognição.

## 7 Considerações Finais

Este trabalho apresentou um estudo sobre a Teoria da Análise Facetada e Mapas Conceituais para representação e recuperação de informação em contextos digitais. Como estudo de caso foi analisada uma ontologia de domínio, chamada *OntoAgroHidro*, criada pela Embrapa. O problema identificado é a necessidade de uma interface para navegação e recuperação de informação que ofereça aos seus usuários abstrações sobre as propriedades lógicas da ontologia e que favoreça a cognição humana sobre o domínio.

Como experimento inicial avaliou-se o protótipo da ferramenta desenvolvida por Silva (2013), onde uma amostra da ontologia foi mapeada

manualmente para a estrutura de catálogos da ferramenta *Facetlog*. A partir dos resultados obtidos, acredita-se que seja possível implementar uma solução computacional que possa facilitar a criação do modelo conceitual para navegação na classificação facetada e mapas conceituais em todo o domínio representado pela *OntoAgroHidro*, permitindo, assim, facilitar o aprendizado humano sobre o domínio, bem como o compartilhamento de informação.

Como trabalhos futuros pretende-se realizar a implementação de uma solução que possa mapear computacionalmente a estrutura OWL da ontologia para uma estrutura baseada na ferramenta *Facetlog*. Além disso, considera-se necessário um estudo com usuários para avaliar o uso da interface e o seu potencial em contribuir com a cognição humana.

## Notas

- (1) [http://www.w3.org/standards/techs/owl#w3c\\_all](http://www.w3.org/standards/techs/owl#w3c_all)
- (2) <http://protege.stanford.edu/>
- (3) <http://facetlog.com/>

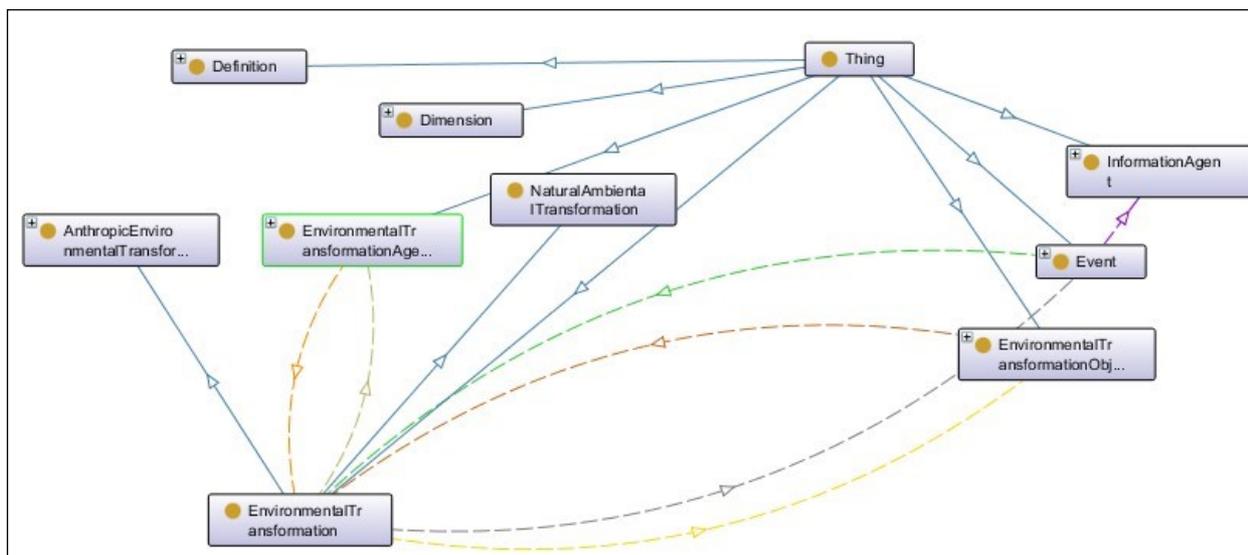
## Referências

- Almeida, M. B. (2013). Revisiting ontologies: A necessary clarification. *Journal of the American Society for Information Science and Technology*, 64(1), 1682-1693.
- Almeida, M. B., & Bax, M. P. (2003). Uma visão geral sobre ontologias: pesquisa sobre definições, tipos, aplicações, métodos de avaliação e de construção. *Ciência da Informação*, 32 (3), 7-20.
- Amoretti, M. S. M., & Tarouco, L. M. R. (2000). Mapas conceituais: modelagem colaborativa do conhecimento. *Revista Informática na Educação: Teoria & Prática*, 3(1), 67-71.
- Baeza-Yates, R., & Ribeiro-Neto, B. (1999). *Modern Information retrieval*. New York: Addison- Wesley.
- Barbosa, A. P. (1972). Classificações facetadas. *Ci. Inf. Rio de Janeiro*, 1(2), 73-81.
- Berti-Junior, D. W., Andrade, I. A., & Cervante, B. M. N. (2011). Mapas Conceituais uma ferramenta tecnológica aplicada a organização da informação e do conhecimento. *XXIV Congresso Brasileiro de Biblioteconomia, Documentação e Ciência da Informação*.
- Bonacin, R., Nabuco, O. F., & Junior, I. P. (2013) Conceptualizing the Impacts of Agriculture on Water Resources: experiences and ontology engineering challenges. *Fifth International Conference on Management of Emergent Digital EcoSystems*. ACM Press, 262-269.
- Bonacin, R., Nabuco, O. F., & Junior, I. P. (2015). Ontology models of the impacts of agriculture and climate changes on water resources: Scenarios on interoperability and information recovery. *Future Generation Computer Systems*.
- Broughton, V. (2006). The need for a faceted classification as the basis of all methods of information retrieval. *Aslib Proceedings*, 58 (1/2), 49-72.
- Campos, M. L. A. (2001). *A organização de unidades do conhecimento em hiperdocumentos: o modelo conceitual como um espaço comunicacional para realização da*

- autoria. 198 p. Tese (Doutorado em Ciência da Informação), Universidade Federal do Rio de Janeiro, Rio de Janeiro.
- Campos, M. L. A. (2010). O papel das definições na pesquisa em ontologia. *Perspectivas em Ciência da Informação*, 15(1), 10-20.
- Costa, A. (2012). *Taxonomia para ferramentas de autoria: apoio na elaboração de material didático*. Dissertação (Mestrado em Computação Aplicada). Universidade Vale do Itajaí, Vale do Itajaí.
- Duarte E. A., & Cerqueira, R. F. P. de. (2007). Análise facetada: um olhar face a modelagem conceitual. *Revista Digital de Biblioteconomia e Ciência da Informação*, 4(2), 39-52.
- Gil, A. C. (1991). *Como elaborar projetos de pesquisa*. (3. ed.). São Paulo: Atlas.
- Guarino, N. (1998). Formal Ontology and Information Systems. *Formal Ontologies in Information Systems*, N. Guarino (Ed.), IOS Press, pp. 3 -15.
- Guizzardi, G., FALBO, R., & Guizzardi, R. S. S. (2008) A importância de Ontologias de Fundamentação para a Engenharia de Ontologias de Domínio: o caso do domínio de Processos de Software. *Revista IEEE América Latina*, 6(3); 244-251.
- Guizzardi, G. (2000). *Desenvolvimento para e com reuso: Um estudo de caso no domínio de vídeo sob demanda*. Dissertação (Mestrado em Informática). Universidade Federal do Espírito Santo, Vitória.
- Hearst, M. A. (2008). Uis for faceted navigation: Recent advances and remaining open problems. *Proceedings of the Second Workshop on Human-Computer Interaction and Information Retrieval*, 13-17.
- Lima, G. A. B. O. (2004). *Mapa hipertextual (MHTX) um modelo para a organização hipertextual de documentos*. 2004. 199f. Tese (Doutorado em Ciência da Informação). Universidade Federal de Minas Gerais, Belo Horizonte.
- Macêdo, L. N. (2007). Avaliação de um Objeto de Aprendizagem com Base nas Teorias Cognitivas. *XIII Workshop em Informática Educativa*. Rio de Janeiro, RJ,
- Maculan, B. C. M. dos S. (2011). *Taxonomia facetada navegacional: construção a partir de uma matriz categorial para trabalhos acadêmicos*. Dissertação (Mestrado em Ciência da Informação). Universidade Federal de Minas Gerais, Belo Horizonte.
- Marconi, M. A.; Lakatos, E. M. (2003). *Fundamentos de metodologia científica*. (7. ed.). São Paulo: Atlas.
- Moreira, M.A. (2010). *Mapas conceituais e aprendizagem significativa*. São Paulo: Centauro Editora.
- Novak, J. D.; & Cañas, A. J. (2008). The theory underlying concept maps and how to construct and use them. *Florida Institute for Human and Machine Cognition Pensacola*.
- Pontes, F. V., & Lima, G. A. B. (2012). Organização do conhecimento em bibliotecas digitais de teses e dissertações: uma abordagem baseada na classificação facetada e taxonomias dinâmicas. *Encontro Nacional de Pesquisa em Ciência da Informação*, Rio de Janeiro.
- Sales, L. F.; Campos, M. L. A.; Gomes, H. E. (2008). Ontologias de domínio: um estudo das relações conceituais. *Perspectivas em Ciência da Informação*, 13, 62-76.
- Silva, M. F. (2013). *Proposta de Modelo de Colaboração para Catálogo Web Facetado*. Tese (Doutorado em Ciência da Informação), Universidade Federal de Minas Gerais, Belo Horizonte.
- Silva, M. B., & Neves, D. A. B. (2008). A aplicação da teoria da classificação facetada em banco de dados, através da modelagem conceitual. XII ENANCIB, *Encontro Nacional de Pesquisa em Ciência da Informação*, Brasília.
- Souza, B. P. G. (2006). *O uso de mapas conceituais como ferramenta no planejamento de aulas*. Monografia (Curso de Licenciatura em Química). Universidade Federal de Minas Gerais, Belo Horizonte.
- Tavares, R. (2007). Construindo mapas conceituais. *Ciência e Cognição*. 12(1). 72-85.
- Tristão, A. M. D., Fachin, G. R. B., & Alarcon, O. E. (2004) Sistema de classificação facetada e tesauros: instrumentos para organização do conhecimento. *Ciência da Informação*, 33(2), 161-171.

## Anexo 1

Representação das principais classes da OntoAgroHidro na ferramenta Protégé



## Anexo 2

Interface de um catálogo na ferramenta Facetlog

---

# Abordagens cognitivas da ciência da informação no Brasil: mapeamento conceitual

*Cognitive approaches of information science in Brazil: conceptual mapping*

---

Dulce Amélia de Brito Neves (1), Mariângela Spotti Lopes Fujita (2)

(1) Universidade Federal da Paraíba (UFPB), Centro de Ciências Sociais Aplicadas, Departamento de Ciência da Informação, Cidade Universitária, CEP: 58051-900, João Pessoa, Paraíba, Brasil, damelia1@gmail.com.br

(2) Universidade Estadual Paulista "Júlio de Mesquita Filho (UNESP), Campus de Marília, Av. Hygino Muzzi Filho, 737, Bairro Mirante, CEP: 17.525-000 - Marília, São Paulo, Brasil, fujita@marilia.unesp.br

## Resumen

A pesquisa teve como objetivo geral: Analisar na produção científica no Brasil relacionada a Cognição e a Ciência da Informação no período de 2003 a 2012. Elaborou-se um levantamento nos periódicos de Ciência da Informação, com acesso online buscando-se a produção científica sobre Cognição e Ciência da Informação no Brasil em periódicos das áreas entre os anos de 2003 e 2012, tendo em vista que os números de alguns periódicos estão com sua publicação atrasados. Os autores foram identificados bem como a concentração da produção científica: Ciências Cognitivas e CI, sociocognição de catalogadores, leitura de indexadores, metacognição na CI; mapear na literatura nacional. Os resultados mostram que houve uma concentração nos seguintes periódicos: Encontros Bibli publicou a última vez em 2010; Datagrama Zero, em 2009; Ciência da Informação em 2008; e Perspectivas em Ciência da Informação em 2011. O tema cognição-ciência da informação, temos um total de 27 artigos em 10 anos, escritos por 27 autores. Existe um pequeno grupo de pesquisadores atuando na área e 55,5% do total de publicações, esta foi a única publicação na temática e datam, em sua maioria de 2007, 2008 e 2009, com apenas uma publicação de 2011, o que demonstra que não houve uma apropriação do tema em sua produção científica como eixo central para esses autores.

**Palavras-chave:** Cognição, Ciência da Informação; Mapas conceituais.

## 1. Introdução

A cognição tem sido alvo de atenção de estudiosos, pensadores e filósofos desde a antiguidade clássica. Conceituada como o ato ou processo de conhecer, envolvendo percepção, memória, raciocínio, entre outras atribuições mentais, o termo cognição teve origem em escritos de Platão e Aristóteles.

Nas últimas décadas ocorreu um incremento nos estudos sobre a cognição humana, princi-

## Abstract

This research purpose identifies the scientific communication in Brazil related to Cognition and Information Science theoretical- methodological options adopted by the authors. The purpose of studies of cognitive related, on the long search for understanding the human mind and the processes related to understanding, apprehension of knowledge. The theme has approach in Information Science holds, but there are no studies that indicate common points between the two disciplines and how they interrelate. This project was conceived by observing the incidence of research addressing cognition and its relation to Information Science. The first step is select publications in databases covering scientific literature on Cognition and Information Science in Brazil in journals areas between the years 2002 and 2012, to identify authors and concentration of scientific literature, these authors relate to the theories discussed by them and methodologies used in their studies. As a technique for collecting data using content analysis on these studies in order to map the theories and methodologies and collate the data obtained in studies looking for convergence. The empirical analysis of the categories will be based on conceptual mapping and bibliometric studies.

**Keywords:** Cognition, Information Science; Conceptual Maps.

palmente após o surgimento do computador e da modelagem computacional. A partir dessa interrelação, as pesquisas desenvolvidas sobre a cognição humana têm buscado apreender o modo como as pessoas, pensam, interpretam e percebem o mundo (Vygotsky 1998, STERNBERG 2000, NEVES 2006).

A ascensão de estudos de cunho cognitivos está relacionada, por um lado, a longa busca pela compreensão da mente humana e os processos ligados a compreensão, apreensão de conhecimento. Por outro lado, o fortalecimento

dos estudos ligados a sistemas inteligentes na TI e a robótica em particular.

A temática já ocupa destaque na CI, mas ainda não existem estudos que indiquem pontos comuns entre as duas disciplinas e como estas se interrelacionam. A presente pesquisa foi idealizada a partir da observação da incidência de pesquisas que abordam cognição e sua relação com a CI.

Apesar de pesquisas sobre Cognição e CI estarem em evidência por pesquisadores em nível internacional, observamos maior incidência de estudos no Brasil no período compreendido entre 2003 e 2013. E nos questionamos: Há evolução na produção científica brasileira associada a Cognição e Ciência da Informação? Quais as abordagens teórico-metodológicas usadas pelos autores em seus estudos?

A pesquisa teve como objetivo geral: Analisar na produção científica no Brasil relacionada a Cognição e a Ciência da Informação no período de 2003 a 2012, tendo em vista que os números de alguns periódicos estão com sua publicação atrasados e só poderão ser atualizados em anos posteriores.

Nesse sentido foram desenvolvidos os seguintes objetivos específicos: identificar a comunicação científica de pesquisadores brasileiros que trabalham com Cognição e CI no período de 2003 a 2013; verificar na comunicação científica literatura nacional as seguintes palavras-chave: Ciências Cognitivas e CI, sociocognição de catalogadores, leitura de indexadores, metacognição na CI; mapear na literatura nacional as seguintes variáveis: periódicos científicos em Ciência da Informação no Brasil na temática Cognição - Ciência da Informação; concentração da produção em Cognição-Ciência da Informação em Periódicos científicos brasileiros em Ciência da Informação; autores publicando sobre o tema cognição-ciência da informação; quantidade de artigos publicados por autor e último ano de publicação na temática.

## 2. Cognição e ciência da informação

A Ciência Cognitiva é uma área de estudos interdisciplinares, que se interrelaciona com a Psicologia Cognitiva, Ciência da Computação, Sistemas de Informação, Inteligência Artificial, Neurociências, Linguística e CI, entre outras. A partir dessa interrelação, as pesquisas desenvolvidas sobre a cognição humana têm buscado apreender o modo como as pessoas, pensam, interpretam e percebem o mundo.

Nas últimas décadas ocorreu um incremento nos estudos sobre a cognição humana, princi-

palmente após o surgimento do computador e da modelagem computacional. Com o aumento da produção científica, os estudos bibliométricos se tornaram essenciais, oferecendo informações relevantes sobre qualquer domínio, ao identificar, evidenciar e visualizar o conhecimento científico construído de uma temática, disciplina ou área do conhecimento.

Tendo em vista sua origem interdisciplinar, a CI teve sua emergência ligada às Ciências Sociais Aplicadas e a tecnologia da informação. Sua principal função, como afirma Neves (2006 p. 40),

é produzir conhecimentos que contribuam para a solução de problemas relacionados à organização de sistemas de informação especializados na incorporação, sistematização, disseminação e recuperação da informação.

Saracevic (1996, p. 47), afirma que a

Ciência da Informação é um campo dedicado às questões científicas e as práticas profissionais voltadas para os problemas da comunicação do conhecimento e de seus registros entre os seres humanos, no contexto social, institucional ou individual, do uso e das necessidades de informação.

Para Capurro (1991, p. 3-4)

Os seres humanos são processadores biológicos de informação. A informação é realidade duplamente codificada... A ciência da informação pretende estudar a informação em si mesma, ou seja, contribuir para sua análise e sua construção.

Em nosso ponto de vista, existe uma convergência entre os conceitos de Saracevic (1996) e o de Capurro (1991). Enquanto Saracevic (1996) enfatiza a CI como um campo do conhecimento que lida com a tríade pertinente ao seu âmbito produzindo conteúdo referente a organização, representação e disseminação da informação em um contexto social, Capurro (1991) volta-se aos seres vivos que processam informação, cabendo a CI o papel de conhecimento aglutinador e mediador dos conteúdos informacionais, viabilizando a construção de uma mediação a partir do conhecimento e aproximação com estes seres. É precisamente nesta convergência que são incrementados estudos que apontem novos modos ao buscar uma maior interação aproximando-se dos usuários da informação a CI.

O objetivo da CI é realizar a comunicação da informação desejada entre o ser humano que a necessita e aquele que a gerou. Como tal a CI é uma disciplina que lida com aspectos da cognição humana e seus processos. Existem, entretanto, diversas abordagens epistemológicas e filosóficas de tais processos. Para citar um dos mais importantes o processo pode ser do visto

do ponto de vista do pragmatismo, racionalismo, hermenêutica e fenomenologia, materialismo e ponto de vista cognitivo na mente.

Nesse caso, a função primordial da CI é articular sistemas de informação que operem no nível físico a diversidade comportamental apresentada pelos sujeitos na busca da informação. Para desempenhar esta função a Ciência da Informação dialoga basicamente com as áreas que envolvem: comunicação dos/entre os seres humanos, o conhecimento, a informação e *seus registros, as necessidades e o uso da informação, os contextos sociais, institucionais e individuais*, como enfatiza Lima (2003, p. 77).

Como ciência a CI vem buscando novos percursos teórico-metodológicos, em sua relação com a TI e nos enfoques cognitivos das pesquisas voltadas para três eixos: na compreensão das mudanças tecnológicas e a suas funções na comunicação da informação; no papel social e econômico da informação e na a efetivação do diálogo interdisciplinar.

Na literatura internacional, podemos observar que existem pontos de concentração nas pesquisas voltadas a relação entre cognição e CI gerando as seguintes palavras-chave na interrelação: estudos de usuário, cognição no desempenho dos indexadores, abordagem sociocognitiva, cognição profissional.

Na Psicologia Cognitiva a teoria do processamento da informação originou vários estudos voltados ao detalhamento dos processos cognitivos relacionados a solução de problemas (HUNT, 1980; NEWELL; SIMON, 1972; STERNBERG, 2000). Segundo Neves (2006, p.41)

os psicólogos do processamento da informação estudam as capacidades intelectuais humanas, analisando a maneira como as pessoas solucionam as difíceis tarefas mentais para construir modelos artificiais que têm por objetivo compreender os processos, estratégias e representações mentais utilizadas pelas pessoas no desempenho destas tarefas.

Estas novas abordagens geram em si novos termos ou adaptação de outros que, muitas vezes, necessitam de adequação linguística e revisão conceitual em língua portuguesa a fim de evitar redundância pelo mau uso de um termo ou conceito.

*A pesquisa empírica com o enfoque cognitivo na Ciência da Informação*, como enfatiza Neves (2006) coloca em evidência estudos voltados à recuperação da informação. Isto se deve a desdobramentos do modelo, denominado por Jacob e Shaw (1998) de “quase-paradigma”, do *ponto de vista* cognitivo, apresentado inicialmente por De Mey (1982, p. 4) ao afirmar que:

O ponto de vista cognitivo da ciência da informação implica que cada ato de processamento da informação, seja ele perceptivo ou simbólico, é mediado por um sistema de categorias e conceitos os quais, para o mecanismo de processamento da informação, constituem um modelo de mundo.

A temática vem ocupando destaque na CI com estudos de autores que observaram ser a cognição um caminho para compreender os processos cognitivos relacionados a interação dos usuários, dos profissionais em um sistema de informação.

Já em 1977, Brooks desenvolvia pesquisas relacionadas ao ponto de vista cognitivo na busca por um modelo mental do usuário em sua relação com os sistemas de informação, com foco na Representação da Informação (RI).

Lima (2003, p. 81) sintetiza, com clareza, a repercussão do ponto de vista cognitivo na CI:

A compreensão de que o processamento da informação desempenhado pelo sistema simula o processamento mental que o indivíduo faz para entender o mundo é compartilhada por vários autores... que tem sido corroborado por diversos pesquisadores cognitivistas e utilizado como base para suas teorizações (Allen, 1991; Ingwersen, 1982; Ingwersen, 1996; Jacob & Shaw, 1998; Ellis, 1992; Daniels, 1986), cujo ponto comum tem sido os aspectos cognitivos em relação ao usuário e sua interação com o sistema, principalmente na RI.

Observamos que a pesquisa com enfoque cognitivo na CI mostra uma tendência a estudos relativos a RI como observado pelos autores citados. Hjørland (2013) publicou uma revisão de seis abordagens diferentes na interação entre sistemas comerciais de informação e seus usuários, enfatizando o marketing que objetiva aquisição de mais usuários.

Além disto, os pesquisadores da CI procuram encontrar maior aporte no entendimento das funções cognitivas não apenas dos indivíduos em interação com os sistemas de recuperação da informação, mas buscam enfatizar os profissionais da informação, como indexadores, catalogadores, bibliotecários de referencia.

Observamos na literatura nacional da área convergência de pesquisas que geram as seguintes palavras-chave: Ciências Cognitivas e CI, sociocognição de catalogadores, leitura de indexadores, metacognição na CI e todos os pontos de interrelação voltados àqueles que desempenham suas funções nos sistemas de informação em constante relação com os usuários, como também, estudos de usuários.

Tendo em vista a evidencia pesquisas relacionadas a cognição e CI internacionalmente, sentimos a necessidade de conhecer mais efetiva-

mente estas pesquisas no âmbito da América Latina a fim de identificar a ocorrência de focos de investigação. Além disso, tentaremos identificar teorias e metodologias abordadas nestas pesquisas buscando uma relação particular entre cognição e CI.

A metodologia traçada para a presente pesquisa visou atender as questões relacionadas aos objetivos norteadores da pesquisa. Nesse sentido, descrevemos cada etapa e a ação que será desenvolvida.

Em um primeiro momento pretendemos apenas abranger a comunicação científica dos periódicos das áreas da cognição e da CI publicados na América Latina.

A pesquisa foi elaborada a partir de um levantamento nos periódicos de Ciência da Informação, com acesso online buscamos a produção científica sobre Cognição e Ciência da Informação no Brasil em periódicos das áreas entre os anos de 2003 e 2012, tendo em vista que os números de alguns periódicos estão com sua publicação atrasados e só poderão ser atualizados em anos posteriores.

### 3. Mapeamento conceitual

O mapa conceitual é uma técnica de análise que procura representar graficamente o conhecimento ou identificar relações entre conceitos através de um esquema. Como representações gráficas indicam as relações existentes entre conceitos, que são conectados através de palavras-chave.

Estes mapas foram propostos por David Ausubel como instrumento da aprendizagem significativa a ferramenta didática *Mapa Conceitual* pode servir para tornar mais significativa a aprendizagem. Também é empregado para compreender os processos mentais de uma pessoa durante a resolução de um problema, busca-se identificar o comportamento cognitivo dos indivíduos representando.

Em nossa pesquisa o mapeamento será usado como técnica para a identificação e análise dos dados coletados através da identificação das palavras-chave identificadas nas comunicações científicas, como também serão elaborados mapas dos autores e dos periódicos. Serão consideradas as seguintes variáveis na construção dos mapas: periódicos científicos em Ciência da Informação no Brasil na temática Cognição - Ciência da Informação; concentração da produção em Cognição-Ciência da Informação em Periódicos científicos brasileiros em Ciência da Informação; autores publicando sobre o tema cognição-ciência da informação; quantidade de

artigos publicados por autor e último ano de publicação na temática.

A escolha dos anos 2003 a 2012 deveu-se a uma observação anterior quando percebemos a crescente publicação de artigos com o tema Cognição e CI, tornando-se evidente a necessidade de um estudo mais aprofundado.

### 4. Análise dos dados da pesquisa: mapeamento

Os periódicos identificados com acesso online em Ciência da Informação no Brasil no período de 2003 - 2012 foram: BrazilianJournalofInformation Science, DataGramaZero, Em questão, Encontros Bibli, Informação & Informação, Informação & Sociedade: estudos, Perspectivas em Ciência da Informação, Revista ACB, Revista Brasileira de Biblioteconomia Documentação, Revista **Ciência da Informação**, Revista Eletrônica Informação e Cognição, Revista Iberoamericana de Ciência da Informação e Transinformação.

Nestes periódicos foram efetuadas buscas pelas seguintes palavras-chave: Ciências Cognitivas e CI, sociocognição de catalogadores, leitura de indexadores, metacognição na CI. Além disso, foram elaborados mapas conceituais: Periódicos científicos em Ciência da Informação no Brasil na temática Cognição - Ciência da Informação; Concentração da produção em Cognição-Ciência da Informação em Periódicos científicos brasileiros em Ciência da Informação; Autores publicando sobre o tema cognição-ciência da informação; Quantidade de artigos publicados por autor e último ano de publicação na temática cognição em periódicos brasileiros em Ciência da Informação de 2003 a 2012, que passaremos a analisar na Figura 1 (Anexo 1).

No mapa apresentado na Figura 1 observamos que as revistas onde se concentram as publicações já não publicam no tema há pelo menos dois anos: Encontros Bibli publicou a última vez em 2010; Datagrama Zero, em 2009; Ciência da Informação em 2008; e Perspectivas em Ciência da Informação em 2011.

A produção se concentra em periódicos de Qualis B1 (65%), mas detém 21% de sua produção em revista de Qualis A1, uma revista especificamente, a Perspectivas em Ciência da Informação, se destaca entre os periódicos com Qualis A1 em termos de publicação no assunto, 14% da produção, como vemos na Figura 2 (Anexo 2).

Com relação aos autores e a produção no tema cognição-ciência da informação, temos um total de 27 artigos em 10 anos, escritos por 27 auto-

res. O que mostra um pequeno grupo de pesquisadores atuando na área (Figura 3, Anexo 3).

Observa-se que a Professora Mônica Nassif possui uma produção semelhante à da Professora Dulce Neves, sendo que a publicação da Professora na temática em questão parou em periódicos nacionais desde 2008. Segue quantitativamente a produção das Professoras Mariângela Fujita, com última publicação em 2010 (no período dessa pesquisa); a professora Aida Varela, que publicou em 2010; e Ludmila Venâncio, com última publicação em 2008. Além disso, em 55,5% do total de publicações, esta foi a única publicação na temática e datam, em sua maioria de 2007, 2008 e 2009, com apenas uma publicação de 2011, o que demonstra que não houve uma apropriação do tema em sua produção científica como eixo central para esses autores, pelo menos não na publicação em periódicos científicos nacionais. Quanto aos periódicos, ressalta-se, entretanto, que as publicações têm diminuído no tema em periódicos nacionais, mesmo nos principais publicadores. No entanto, é importante observar o atraso na publicação dos periódicos, pois às vezes os números saem com o intervalo muito longo, de um ano ou mais.

Nota: o ano da última publicação no tema do autor corresponde a cada grupo de cores, identificado acima da primeira coluna, os números anteriores aos nomes tratam-se de números de ordem, sem interferência no resultado do mapa.

## 5. Considerações finais

Acreditamos que o objetivo que norteou nossa pesquisa “analisar na produção científica no Brasil relacionada a Cognição e a Ciência da Informação no período de 2003 a 2012” foi totalmente atingido. Nos limitamos a este período tendo em vista que os números de alguns periódicos estão com sua publicação atrasados e só poderão ser atualizados em anos posteriores.

Além disso, identificamos a comunicação científica de pesquisadores brasileiros em periódicos nacionais que trabalham com Cognição e CI no período de 2003 a 2012, verificamos na comunicação científica literatura nacional as seguintes palavras-chave: Ciências Cognitivas e CI, sociocognição de catalogadores, leitura de indexadores, metacognição na CI.

O mapeamento da literatura nacional foi efetuado tendo em vista as seguintes variáveis de identificação dos periódicos científicos em Ciência da Informação no Brasil na temática Cognição - Ciência da Informação, a concentração da produção em Cognição-Ciência da Informação em Periódicos científicos brasileiros em Ciência

da Informação, os autores que publicaram sobre o tema cognição-ciência da informação, a quantidade de artigos publicados por autor e último ano de publicação na temática.

Os mapas evidenciam a crescente publicação do tema Cognição e CI foi estudado no período de 2003 a 2012 mostrou que o tema, apesar de ser recente na literatura da CI foi bem focalizado,

Apesar de termos atingido o propósito da pesquisa, estamos longe e damos por finalizada, pois restam muitas inquietações e houve o surgimento de outras.

Tendo em vista o envolvimento de questões conceituais e epistemológicas que perpassam os aspectos relacionados a cognição humana e sua interação com sistemas de organização do conhecimento, acreditamos que há uma necessidade emergencial a agregação de mapeamento que interrelacione e proporcione concretude a relação subjetiva entre cognição e CI.

Acreditamos que ao identificarmos teorias e metodologias adotadas pelos autores e elaborarmos os mapas conceituais será explicitada uma rede de conceitos possibilitando a visualização tendo em vista que os mapas conceituais estabelecem uma representação e indicam relações entre conceitos na forma de proposições.

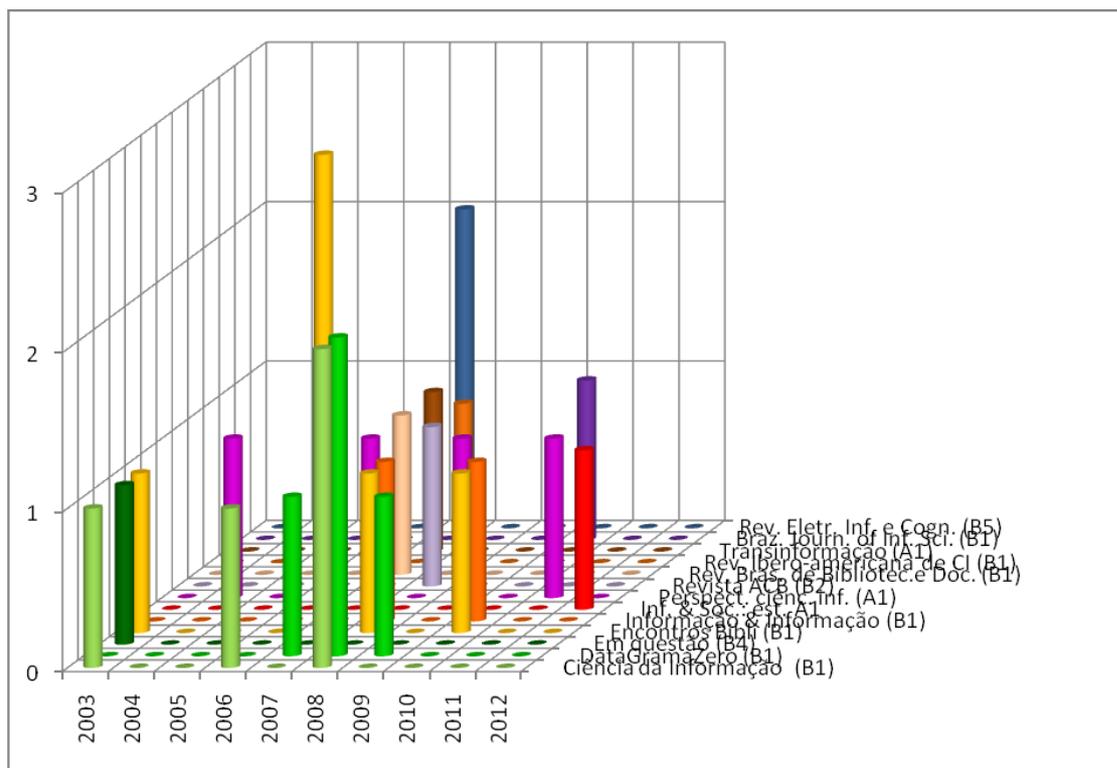
Os resultados parciais e finais serão divulgados em periódicos nacionais e internacionais, como também em congressos, encontros, dentre outros. Além disso, acreditamos que haverá desdobramentos na pesquisa que poderão abranger mapeamento da produção científica em outros âmbitos geográficos.

## Referencias

- BELKIN, N.J. (1980). Anomalous states of knowledge as a basis for information retrieval. **Canadian Journal of Information Science**, 5, p 133-143.
- \_\_\_\_\_. (1990). The cognitive viewpoint in information science. **Journal of Information Science**, v.16, p.11-15.
- \_\_\_\_\_.; Oddy, R.N.; Brooks, H.M. (1982). ASK for Information Retrieval: Part I. Background and Theory. **Journal of Documentation**, v. 38, n. 2, 61-71.
- BROOKES, B.C. (1977). The developing cognitive view in information science. En: International Workshop on the Cognitive Viewpoint, CC-77, 195-203.
- \_\_\_\_\_. (1980). The foundations of information science: Part I: Philosophical Aspects. **Journal of Information Science**, 2, 125-133.
- BUCKLAND, M. K. (1991). Information as thing. **Journal of the American Society for information Science**, v. 42. n. 5. p. 351-360.
- CAPURRO, Rafael (1991). Foundations of information science: review and perspectives. In: INTERNATIONAL CONFERENCE ON CONCEPTIONS OF LIBRARY AND INFORMATION SCIENCE, 1991, Tampere. Electronic Proceedings... Tampere: University of Tampere. Dis-

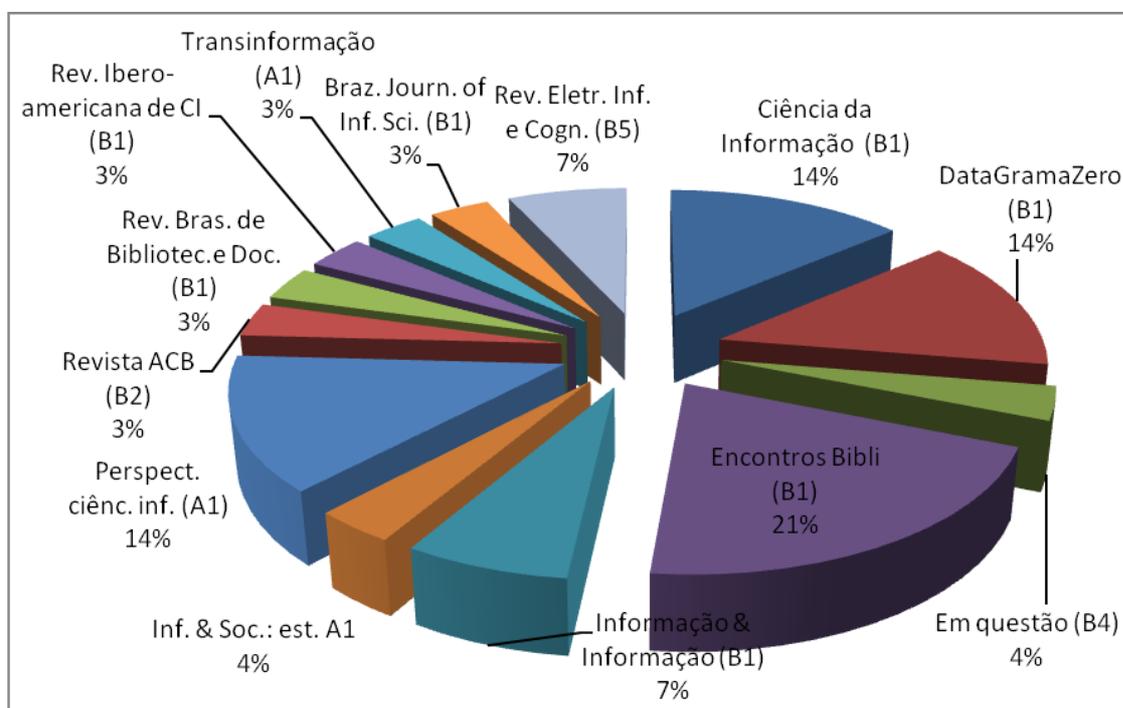
- ponível em: <<http://www.capurro.de/tampere91.htm>>. Acesso em: 2013
- \_\_\_\_\_. (2003). Epistemologia e ciência da informação. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO - ENANCIB, 5., 2003. Belo Horizonte. **Anais...** Belo Horizonte: UFMG.
- \_\_\_\_\_, HJØRLAND, B. (2003). The concept of Information. **ARIST**, v.37, Chapter 8, p. 343-411.
- DE MEY, M. (1982). The cognitive paradigm: an integrated understanding of science development. Chicago: University of Chicago.
- HJØRLAND, B. (2013) User-based and cognitive approaches to knowledge organization: a theoretical analysis of the research literature. **Knowledge organization**, v. 40, n. 1, p. 11-27.
- \_\_\_\_\_. (2000). Library and information science: practice, theory, and philosophical basis. **Information Processing and Management**, v 36, p. 501-531.
- \_\_\_\_\_. (2000a). Documents, Memory Institutions and Information Science. En: *Journal of Documentation*, v. 56, n 1, 27-41.
- \_\_\_\_\_. (2003) Epistemology and the Socio-Cognitive perspective in Information Science. *Journal of the American Society for Information Science and Technology*, v. 53, n. 4, 257-270.
- \_\_\_\_\_. (1998). Theory and Metatheory of Information Science: A New Interpretation. **Journal of Documentation**, v. 45, n. 5, 606-621.
- \_\_\_\_\_, ALBRECHTSEN, Hanne (1995). Toward a new horizon in information science: domain-analysis. **Journal of the American Society of Information Science**, v.46, n.6, p. 400 – 425.
- INGWERSEN, Peter (1992). **Information Retrieval Interaction**. London: Taylor Graham.
- \_\_\_\_\_. (1999). Cognitive Information Retrieval. En: Martha E. Williams (Ed.): *Annual Review of Information Science and Technology (ARIST)*, Medford, NJ: Information Today Inc., Vol. 34.
- HUNT, E. (1980). Intelligence as an information processing concept. *British Journal of Psychology*, v. 71, p. 449-474.
- JACOB, Elin K.; SHAW, D. (1998). Sociocognitive perspectives on representation. *Annual Review of Information Science and Technology*, v. 33, p. 131-185.
- LIMA, Gercina B. (2003). Interfaces entre a ciência da informação e a ciência cognitiva. **Ciência da Informação**. v.32, n.1, p. 77-87.
- NEWELL, A.; SIMON, H. A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall.
- NEVES, Dulce Amélia de B. (2012). Representação temática da informação e mapas cognitivos: interações possíveis. **Informação & sociedade: estudos**. v.22, p. 39-47, Número Especial.
- \_\_\_\_\_. (2011). **Metacognição, informação e conhecimento: pensando em como pensar**. Recife: Néctar. ISBN 978-85-60323-34-0
- \_\_\_\_\_. (2011). Mapas conceituais em pesquisa sobre representação da informação. **Páginas A&B: Arquivos e Bibliotecas**, Porto/Portugal. n. 2-7, p. 73 – 89. ISSN 087356700205.
- \_\_\_\_\_. (2006). Ciência da informação e cognição humana: uma abordagem do processamento da informação. **Ciência da Informação**, Brasília, v. 35, n. 1, p. 39-44, jan./abr. Disponível em: <http://www.ibict.br/cienciadainformacao/viewissue.php?id=39>
- RANGANATHAN, S. R. (1967). **The Five Laws of Library Science**. Bombay: Asia Publishing House.
- SARACEVIC, Tefko (1996). Ciência da informação: origem, evolução e relações. *Perspectivas em Ciência da Informação*, Belo Horizonte, v. 1, n. 1, p. 41-62, jan./jun.
- STERNBERG, Robert J. (2000). **Psicologia cognitiva**. Porto Alegre: Artmed.
- VAKKARI, Pertti (2003). Task-Based Information Seeking. En: B. Cronin (Ed.): **Annual Review of Information Science and Technology (ARIST)**, Medford, NJ: Information Today, Vol. 37, 413-464.
- VYGOTSKI, Lev S.; LURIA, Alexander R.; LEONTIEV, Alexis N. (1998). **Linguagem, desenvolvimento e aprendizagem**. São Paulo: Ícone.
- WERSIG, Gernot (1993). Information science: the study of postmodern knowledge usage. **Information Processing & Management**. v. 29, n. 2, p. 229-239, mar.
- \_\_\_\_\_. (1979). The Problematic situation as a basic concept of information science in the framework of social sciences: a replay to N. Belking: In: *International Federation for Documentation. Theoretical problems of information*. Moscou, FID. p. 48-57.
- WERSIG, G., NEVELING, U. (1975). The phenomena of interest to information science. **The Information Scientist**. v.9, n.4.

## Anexo 1



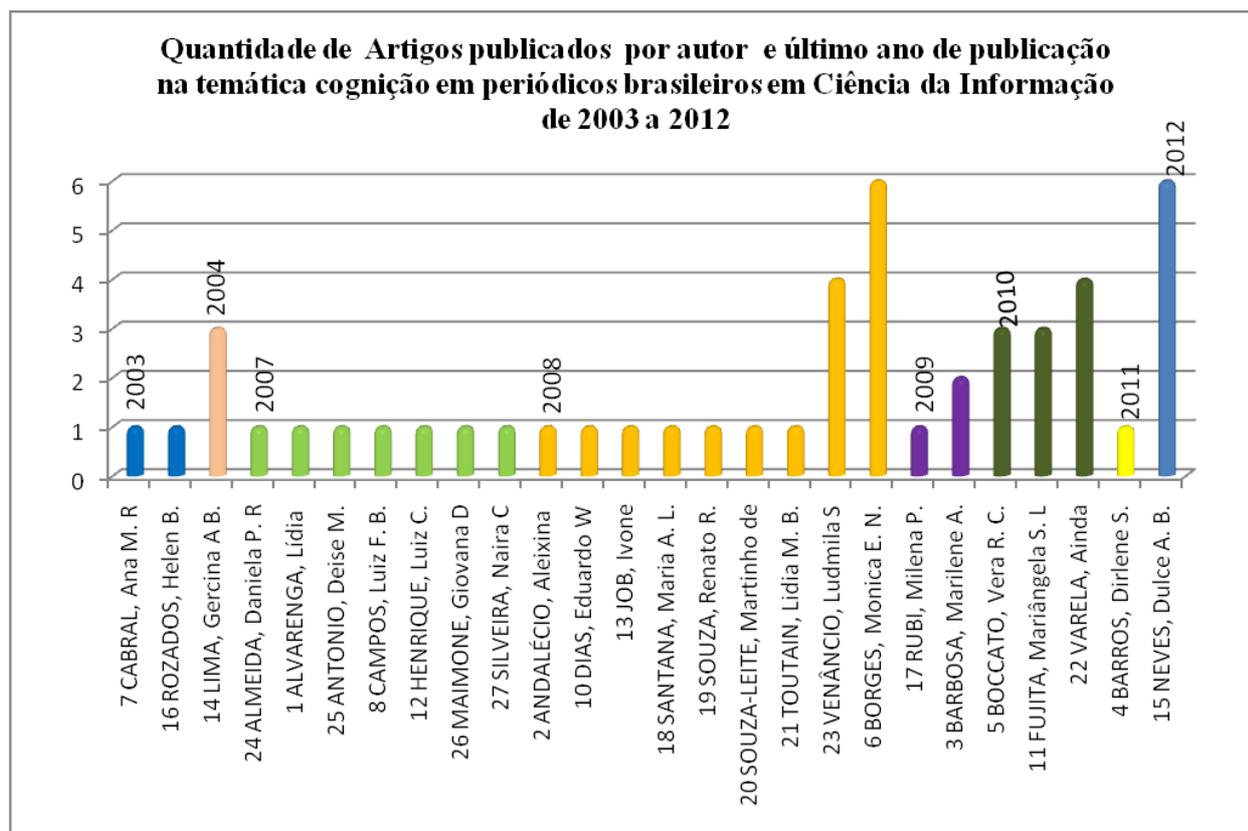
**Figura 1:** Periódicos científicos em Ciência da Informação no Brasil na temática Cognição - Ciência da Informação quantidade de artigos/ano em no período de 2003 a 2012. Fonte: Dados da Pesquisa, 2013

## Anexo 2



**Figura 2:** Concentração da produção em Cognição-Ciência da Informação em Periódicos científicos brasileiros em Ciência da Informação no período de 2003 a 2012. Fonte: Dados da Pesquisa, 2013.

## Anexo 3



**Figura 3:** Quantidade de artigos publicados por autor e último ano de publicação na temática cognição em periódicos brasileiros em Ciência da Informação de 2003 a 2012. Fonte: Dados da Pesquisa, 2013.

---

# Organização da Informação no ambiente Web: produzindo conhecimento a partir de grandes volumes de dados

*Organization and information in the Web: producing knowledge based on large volumes of data*

---

Cláudio José Silva Ribeiro

UNIRIO – Universidade Federal do Estado do Rio de Janeiro, Av. Pasteur, 458/sala 413, CEP 22290-240  
– Rio de Janeiro – RJ – Brasil – claudio.ribeiro@unirio.br

## Resumen

O tema Big Data tem despertado interesse nos profissionais que trabalham com a Gestão da Informação, pois trata de insumo essencial no processo de criação do conhecimento. O estudo deste insumo sempre esteve evidenciado no âmbito da Ciência da Informação. Este relato apresenta o tema e explora os fundamentos que auxiliam no entendimento da abordagem de Big Data. Discute a explosão informacional e a avalanche de dados, chegando aos elementos que compõem o tema. Aborda os 5 Vs do Big Data e as fases de Discovery, Data Preparation, Model Planning e Analytics, bem como as características necessárias para desempenhar o papel de Cientista de Dados. O processo de pesquisa se baseia em um estudo exploratório desenvolvido por meio de pesquisa bibliográfica com análise documental e estudo de caso. O marco empírico para análise dos elementos do domínio partiu da análise da relevância de itens publicados no site de Dados Abertos da Dataprev, onde foram extraídos os conjuntos de dados e as facetas identificadas. Apresenta os resultados preliminares sobre estudos nas fases iniciais de projetos de Big Data, de forma a viabilizar alternativas para representação e disseminação dos grandes volumes de dados presentes na Internet. Ao final, reúne alguns aspectos ligados ao perfil do profissional que está participando destes projetos.

**Palabras clave:** Big Data. Gestão da Informação. Representação da Informação. Facetas. Dados Abertos. Dataprev.

## 1. Introdução

O uso de informação e conhecimento tem impulsionado projetos de investigação, tanto em ambientes acadêmicos quanto em organizações empresariais. Na visão de Castells (1999), nós presenciamos o surgimento de uma economia “informacional e global”:

A economia é informacional porque a produtividade e a competitividade de unidades ou agentes

## Abstract

The Big Data theme has been awakening the interest of professionals who work in the field of Information Management, as it deals with the essential input in the knowledge creation process. This input has always been a focus of study in the sphere of Information Science. This report presents the theme and explores the underpinnings of Big Data in order to achieve a better understanding of this approach. It discusses the information explosion and avalanche of data to single out the elements that comprise the theme. It analyzes the 5 V's of Big Data and the phases of Discovery, Data Preparation, Model Planning and Analytics, as well as the characteristics necessary to perform the role of Data Scientist. The research process was based on an exploratory study using a bibliographical survey, in addition to documentary research and a case study. The empirical framework for analyzing the elements of the domain was developed by examining the relevance of the items published at the Dataprev's (Brazil's National Social Security Institute's Information Technology organization) Open Data website, from which the data sets and facets identified were extracted. This study presents the preliminary results of studies in the initial phases of Big Data projects aimed at developing ways of representing and disseminating the large volumes of data found on the Internet. It ends by presenting some features of the profiles of professionals who are engaged in these projects

**Keywords:** Big Data. Information Management. Information Representation. Facets. Open Data. Dataprev.

nessa economia (sejam empresas, regiões ou nações) dependem basicamente de sua capacidade de gerar, processar e aplicar de forma eficiente à informação baseada em conhecimentos. É global porque as principais atividades produtivas, o consumo e a circulação, assim como seus componentes (capital, trabalho, matéria-prima, administração, informação, tecnologia e mercados) estão organizados em escala global, diretamente ou mediante uma rede de conexões entre agentes econômicos. É informacional e global porque, sob novas condições históricas, a produtividade é gerada, e a con-

corrência é feita em uma rede global de interação (Castells, 1999, p. 87).

Por outro lado, percebe-se que o uso de dados e informação sempre esteve evidenciado como objeto de estudos para a área de Ciência da Informação. As discussões sobre esses insumos essenciais na criação do conhecimento foram impulsionadas pela revolução científica e tecnológica que se seguiu à Segunda Guerra Mundial. No bojo dessas discussões, também marcadas pela necessidade de ajustes tecnológicos (Bush, 1945), surgiram novos conceitos como o “Caos Documental” ou ainda a “Explosão Informacional” (Ribeiro, 2008). Os profissionais envolvidos com pesquisas na área de CI sempre estiveram presentes em trabalhos ligados aos aspectos informacionais, mas sofrendo forte influência também pelos apelos tecnológicos (Saracevic, 1996a).

Na visão contemporânea trazida pelo Prof. Aldo Barreto:

A chegada de uma sociedade eletrônica de informação modificou a delimitação de tempo e espaço dos conteúdos em relação aos receptores. Mas, a explosão de informação de que discorriam Vannevar Bush e Dereck de Solla Price no ambiente de pós guerra mundial foi minorada pelo computador no tempo possível. (Barreto, 2014, *online*)

Partindo de categorização sugerida por Barreto (2014), pode-se afirmar que estamos no “tempo do ciberespaço”, onde a colaboração, a instantaneidade e a fluência digital cada vez mais se fazem presentes na Ciência da Informação. Surge a noção da “avalanche de dados”, pois os dispositivos móveis, em especial os celulares e tablets, que atuam como geradores e consumidores de dados e informação estão presentes no nosso cotidiano. Esse movimento tem sido conhecido como Big Data e vem despertando o interesse por todas as pessoas que tem algum envolvimento com atividades para Gestão da Informação.

A possibilidade de fazer análise deste vasto volume de dados, vem ocasionando um incremento nos debates sobre este campo. Neste sentido, um exemplo recente sobre o uso dessa abordagem vem do esporte. Em reportagem veiculada por diferentes meios de comunicação sobre a seleção da Alemanha, vencedora da Copa do Mundo no Brasil, houve registro sobre a utilização de solução que permitisse a análise de dados sobre jogos, jogadores, táticas de jogo e a possibilidade de avaliar comentários especializados. Ou seja, o uso de dados e informação em grande quantidade como forma de melhorar o desempenho do time alemão. (InfoExame, 2014).

Desde Castells, percebe-se que a economia informacional propõe o uso de dados e informação na obtenção de resultados, logo, depreende-se que não estamos diante de uma novidade em pesquisas no campo da informação. A área de CI tem promovido estudos em temas que exploram esta temática, em especial em abordagens para a gestão de ativos de informação na WEB em níveis gerenciais, táticos e operacionais. (Ribeiro, 2008).

Dentro deste contexto, este relato apresenta alguns resultados preliminares sobre o envolvimento de profissionais da informação nas fases iniciais de projetos de Big Data, de forma a viabilizar alternativas para representação e disseminação dos grandes volumes de dados presentes na Internet.

## 2. A avalanche que incrementa o volume de dados

A maior disponibilidade de computadores, editores de texto, programas de correio eletrônico, a multiplicidade de redes sociais e de colaboração, a computação móvel, programas que fomentam a troca de mensagens, o forte uso do serviço Web e de localização geográfica, facilitaram em muito a criação ou obtenção de novos objetos portadores de informação. A tecnologia disponível na atualidade é extremamente adequada para produzir e armazenar múltiplas informações com vasto volume, contudo, nem sempre a organização destas informações permite a recuperação imediata de determinado conteúdo.

Cabe ressaltar aqui um outro aspecto: a forte relação entre dados e informação, que remonta a debates originados em várias áreas do conhecimento. A visão adotada nesta pesquisa é trazida por Davenport (1998), onde dados são simples observações sobre o estado do mundo, sendo facilmente estruturados, quantificados e transferíveis. Davenport continua e registra que a informação pode ser entendida como um conjunto de dados dotados de relevância e propósito, requerendo análise e mediação para se obter consenso em relação ao seu significado.

Neste contexto é possível afirmar que dados são necessários para a geração de informação e quanto maior for a quantidade de dados, maior será a quantidade de informação a ser tratada.

Ademais, como observado na introdução deste relato, a crescente utilização de meios de comunicação, com alto grau de mobilidade e fazendo uso da Web, contribuem para o aumento desse volume de dados. Percebe-se que o caos

documental e a explosão informacional continuam presentes, mas agora caracterizados como a “avalanche de dados”, pois os dispositivos móveis e as câmeras de vídeo incrementam cotidianamente o volume de dados produzido e consumido pela sociedade. (Heath; Bizer, 2011; Shiri, 2014).

Estes aspectos endossam a previsão feita por Alvin Tofler, quando este autor observou as mudanças no comportamento da sociedade pós-moderna, em seu livro publicado em 1984. O termo *prosumer* (1), forjado por Tofler, definia um novo perfil de interação da sociedade de consumo (Tofler, 1984).

Xie, Bagozzi e Troye (2008) e Fonseca, Gonçalves, Oliveira e Tinoco (2008) convalidam o surgimento deste novo perfil quando observam que esta sociedade passou a ser composta por indivíduos que possuem muita informação e estão constantemente buscando benefícios para seu próprio consumo. Consequentemente, é possível inferir que esta forma de interação começou a ser impulsionada a partir do incremento do uso dos dispositivos móveis que dão acesso imediato a uma miríade de informações

Em nível mundial, estima-se que a expansão das fontes de dados tenha crescimento de aproximadamente 50 vezes nos próximos 10 anos. (Emc, 2014).

Robredo (2011) também constata o crescimento e mostra que a quantidade de registros digitais passou de 0,28 ZB (zetabytes) (2) em 2007 para 1,2 ZB em 2010 e espera-se um enorme salto para 35 ZB em 2020.

Fruto deste cenário rico em volume e variedade de fontes, tem surgido uma disciplina que, apesar de não ser apenas um tema essencialmente tecnológico, vem sendo impulsionado pelos projetos de tecnologia: a vertente de Big Data.

### 3. Tratando a necessidade de informações para grandes volumes de dados

Compreender as necessidades de informação, parte do pressuposto que deverão ser mapeados os desejos e anseios dos usuários participantes de um domínio ou contexto específico. Nesta direção, procura-se estudar o que leva os usuários a formular perguntas, bem como são formuladas as respostas a estas perguntas. (HjØrland; Albrechtsen, 1995).

No entanto, tratar grandes volumes de dados torna complexa a compreensão das necessidades, uma vez que pode-se estar diante de múltiplas facetas contendo diferentes informações.

A geração e coleta de muitos dados, além dos objetos portadores de informação que possuem formatos variados, nos remete para um contexto onde existem problemas que desconhecíamos por completo. A necessidade de gerir estes altos volumes repousa na abordagem conhecida como “Big Data” (Fox; Hendler, 2011).

Para dar conta dessa abordagem será preciso explorar mais alguns aspectos: a noção de Big Data, a visão de *Analytics* e o novo perfil que está surgindo para atuar neste campo, o Cientista de Dados.

#### 3.1. Fundamentos sobre Big Data

A World Wide Web reúne fontes de diversas naturezas e, conseqüentemente, apresenta aos pesquisadores vários desafios para trabalhar com a imensa coleção de dados e informação em estoques. Barreto (2014, *online*) corrobora esta noção quando observa que “[...] com a condição online os estoques e os fluxos de informação, renomeados para ‘Big Data’, são multidirecionados [...]”. Vencer os desafios originados pela avalanche de dados, exige para o profissional que trabalha com organização de informação competências diferentes no tratamento deste insumo, pois na visão de Cavalcanti (2014, *online*) “Para entender este novo mundo precisamos de novos óculos, de um novo pensamento (precisamos evoluir do pensamento cartesiano para um pensamento mais complexo) [...]”.

Os projetos que usam a noção de Big Data precisam cobrir diferentes fontes e naturezas de dados, que podem ser entendidos como elementos ligados ao domínio da informação. Já para elementos ligados à tecnologia, estes projetos devem ser desenvolvidos fazendo uso de processamento em nuvem (3) e com tecnologias específicas, tais como processamento de rotinas em paralelo e ferramentas para otimização como *Hadoop* (4) e *MapReduce* (5), *HDFS* (6), além de abordagens de *MachineLearning* (7). (Davenport, 2014).

A noção de Big Data contempla o trabalho com alto volume de dados, com um processo de disseminação e atualização ocorrendo em grande velocidade. Além disto, tudo isto pode estar presente em múltiplas fontes e com alta complexidade para sua obtenção e análise (Dumbill, 2012).

Em artigo sobre o uso de Big Data com processamento em nuvem, Hashem et. al. (2014) complementam a noção, quando apresentam um conjunto de categorias que permite compreender as características de projetos com Big Data. A compreensão das fontes disponíveis na

Web de Dados, o originadas em projetos de colaboração que façam o uso de *social media*, além de levantar as características de formato dos dados e do processamento e/ou transformação destes, são aspectos que contribuem para o incremento de projetos com esta abordagem.

Inicialmente a noção de Big Data se apresentou apoiada em 3 características básicas, que são utilizadas para observar os diferentes aspectos envolvidos. São elas: o Volume, a Variedade e a Velocidade, caracterizando os 3 V's de Big Data. (Dumbill, 2012; Mello, Leite e Martins, 2014). No entanto, fruto do desenvolvimento de outras pesquisas conduzidas por empresas ligadas a projetos em Tecnologia da Informação, novas visões foram agregadas, sendo incorporado então um outro V para cobrir a parcela sobre a Veracidade dos dados. (Zikopoulos et al., 2013).

Power (2014) convalida a visão de Zikopoulos et. al., observando que Big Data se apoia nos 4 V's, mas, de forma análoga, surgiu uma quinta característica essencial no tratamento dos dados e informação: o Valor. (Ibm, 2015). Segundo uma visão mais empresarial promovida por empresas de aconselhamento estratégico em Tecnologia da Informação (8), o ambiente de Big Data repousa sobre a possibilidade de transformar dados em informação que agregue valor aos negócios.

Assim, apoiado nas abordagens enunciadas anteriormente, pode-se dizer que para trabalhar com projetos de Big Data, será necessário observar as 5 dimensões presentes na Figura 1.

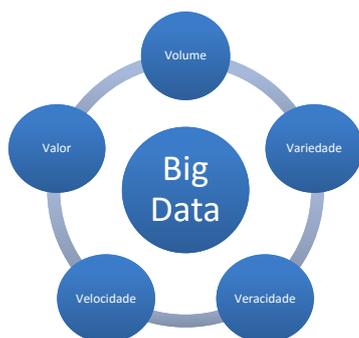


Figura 1: Os 5 V's girando em torno de Big Data (Elaborado pelo autor a partir de Hashem et. al. (2014) e Ibm (2015))

Duas das cinco dimensões presentes na Figura 1 correspondem a capacidade de tratar o grande quantitativo de dados presentes no cotidiano (Volume e Variedade). Como pode ser observado anteriormente (ver subseção 1), o uso de celulares e smartphones pelos indivíduos da nossa sociedade impulsionou o tema, na medi-

da em que estes dispositivos podem funcionar com sensores que capturam e enviam informação. São cerca de 6 bilhões de dispositivos (OnuBr, 2014) gerando, consumindo, processando e (re)utilizando dados e informações.

Além destes dispositivos existem também câmeras de monitoramento em prédios, lojas, ruas e avenidas espalhadas pelas cidades. Este novo comportamento que vem sendo praticado pela sociedade, deixaria G. Wells, autor de 1984, impressionado com a possibilidade de observação que vem tomando a sociedade. A *Surveillance Society* (Mann; Nolan; Wellman, 2003) também vem incrementando o cotidiano com outra enxurrada de informações. Estima-se que a quantidade de vídeos produzidos diariamente ultrapassa a produção dos primeiros 50 anos de vida da televisão (Davenport, 2014).

Adiciona-se a esse cenário, uma vasta coleção de outras fontes e formas para geração de unidades documentárias. O crescimento do uso de documentos digitais e páginas Web nas organizações, recursos estes estruturados por meio de ferramentas para Gestão de Conteúdo (Ribeiro, 2008), bem como o desenvolvimento de propostas de uso da Web of Data e Linked Data (Ribeiro; Almeida, 2011) também têm contribuído para um aumento em Volume e Variedade de dados e informação.

A dimensão da Velocidade corresponde a característica mais perceptível para os usuários. A melhoria dos canais de transmissão, com redes em fibra ótica, emissores de sinais de alta capacidade e uso de satélites, traz reflexo direto no nosso cotidiano. Pode-se afirmar que o desenvolvimento da tecnologia de processamento, incluindo o uso de processadores, canais e hardware de armazenamento, dobra a cada período de 2 anos, o que contribui para incrementar o volume das unidades documentárias (Florissi, 2012).

A dimensão da Veracidade esta ligada à qualidade da informação (Abib, 2010). Esta é essencial para que os usuários interessados (executivos, gestores públicos e a sociedade em geral) usem e (re)usem os dados de maneira apropriada e real, gerando informações críveis e compartilhando significados. (Assis; Moura, 2011).

A dimensão do Valor é uma característica intrínseca dos dados, pois corresponde à capacidade desses dados gerarem informações que agreguem valor ao processo de tomada de decisão para os usuários. Essa capacidade está relacionada com formas de inovação e melhoria em processos de tomada de decisão, além de contribuir para a obtenção de vantagens econômico-

cas a partir destes grandes volumes de dados. (Power, 2014; Ibm, 2015).

### 3.2. Big Data Analytics

Trabalhar com as dimensões já apresentadas, pressupõe a possibilidade de analisar os dados e informações geradas. Neste sentido, sabe-se que a capacidade do indivíduo para identificar e representar requisitos e necessidades de informação é limitada. Esta capacidade depende, principalmente, da correlação entre os diferentes níveis de abstração e granularidade das representações no domínio que está sendo representado (Goguen; Jirotko, 1994).

Aliado a isto, é possível afirmar que estes indivíduos interpretam e representam de maneira diferenciada, pois a capacidade de percepção do cérebro humano estrutura a realidade e as conexões de informação segundo diferentes caminhos (Kent, 1998).

Assim, pode-se supor que para a execução da tarefa de análise desta vasta coleção, o ser humano precisa lançar mão de técnicas e ferramentas que reduzam a complexidade do domínio que está sob análise. A atividade de *Analytics* se caracteriza pela possibilidade de se realizar análise preditivas nos conjuntos de dados, fazendo uso de abordagens estatísticas (Siegel, 2013).

A separação em categorias e a posterior reunião em conjuntos, possibilita a execução de minerações (*mining*) (9) em determinados agrupamentos de grandes coleções, preparando assim os conjuntos de dados para serem carregados nos chamados lagos de dados (*data lake*). Este esforço de preparação é entendido como etapa de *Discovery* e precede a carga no *data lake*. (Oliveira, 2013; Tavares, 2014).

Na etapa de *Discovery* é possível utilizar outras abordagens para separação, limpeza dos dados e categorização, com a utilização, por exemplo, das fontes de origem do conjuntos (proveniência). Um dos objetivos finais para a etapa de *Discovery* é a obtenção de representações para os conjuntos carregados no lago de dados (*data preparation* e *model planning*). A Figura 2 representa de forma esquemática o ciclo para o trabalho com *Discovery*.

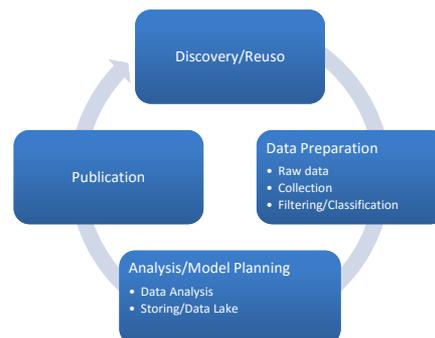


Figura 2: Ciclo de Discovery (Elaborado pelo autor a partir de Tavares (2014) e Khan et. al. (2014))

Este processo demanda uma nova arquitetura de análise, uma vez que os dados podem ser obtidos a partir de fontes diversas e originários de tecnologias diferentes. Seymour (2014) convalida esta percepção quando afirma que para analisar uma grande quantidade de dados, muitas vezes gerados de forma concomitante e em paralelo, pode existir uma grande dificuldade de correlacioná-los.

Ademais, Sathi (2013) destaca que em função destas múltiplas fontes, o trabalho com os processos de negócio de uma organização começa a prescindir de um profissional que saiba executar tarefas ligadas ao esforço de *Analytics*. Assim se torna possível desenvolver novos produtos e serviços para os clientes e/ou usuários.

### 3.3. O cientista de dados e o esforço de Analytics

A partir das observações apresentadas nas subseções anteriores é possível verificar a necessidade de desenvolver um novo perfil profissional que reúna diferentes disciplinas no trato de dados e informação. Decorrente desta inquietação, surgiu então o perfil denominado Cientista de Dados (*Data Cientist*). Davenport e Patil (2012) destacam que este profissional deve possuir, dentre outras características, a capacidade de aplicar ferramentas analíticas e algoritmos, com o intuito de gerar previsões para o desenvolvimento de produtos e serviços.

Na visão de Oliveira (2013), este profissional deve reunir conhecimentos em disciplinas ligadas à matemática e à estatística, principalmente pela necessidade de tratar os múltiplos conjuntos de dados. O uso de modelos matemáticos, a aplicação de técnicas de regressão, a capacidade de correlacionar variáveis e a formulação de hipóteses, também são características desejáveis para este profissional.

Oliveira continua e complementa observando que o perfil também deve incorporar caracterís-

ticas de trabalho em colaboração apoiado em amplo processo de investigação, comunicação e criação.

Neste sentido, incorpora-se também ao perfil a descoberta de requisitos e necessidades de informação em múltiplos contextos. Com o envolvimento de diferentes personagens (usuários, clientes, parceiros de negócio, informações de mercado, redes sociais, feeds de notícias, dentre outros), o Cientista de Dados deve incrementar sua capacidade de criação e colaboração (Breitman, 2014).

Portanto, para o Cientista de Dados e Informação descobrir requisitos, envolver diferentes personagens e conduzir trabalhos em colaboração, propõe-se que ele parta de algumas orientações para nortear o trabalho etapa de *Discovery* (Oliveira; 2013):

- a) A necessidade de conhecer o contexto dos dados e informações, identificando todos os elementos pertinentes (fontes, descrições e correlações);
- b) A necessidade lançar visões diferenciadas para formular propostas de organização dos dados e informações;
- c) A necessidade de propor categorizações para auxiliar projetos de conjuntos de dados (*clusters*);

Para desenvolver projetos em Big Data, Shiri (2014) apresenta uma abordagem que se aproxima e convalida as orientações de Oliveira, pois propõe observar o contexto segundo as facetas que auxiliam na compreensão dos elementos: Contexto/ambiente, Análise/Analytics, Metadados, Atividades e Operações, Tipos de Dados e Pessoas (Figura 3).

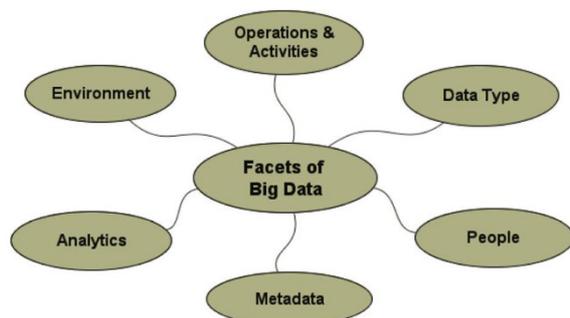


Figura 3: Facetas utilizadas em projetos de Big Data (Shiri; 2014)

A visão adotada por esta investigação procurou evidenciar a faceta *Data Type*. No âmbito desta faceta, Shiri continua e propõe que o processo de análise contemple também as subfacetas:

*nature, context, creator, processing, publication, structure, format e access* (Figura 4).

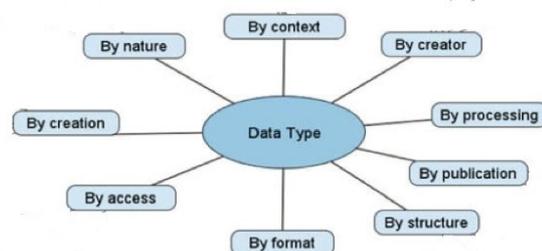


Figura 4: Subfacetas da faceta *Data Type* utilizadas em projetos de Big Data (Shiri; 2014).

#### 4. Metodologia

O recorte deste estudo contempla o uso de recursos ligados a organização do conhecimento para auxiliar no desenvolvimento de projetos de Big Data, objetivando especialmente a etapa inicial de *Discovery*.

Trata-se de um estudo exploratório e tendo em vista o recorte adotado, buscou-se identificar os conjuntos de dados e de informação que possuem correlação com a temática de Acidentes de Trabalho. Esta fase foi desenvolvida, por meio de Análise da Relevância dos dados disponíveis no portal de Dados Abertos da Dataprev (10). As questões sobre a relevância da informação tratadas por Saracevic (1970; 1996b) foram utilizadas para auxiliar na definição dos conjuntos de interesse. Os itens relevantes, em geral, atenderam a questões formuladas nessa pesquisa e percebidas como de interesse para o cliente e/ou usuário. Estas questões foram desenvolvidas com o apoio de critérios, sintetizados a seguir (Ribeiro, 2012):

- As informações ou os conjuntos de dados selecionados podem ocasionar sentimentos de excitação e satisfação aos clientes e/ou usuários;
- Especificidade de pesquisas propostas para os clientes e/ou usuários estão contempladas nos conjuntos selecionados;

Oliveira complementa e corrobora a abordagem com a proposta das duas primeiras tarefas para executar o trabalho de *Discovery* na atividade de *Analytics*, são: a identificação das informações que serão necessárias para possibilitar interpretações sobre os resultados, correlações, implicações e causalidade; e determinar o(s) tipo(s) de organização para estas informações (classificações e criação de *clusters* de dados).

Após o recorte dos dados no domínio, partiu-se para a análise da faceta evidenciada e suas

subfacetas, com o intuito de auxiliar no processo para categorização dos conjuntos de dados. Conforme Shiri (2014) observa, o uso da análise de facetas possibilita delinear características, atributos e diferentes aspectos de um ambiente complexo. Ademais, em relação ao conteúdo analisado, para Li e Belkin (2008), um esquema de classificação pode refletir a possibilidade de compreender melhor os aspectos ligados ao processo de obtenção deste conteúdo.

## 5. O experimento sobre os dados de acidente de trabalho

A Previdência Social é detentora de muitos dados e informações da sociedade brasileira. Com o esforço para publicação de dados em formato aberto para poderem ser reutilizados pela sociedade (Ribeiro; Almeida, 2011), surgiram muitas possibilidades de reuso, tanto em projetos de pesquisa específicos quanto em trabalhos acadêmicos (Rodrigues, 2012; Campos; Campos; Carvalho; Lima, 2012; Germano, 2013).

Esta investigação está sendo desenvolvida no âmbito de um projeto de pesquisa em curso na Unirio e está apoiada em trabalhos anteriores desenvolvidos pelo autor. O objetivo deste estudo foi investigar um conjunto de ações que viabilizem a participação do profissional da informação em projetos de Big Data, em especial na atividade de *Analytics*. Assim, a delimitação do campo empírico se deu a partir da questão: é possível analisar os dados abertos sobre Acidentes de Trabalho disponíveis no portal da Previdência e organizá-los segundo os princípios da Organização do Conhecimento?

Neste sentido, partiu-se do pressuposto que a elaboração de modelo conceitual de dados é útil, pois permite avaliar, inclusive, a possibilidade de realizar integrações de dados e associações. Para Kent (1998), o modelo conceitual de dados nos ajuda a compreender melhor as diferentes visões da realidade. Ademais, Sukumar e Ferrel (2013) convalidam esta estratégia quando afirmam que conhecer estrutura dos dados e seus relacionamentos, pode facilitar o trabalho de análise de conjuntos de dados.

Assim, a investigação partiu da identificação dos dados e informações relevantes no Portal de Dados abertos da Dataprev e, neste sentido, foram identificados dois conjuntos de dados que poderiam ser úteis para responder a questão de partida: Acidentes de Trabalho por CBO (11) e Acidentes de Trabalho por CNAE 2.0 (12). A representação a seguir (Figura 5) foi gerada a partir de uma extensão do modelo obtido em Ribeiro e Almeida (2011).

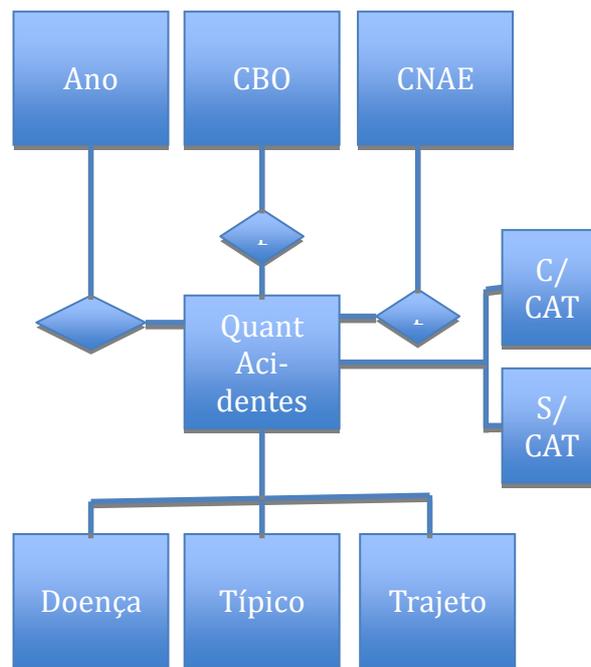


Figura 5: Modelo de Dados dos conjuntos escolhidos (elaborado pelo autor)

Cabe registrar que o modelo de dados utilizado permitiu representação semântica na medida em que este modelo apresenta elementos (entidades e atributos) e relações em um determinado domínio (Campos, 2001). Campos, Campos, Carvalho e Lima (2012) reafirmam esta estratégia de modelagem conceitual, pois a mesma permite uma boa forma de efetuar a representação do conhecimento em um domínio.

A partir do modelo de dados buscou-se o entendimento das dimensões propostas por Shiri. Adotou-se a faceta *Data Type*, pois o objetivo dessa pesquisa foi explorar os conjuntos de dados disponíveis. As subfacetas foram analisadas para cada conjunto de dados presente no modelo de dados, levando-se em conta a origem, o formato, o processamento e o contexto, conforme conteúdo disponível no portal da Dataprev (10) (15).

A subfaceta *context* foi incrementada com a descrição do conjunto de dados, o que permitiu uma análise mais apurada do significado do conjunto. A subfaceta *processing* foi omitida no processo de análise, pois a forma de geração do conjunto não estava disponível no portal no momento da pesquisa.

A tabela a seguir apresenta a sistematização dessa análise:

| Conjunto de Dados        | Creator      | Format | Context   |
|--------------------------|--------------|--------|---|
| Quant. Acidentes-Doença  | Gov/INSS     | JSON   | Acidentes ocasionados por qualquer tipo de doença profissional peculiar a determinado ramo de atividade constante na tabela da Previdência Social.  |
|                          |              | XML    |   |
|                          |              | CSV    |   |
| Quant. Acidentes-Típico  | Gov/INSS     | JSON   | Acidentes decorrentes da característica da atividade profissional desempenhada pelo acidentado;   |
|                          |              | XML    |   |
|                          |              | CSV    |   |
| Quant. Acidentes-Trajeto | Gov/INSS     | JSON   | Acidentes ocorridos no trajeto entre a residência e o local de trabalho do segurado e vice-versa;   |
|                          |              | XML    |   |
|                          |              | CSV    |   |
| Quant. Acidentes-c/CAT   | Gov/INSS     | JSON   | Corresponde ao número de acidentes cuja Comunicação de Acidentes do Trabalho – CAT foi cadastrada no INSS. Não são contabilizados o reinício de tratamento ou afastamento por agravamento de lesão de acidente do trabalho ou doença do trabalho, já comunicados anteriormente ao INSS.   |
|                          |              | XML    |   |
|                          |              | CSV    |   |
| Quant. Acidentes-s/CAT   | Gov/INSS     | JSON   | Corresponde ao número de acidentes cuja Comunicação de Acidentes do Trabalho – CAT não foi cadastrada no INSS. O acidente é identificado por meio de um dos possíveis nexos: Nexo Técnico Profissional/Trabalho, Nexo Técnico Epidemiológico Previdenciário – NTEP ou Nexo Técnico por Doença Equiparada a Acidente do Trabalho. Esta identificação é feita pela nova forma de concessão de benefícios acidentários |
|                          |              | XML    |   |
|                          |              | CSV    |   |
| CNAE                     | Gov/IBGE     | JSON   | Classificação Nacional de Atividade Econômica (99 subclasses)   |
|                          |              | XML    |   |
|                          |              | CSV    |   |
| CBO                      | Gov/MT E     | JSON   | Código Brasileiro de Ocupações (192 subgrupos)  |
|                          |              | XML    |   |
|                          |              | CSV    |   |
| Ano                      | Gov/Dataprev | JSON   | Ano   |
|                          |              | XML    |   |
|                          |              | CSV    |   |

Tabela 1: Análise das subfacetadas escolhidas (elaborado pelo autor)

A análise das subfacetadas presentes na Tabela 1 apoiou a execução da etapa de *Discovery*, por

meio da exploração e organização com o conteúdo levantado.

A descoberta de características e similaridades nas subfacetadas permitiram a agregação e/ou sumarização dos conjuntos de dados. A identificação tanto da origem (*creator*) quanto do formato (*format*), permitem propor processos de carga no *data lake* (ver 3.2) mais ágeis, além de viabilizar discussões ligadas à gestão e ao licenciamento dos dados (13). O contexto (*context*) permite uma análise sobre os itens de dados que compõem o conjunto, possibilitando um entendimento da informação gerada. Esta dimensão também pode determinar mudança na estratégia da carga no *data lake*, na medida em que o cientista pode identificar características de similaridade na geração da informação.

Outro aspecto importante é a possibilidade de identificar correlações entre elementos. A partir do campo *context* do conjunto de dados Quant. Acidentes-Trajeto, identifica-se a possibilidade de correlacionar os locais mais perigosos em relação aos acidentes, por meio de coordenadas GPS (14) e gerando um mapa visual com recursos do *Google Maps*.

De forma análoga, o campo *context* do conjunto de dados Quant. Acidentes-Doença, indica a possibilidade de correlacionar doenças (por meio de seus códigos de doença) com informações sobre investimentos em saúde.

A dimensão do processo de geração (*processing*) foi a mais prejudicada no processo de análise, tendo em vista a inexistência de conteúdo levantado. Todavia, percebe-se que esta é uma característica fundamental para complementar a semântica dos conjuntos de dados e que poderá, em conjunto com o contexto, ser útil na definição de vocabulários específicos para a descrição dos conjuntos de dados. O uso destes vocabulários pode auxiliar na descrição a identificação da estrutura semântica dos elementos sob análise (Ribeiro; Vieira, 2014)

## 6. Considerações Finais

O uso das abordagens ligadas ao tema Big Data ainda carece de investigação pela área de Ciência da Informação. O uso dessas visões poderá auxiliar na melhoria da oferta de serviços de informação. Conhecer o ambiente de dados e informação, que deve ser originário de diferentes fontes, efetuar a organização de conjunto de dados (categorizá-los?), realizar entrevistas junto aos clientes e/ou usuários e desenvolver os modelos (tanto estruturais quanto matemáticos), contribuirá para o projeto desses serviços. Shiri (2014) convalida esta percepção quando

afirma que a Ciência da Informação e em especial os métodos de organização do conhecimento, são terrenos sólidos e adequados para desenvolver estudos em Big Data.

Para Minelli, Chambers e Dhiraj (2013), o momento que vivemos é especial, pois a contínua redução do custo dos equipamentos, além do uso de novos softwares e ferramentas para apoiar os processos de gestão de dados e informação, têm contribuído para um momento especial no tratamento da informação.

Quando lançamos um olhar na direção das bibliotecas e unidades de informação, é possível perceber que tanto os estoques existentes quanto as demandas por outras fontes têm crescido de forma exponencial, pois os ativos de informação de interesse para os usuários estão armazenados em diferentes bases de dados, usando bancos de dados e plataformas de computação heterogêneas (Meletiou, Katsirikov; 2009). Adicione-se a isso, a presença cada vez maior de redes sociais e aplicações de colaboração nos serviços das unidades de informação, ocasionando mudanças no comportamento dos usuários (De Jesus, Da Cunha; 2012; Santos, Da Rocha; 2012). Em suma, compreender e participar de esforços com o uso de Big Data pode ser valioso para auxiliar na estruturação de unidades de informação.

Por fim, Shiri (2014) observa que a noção de Big Data está sendo incorporada gradualmente no ambiente de informação digital, logo, é essencial que sejam realizadas investigações que contemplem o uso de outras facetas. Em especial no que diz respeito à faceta de Metadados, em decorrência de alguns aspectos emergentes como: o intenso uso de repositórios, a necessidade de criação de mecanismos para acesso eficiente, além da colaboração entre pesquisadores e instituições no ambiente de *Web of data*, bem como na necessidade de intercâmbio de informações.

Projetos de Big Data são desenvolvidos com os objetivos de criar novos produtos, compreender necessidades dos clientes e seus comportamentos, bem como perceber novos mercados. Para isto, é necessário desenvolver teorias para tratar com clientes e usuários, construindo hipóteses e identificando dados e informações relevantes. Este processo deve ser repetido e refinado, de acordo com os experimentos realizados e as respostas obtidas (Marchand; Peppard, 2013).

Espera-se que este movimento de pesquisa na área da Ciência da Informação, ilumine o caminho a ser trilhado e possibilite que outros pesquisadores interessados possam se engajar

nesta discussão, levando este tema para além da tecnologia.

## Notas

- (1) Neologismo criado por Alvin Tofler que junta os conceitos de produtor e consumidor.
- (2) A unidade Zettabyte é equivalente, em valores aproximados, a 1.000 Exabytes, ou a 1.000.000 de Petabytes, ou ainda, a 1.000.000.000 de Terabytes.
- (3) Tradução livre de *Cloud Computing*
- (4) Também conhecida como *Apache Hadoop* é tecnologia *open source* desenvolvida pela *Google* e *Yahoo* para processar muitos dados em servidores, usando a noção de processamento em paralelo e uso de *clusters* (conjuntos) de computadores no processamento. Possui versões customizadas desenvolvidas por fabricantes de software e hardware.
- (5) *MapReduce* é a proposta de arquitetura que deu origem à tecnologia de *Hadoop*. Usa a estratégia de dividir para conquistar, ou seja, distribui e aloca um problema muito grande em *clusters* de armazenamento, usando registros serializáveis do tipo <chave, valor>.
- (6) Sigla de *Hadoop Distributed File System*. Estrutura de armazenamento de arquivos que utiliza blocos de até 64 Megabytes, que são muito menores do que os blocos de particionamento tradicionais, possibilitando a transmissão dos blocos em modalidade de *streaming*.
- (7) *Machine Learning* trata o uso de algoritmos que identificam o melhor modelo para ser aplicado ao conjunto de dados.
- (8) Empresas de aconselhamento estratégico é como são conhecidas as empresas de consultoria em Tecnologia da Informação, que prestam serviços de elaboração de relatórios sobre o uso e adoção de novas tecnologias.
- (9) A noção de *mining* de dados passa pela extração e análise de grandes volumes de informação em busca de padrões e comportamentos.
- (10) Disponível em <http://dadosabertos.dataprev.gov.br>
- (11) Classificação Brasileira de Ocupações
- (12) Classificação Nacional de Atividade Econômica
- (13) O licenciamento para reuso destes conjuntos de dados precisa ser feito com a instituição provedora e deve se apoiar nos 4 tipos de licenciamento disponível: ODC-BY, ODC-ODBL, ODC-DBCL e ODC-PDDL (Open Data Commons, 2015).
- (14) GPS – *Global Positioning System* – Sistema Global que identifica coordenadas geográficas das localidades.

(15) A descrição da subfaceta *context* foi extraída do Anuário Estatístico Ano 2013. Pag. 543. Disponível em <http://www.previdencia.gov.br/estatisticas/>

## Referencias

- Abib, G. (2010). A qualidade da informação para a tomada de decisão sob a perspectiva do sensemaking: uma ampliação do campo. *Ciência da Informação*, 39(3). 73-82.
- Assis, J. Moura, M. A. (2011). A Qualidade da Informação na WEB: uma abordagem semiótica. *Informação & Informação*. 16(3), 96-110.
- Brietman, K. (2014). Big Data Seen from the Clouds. 2014. Palestra apresentada no 2o. EMC Summer School on Big Data. EMC/NCE/UFRJ. Rio de Janeiro. 2014. Disponível em [http://2014.emcbigdataschool.nce.ufrj.br/images/presentations/Big\\_Data\\_Summer\\_School\\_Karin.pdf](http://2014.emcbigdataschool.nce.ufrj.br/images/presentations/Big_Data_Summer_School_Karin.pdf) Acesso em Maio de 2015.
- Campos, M. L. A. (2001). *A organização de unidades de conhecimento: o modelo conceitual como espaço comunicacional para a realização da autoria*. (Doutorado em Ciência da Informação). ECO/IBICT, Universidade Federal do Rio de Janeiro, Rio de Janeiro.
- Campos, L. M.; Campos, M. L. A.; Carvalho, M. G. P.; Lima, D. V. M. (2012). Dados abertos interligados e o espaço do profissional de informação: Uma aplicação no domínio da enfermagem In: (2012) *Encontro Nacional de Pesquisa em Ciência da Informação*, Rio de Janeiro: FIOCRUZ. Disponível em: <http://www.eventosecongressos.com.br/metodo/enancib2012/anais/index.php>. Acesso em: Maio de 2015.
- Ciarini, A. E. M. (2013). Research on Big Data and Opportunities. Palestra apresentada no 1o. EMC Summer School on Big Data. EMC/NCE/UFRJ. Rio de Janeiro.
- Davenport, T. H. (2014). *Big Data at Work: Dispelling the Myths, Uncovering the Opportunities*. Harvard Business Review Press Books.
- Davenport, T. H.; Patil, D.J. (2012). Data Scientist: The Sexiest Job of the 21st Century. *Harvard Business Review* 90 (10-October), 70–76.
- De Jesus, D. L.; Da Cunha, M. B. (2012). Produtos e serviços da web 2.0 no setor de referência das bibliotecas. *Perspectivas em Ciência da Informação*, 17 (1), 110-133.
- Dumbill, E. (2012). What is Big Data? In: (Editor) (eds). *O'Reilly Media Inc. Big Data Now: current perspectives*. O'Reilly Media:California. 2012. Disponível em: <http://www.oreilly.com/data/free/files/big-data-now-2012.pdf>. Acesso em: Maio 2015.
- EMC. Brazil country brief. (2014). *The Digital Universe of opportunities*. Disponível em: <http://www.emc.com/collateral/analyst-reports/idc-digital-universe-2014-brazil.pdf>. Acesso em: Maio 2015.
- Ferreira, A. M. J. F. C.; Santana, R. C. G.; Vidotti, S. A. B. G. (2012). Second life: perspectivas para potencializar o acesso a dados públicos. In: (2012). *Encontro Nacional de Pesquisa em Ciência da Informação*, Rio de Janeiro: FIOCRUZ. Disponível em: <http://www.eventosecongressos.com.br/metodo/enancib2012/anais/index.php>. Acesso em: Maio de 2015.
- Florissi, P. (2014). Big Data: Challenges and Opportunities. Palestra apresentada no 2o. EMC Summer School on Big Data. EMC/NCE/UFRJ. Rio de Janeiro.
- Florissi, P. EMC On Big Data. (2012) . Disponível em: <https://www.carecorenational.com/healthcaresummit/powrpoints/PatriciaFlorissiPhD.pdf>. Acesso em: Maio de 2015.
- Fonseca, M. J.; Gonçalves, M. A.; Oliveira M. O. R.; Tinoco, M. A. C. (2008). Tendências sobre as comunidades virtuais da perspectiva dos prosumers. *RAE Eletrônica*. 7 (2). Disponível em <http://rae.fgv.br/rae-eletronica/vol7-num2-2008/tendencias-sobre-comunidades-virtuais-perspectiva-prosumers>. Acesso em: Maio de 2015.
- Fox, P.; Hendler, J. (2011). Changing the Equation on Scientific Data Visualization. *Science* 331, 705. Disponível em: [http://data2discovery.org/dev/wp-content/uploads/2013/05/Fox-and-Hendler\\_Visualization\\_Science-2011-Fox-705-8.pdf](http://data2discovery.org/dev/wp-content/uploads/2013/05/Fox-and-Hendler_Visualization_Science-2011-Fox-705-8.pdf). Acesso em: Maio de 2015
- Germano, E. C. (2013). *Modelos de negócios adotados para o uso de dados governamentais abertos: estudo exploratório de prestadores de serviços na cadeia de valor dos dados governamentais abertos*. 2013. Dissertação (Mestrado em Administração) - Faculdade de Economia, Administração e Contabilidade, University of São Paulo, São Paulo, 2013. Disponível em: <http://www.teses.usp.br/teses/disponiveis/12/12139/tde-10012014-155226/>. Acesso em: Maio de 2015.
- Goguen, J. A. (1994). Requirements Engineering as the reconciliation of social and technical issues. In: *Jirotko, M. e Goguen, J. A. (1994) (Ed.). Requirements Engineering - Social and Technical Issues*. London, 1994. p.165-199.
- Gopintah, M. A.; Das, P. (1997). Classification and representation of Knowledge. *Library Science with a Slant to Documentation and Information Studies*. 34 (2). 85-90.
- Hashem et. al. The rise of "Big Data" on cloud computing: Review and open research issues. *Information Systems*, 47 (2015) 98-115
- Heath T.; Bizer C. (2011). *Linked Data: Evolving the Web into a Global Data Space (1st edition)*. Synthesis Lectures on the Semantic Web: Theory and Technology. Morgan & Claypool Publishers.
- Hjørland, B.; Albrechtsen, H. (1995). Toward a new horizon in information science: Domain-analysis. *Journal of the American Society for Information Science*, 46(6), p. 400-425.
- Ibm (2015). IBM BigData e Analytics Hub. *Why only one of the 5 Vs of big data really matters* Disponível em: <http://www.ibmbigdatahub.com/blog/why-only-one-5-vs-big-data-really-matters>. Acesso em: Maio 2015.
- InfoExame (2014). Solução de Big Data é um dos segredos da seleção alemã na copa. Disponível em : <http://info.abril.com.br/noticias/it-solutions/2014/07/solucao-de-big-data-e-um-dos-segredos-da-alemanha-na-copa-do-mundo.shtml>. Acesso em: Maio de 2015.
- Li, Y.; Belkin, N. J. (2008). A faceted approach to conceptualizing tasks in information seeking. *Information Processing and Management* .44 .1822–1837.
- Khan, N et. al. (2014). Big data: survey, technologies, opportunities, and challenges.(Report). *The Scientific World Journal*, 14.
- Kent, W. (1998). *Data and Reality*. 1stBooksLibrary. Bloomington.
- Mann, S.; Nolan, J; Wellman, V. Sousveillance: Inventing and Using Wearable Computing Devices for Data Collection in Surveillance Environments. *Surveillance & Society* 1:3 (2003) 331-355
- Marchand, D. A.; Peppard, J. (2013). Why IT Fumbles Analytics. *Harvard Business Review*. Jan-Fev.
- Mattoso, A. Scientific Workflows and Big Data. Palestra apresentada no 1o. EMC Summer School on Big Data. EMC/NCE/UFRJ. Rio de Janeiro..
- Meletiou, A; Katsirikou, A. (2009). Methodology of analysis and interrelation of data about quality indexes of library

- services by using data-and knowledge-mining techniques. *Library Management*. 30 (3), 138-147.
- Mello, R., Leite, L. R., e Martins, R. A. (2014). Is Big Data the Next Big Thing in Performance Measurement Systems? Guan, Y e Liao, H. (eds). *Proceedings of the 2014 Industrial and Systems Engineering Research Conference*.
- Minelli, M.; Chambers, M.; Dhiraj, A. (2013). *Big Data, Big Analytics: Emerging Business Intelligence and Analytic Trends for Today's Businesses*. Wiley CIO Series.
- Oliveira, A. (2013). Data Science and Data Analytics. Palestra apresentada no 1o. EMC Summer School on Big Data. EMC/NCE/UFRJ. Rio de Janeiro.
- ONUBR. (2013). ONU: Dos 7 bilhões de habitantes do mundo 22 de março de 2013. Disponível em <http://www.onu.org.br/onu-dos-7-bilhoes-de-habitantes-do-mundo-6-bi-tem-celulares-mas-25-bi-nao-tem-banheiros/> Acesso em: Maio de 2015.
- Open Data Commons (2015). Disponível em: <http://www.opendatacommons.org/about>. Acesso em : junho de 2015.
- Porto, F. (2013). Big Data in Astronomy: The LIneA-DEXL case. Palestra apresentada no 1o. EMC Summer School on Big Data. EMC/NCE/UFRJ. Rio de Janeiro.
- Power, D. J. (2014). Using 'Big Data' for analytics and decision support. *Journal of Decision Systems*, 3 (2).
- Ribeiro, C. J. S. (2008). *Diretrizes para o projeto de portais de informação: uma proposta interdisciplinar baseada na Análise de Domínio e Arquitetura da Informação*. 298 f. Tese (Doutorado em Ciência da Informação) – Convênio UFF/IBICT, Rio de Janeiro.
- Ribeiro, C. J. S. (2012). Entendimento de requisitos de sistema com abordagem orientada ao domínio. *Data-GramaZero - Revista de Informação*. (Abril 2012), 13 (2). Disponível em [http://www.dgz.org.br/abr12/Art\\_01.htm](http://www.dgz.org.br/abr12/Art_01.htm). Acessado em: Maio de 2015.
- Ribeiro, C. J. S.; Almeida, R. F. (2011). Dados Abertos Governamentais (Open Government Data): Instrumento para Exercício de Cidadania pela Sociedade. In: *Elmira Simeão, Jorge Henrique Cabral Fernandes, Isa Maria Freire. (Org.). XII Enancib - Políticas de Informação para a Sociedade* Brasília: Thesaurus.
- Ribeiro, C. J. S.; Pereira, D. V. (2014). El proceso de publicación e los datos gubernamentales abiertos acerca de la clase de la Seguridad Social Brasileña de Vocabulario Controlado del Gobierno Electrónico (VCGE). In: *Vanderkast E. J. S. (ed) (2014). El acceso a la información gubernamental : experiencias y expectativas*. Cidade do México: UNAM, Instituto de Investigaciones Bibliotecológicas y de la Información
- Robredo, J. (2011). Do documento impresso à informação nas nuvens: reflexões. *Liinc em Revista*, 7 (1), 19-42. Disponível em: <http://www.ibict.br/liinc> . Acesso em maio de 2015.
- Rodrigues, F. A. (2012). *Mapeamento de tecnologias informacionais sobre dados abertos em saúde pública: destino de repasses financeiros federais*. Dissertação (Mestrado em Ciência da Informação). Universidade Estadual Paulista. Faculdade de Filosofia e Ciências.
- Santos, E. L.; Da Rocha, S. M. (2012). O blog como ferramenta de comunicação entre a biblioteca e seus usuários: a experiência da biblioteca Lydio Bandeira de Mello, da Faculdade de Direito da Universidade Federal de Minas Gerais. *Encontros Bibli*, 17 (33). 134-152.
- Santos, I. H. R. (2014). BigData Research and Development at Petrobras. 13 de maio de 2014. Palestra apresentada no 2o. EMC Summer School on Big Data. EMC/NCE/UFRJ. Rio de Janeiro. 2014. Disponível em: [http://2014.emcbigdataschool.nce.ufrj.br/images/presentations/Ismael\\_BigDataTOOL\\_SummerSchool\\_v2.pdf](http://2014.emcbigdataschool.nce.ufrj.br/images/presentations/Ismael_BigDataTOOL_SummerSchool_v2.pdf). Acesso em: Maio de 2015.
- Sarecivic, T. (1970). The concept of "relevance" in Information Science: an historical review. In: *Saracevic, T (Ed.) Introduction to Information Science*. New York: R. R. Bowker, 111-154.
- Saracevic, T. (1996a). Ciência da informação: origem, evolução e relações. *Perspectivas em Ciência da Informação*, 1 (1). 41-62..
- Saracevic, T. (1996b). Relevance reconsidered 1996. Information Science: Integration in Perspectives. In: *International Conference of Library and Information Science (COLIS, 2)*. Copenhagen, Denmark, 14-17 Oct.
- Sathi, A. (2013). *Big Data Analytics: Disruptive Technologies for Changing the Game*. Mc Press. 2013.
- Seymour, C. (2014). The State of Big Data. *EContent-Mag.com*. Jan-Feb. 26-27.
- Siegel, E. (2013). *Predictive Analytics*. Wiley. New Jersey.
- Sukumar, S. R.; Ferrel, R. K. (2013). 'Big Data' collaboration: Eploring, recording, and sharing enterprise knowledge. *Information Services & Use*. 33. 2013.
- Tavares, E. BIG DATA in Business. 12 de maio de 2014. Palestra apresentada no 2o. EMC Summer School on Big Data. EMC/NCE/UFRJ. Rio de Janeiro. 2014. Disponível em: [http://2014.emcbigdataschool.nce.ufrj.br/images/presentations/Apresentacao\\_Elaine\\_Tavares.pdf](http://2014.emcbigdataschool.nce.ufrj.br/images/presentations/Apresentacao_Elaine_Tavares.pdf) . Acesso em: Maio de 2015.
- Tofler, A. *Third Wave: The classic study for tomorrow*. Bantam Books: New York. 1984
- Xie, C.; Bagozzi, R.; Troye, S. (2008). Trying to prosume: toward a theory of consumers as co- creators of value. *Journal of the Academy of Marketing Science*, 36, 62-78.
- Wurman, R. S. *Ansiedade de Informação 2*. São Paulo: Editora de Cultura, 2005.
- Zikopoulos, P., deRoos, D., Parasuraman, K., Deutsch, T., Giles, J., & Corrigan, D. (2013). Harness the Power of Big Data: *The IBM Big Data Platform*. New York: McGraw Hill.

---

# A representação da imagem do bibliotecário nas histórias em quadrinhos (HQs): um estudo descritivo

*The representation of the librarian's image in the comics: a descriptive study*

---

**Fernando Bittencourt dos Santos (1), Maria Rosa Gomes Dória (2)**

(1) Universidade Federal de Sergipe, Cidade Universitária Prof. José Aloísio de Campos, Av. Marechal Rondon, S/N, Jardim Rosa Elze. Centro de Ciências Sociais e Aplicadas I - Andar Térreo São Cristóvão/SE, Brasil - CEP: 49100-000, email: fernandoubatuba@hotmail.com

## Resumo

A representação da informação é uma das áreas mais importantes da Biblioteconomia, tendo em vista que os profissionais que trabalham diretamente com esta última precisam dominar os processos de análise e síntese na leitura de um determinado documento, visando sempre atender as necessidades informacionais dos usuários, quando estes buscam a informação que necessitam em uma unidade de informação, seja ela física ou virtual, de modo a recuperá-las com eficácia e no menor tempo possível. Dentro desta perspectiva e considerando que os quadrinhos são formas de representação da informação dentro de um contexto imagético-textual, o presente trabalho apresenta como objetivo geral: 1) a análise da representação do bibliotecário no universo das histórias em quadrinhos, a partir da leitura do Blog Bibliocomics. Constituem-se objetivos específicos: 2) abordar a representação da informação no contexto da Biblioteconomia, bem como no universo dos quadrinhos; 3) caracterizar as histórias em quadrinhos e suas contribuições para a área de Biblioteconomia, no que concerne a divulgação desta área e 4) discutir se as imagens condizem com os aspectos reais do profissional bibliotecário. Quanto à metodologia, no que concerne aos objetivos, procedimentos e abordagem do problema, trata-se de uma pesquisa descritiva, bibliográfica e qualitativa, respectivamente, onde foram analisados nove quadrinhos. Como resultados, constatamos que os quadrinhos analisados são capazes de divulgar características inerentes ao bibliotecário, delineando aspectos culturais e sociais deste profissional, aspectos estes que podem ser positivos e negativos, sob uma perspectiva cômica e reflexiva na leitura e interpretação dos mesmos.

**Palavras chave:** História em quadrinhos. Representação da informação. Bibliotecário. Blog Bibliocomics.

## 1. Introdução

A representação da informação é uma das áreas mais importantes da Biblioteconomia, tendo em vista que os profissionais que trabalham diretamente com esta última precisam

## Abstract

The representation of information is one of the most important areas of librarianship, considering that the professionals who work directly with the latter need to master the process of analysis and synthesis in the reading of a document, aiming to meet the information needs of users when they seek the information they need in a unit of information, whether physical or virtual it in order to retrieve them effectively and in the shortest time possible. From this perspective, and considering that comics are forms of information representation within a imagetic-textual context, this paper presents the general objective: 1) the analysis of the librarian representation in the universe of comics, from the reading of Blog Bibliocomics. Constitute specific objectives: 2) address the representation of information in the context of Librarianship as well as in the world of comics; 3) characterize the comics and his contributions to the Library area, regarding the disclosure of this area and 4) discuss whether the images are consistent with the real aspects of the librarian. Regarding the methodology, regarding the objectives, procedures and approach to the problem, it is a descriptive, bibliographical and qualitative, their mind-where nine comics were analyzed. As a result, we find that the comic analyzed are able to disclose characteristics of the librarian, outlining cultural and social aspects of this professional, aspects that can be positive and negative, in a comical and reflective perspective on reading and interpretation.

**Keywords:** Comics. Representation of information. Librarian. Blog Bibliocomics.

dominar os processos de análise e síntese na leitura de um determinado documento, visando sempre atender as necessidades informacionais dos usuários, quando este último busca a informação que necessita em uma unidade de informação, seja ela física ou virtual, de modo a

recuperá-las com eficácia e no menor tempo possível.

Lima e Álvares (2012, p.35) contribuem com a afirmação anterior, assinalando que:

A representação da informação é importante objeto de trabalho dos arquivistas, bibliotecários, museólogos, dentre outros profissionais da informação. Eles precisam dominar os processos de análise e síntese dos documentos para gerar formas de representação que permitam identificá-los e recuperá-los a partir dos seus atributos e características principais.

Dentro desta perspectiva e considerando que os quadrinhos são formas de representação da informação dentro de um contexto imagético-textual, sendo que McCloud (2005, p.20) assinala o conceito de histórias em quadrinhos como: “Imagens pictóricas e outras justapostas em sequência deliberada destinadas a transmitir informações e/ou a produzir uma resposta no espectador”, o presente trabalho de pesquisa apresenta como objetivo geral a análise da representação do bibliotecário no universo das histórias em quadrinhos, a partir da leitura do Blog Bibliocomics (1), este último, representa muitas vezes com humor, o universo da Biblioteconomia.

Constituem-se objetivos específicos: abordar a representação da informação no contexto da Biblioteconomia, bem como no universo dos quadrinhos; caracterizar as histórias em quadrinhos e suas contribuições para a área de Biblioteconomia, no que concerne a divulgação desta área e discutir se as imagens condizem com os aspectos reais do profissional bibliotecário.

Apesar de o blog ser direcionado para a Biblioteconomia, a sociedade em geral pode ter acesso às informações ilustradas em seus quadrinhos, tendo em vista que o referido blog é de livre acesso, e ao fazer a leitura destes o indivíduo pode produzir uma imagem do mundo biblioteconômico de acordo com as informações que estão contidas nestas tirinhas de humor, podendo esta ser positiva ou negativa da profissão, que condizem ou não com a realidade desta área do conhecimento.

O tema “Histórias em Quadrinhos” vem sendo abordado com frequência na área da Ciência da Informação. E o estudo destes, representando o contexto da Biblioteconomia é de grande relevância social e científica, pois não estará abordando somente o quadrinho em si, mas a representatividade de uma determinada ciência, ilustrada em um importante recurso informativo.

Desta forma, temos a seguinte questão de pesquisa que norteia este trabalho: Como o profes-

sional bibliotecário está representado nos quadrinhos do blog Bibliocomics?

Os quadrinhos disponíveis no blog Bibliocomics apresentam vários elementos do universo biblioteconômico, tais como: a leitura, o livro, a internet, bem como as tecnologias da informação e comunicação (TIC), entre outros. Porém, nos limitamos a explorar o elemento bibliotecário, devido à representatividade e grande pertinência deste profissional para a área da Biblioteconomia.

## 2. Metodologia

Quanto aos objetivos, esta pesquisa se configura como descritiva, sendo que Cerro, Bervian e Silva (2007, p.61) assinalam que está última:

[...] observa, registra, analisa e correlaciona fatos ou fenômenos (variáveis) sem manipulá-los. Procura descobrir, com a maior precisão possível, a frequência com que um fenômeno ocorre, sua relação e conexão com outros, sua natureza e suas características. Busca conhecer as diversas situações e relações que ocorrem na vida social, política, econômica e demais aspectos do comportamento humano, tanto do indivíduo tomado isoladamente como de grupos e comunidades mais complexas.

Dentro desta perspectiva, descreveremos as características dos quadrinhos selecionados no Blog Bibliocomics, visando atender os objetivos propostos neste trabalho. Quanto à abordagem do problema, esta pesquisa se configura como qualitativa. O universo da pesquisa foi composto por 9 (nove) quadrinhos referentes ao profissional bibliotecário.

### 2.1 Análise dos Quadrinhos

A análise serve para refletir o que está sendo apresentado nos quadrinhos, os quais têm o intuito de entreter e ao mesmo tempo transmitir uma mensagem relevante para o público leitor desses últimos. Os quadrinhos selecionados foram analisados individualmente, levando em consideração os aspectos e características inerentes ao bibliotecário, no que condiz aos quadrinhos em relação às imagens deste profissional, sendo que a análise destes possibilitará uma visão sobre este último, representado nos quadrinhos a seguir:



Figura 1: Vida de bibliotecária fora do ambiente de trabalho

O quadrinho acima, demonstra que o estereótipo da bibliotecária perpassa seu ambiente de trabalho, ela aparece no ambiente familiar com as mesmas características que são associadas a ela, coque no cabelo, óculos e criteriosa com prazos, não perdoando financeiramente a atraso do que quer que seja. Esses estereótipos do bibliotecário,

[...] independente de serem certas ou erradas as imagens guiam comportamentos e isso implica afirmar, que a visão que uma pessoa ou grupo social tem sobre o bibliotecário ao ser partilhado socialmente, influencia e dirige a maneira de pensar e agir de outras pessoas em relação a este profissional. (Santos, 2011, p. 4)



Figura 2: O usuário curioso e a Bibliotecária

No quadrinho acima, vemos a indignação da bibliotecária, em decorrência do comportamento inadequado do usuário quando esta última está guardando um livro na estante. Embora tenhamos a representação cômica na leitura do quadrinho, dentro do contexto da Biblioteconomia e no fazer dos profissionais da informação, estes últimos devem aguçar a curiosidade dos usuários, levando-se em conta a multiplicidade de fontes e serviços de informação que os auxiliam

na busca e recuperação da informação, e devem despertar o gosto e a curiosidade pela leitura, sendo ideal que esta comece na mais tenra idade. Para Hillesheim e Fachin (2004) cabe aos profissionais da área, aproveitarem esses momento de curiosidades e despertar em cada um o hábito da leitura, o uso pela biblioteca, a pesquisa e busca de informação e o reconforto se sentir-se saciado, de ter descoberto, de ser informado.

Na análise do quadrinho do Anexo 1, além de apresentar o lazer, representado pelo futebol, temos também dois equipamentos culturais: o cinema e a biblioteca, bem como a figura de uma possível bibliotecária, sem perfil para a área do Serviço de Referência e Informação, tendo em vista a sua insatisfação com uma usuária.

Dentro dessa perspectiva e considerando os aspectos psicológicos que influenciam o trabalho do bibliotecário quando este lida diretamente com o usuário, na qual vemos ilustrado no quadrinho, o balcão de referência, Witter (1986, p.33) é enfática, ao afirmar que:

As relações psicológicas entre usuários e bibliotecários são complexas e influenciadas por variáveis de cada um deles, além de estarem sujeitas às influências do ambiente. O comportamento do usuário tem alguma influência no do bibliotecário e vice-versa, podendo conduzir a um relacionamento muitas vezes não positivo para ambos. Para tornar essas relações mais adequadas, é necessário planejar, para modificar não só o comportamento de usuários e de bibliotecários, mas também o currículo usado para formação do bibliotecário.

Com relação ao Quadrinho do Anexo 2, na entrevista de referência o usuário nem sempre sabe o quer, e aí cabe ao bibliotecário tentar compreender e atender a necessidade informacional deste. Neste serviço, a orientação e auxílio ao usuário "compreende as atividades exercidas face a face com o usuário visando facilitar o acesso às coleções e a facilidade na utilização dos demais recursos que a biblioteca ofereça, como os catálogos, bases, acesso à Internet etc". (Maciel; Mendonça, 2006, p. 36)

Ao analisar o Quadrinho do Anexo 3, constatamos que o personagem que representa o bibliotecário não tem o hábito de ler, em decorrência da falta de tempo, contrariando assim uma das missões desse profissional: o incentivo a leitura.

No que concerne ao leitor-bibliotecário, Moura (2006, p. 30) aborda que:

A leitura, atividade fundamental no processo de trabalho desses profissionais, pressupõe um articulado movimento de interação entre o leitor e o texto a ser incorporado ao sistema. E isso não

ocorre sem o grau de subjetividade de corrente do processo interpretativo, visto que a trajetória dos bibliotecários, no mundo da leitura é talhada pela profissão.

Na análise do quadrinho do Anexo 4, constatamos mais uma vez que o estereótipo do bibliotecário perpassa o ambiente de trabalho, mantendo suas características até mesmo quando este se aposenta, independente se está cuidando de livros ou realizando outra atividade, ela aparece pedindo silêncio, com o famigerado coque no cabelo e óculos fundo de garrafa.

Dentro desta assertiva, “neste sentido, o imaginário coletivo tem efetuado a construção de estereótipos do bibliotecário a despeito de suas características e atitudes julgadas como verdadeiras e compartilhadas por determinados grupos.” (Santos, 2011, p.4).

Já no quadrinho do Anexo 5, o “silêncio”, dentro do âmbito do senso comum, é sempre associado à profissão de Biblioteconomia e no caso o quadrinho acima demonstra que fora do universo das unidades de informação, o comportamento do bibliotecário é diferente, o que por fim, acaba não desconstruindo a imagem equivocada deste profissional em sua atuação em unidade de informação. Desta forma, Silveira (2008, p. 87, grifo nosso) assinala que “De Alexandria ao início do século XX, a atividade dos bibliotecários se caracterizou pelo silêncio, pela solidão e pelas arduas práticas de organização do conhecimento.”

No quadrinho do Anexo 6, o personagem apresenta o bibliotecário como o profissional que pode ajudar nas pesquisas na biblioteca, no entanto, ao mesmo tempo ele reclama da qualidade do serviço prestado por este profissional. Entretanto este personagem está equivocado, pois ele não tem discernimento sobre a obra que ele está solicitando e reclama pela falta de ajuda do profissional da informação. Nesse contexto, considera-se que “[...] as pessoas não só precisam encontrar o que desejam, mas precisam desejar o que, de fato, precisam.” (Milanesi, 2003, p. 228).

O Bibliotecário de Referência, como aquele profissional considerado o mediador entre a informação e o usuário e que mantém o contato mais próximo com este último, nunca deve dizer “Não” ao usuário, quando este necessita de uma informação, mesmo que este usuário não saiba ao certo o que necessita ou esteja desinformado, sendo que o bibliotecário deve orientá-lo e direcioná-lo a outra fonte de informação, caso a sua biblioteca não possua a obra ou corrigi-lo quando a informação apresentada está errada.

Na análise do último quadrinho (Anexo 7), este último faz alusão de forma cômica, ao trabalho do bibliotecário, como algo moroso e silencioso, sem muita dinamicidade, necessitando este último, de uma mudança de postura, representado no caso, pelos “momentos de loucura.” A necessidade de uma postura profissional dinâmica e criativa, pode ser o diferencial de uma unidade de informação, tornando-a mais atrativa, em detrimento ao fazer profissional do bibliotecário quando este se limita aos serviços meio (formação, desenvolvimento e organização de coleções).

Desta forma, Costa, Ramalho e Silva (2003, p. 153) asseveram que “[...] num tempo de estruturas e instituições sociais mutantes, exigentes de perfis e competências profissionais dinamicamente transformados, o profissional bibliotecário deve estar atento a estes novos paradigmas.”

### 3. Considerações finais

A representação da informação se configura como uma das principais atividades do fazer biblioteconômico. Um documento representado de forma adequada em um sistema de recuperação da informação impresso ou eletrônico, otimiza o processo de busca e recuperação da informação pelo usuário.

A Biblioteconomia é movida por informação, a qual está presente em diversos suportes e formatos. Entretanto, o livro é um dos elementos mais associados a esta área do conhecimento, pois este sempre esteve inserido na sociedade, seja em suporte mineral, animal ou vegetal, nos suportes impressos ou em meio eletrônico, ele perpassa todas as épocas e se adéqua as tecnologias vigentes.

Os quadrinhos são veículos de disseminação da informação, de extrema importância no campo da representação, pois o mesmo está presente na sociedade desde a época em que os homens viviam em cavernas e nelas faziam desenhos de suas caçadas e outras experiências, representando assim o seu cotidiano.

Apesar das histórias em quadrinhos terem enfrentado preconceitos pelas pessoas mais letradas, que consideravam os quadrinhos uma leitura desqualificada, não digna de admiração por intelectuais de determinada época, hoje os quadrinhos são fontes de informação importantes no processo de formação de leitores e na disseminação da informação, de forma prazerosa e lúdica, sobre um determinado assunto, inserido no contexto da vida cotidiana, representando diferentes áreas do conhecimento, como é o caso dos quadrinhos do Bibliocomics, que re-

presentam vários elementos das áreas de Biblioteconomia, Documentação e Ciência da Informação.

A análise dos quadrinhos do Bibliocomics nos mostra que estes são capazes de divulgar características inerentes ao bibliotecário, delineando aspectos culturais e sociais deste profissional.

Contudo, no que se referem aos quadrinhos do Bibliocomics, eles abordam o bibliotecário da seguinte forma:

Os quadrinhos focam-se no perfil do profissional bibliotecário, baseando-se nos estereótipos tradicionalmente conhecidos, nos quais infelizmente vem caracterizando há décadas este profissional, sendo este último representado como uma senhora ou um senhor austero, idoso, com óculos fundo de garrafa e que vive pedindo silêncio. No entanto, esta visão da profissão de bibliotecário está mudando, tendo em vista que nos quadrinhos é possível identificar a imagem do bibliotecário com o perfil mais descontraído e preocupado em atender de forma satisfatória o usuário, que nem sempre sabe ao certo o que procura em uma unidade de informação. A imagem estereotipada não condiz o perfil atual que esse profissional assume.

Notamos também que na atuação deste profissional na prática, deve-se ter um perfil adequado para se trabalhar nos serviços fins da biblioteca, pois este último corresponde a todas as atividades do Serviço de Referência e Informação (SRI) da biblioteca, a saber: atendimento, circulação, comutação bibliográfica, educação e estudo de usuários, divulgação dos produtos e serviços de informação, entre outros. O Serviço de Referência e Informação (SRI) é o cartão de visita de toda a unidade de informação, sendo que o profissional que trabalha nesta área deve ter facilidade em lidar com o público, que necessita de informação rápida e confiável.

No entanto, esta pesquisa instigou nos autores, perspectivas de estudos futuros, com a pretensão de aprofundamento no que diz respeito ao perfil do bibliotecário representado em charges ou filmes, observando a sátira e a crítica em relação às características comuns e estereotipadas dos bibliotecários em relação ao perfil atual destes profissionais.

## Notas

(1) Blog elaborado pelo Bibliotecário Alexandre Meireiros. Fonte: <http://bibliocomics.blogspot.com.br/>

## Referencias

- Cervo, A. L.; Bervian, P. A.; SILVA, R. (2007) *Metodologia científica*. 6.ed. São Paulo: Pearson Prentice Halls.
- Costa, Luciana Ferreira da; Ramalho, Francisca Arruda; Silva, Alan Curcino Pedreira da. Pela (IN) Formação Profissional: necessidades e perspectivas dos estudantes de graduação em Biblioteconomia/UFPB, em seu processo de conclusão. *Informação e Sociedade*. 13:2, (Jul./Dez.2003). 151-172
- Lima, José Leonardo Oliveira; Alvares, Lilian (Org.) (2012) Organização e representação da informação e do conhecimento. In: Alvares, Lilian (Org.). *Organização da informação e do conhecimento: conceitos, subsídios interdisciplinares e aplicações*. São Paulo: B4 Editores. Cap. 1. p. 21-48.
- Maciel, Alba Costa; Mendonça, Marília Alvarenga Rocha (2006). *Bibliotecas como organizações*. Rio de Janeiro: Interciência; Niterói: Intertexto.
- McCloud, Scott (2005). *Desvendando os quadrinhos*. São Paulo: Makron Books.
- Milanesi, Luís (2003). *A casa da invenção: biblioteca centro cultural*. 4.ed. São Paulo: Ateliê Editorial.
- Moura, Maria Aparecida (2006). Leitor- bibliotecário: interpretação, memória e as contradições da intersubjetividade em processos de representação informacional. In: Naves, Madalena Martins Lopes; Kuramoto, Hélio. (org.) *Organização da informação princípios e tendências*. Brasília, 2006. Cap. 2
- Santos, Maricelia Ferreira dos Santos (2011). Tirinhas de Humor: a imagem do profissional Bibliotecário nas histórias em quadrinhos. In: XXIV Congresso Brasileiro de Biblioteconomia, Documentação e Ciência da Informação. Maceió, *Sistemas de informação, multiculturalidade e inclusão social*. Disponível em: <<http://febab.org.br/congressos/index.php/cbbd/xxivqpapier/download>> Acesso em: 10 junho 2015.
- Silveira, Fabrício José Nascimento (2008). O bibliotecário como agente histórico: do humanista ao moderno profissional da informação. *Informação e Sociedade: Estudos*. 18:3 (Set./Dez.2008.) 83-94.
- Witter, Geraldina Porto (1986). Aspectos psicológicos no relacionamento bibliotecário e usuários. *Ciência da Informação*, (Jan./Jun.1986). 33-37.

## Anexo 1



Figura 3: Em busca de um perfil para o Serviço de Referência

## Anexo 2

## DA PESQUISA



Figura 4: A entrevista de referência

## Anexo 3

## CROCK e os legionários

Rechin &amp; Wilder



Figura 5: O Bibliotecário não leitor

## Anexo 4

## URBANO, o aposentado

A. Silvério



Figura 6: Vida de aposentada

## Anexo 5

## CROCK e os legionários

Rechin &amp; Wilder



Figura 7: Bibliotecário de férias

## Anexo 6

## GENTE COMO A GENTE

Mark Cullum



Figura 8: O problema da desinformação do usuário

## Anexo 7

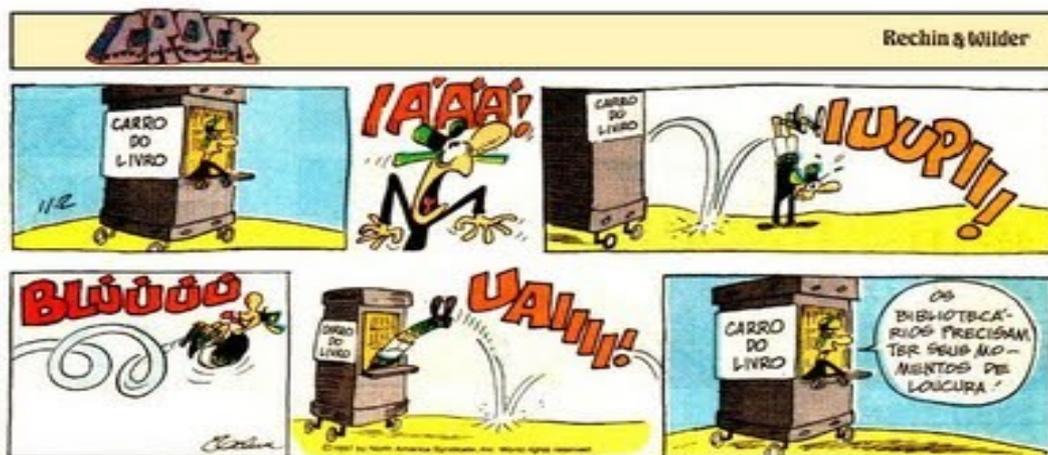


Figura 9: Bibliotecário extravasando as energias

---

---

# Estudo da Produção e Divulgação Científica do Instituto Federal de Educação, Ciência e Tecnologia de São Paulo: Uso do software livre Script Lattes

*Study of Production and Scientific Disclosure at the Federal Institutions of Education, Science and Technology of São Paulo: Using Free software Script Lattes*

---

**Santos, Marcel Pereira (1), Costa, Luzia Sigoli Fernandes (2)**

- (1) Mestre pelo Programa de Pós-Graduação em Ciência, Tecnologia e Sociedade da UFSCar – Campus São Carlos, São Carlos – Brasil. Bibliotecário-Documentalista do Instituto Federal de Educação, Ciência e Tecnologia de São Paulo – Campus Araraquara, São Paulo – Brasil. marcelantos@ifsp.edu.br  
Ramal de Acesso Eng. Heitor de Souza Pinheiro, s/n, Jardim dos Manacás, Araraquara – São Paulo, Brasil, CEP. 14801-600, telefone 55 16 3303-2341
- (2) Docente do Departamento de Ciência da Informação e do Programa de Pós-Graduação em Ciência, Tecnologia e Sociedade da UFSCar – Campus São Carlos, São Carlos – Brasil. luziasigoli@ufscar.br  
Rodovia Washington Luiz Km 235, São Carlos - São Paulo, Brasil, CEP. 13.565-905, telefone 55 16 3351-8351

## Resumo

No Brasil os Institutos Federais de Educação, Ciência e Tecnologia foram criados em 23 de dezembro de 1909 com a denominação "Escolas de Aprendizes Artífices", destinadas ao ensino profissional, primário e gratuito. Em 2008, transformaram-se em "Institutos Federais de Educação, Ciência e Tecnologia". Assim, com *status* de Universidade, trabalhando com a tríade "ensino-pesquisa-extensão". Atualmente, encontram-se dificuldades para pesquisar, localizar e principalmente centralizar da produção científica do Instituto Federal de São Paulo (IFSP), objeto de estudo. O objetivo foi realizar o levantamento da produção científica dos docentes do IFSP que atuam no ensino superior, através da utilização da bibliometria para o levantamento dos indicadores. Os indicadores foram apontados através de coleta realizada com o software livre *scriptLattes* que registra toda a produção dos currículos *Lattes* existentes na plataforma *Lattes*. A transformação em IFSP ainda não impactou no desenvolvimento da pesquisa institucional. Ressalta-se que o software *scriptLattes* foi uma ferramenta crucial para o a extração dos dados dos pesquisadores, de maneira que contribuiu amplamente com o desenvolvimento da pesquisa, neste sentido, investigações e desenvolvimentos acerca de sistemas de informação livres possibilitam que um número cada vez maior de pesquisadores possam ter acesso as tecnologias que certamente auxiliarão nas pesquisas científicas.

**Palavras chave:** Instituto Federal. Produção Científica. Indicadores. Bibliometria.

## Abstract

In Brazil, Federal Institutes of Education, Science and Technology were created in December 23, 1909, under the designation "Schools of artificer apprentices", and were destined to vocational, primary and free education. In 2008, they have turned into "Federal Institutes of Education, Science and Technology". Thereby, with University status, working with the "teaching-research-extension" triad. Currently there are challenges in searching, locating and centralizing the scientific production of the Federal Institute of São Paulo (IFSP), object of this study. The objective was to survey the scientific production of IFSP teachers who work in higher education, through the use of bibliometrics for the survey of the indicators. The indicators were identified through the collection performed by *scriptLattes*, a free software which lists all *Lattes* Curricula production in *Lattes* platform. The transformation into IFSP has not yet impacted on the development of the institutional research. It is noteworthy that *scriptLattes* software was a crucial tool for the extraction of data from researchers, so that it greatly contributed to the development of the research. In this sense, research and developments about free information systems enable a growing number of researchers to access the technologies which will surely assist in scientific researches.

**Keywords:** Federal Institute. Scientific Production. Indicators. Bibliometrics.

## Introdução

Os Institutos Federais de Educação, Ciência e Tecnologia, foram criados em 23 de dezembro de 1909, através do da publicação do Decreto nº 7.566, com a denominação “Escolas de Aprendizizes Artífices” (Fig. 1), o ensino oferecido foi de nível profissional e primário, os cursos eram gratuitos.



Figura 1. *Escola de Aprendizizes e Artífices*

Durante sua trajetória até o momento, os Institutos Federais passaram por várias transformações, em 1937 transformou-se em “Liceus Profissionais”, com cursos destinados ao ensino profissional de todos os ramos e graus, em 1942 passou a denominar-se “Escolas Industriais e Técnicas”, em 1959 “Escolas Técnicas Federais”. No ano de 1994 a Lei nº 8.948, transformou as Escolas Técnicas Federais em “Centros Federais de Educação Tecnológica – CEFETs”. A situação atual teve seu marco no ano de 2008, com a publicação da Lei nº 11.892, cujos Centros Federais de Educação Tecnológica transformaram-se em “Institutos Federais de Educação, Ciência e Tecnologia”, os popularmente conhecidos **Institutos Federais**, que passaram a ter status de universidade, portanto, desde 2008 trabalhando com a tríade: ensino, pesquisa e extensão.

O estudo tem como foco o Instituto Federal de Educação, Ciência e Tecnologia de São Paulo (IFSP), que atualmente possui trinta campi, possui aproximadamente 25 mil alunos matriculados.

Atuam no Ensino Médio (50% das vagas), nas licenciaturas (20% das vagas) e nos cursos superiores de tecnologia ou bacharelados tecnológicos (30% das vagas). Possuem estrutura multicampi e uma territorialidade definida, assumindo compromisso de intervenções em suas respectivas regiões, identificando problemas e criando soluções tecnológicas para o desenvolvimento sustentável, com inclusão social [...] Os IFs buscam responder aos anseios da sociedade no tocante às necessidades de investimentos na EPT. Com características peculiares, objetivam ofertar condições para a construção e desenvolvimento do saber, abrangendo as especificidades e territorialidade locais de

cada Instituto, na tentativa de um evoluir científico-tecnológico sustentável e consciente (SANTOS, 2012, p. 43-45).

Com uma história de cento e três anos, onde durante um centenário desta instituição (de 1909 a 2008) o IFSP possuía como foco principal o ensino profissional, distante das atividades de pesquisa e de extensão, atualmente mesmo regulamentada em seu Plano de Desenvolvimento Institucional (PDI), a pesquisa é uma vertente ainda pouco explorada e divulgada, pois se encontra em fase de implantação, divulgação e solidificação, portanto, objetivou-se discutir a importância da divulgação científica como ferramenta somática no processo de desenvolvimento da pesquisa institucional. Demonstrar a importância da divulgação científica para a comunidade científica e para sociedade de maneira geral.

Portanto, para a consolidação dos produtos oriundos das práticas de pesquisa dentro do IFSP, utilizou-se os conceitos acerca da divulgação científica, uma prática que proporciona a comunidade científica divulgar seus resultados, cujos resultados dessas práticas são fundamentais para a ampliação do conhecimento, compreensão e consolidação.

Ressalta-se que o IFSP ainda se encontra em fase de expansão e estruturação física e de recursos humanos, portanto é fundamental que sua comunidade científica se organize, mapeie ações já consolidadas nos diversos campi e utilizem amplamente da divulgação científica e suas vertentes para assim consolidar, repassar e divulgar tanto para comunidade interna quanto externa a pesquisa institucional.

### 1.1. O Instituto Federal de São Paulo (IFSP)

A Educação Profissional e Tecnológica (EPT), atualmente assim denominada, recebeu destaque no ano de 2009, diferentes atividades estão sendo desenvolvidas para que esta seja cada vez mais valorizada e reconhecida. Há por parte do Governo Federal iniciativas, tais como a realização de fóruns, eventos, e a expansão da Rede Federal de Educação Profissional e Tecnológica, que completou seu centenário no ano de 2009. O histórico disponível no site do IFSP (IFSP, 2011) apresenta a trajetória da instituição que compõe esta Rede Federal de Educação Profissional e Tecnológica, também conhecida como Rede Federal (Figura 2).



Figura 2. Reordenação do Instituto Federal de Educação, Ciência e Tecnologia

O IFSP foi criado inicialmente como Escola de Aprendizes e Artífices em 23 de setembro de 1909 com projeto inicial para formação de operários e contramestres, no período foram criadas no Brasil um total de dezenove escolas de Aprendizes e Artífices.

Dos anos 60 a 90 o IFSP era denominado Escola Técnica Federal de São Paulo, ministrava cursos técnicos de nível médio em Mecânica e Edificações, posteriormente passou a ministrar cursos técnicos em Eletrotécnica, Telecomunicações, Processamento de Dados e Informática Industrial. Em 1987 foi inaugurada a segunda Escola Técnica Federal de São Paulo no município de Cubatão (litoral paulista) e em 1996 foi inaugurada a terceira escola no município de Sertãozinho (interior do Estado de São Paulo). A partir de 2000 a instituição sofreu novas mudanças, transformando-se em CEFET-SP, acarretando sua reformulação e expansão.

A Lei nº 11.892 de 28 de dezembro de 2008, criou a Rede Federal de Educação Profissional, Científica e Tecnológica e também os Institutos Federais de Educação, Ciência e Tecnologia, esta Lei foi sancionada pelo Presidente da República, Luiz Inácio Lula da Silva:

Art. 1º. Fica instituída, no âmbito do sistema federal de ensino, a Rede Federal de Educação Profissional, Científica e Tecnológica, vinculada ao Ministério da Educação e constituída pelas seguintes instituições:

I - Institutos Federais de Educação, Ciência e Tecnologia - Institutos Federais;

II - Universidade Tecnológica Federal do Paraná - UTFPR;

III - Centros Federais de Educação Tecnológica Celso Suckow da Fonseca - CEFET-RJ e de Minas Gerais - CEFET-MG;

IV - Escolas Técnicas Vinculadas às Universidades Federais.

Parágrafo único. As instituições mencionadas nos incisos I, II e III do caput deste artigo possuem natureza jurídica de autarquia, detentoras de autonomia administrativa, patrimonial, financeira, didático-pedagógica e disciplinar.

Art. 2º. Os Institutos Federais são instituições de educação superior, básica e profissional, pluricurriculares e multicampi, especializados na oferta de educação profissional e tecnológica nas diferentes modalidades de ensino, com base na conjugação de conhecimentos técnicos e tecnológicos com as suas práticas pedagógicas, nos termos desta Lei (Brasil. Presidência da República, 2008).

A transformação dos CEFETs em Institutos Federais teve como base a necessidade de adequação à nova realidade da Rede Federal, assim como o estatuto atribuído aos Institutos Federais, tendo estes plena autonomia administrativa, financeira e pedagógica, equiparando-se às universidades federais, ou seja, os Institutos Federais trabalhariam além da vertente ensino, também as vertentes da pesquisa e da extensão.

No seu quadro, o IFSP possui: 980 docentes; 570 técnico-administrativos e 17.202 discentes (Tabela I).

| Nº de Discentes | Docentes       |                          | Técnico-administrativos |                          |
|-----------------|----------------|--------------------------|-------------------------|--------------------------|
|                 | Nº de Docentes | Relação Discente/Docente | Nº de Tec-adm           | Relação Discente/Tec-adm |
| 18.828          | 1.105          | 17,04                    | 598                     | 31,48                    |

Tabela I. Relação de alunos/professores e alunos/técnico-administrativos

Como parte das iniciativas do Governo Federal em torno da EPT em 2009, foi criada a Lei nº 11.940, de 19 de maio de 2009 que:

Art. 1º define em todo o território nacional o ano de 2009 como o ano da educação profissional e tecnológica;

Art. 2º estabelece o dia 23 de setembro como o dia nacional dos profissionais de nível técnico (Brasil. Presidência da República, 2009).

De acordo com Pacheco:

Os Institutos Federais de Educação Ciência e Tecnologia são a síntese daquilo que de melhor a Rede Federal construiu ao longo de sua história e das políticas de educação profissional e tecnológica do governo federal. São caracterizados pela ousadia e inovação, necessárias a uma política e um conceito que buscam antecipar aqui e agora as bases de uma escola contemporânea do futuro e comprometida com uma sociedade radicalmente democrática e socialmente justa. (Pacheco, 2008, p. 8).

A EPT está na agenda do Governo Federal, tem-se no momento a Expansão da Rede Federal de Educação Profissional, Científica e Tecnológica, somada a tantas outras ações e iniciativas para se discutir a EPT no Brasil, enxerga-se este um propício momento para discutir o papel da pesquisa e da extensão no IFSP.

## 1.2. A Produção Científica no IFSP

A mudança de CEFET para Instituto Federal é recente conforme explanado anteriormente, o Estatuto foi aprovado em 31 de agosto de 2009 de acordo com a Resolução nº 1 do Conselho Superior do IFSP, no Estatuto verifica-se:

Art. 4º O IFSP tem as seguintes finalidades e características: [...]

V. constituir-se em centro de excelência na oferta do ensino de Ciências, em geral, e de Ciências Aplicadas, em particular, estimulando o desenvolvimento de espírito crítico, voltado à investigação empírica. [...]

VII. desenvolver programas de extensão e de divulgação cultural científica e tecnológica;

VIII. realizar e estimular a pesquisa, a produção cultural, o empreendedorismo, o cooperativismo e o desenvolvimento científico e tecnológico;

IX. promover a produção, o desenvolvimento e a transferência de tecnologias sociais, notadamente as voltadas à preservação do meio ambiente. [...]

Art. 5º. O IFSP tem os seguintes objetivos: [...]

III. realizar pesquisas, estimulando o desenvolvimento de soluções técnicas e tecnológicas, estendendo seus benefícios à comunidade (IFSP, 2009).

Ao ser denominado Instituto Federal, o IFSP tem atribuída características e *status* de uma universidade federal, portanto, deverá se solidificar nas três vertentes: Ensino, Pesquisa e Extensão.

De acordo com o Plano de Desenvolvimento Institucional (PDI) do IFSP (2008) as atividades de pesquisa têm como objetivo formar recursos humanos para a investigação, a produção, o empreendedorismo e a difusão de conhecimentos culturais, artísticos, científicos e tecnológicos, sendo desenvolvidas em articulação com o ensino e a extensão, ao longo de toda a formação profissional.

Uma concepção que leve em conta a pesquisa e a produção de conhecimento como sustentáculo do processo educacional deve refletir sobre a natureza dessa pesquisa, a consecução de ações pedagógicas interdisciplinares, bem como as possibilidades de financiamento. A pesquisa está inserida dentro de determinados parâmetros históricos, sociais e culturais que forjam uma determinada concepção de mundo, que deve ser refletida pelos agentes que compõe a Instituição. Conforme Chizzotti (2001 apud Rodrigues, 2003), a pesquisa pode ser considerada como um esforço metódico em busca de informações para produzir conhecimentos no-

vos, ampliar a compreensão do mundo e auxiliar na solução dos problemas concretos que as pessoas e a sociedade enfrentam sobre questões não respondidas, inquietações que emergem da prática ou da reflexão, assim, deve constituir-se em uma atitude cotidiana e desmistificada, como também aponta Demo (1996).

Em um mundo dinâmico como o que vivemos, “acrescido constantemente de novas descobertas, de partilhas intensas, percepções discrepantes, posições divergentes e ações contraditórias”, os indivíduos obrigam-se a buscar novas explicações e a alcançar uma compreensão mais abrangente dos fatos e dos atos, quando das realidades individual e social em que cada um vive. [...] quando propicia o prazer da descoberta e a importância do conhecer, quando provoca a observação, mobiliza a curiosidade, move a busca de informações, esclarece dúvidas e a orienta as ações, o ensino ganha novo significado (Chizzotti, 2001, p. 106 apud Rodrigues, 2003, p. 367-368).

Portanto, discutir o papel da pesquisa no interior do IFSP é uma forma também de discutir qual o seu relacionamento com a sociedade, com o aprimoramento da ciência e, consequentemente, com o desenvolvimento do país.

Baseados no PDI, os norteadores da pesquisa dentro do IFSP (2008) são:

Pesquisa sintonizada com o Plano de Desenvolvimento Institucional;  
Pesquisa com função estratégica, perpassando todos os níveis de ensino;  
Desenvolvimento de pesquisas para o atendimento de demandas sociais do mundo do trabalho e da produção, com impactos nos arranjos produtivos locais e contribuição para o desenvolvimento local, regional e nacional;  
Estímulo à pesquisa comprometida com a inovação tecnológica e a transferência de tecnologia para a sociedade;  
Estímulo à pesquisa pura.

De acordo com Santos (2012):

percebe-se a oportunidade que os IFs possuem de propiciar ao país crescimento científico, tecnológico e social, por intermédio da realização de pesquisas aplicadas à sociedade, atividades de pesquisa, atividades de extensão, crescimento da produção científica, propostas sustentáveis para soluções de problemas (Santos, 2012, p. 45).

Entretanto, vale ressaltar que dentro da vertente pesquisa, obtêm-se como um de seus produtos a produção científica, que de acordo com Moura et al. (2002) caracteriza-se como um vetor importante para consolidação do conhecimento nas áreas do saber. A universidade é, portanto, o *locus* por excelência onde essa produção é gerada, advinda das pesquisas e estudos desenvolvidos no meio acadêmico, nos vários campos do conhecimento e tendo os Institutos

Federais equivalência às universidades, este também deverá ser o cerne, o ventre de produções científicas e tecnológicas relacionadas à EPT, às suas produções, avanços e inovações científicas e tecnológicas, tendo os Institutos Federais o princípio básico de estimular e auxiliar o desenvolvimento local de seu entorno.

## Procedimentos metodológicos

Para análise da produção científica realizada dentro do IFSP foram selecionados os docentes que atuam diretamente junto ao ensino superior, cuja amostra é de 427 professores (Tabela II). Tal seleção justifica-se devido ao estreitamento das relações do ensino superior com as práticas e programas da pesquisa científica, um dos objetos de estudo deste trabalho.

|                | Doutores | Mestres | Especialistas | Graduados | Total |
|----------------|----------|---------|---------------|-----------|-------|
| Nº de Docentes | 123      | 197     | 103           | 04        | 427   |
| %              | 29       | 46      | 24            | 1         | 100   |

Tabela II. Grau de especialização dos docentes que atuam no ensino superior

O mapeamento da produção científica do IFSP foi baseado no banco de currículos da Plataforma *Lattes*. A Plataforma *Lattes* representa a experiência do CNPq na integração de bases de dados de Currículos, de Grupos de pesquisa e de Instituições em um único Sistema de Informações (Conselho Nacional de Desenvolvimento Científico e Tecnológico, [s. d.]).

Para coleta dos dados foi utilizado o *software scriptLattes*. De acordo com Mena-Chalco e Cesar Jr. (2009), o *scriptLattes*, um sistema de código aberto, foi projetado para fornecer dados através de relatórios criados automaticamente. O *scriptLattes* pode ser adquirido gratuitamente através do site:

<http://scriptLattes.sourceforge.net>

Dado um grupo de pesquisadores cadastrados na plataforma *Lattes*, o *scriptLattes* coleta seus currículos *Lattes* do site do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), extrai as informações de interesse, elimina as produções científicas redundantes e cria relatórios sobre a produção, relatórios de supervisões acadêmicas, bem como gráficos.

Foi efetuada uma coleta teste dentre o grupo de docentes selecionados. Dos 427 docentes, 42 foram consultados para verificar o nível de atualização dos currículos *Lattes*. Assim, foi constatado que dos 42 docentes, 02 não possuíam currículos cadastrados; dos 40 que possuíam, 38 apresentavam atualizações entre o ano de 2012 e 2013.

Desta forma, com o ano de 2011 como limite para coleta dos currículos, garantiu-se que o total dos currículos *Lattes* dos docentes representaria o quadro real da situação da produção científica do IFSP.

De acordo com Mena-Chalco e Cesar Jr. (2009), o sistema operacional indicado para trabalhar com o *scriptLattes* é o *Linux*, que facilitará na execução do *software*.

Após a instalação, foi necessário a confecção de uma lista com os códigos chaves de todos os currículos *Lattes* dos docentes que constam na amostra de pesquisa, cujo código (Figura 3) é disponibilizado com o simples acesso ao currículo junto a Plataforma *Lattes*.



Figura 3. Ilustração do código chave do currículo *Lattes*

Da amostra inicial com 427 docentes, foram coletados 394 códigos chaves. Portanto, 33 docentes (7%) não possuem currículos registrados na Plataforma *Lattes*.

A coleta foi realizada manualmente. Nas instâncias em que o currículo não foi encontrado, verificou-se inicialmente se a grafia do nome estava correta. A verificação ocorreu através do Portal da transparência do Governo Federal, onde todos os servidores do IFSP estão registrados. Também se utilizou a opção de busca avançada, com as tentativas de busca com vínculo institucional; somente os primeiros nomes; área de atuação. Desta forma buscou-se restringir as chances do currículo estar cadastrado e não ser encontrado.

A coleta através do *scriptLattes* ocorreu às 21h56min do dia 17 de novembro de 2013, com duração de 33 minutos, com a utilização de uma conexão de internet banda-larga.

## Resultados

Para realização da análise dos resultados, os docentes do IFSP foram divididos em 2 grupos: docentes com ingresso na instituição no período de 1977 a 2007 (148 docentes – 37,6%) e docentes com ingresso entre 2008 e 2012 (246 docentes – 62,4%).

Esta divisão foi realizada com o objetivo de apontar as diferentes percepções dos docentes com maior tempo na instituição em relação aqueles com menor tempo.

Também se objetivou investigar se as publicações dos docentes com pouco tempo de instituição são produtos das pesquisas desenvolvidas no ambiente da própria instituição ou em outras, como é o caso de docentes que são alunos de pós-graduação.

Na coleta do *ScriptLattes*, estão inclusos no "Total da produção bibliográfica":

- Artigos completos publicados em periódicos;
- Trabalhos completos publicados em anais de congressos;
- Resumos expandidos publicados em anais de congressos;
- Resumos publicados em anais de congressos;
- Apresentações de trabalho.

Os gráficos em específico no período de 2008-2012 não foram descritos os índices de publicações dos anos 2006, 2007 e 2008, pois, nesse período os docentes não faziam parte do quadro de servidores do IFSP, desta forma, nenhuma de suas publicações foi realizada dentro da instituição.

O gráfico 1 demonstra uma tendência de crescimento no período de 2009 a 2011 (após a transformação em Instituto Federal) em que a média de publicações por docente passou de 1,14 no ano de 2009 para 1,39 no ano de 2011, crescimento de 21,9%. Mas, devido às várias formas de publicações, deve-se ter cautela com a análise deste crescimento.

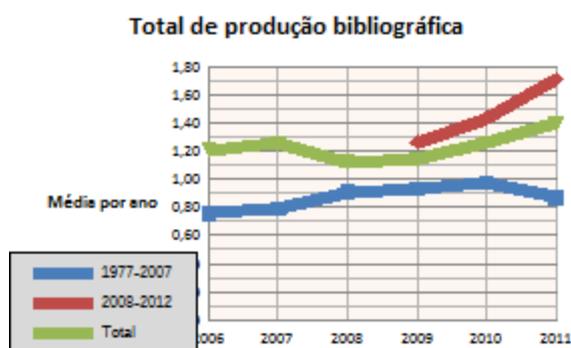


Gráfico 1. Total de produção bibliográfica

Destaca-se como fator negativo a queda de publicações no ano de 2011 entre os docentes que ingressaram no período de 1997-2007. Observa-se que a média de publicações por docentes foi 0,97 no ano de 2010 e 0,87 no ano de 2011, queda de 11,5%.

Referente aos artigos completos publicados em periódicos, os índices do gráfico 2 apresentam uma tendência negativa, pois no ano de 2011 ocorreu queda do total de publicações, com destaque para os docentes que ingressaram no ano de 2008-2012, em que a média de publicações por docente foi de 0,31 no ano de 2010 para 0,23 no ano de 2011, queda de 25,8%.

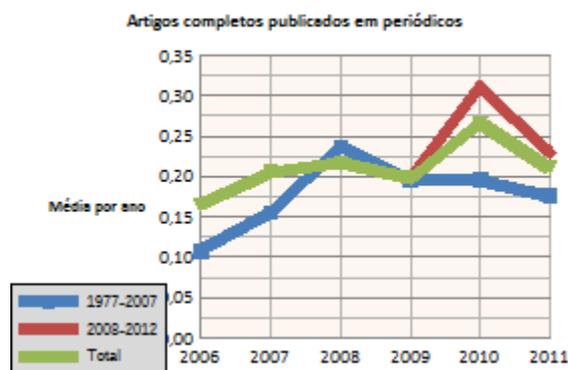


Gráfico 2. Artigos completos publicados em periódicos

Referente aos resumos publicados em anais de congressos, os índices do gráfico 3 apontam um crescimento após a transformação em IFSP no ano de 2008. Mas, em específico aos docentes com menor tempo de ingresso (2008-2012), principalmente no ano de 2011, neste mesmo ano, aponta uma queda entre os docentes com maior tempo de IFSP (1977-2007).

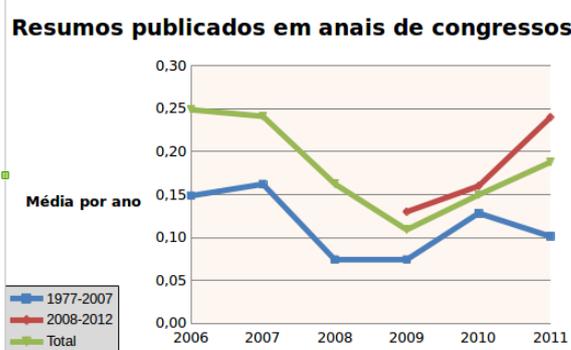


Gráfico 3. Trabalhos completos publicados em anais de congressos

Referente aos resumos expandidos publicados em anais de congressos, os índices do gráfico 4 apontam que, mesmo com o crescimento da média de publicação no ano de 2011 apresentado por docentes com menor tempo de ingresso (2008-2012), este tipo de publicação não é uma prática com ênfase dentro do IFSP. Neste mesmo ano, os docentes com maior tempo de ingresso (1977-2007) não apresentaram nenhuma publicação.

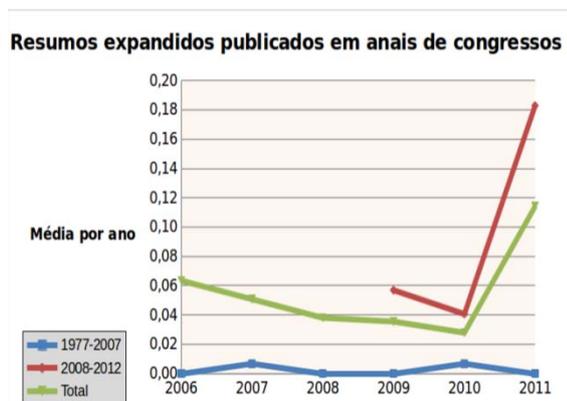


Gráfico 4. Resumos expandidos publicados em anais de congressos

Referente aos resumos publicados em anais de congresso, os índices do gráfico 5 demonstram que esta prática foi menos explorada após a transformação em IFSP, com queda no número de publicação realizada por todos os docentes analisados.

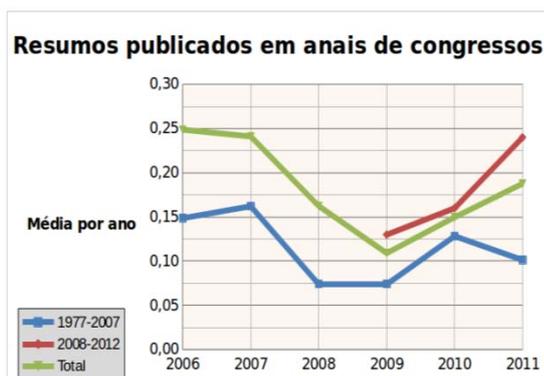


Gráfico 5. Resumos publicados em anais de congressos

Referente à apresentação de trabalho em eventos, os índices do gráfico 6 demonstram que, após a transformação em IFSP, ocorreu um crescimento na média de publicação por docentes que ingressaram recentemente (2008-2012) e uma queda dentre os docentes com maior tempo de ingresso (1997-2007). Verifica-se que não ocorreram grandes modificações do número total de publicações após o ano de 2008.



Gráfico 6. Apresentação de trabalho em eventos

## Considerações finais

Após as análises realizadas, conclui-se que a produção científica dentro do IFSP sofreu poucas modificações entre o período analisado antes e depois da transformação de CEFET-SP em IFSP que ocorreu em 29 de dezembro de 2008 através da sanção da Lei 11.892. O tempo de transformação da instituição ainda é pequeno e a reflexão sobre a mudança está em início. A instituição é centenária, e não será em apenas 5 anos que todas as mudanças serão percebidas e desenvolvidas pelo IFSP. Há a necessidade de tempo de maturação.

Ao analisar os indicadores da produção bibliográfica do IFSP entre os anos de 2009 e 2011, demonstrou-se que a tendência foi de crescimento, pois a média de publicações por docente passou de 1,14 no ano de 2009 para 1,39 no ano de 2011, apresentando um crescimento de 21,9%. Entretanto, mesmo apresentando um cenário positivo, dois fatores podem ser os responsáveis: as publicações serem frutos dos trabalhos realizados em programas de pós-graduação em outras instituições; e o crescimento das ofertas das bolsas de iniciação científica que ocorreu dentro do IFSP.

Quanto à metodologia da pesquisa, as ferramentas utilizadas demonstraram-se eficientes para coleta das informações. Entretanto, para que a coleta da produção bibliográfica realizada através do software *ScriptLattes* ocorresse de forma satisfatória, foi fundamental que os docentes tivessem seus currículos atualizados e, para garantir esta atualização de maneira consistente, o período de coleta foi fundamental (o ano limite para coleta foi 2011 e, assim, próximo a 2 anos da data da coleta).

Ressalta-se a importância da utilização do software *ScriptLattes*, uma ferramenta gratuita que proporcionou todas as condições para que se realizasse o levantamento e análise dos indicadores relacionados na pesquisa.

A utilização desta ferramenta indica que softwares livres podem e devem ser utilizados na contribuição do avanço das pesquisas científicas.

## Referencias

- Brasil. Presidência da República. Lei nº 11.182, de 29 de dezembro de 2008. Institui a Rede Federal de Educação Profissional, Científica e Tecnológica, cria os Institutos Federais de Educação, Ciência e Tecnologia. Brasília, DF, 29 dez. 2008. Disponível em: <[http://www.planalto.gov.br/ccivil\\_03/\\_ato2007-2010/2008/lei/111892.htm](http://www.planalto.gov.br/ccivil_03/_ato2007-2010/2008/lei/111892.htm)>. Acesso em: 20 set. 2010.
- Brasil. Presidência da República. Lei nº 11.940, de 19 de maio de 2009. Estabelece 2009 como Ano da Educação Profissional e Tecnológica e o dia 23 de setembro como o Dia Nacional dos Profissionais de Nível Técnico. Brasília, DF, 19 maio 2009. Disponível em: <<http://www.leidireto.com.br/lei-11940.html>>. Acesso em: 19 set. 2010.
- Conselho Nacional de Desenvolvimento Científico e Tecnológico. Sobre a plataforma *Lattes*. Brasília-DF, [s.d.]. Disponível em: <<http://Lattes.cnpq.br/conteudo/aplataforma.htm>> Acesso em: 10 jul. 2012.
- Demo, P. *Pesquisa: princípio científico e educativo*. 5. ed. São Paulo: Cortez, 1996. 120 p.
- Instituto Federal de Educação, Ciência e Tecnologia de São Paulo (IFSP). Conselho Superior. Plano de Desenvolvimento Institucional. São Paulo, 2008. Disponível em: <[www.ifsp.edu.br](http://www.ifsp.edu.br)>. Acesso em: 20 set. 2010.
- Instituto Federal de Educação, Ciência e Tecnologia de São Paulo (IFSP). Histórico da Instituição. São Paulo, 2011. Disponível em: <[http://189.108.236.229/internet/index.php?option=com\\_content&view=article&id=73&Itemid=59&limitstart=2](http://189.108.236.229/internet/index.php?option=com_content&view=article&id=73&Itemid=59&limitstart=2)>. Acesso em: 20 jul. 2011.
- Moura, A. M. S; Mattos, C. V; Silva, D. C. Acesso e recuperação da produção científica pela biblioteca universitária: os Anais de Eventos. In: Seminário Nacional De Bibliotecas Universitárias, 12., Rio de Janeiro, 2002. *Anais...* Rio de Janeiro: UFF, 2002.
- Pacheco, E. *Os Institutos Federais: uma revolução na educação profissional e tecnológica*. 2008. Disponível em: <[http://www.ufpel.edu.br/cavq/noticias/arq/1\\_cartilha\\_institutos.pdf](http://www.ufpel.edu.br/cavq/noticias/arq/1_cartilha_institutos.pdf)>. Acesso em: 06 jan. 2011.
- Rodrigues, M. E. F. A pesquisa no ensino e o ensino na pesquisa. *Transinformação*, Campinas, v. 15, n. 3, p. 363-372, set./dez. 2003. Disponível em: <<http://revistas.puc-campinas.edu.br/transinfo>>. Acesso em: 04 jul. 2012.
- Santos, C. A. S. *As unidades de informação dos Institutos Federais no apoio ao desenvolvimento da ciência e da tecnologia: um estudo de percepção sociocognitiva com o uso do protocolo verbal em grupo*. 2012, 248 f. Dissertação (Mestrado em Ciência, Tecnologia e Sociedade)-Centro de Educação e Ciências Humanas, Universidade Federal de São Carlos, 2012.

---

# Cognição, neuromarketing e neuroeconomia: entendendo as decisões dos atores baseados em informações

*Cognition, neuromarketing and neuroeconomics: understanding the decisions of information-based actors*

---

**Roberto Vilmar Satur (1), Dulce Amélia de Brito Neves (2), Emeide Nóbrega Duarte (3)**

(1) Universidade Federal de Paraíba, Dirección postal Rua Itamar Neiva Monteiro, n. 32, Bairro Água Fria, João Pessoa - Paraíba - Brasil, Cep 58074.000), Correo electrónico (robertosatur@yahoo.com.br)

(2) Dirección postal Rua Rua Saffa Said Abel da Cunha, n.38, Bairro - Tambauzinho, João Pessoa - Paraíba - Brasil, Cep 58042-220), Correo electrónico (damelia1@gmail.com)

(3) Dirección postal (Rua Aderbal Maia Paiva, s/n, Bairro: Portal do Sol, João Pessoa - Paraíba - Brasil, CEP:58.046-527), Correo electrónico (emeide@hotmail.com)

## Resumo

Este artigo apresenta uma discussão baseada em pesquisa teórica que objetiva compreender e discutir sobre o comportamento dos indivíduos quanto ao uso das informações para o estabelecimento de prioridades e preferências, realização de escolhas e tomada de decisão, contextualizando os estudos cognitivos, além da Neurociência e da Psicologia aplicada à Economia e ao marketing. Considerando o estudo apresentado, é possível perceber a importância dessas novas abordagens, especialmente da Neuroeconomia e do Neuromarketing, para entender o comportamento dos agentes. Isso decorre, especialmente, do fato de as teorias tradicionais não conseguirem explicar, a contento, o comportamento dos indivíduos em relação às suas preferências, às escolhas e às decisões. O estudo aponta avanços e contribuições significativos para se entender o comportamento dos indivíduos no uso da informação e quando precisam tomar decisões.

**Palavras-chave:** Informação para decisão. Cognição. Neuromarketing. Neuroeconomia.

## 1. Introdução

Há algum tempo, a Ciência vem fazendo caminhos entrelaçados, especialmente no que se chama de interdisciplinaridade e transdisciplinaridade, visando melhorar o desempenho em relação à busca de respostas convincentes que inquietam os pesquisadores. No caso da Economia e do Marketing, não é diferente. A permanente busca de entender cada vez mais o comportamento do consumidor ou agente econômico mostrou que a análise puramente racional não dava respostas suficientes. Por essa razão, começou a busca de novas ferramentas, novos recursos e novas ciências. Nesse sentido, a ciência da Cognição recebeu o reforço da

## Abstract

This paper presents a theoretical research-based discussion that aims to understand and discuss the behaviour of individuals in the use of information for establishing priorities and preferences, conducting choices and decision making. It also contextualises cognitive studies, as well as the neuroscience and psychology applied to economics and marketing. As a result, it is possible to realize the importance of these new approaches, especially neuroeconomics and neuromarketing to understand the behaviour of agents. This is especially due to the fact that the traditional theories fail to respond well to individuals' behaviour, preferences, choices and decisions. The study shows progress and significant contributions to the understanding of the behaviour of individuals in the use of information for decision.

**Keywords:** Information for Decision. Cognition. Neuromarketing. Neuroeconomics.

Neurociência, e esse conjunto de estudos começou a ser utilizado na Economia e no Marketing, gerando o que se chama de Neuroeconomia e Neuromarketing.

Quanto à metodologia de estudo, este artigo apresenta uma revisão teórica das ciências da Cognição, da Neuroeconomia e do Neuromarketing, focados na informação, visando entender o ator, o agente, o cidadão e o consumidor e suas reações diante das situações e das possibilidades do mercado, do consumo, do marketing, da informação disponível e do estímulo emocional e, a partir disso, quais as escolhas e as preferências desse ator.

Nessa perspectiva, este artigo começa com uma breve conceituação das Ciências cognitivas, da Neuroeconomia e do Neuromarketing. Num segundo momento, discute sobre a questão das escolhas e das preferências, nas perspectivas da Economia e do Marketing, e, em seguida, sobre a Cognição e a Informação como suportes de escolhas. Finalmente, tece algumas considerações sobre questão da escolha e da informação, na perspectiva do Neuromarketing e da Neuroeconomia, ou seja, da Neurociência, para que se possa entender o comportamento do ator.

## 2. Entendendo comportamentos, escolhas e preferências sob olhar da cognição, neuroeconomia e neuromarketing

O que a Cognição, a Microeconomia e o Marketing têm em comum? O que, efetivamente, contribui para que o ator (indivíduo) faça escolhas e tome decisões? Na verdade, todas essas teorias buscam entender o comportamento dos atores e quais as informações e os estímulos que os motivam em suas escolhas. Segundo Rafael Capurro (2003, p.10),

[...] os conteúdos intelectuais formam uma espécie de rede que existe somente em espaços cognitivos ou mentais, e chama tais conteúdos de informação objetiva [...] sujeito cognoscente possuidor de "modelos mentais" do "mundo exterior" que são transformados durante o processo informacional. [...] a busca de informação tem sua origem na necessidade ("need") que surge quando existe o mencionado estado cognitivo anômalo, no qual o conhecimento ao alcance do usuário, para resolver o problema, não é suficiente.

Assim, o sujeito procura processar e significar a informação, de modo que ela possa subsidiá-lo nas decisões e nas escolhas. Para Mônica Borges *et al.* (2004, p.85)

[...] a informação necessariamente deveria ser representada na mente do sujeito, pois a mente não "pertence" ao mesmo mundo real objetivo, assim tudo que resta é representá-lo. No entanto, apesar de ser visto como um "processador de informação", o cérebro humano era considerado "o melhor" processador de informação que existe.

Esforços das teorias tradicionais para entenderem o comportamento dos atores historicamente têm sido empreendidos haja vista, por exemplo, que a microeconomia sempre estudou o consumidor. O problema é que essa teoria enveredou para modelos matemáticos racionais, visando entender esse comportamento. Enquanto isso, o marketing também estuda exaustivamente o consumidor, especialmente o seu comportamento. Nesse caso, o problema é que focou sobremaneira em pesquisas empíricas sobre

esses cidadãos e suas reações. Assim, apesar de todos os esforços, os resultados ainda eram considerados insuficientes. Além do mais, o próprio cidadão foi aprimorando e mudando suas formas de exercitar suas escolhas, preferências e decisões. Fabian Furlaneto *et al.* (2013, p. 106), cita Cardoso e Pinto (2010):

[...] o ato de consumir tem se tornado mais complexo ao longo dos tempos, evidenciando frequentemente aspectos que superam a "ida as compras", cuja importância vem se tornando gradativamente menor conforme disparam as compras pela internet. [...], o consumo de produtos pode ser entendido tanto como uma atitude utilitária, racional, como enquanto atividade hedônica, envolvendo experiências emocionais e/ou sensoriais. [...] dessas dimensões no domínio das compras virtuais, em que o impulso para o consumo se faz sensível a variáveis tão diversas como a luminosidade da tela do computador ao horário.

Para entender as reações dos atores, os pesquisadores também foram aprimorando suas teorias e buscando suportes complementares na interdisciplinaridade, como pode ser visto na citação que segue de Moreira *et al.* (2011, p. 99):

Dentro do arcabouço teórico da economia e administração, ramos como a microeconomia e o marketing foram os primeiros a se interessarem pelo tema. Entretanto, as simples modelagens matemáticas da microeconomia ou a observação empírica do marketing não garantem a compreensão absoluta de todas as decisões. Pela carência de maiores detalhes sobre o real comportamento humano, novas áreas de estudos emergiram. Aliando conhecimentos de economia e marketing aos da psicologia, neurociência e outras ciências cognitivas, surgiram a Neuroeconomia e o Neuromarketing. [...] duas novas fronteiras de estudo [...] que [...] prometem revolucionar a compreensão acerca das preferências e escolhas dos indivíduos.

Essa tendência neurocientífica nas pesquisas em marketing e economia "segundo Furlaneto *et al.* (2013, p.106) representam o ponto mais alto no desenvolvimento técnico em sentido à maximização informacional". No entanto são campos novos.

Segundo Moreira *et al.* (2011) a Neuroeconomia surgiu na última década do Século XX. Shizgal e Conover (1996), e Platt e Glimcher (1999), mesmo sem citarem o termo neuroeconomia, foram os primeiros a mencionar o substrato neurobiológico do comportamento de escolhas e relacioná-lo à teoria de escolhas do consumidor da microeconomia. Já Camerer, Lowenstein e Prelec começam a estudar e a desenvolver essa nova teoria, na perspectiva da economia comportamental, e ligá-la à construção da abor-

dagem neurobiológica para compreender fenômenos econômicos que resultam no termo neuroeconomia.

Ferreira *et al.* (2014, p.8), cita Lowenstein, Camerer e Prelec (2008) que afirmam,

[...] os alicerces da teoria econômica foram construídos supondo que os detalhes sobre o funcionamento do cérebro, a “caixa-preta” fundamental, nunca seriam conhecidos, entretanto, nas duas últimas décadas, a economia comportamental, que nada mais é que a importação de ideias da psicologia para a economia vem modificando essa crença. Algumas ideias estão provando que o estudo do cérebro e do sistema nervoso começa a possibilitar a medição direta dos pensamentos e sensações, o que tende a mudar nossa compreensão da relação entre mente e ação e, assim, a interferir diretamente na economia, surgindo assim a “Neuroeconomia”.

Recorrendo aos escritos de Hill e Martin (2014) pode-se dizer que essa nova forma de estudar o comportamento dos atores significa que a teoria econômica passou a aceitar que a natureza multifacetada do consumo inclui aspectos sensoriais, emocionais e experiências de uso em seu comportamento de escolhas e decisões. O Marketing também se rendeu a esses estudos, sobretudo considerando que, por exemplo, existem consumidores de marca, que a seguem, compram-na e endeusam-na e lhe dão atributos e valores muito além do que é efetivamente o próprio produto da referida marca. Ou seja, quem compra um produto ou serviço de determinada marca parece estar comprando sua autodefinição, e os argumentos para tal variam, mas tendem a ficar entre algum ou alguns dos que seguem: gera emoção, é sexy, “todo mundo” o tem, é a moda do momento, somente a elite pode comprar, é inteligente e é a última tecnologia. O desenvolvimento da moral (cognitiva) é importante para impedir, inclusive, que algumas tentadoras “táticas de marketing” feitas por profissionais que violam inclusive questões éticas acabem influenciando o consumidor, especialmente em seus pontos vulneráveis.

O Neuromarketing, que, juntamente com a Neuroeconomia, vem estudando todas essas questões, surgiu, segundo Moreira *et al.* (2011), igualmente, no final do Século XX. No entanto, não há consenso sobre seu efetivo criador. Alguns atribuem tal fato a David Lewis, que, desde a década de 1980, estudam atividades cerebrais em resposta a comerciais de televisão, fato que o tornou famoso e um sucesso como consultor na área, atendendo a empresas como a British Airways, a Hewlett a Packard, a IBM, a Mcdonalds e a Toyota. Outros apontam Gerald Zaltman, por ter feito escaneamento cerebral de

voluntários relacionando suas reações às suas preferências por produtos e marcas. Também há quem atribua tal criação a Ale Smidts, que adotou o uso de tomografia de ressonância magnética funcional e outras ferramentas para fazer estudos de marketing e comportamento das pessoas quanto a escolhas e a estímulos.

Atualmente, segundo Furlaneto *et al.* (2013) além da ressonância, outras técnicas, como eletroencefalografia (EEG), registros eletrofisiológicos, estudos de rastreamento ocular (já que a retina é parte do sistema nervoso), também são adotadas nesses estudos. Por exemplo, a partir dos estudos de rastreamento ocular (*eye tracking*), foi possível melhorar o *design* dos veículos e certos direcionamentos de propagandas.

Estudos de marca e estímulos do cérebro sempre intrigaram e instigaram estudiosos. Segundo Moreira *et al.* (2011) usando por referência Pirouz (2004 e Botelho (200\*) talvez o mais famoso desses estudos foi feito pelos pesquisadores de *Baylor College of Medicine* (EUA), onde os participantes receberam refrigerantes da Coca-cola e da Pepsi, enquanto suas imagens cerebrais eram registradas. Quando tomaram os dois refrigerantes sem identificação da marca, as preferências se dividiram. Na segunda experiência, já com as marcas identificadas, as preferências pela Coca-cola subiram para 75%. As amostras sem identificação ativaram a região do cérebro responsável pela recompensa na metade dos participantes, enquanto nas amostras identificadas a região acionada foi o hipocampo, que está relacionado à memória, ou seja, as escolhas não eram pelo que mais agradava (recompensa), mas pelo que socialmente considerava melhor ou a que estava acostumado.

Outros estudos reafirmaram esses experimentos e fizeram novas constatações dessa ordem. Foram ativadas diversas regiões do cérebro, na hora da escolha, tanto nas relacionadas à recompensa social quanto as áreas límbicas que apresentavam diferenças entre gêneros. Inclusive se perceberam influências relacionadas aos aspectos hormonais e às escolhas intertemporais. Também houve outras descobertas sobre as diferenças quanto ao comportamento e às reações de homens e mulheres, o que se pode relacionar ao formato e ao tamanho de algumas partes do cérebro e à ação dos hormônios sexuais. Para Moreira *et al.* (2011, p.106-7):

Os hormônios são mensageiros bioquímicos que interferem nos mecanismos do cérebro fazendo-o agir dessa ou daquela maneira [...]. diferenças de dimensão, assimetria e lateralização dos cérebros masculinos e femininos fazem com que homens e

mulheres tenham uma percepção distinta e acessem áreas distintas na hora das escolhas.

Assim, um dos conceitos muito válidos, em se tratando de Neuromarketing, pode ser o que foi construído por Braidot (2005) e rerepresentado por Furlaneto *et al.* (2013, p.109):

O Neuromarketing pode ser entendido como uma disciplina moderna, produto da convergência das neurociências ao marketing. [...] O Neuromarketing se apoia em tecnologias de diagnóstico para poder identificar padrões de atividades cerebrais que revelam os mecanismos internos dos indivíduos, enquanto expostos a determinados estímulos.

Devido aos estudos citados acima, fica mais fácil entender que não é plenamente verdadeira a ideia de que os atores fazem escolhas ótimas baseadas na racionalidade, como afirma a teoria tradicional. Ou seja, para um ator, também pode ser considerada uma escolha ótima aquela que mistura emoção, racionalidade, cognição, confiança, entre outros.

Apesar de todas essas experiências e importantes contribuições relatadas, constata-se que, na América Latina e no Brasil, os estudos de Neurociência e suas aplicações em Neuroeconomia e Neuromarketing ainda são muito incipientes. Como afirmam "Furlaneto *et al.* (2013, p. 108), o único laboratório de neurociências brasileiro de ponta com linhas de estudo e professores dedicados à neuroeconomia e à neuromarketing é o Laboratório de Neurociências Clínicas da Unifesp". Esse é campo ainda em aberto para estudos.

### 3. Escolhas: sob a visão do marketing e da economia

Os objetivos da Teoria Microeconômica e do Marketing na questão do consumidor "é de entender o comportamento dos indivíduos em tomadas de decisões, entre eles, aqueles relacionados ao consumo". A Teoria Microeconômica, que foi a primeira a surgir com esse contexto, segundo Moreira *et al.* (2011, p.100),

[...] utiliza-se de modelos matemáticos para formular as preferências e escolhas do consumidor. A partir de axiomas iniciais a teoria microeconômica representa matematicamente e graficamente como deveriam ser as preferências dos agentes. Construindo funções utilidade e gráficos como as curvas de indiferença, cria-se um perfil de consumo, que muitas vezes se assemelha mais a um perfil médio do que a verdadeira preferência do consumidor individual.

Para a teoria neoclássica da Economia, pensamento predominante da microeconomia, os agentes levam em conta alguns princípios para

suas decisões e escolhas, que segundo Mankiw *et al.* (2005) e Vera (2013) são: a) os recursos são escassos e as necessidades são ilimitadas; b) para optar por algo, o consumidor renuncia outras e considera o custo da oportunidade disso; c) os agentes aprendem com os erros do passado e maximizam os resultados do futuro, com base em todas as informações disponíveis, devidamente analisadas de forma racional; d) os agentes consideram os incentivos e as facilidades oferecidas, respondendo a elas. Para garantir tudo isso, admite-se que as expectativas sejam transitivas, convexas e contínuas. Para isso, o agente deve ter pleno conhecimento das alternativas disponíveis, calcular todos os possíveis resultados, a partir do universo de escolhas possíveis, e tomar uma decisão que minimize seus riscos.

Vera (2013) cita Simon *et al.* (1997) pois este alerta que esses pontos dificilmente são garantidos na prática, sobretudo porque as informações disponíveis para a decisão não são completas (assimétricas) sobre todas as alternativas disponíveis. Logo, a racionalidade é limitada, e as decisões têm altos níveis de erros. Essa limitada capacidade de adquirir informações e conhecimentos que o ser humano enfrenta é considerada o ponto fraco da suposta racionalidade defendida pelos modelos neoclássicos, e isso leva os agentes a não conseguirem ter critérios e informações plenamente claras para tomar decisões ótimas.

Os temas comportamento, preferências, escolhas e decisões dos atores (especialmente os consumidores) cresceram em importância com as teorias do Marketing, que focam o ator como consumidor e seu comportamento ao decidir o consumo. Segundo Moreira *et al.* (2011) essa é uma visão prática e de casos particulares. Assim, constrói uma visão de preferência que, quase sempre, o próximo indivíduo não tem.

Baseados no que foi apresentado até aqui, podemos dizer que a Microeconomia e o Marketing são áreas distintas, mas têm objetivos comuns, como, por exemplo: estudar o comportamento dos atores, em seus processos de decisão e de escolha, quando têm à disposição muitas opções, informações assimétricas e um ambiente de incertezas, constituindo formas e modelos valiosos como ponto de partida, mas limitados em um contexto mais amplo para, efetivamente, entender-se o indivíduo e seu comportamento cognitivo e emocional de escolhas e preferências. Foi nessa perspectiva que a Economia e o Marketing avançaram desenvolvendo novas áreas de estudo, para tentar desvendar como, nas questões de preferências e de escolhas, "segundo Moreira (2011, p.100), o

cérebro desse indivíduo recebe, processa as informações e o condiciona a uma decisão. Essas duas novas áreas de estudos são denominadas de Neuroeconomia e Neuromarketing”

Damasio citado por Moreira *et al.* (2011) admite, que, possivelmente, os axiomas previstos na teoria do consumidor nem sempre são válidos ou que as “curvas de indiferença” não sejam tão indiferentes ou nem tão curvas, como previsto pela teoria tradicional. É possível afirmar que nem todas as variáveis que influenciam escolhas são fáceis de mensurar, que o indivíduo pode não ser tão racional nas escolhas, como quer prever a teoria tradicional, e, tampouco, que os indivíduos se comportam de forma uniforme. Logo, toda generalização é perigosa.

No entendimento de Moreira *et al.* (2011, p.102),

a diversidade de preferências individuais [...] a concorrência acirrada, a infinidade de novos produtos e serviços, a comodidade da compra pela internet e os apelos sociais que condicionam o comportamento humano.

São este apenas alguns exemplos que mostram que o Marketing também precisa estar se atualizando constantemente para dar conta de toda a complexidade que envolve o ambiente de escolhas dos consumidores. Para Bogazzi (1974) *apud* Burgueti (2004) é preciso ouvir vozes internas e externas ao indivíduo, interpretar culturas, entender entornos e contornos e a mente humana e sua capacidade de cognição, que o conduz a fazer escolhas e a tomar decisões. É preciso entender que o ser humano não faz somente trocas utilitaristas racionais, mas também simbólicas, ou seja, é mais complexo do que se imagina.

É preciso ingressar em novas fronteiras desse entendimento, as quais têm relação direta com a cognição, a emoção e outras variáveis humanas, ou seja, elas têm nome e sobrenome: Neuroeconomia e Neuromarketing. Por essa razão, o próximo passo será o de estudar um pouco mais sobre Cognição, Neuroeconomia e Neuromarketing, buscando identificar o que, de fato, consegue prender a atenção dos indivíduos e que será, efetivamente, o suporte que balizará as decisões de escolha.

#### 4. Cognição e informação como suportes para a escolha

A cognição é considerada como consciência, logo, se a pessoa consegue identificar os motivos de determinada reação ou ação, significa que está consciente da ação tomada. Segundo Neves (2012, p.46),

A cognição humana apresenta-se como uma reprodução dos conhecimentos estocados na memória. Todos os dias, desempenhamos tarefas que se repetem por longo tempo, possibilitando a automatização de gestos de procedimentos. Usamos diversas estratégias que nos permitem desempenhar nosso papel na sociedade. O modo como desenvolvemos as atividades cotidianas pode ser expresso em uma rede de conceitos em forma de mapas. Os mapas cognitivos representam o raciocínio humano dando visibilidade ao pensamento estruturando as ideais viabilizando a emergência da subjetividade.

É pelos modelos mentais que trabalhamos as informações de forma racional, afetiva ou emocional. Para Ferreira *et al.* (2014, p.12):

Os modelos mentais têm um papel crucial tanto no nosso pensamento, quanto em nossas ações, e que o contexto social reflete como causa ou efeito a forma que vemos e nos deparamos com a realidade, isso abre ou limita nossas possibilidades de ação.

Tal perspectiva remete ao estudo da cognição. Neves (2006) afirma que a ciência que estuda a cognição tem como principais modelos mentais: os esquemas (ex.: procedimentos para fazer funcionar um aparelho), os planos (ex.: o que fazer para vencer uma partida de xadrez), os roteiros ou *scripts* (ex.: para ir almoçar em um restaurante, é preciso fazer vários procedimentos até chegar lá: vestir-se, tirar o carro da garagem, deslocar-se no trânsito, procurar estacionamento, desembarcar, dirigir-se ao restaurante, selecionar uma mesa e, finalmente, almoçar. Depois, é necessário fazer o procedimento normalmente inverso para o retorno para casa), e as superestruturas ou esquemas textuais (ex.: estudamos exaustivamente algo, todas as suas alternativas e as comparamos para, depois, decidir). São esquemas estudados separadamente, mas, na prática do indivíduo, são simultâneos e não divididos.

Com base em uma interpretação de Senge (2000), Ferreira *et al.* (2014, p.4) asseveram que,

na cognição, o termo “modelo mental” refere-se tanto aos mapas tácitos semipermanentes do mundo que as pessoas retêm em sua memória de longo termo, quanto às percepções de curto prazo que as pessoas constroem como parte dos seus processos diários de raciocínio. De acordo com alguns teóricos da cognição, mudanças nos modelos mentais diários de curto prazo, acumulando-se com o tempo, gradativamente se refletem em mudanças nas crenças arraigadas de longa duração.

As práticas de comunicação e de entretenimento exigem várias habilidades de caráter cognitivo. As mudanças ocorridas nesse entendimento

têm levado mais em conta os processos cognitivos dos usuários que envolvem, segundo Regis *et al.* (2012): a) mais participação em atividades de construção colaborativa de conteúdos nos ambientes que permitem interação social; b) crescimento grandioso do volume de informações distribuídas por diversas plataformas, tendo o usuário como investigador e que precisa ser selecionado rigorosamente devido à abundância e à fragmentação existentes; c) necessidade de selecionar e ordenar atividades considerando o excesso de demanda e de estímulos e d) necessidade de aprender linguagens diversas, programas e códigos mediáticos típicos da cultura digital.

A cognição envolve os modelos ou esquemas mentais dos indivíduos para representarem o ambiente social a partir de suas interações e durante a aprendizagem (Neves, 2012). Logo, os ambientes de comunicação, numa perspectiva digital e de cibernética, alteram o ambiente e reconfiguram comportamentos e interações. São três os modelos cognitivos contemporâneos que justificam a comunicação na cibercultura: cognição integrada, cognição encadeada e cognição distribuída. A seguir, discute-se, mais detalhadamente, sobre esses modelos baseados na abordagem de Regis *et al.* (2012):

#### a) cognição integrada

Com as várias tecnologias disponíveis, especialmente diversas plataformas de acesso e a tecnologia digital, favorecidas pelos hiperlinks da Internet, está-se criando o hábito de fazer buscas fragmentadas de informações usando esses diversos meios e fontes. Tal diversidade de alternativas e de equipamentos de acesso gera informações em excesso e fragmentadas, o tempo todo. Assim, fica difícil saber qual informação selecionar e em qual delas se deter. Diante de tanta informação, o olhar é superficial e só nos atemos ao que mentalmente nosso contexto cognitivo considera mais relevante para resolver um problema do momento ou que poderá ser útil no futuro, segundo a cognição de quem visualiza. O restante passa despercebido. Assim, por exemplo, quando se quer comprar um relógio, olha-se para seu formato, sua pulseira, sua cor e outros detalhes. Porém, quando se quer saber a hora, olha-se somente a hora e nem se percebe, por exemplo, que a pulseira, hoje, tem outra cor. No modelo de cognição integrada, a atenção é o fator fundamental, que funciona como forma de selecionar e de direcionar o processo para perceber e agir no ambiente (Regis *et al.* 2012).

William James (1890, *apud* Regis *et al.*, 2012) já dizia que a atenção é a opção de se privar de

algumas coisas para cuidar de modo mais eficaz que outras, portanto, oposto de distração. Nesse caso, é possível focar no conteúdo para encontrar a informação que se quer. Já para Jonathan Crary (2001, *apud* Regis, 2012), a atenção é a capacidade de se desconectar de um ambiente amplo de estímulos e se concentrar em um número pequeno de fenômenos mais detidamente. Ou seja, falam-se da capacidade de integrar os sistemas perceptivos e motor, que leva à atenção seletiva, por meio da qual se podem escolher os conteúdos que interessam em meio a tanta abundância;

#### b) cognição encadeada

Para Regis *et al.* (2012) com as TICs, a Internet e a cibercultura têm aumentado substancialmente as opções e a capacidade de realizar atividades múltiplas, tudo ao mesmo instante, especialmente as práticas socioculturais, rotineiras e diárias e as práticas comunicativas. Somos multimeios, multitarefas, multifuncionais, acumulando cargos e encargos, fazendo sempre mais de uma coisa ao mesmo tempo. Estamos fragmentados ao redor de muitas tarefas, pouco concentrados em alguma específica. Assim, o desvio de atenção é frequente. No passado, ficava-se uma hora, um turno, um dia em uma tarefa que, quando terminada, em outra hora, em outro turno ou, em outro dia, começava-se outra. Na atualidade, faz-se de tudo ao mesmo tempo. Pesquisas demonstram que os trabalhadores de escritório têm conseguido se concentrar apenas três minutos, em média, por tarefa, sem serem interrompidos por outro estímulo externo (chamada telefônica, bate-papo, que sinaliza alguém mandando recado...) e internas (falta de concentração do próprio cidadão, que sai da mesa, que olha outra coisa...). É a geração das atividades múltiplas, fragmentadas, com déficit de atenção, e distração elevada, em que as atividades viram algo mais automatizado, pouco refletido. Maggie Jackson radicaliza ao dizer que a atual geração está com tamanha erosão de atenção que se corre o risco de ser remido de volta para a idade média.

Ser multitarefas nem sempre é possível, porque algumas tarefas exigem sequencialidade, como por exemplo, a inviabilidade de escutar duas músicas ao mesmo tempo. Já outras ficam tão automatizadas no cérebro que é possível realizá-las, fazendo-a em paralelo com outra, como repor mercadorias na gôndola de um supermercado, ouvir música e conversar com o colega do lado. Mas é importante alertar que isso sempre reflete em algum déficit de atenção ou certa perda de foco. Davenport e Beck (*apud* Regis *et al.* 2012) enfatizam que é desafiadora a equação da atenção relacionada à educação, às

relações sociais, ao trabalho e ao consumo, devido a duas perspectivas: 1ª.) o que fazer para conquistar e conseguir manter a atenção de alguém; e 2ª.) como dividir a própria atenção em diversas partes sem que haja perda substancial de foco.

#### c) *cognição distribuída*

Esse tipo de cognição foca na interação dinâmica entre os atores e os fatores materiais, sociais e ambientais, humanos e não humanos. Entende-se isso, Regis *et al.* (2012) argumentam que a cognição é inseparável da interação e da ação com o mundo em que vive de forma concreta, que a cognição se origina dessas interações, desse ambiente múltiplo e coletivo, e que o conhecimento não é produzido pelo indivíduo, mas por essa interação entre os indivíduos e suas ferramentas utilizadas nesse ambiente, portanto é produzido de forma coletiva, colaborativa e integrativa. Assim, todos, humanos e não humanos, auxiliam no processo de desenvolvimento cognitivo, da cultura e da sociedade. Nesse contexto, o cognitivo não inclui somente o cérebro, mas também o corpo, o mundo natural e os dispositivos tecnológicos. Essa mudança ou transformação do sistema cognitivo só pôde acontecer tão intensamente nas últimas décadas, graças à capacidade de interação gerada pelos meios eletrônicos de comunicação, especialmente a Internet.

O mesmo autor (*ibidem*), afirma que os jovens da chamada geração da Internet têm mais facilidade de atuar colaborativamente nesse ambiente on-line, onde compartilham o que sabem, através de postagem de relatos, vídeos explicativos sobre o que fizeram, ou respondendo a indagações de outros internautas, tanto em redes sociais quanto em sites apropriados para isso (wiki, p2p, redes sociais, Youtube...). Compartilham arquivos de aula, de trabalho, músicas, jogos, tudo o que é possível ser compartilhado. Assim, participam da construção colaborativa do conhecimento, naturalmente, e são exemplos práticos de modelos de cognição distribuída no contexto da cibercultura.

#### d) *Reflexões sobre os modelos e a necessidade de ampliá-los*

Nos seus estudos Terres (2011) afirma que o termo cognição é entendido como aquilo que é racional e que o indivíduo sabe a razão de ter sido feito ou de existir. Se ele não sabe por que fez aquilo, pode estar envolvendo afetividade, e não, cognição. Afetividade é apego, amor, carinho ou consideração por alguém ou algo. Essas variáveis também afetam a decisão, a escolha e a atenção.

É importante atentar para estas palavras de Borges *et al.* (2004, p. 89):

[...] a ideia de que a informação – como algo objetivo – é o que gera novo conhecimento e de que existe o fenômeno da transmissão da informação, tal como é adotado nas abordagens tradicionais sobre a cognição humana, não é mais considerada adequada atualmente. Muitos têm sido os estudos no âmbito da neurociência que têm, cada dia mais, apresentado fortes evidências de que o conhecer humano é bem mais complexo e se relaciona com muitos fatores. Basicamente, as concepções cognitivas contemporâneas consideram que somos seres individuais e seres sociais ao mesmo tempo, a todo instante, de forma absolutamente indissociável, o que modifica sobremaneira os conceitos que prevalecem a respeito da informação e do conhecimento.

Referendados em diversos autores e obras segundo Regis *et al.* (2012), atualmente, tem se utilizado o termo cognição ampliada para referenciar essa visão mais macro que envolve habilidades sensomotoras, perceptivas, emocionais e sociais. Todas essas habilidades são demandadas para se fazerem práticas comunicativas, em que os dispositivos técnicos não são apenas ferramentas ou extensões das habilidades humanas, mas também a reconfiguração, de forma dinâmica e completa, do processo cognitivo.

Segundo Lima (2003) *apud* Neves (2006, p. 40),

nas últimas décadas, ocorreu um incremento nos estudos sobre a cognição humana, principalmente após o surgimento do computador e da modelagem computacional. A ciência cognitiva é uma área de estudos interdisciplinares que se inter-relaciona com psicologia cognitiva, ciência da computação, sistemas de informação, inteligência artificial, neurociências e linguística, entre outras (LIMA, 2003). A partir dessa inter-relação, as pesquisas desenvolvidas sobre a cognição humana têm buscado apreender o modo como as pessoas pensam, interpretam e percebem o mundo. Estudos sobre a natureza e o desenvolvimento cognitivo dos seres humanos estão voltados para quatro teorias principais do desenvolvimento cognitivo: a de Piaget, a neopiagetiana, a de Vygotsky e a abordagem do processamento da informação.

Neves (2006) acrescenta que, na visão cognitiva de Piaget, a cognição humana ocorre com a adaptação biológica. Assim, o conhecimento se constrói aos poucos, com o desenvolvimento de estrutura cognitiva e estágios de desenvolvimento da inteligência. Os neopiagetianos enfatizam as habilidades cognitivas de processar e coordenar, que ajudam o indivíduo a diferenciar as informações que interessam para determinado objetivo. Isso requer a mediação e a interação para a resolução de problemas. Já na perspectiva de Vygotsky, o conhecimento se cons-

trói com a interação de indivíduos na sociedade, e isso também envolve mediação e cooperação entre dois ou mais indivíduos, para que possam entender e reelaborar situações e aprendizados e desenvolver seus modelos de comportamento "ideais" para aquela situação e os indicar (modelos) para outras pessoas, quando estiverem naquele contexto. Essa orientação é assimilada, mas não significa que o orientado terá exatamente o mesmo comportamento do orientador, pois ele tende a reelaborar, a seu modo, o comportamento "ideal" para aquela situação, com base no padrão dado, mas não, como fixo. Já a teoria do processamento da informação (abordagem de processamento da informação) é a que avançou para o estudo, ao focar a ciência neurocientífica. Esse aspecto é abordado de forma mais profunda nas outras partes deste estudo.

Com o advento da cibercultura, especialmente da Internet, as TICs têm incentivado a participação mais efetiva de seus usuários na aprendizagem e na construção do conhecimento por meio da interface, dos códigos mediáticos e da interação social online, desafiando os aparatos mentais e os modelos cognitivos a se ampliarem para modelos cognitivos mais completos.

Capurro (2003), que aponta a existência dos paradigmas físicos, cognitivos e sociais, numa sequência lógica evolutiva, também discute sobre essa relação da cognição com a cibernética (ibidem, p.7):

[...] a emergência da consciência ou da identidade pessoal, como também em relação à tendência niveladora de tais teorias com relação a processos cognitivos no mundo natural não humano, que levam a outros novos projetos tecnológicos, como o da inteligência e o da vida artificial. Isso significa um agravo cognitivo da autoconsciência do ser humano [...] vivemos no horizonte de uma *ontologia digital*, entendido o termo ontologia não no seu sentido clássico de um estudo dos seres, nesse caso dos seres digitais, [...] cujas consequências sociais e ecológicas são difíceis de prever.

Outros autores, como é o caso de Borges *et al.* (2004, p. 88-89), também fazem essa relação da cognição com esses novos tempos:

[...] nos anos 60 e 70, surgiu o movimento da "segunda cibernética", ou "cibernética de segunda ordem" [...] A perspectiva que se tem não é a das capacidades computacionais do ser humano, mas sim de seus "comportamentos próprios", auto-reprodutores. É exatamente essa mudança de paradigma que vem sendo proporcionada pelas abordagens contemporâneas das ciências cognitivas, [...]. Nas abordagens cognitivas contemporâneas, o conhecer humano não se dá pela captação, por parte do organismo, de objetos externos a ele. O meio pode somente desencadear

uma mudança estrutural no organismo, mas não sob a forma de interações instrutivas que determinem (numa relação de causa e efeito) o seu comportamento, ou seja, as mudanças estruturais no organismo. É o que significa para Maturana e Varela (1984), «conhecer é viver, viver é conhecer». O conhecimento, então, é "comportamento adequado" e "ação efetiva" num contexto relacional, no qual cada comportamento, cada ato, somente pode ser visto e compreendido como um ato cognitivo. (*grifo nosso*)

As expressões "Conhecer é viver" e "Viver é conhecer" podem ser entendidas como o fato de que o conhecimento gera aprendizado apreendido e vice-versa. Logo, quem conhece fica mais confiante, mais vivido. O ato de confiar em algo ou em alguém é importante para garantir que as trocas, as escolhas e as decisões aconteçam, mesmo para aquelas que envolvem algum nível de risco ou consequências severas. A pessoa só estará mais segura de que está reduzindo incertezas percebidas se quem lhe ofereceu algo em troca lhe chamou a atenção a tal ponto que cativou e gerou confiança. Isso envolve cognição, conhecimento e afeto, o que leva à atenção e à confiança e resulta em trocas, escolhas e decisões, que geram um novo processo de confiança interpessoal cognitiva, baseada no conhecimento e na experiência (memória, histórico, referências), que, de forma circular, leva a novas razões para confiar.

Assim, para Terres (2011) o que poderia ser uma dicotomia - cognição-afeição – em que a confiança interpessoal cognitiva estaria ligada ao conhecimento acumulado, e a competência e a confiança interpessoal afetiva estariam baseadas, especialmente, na preocupação, no cuidado, na atenção e no interesse, perde importância se vista no conjunto.

Ainda na perspectiva da confiança interpessoal afetiva, que inclui a atenção, pode-se dizer que, para um empreendimento, um indivíduo ou uma ação terem sucesso em suas pretensões, considerando a economia atual, têm que ser hábeis para conquistar atenção (Davenport; Beck *apud* Regis *et al.* 2012). Atenção, confiança e emoção são assuntos que vêm sendo tratados nas abordagens cognitivas contemporâneas, como destaca Borges (2005, p.77-78) com a seguinte concepção:

[...] abordagem cognitiva contemporânea [...] vem sendo crescentemente utilizado na construção de sistemas inteligentes. Além disso, as concepções acerca dos conceitos de informação e de conhecimento mudam bastante nessa perspectiva. Essa concepção cognitiva considera o processo do conhecer como uma dinâmica individual e social ao mesmo tempo. O indivíduo, definido como um ser que possui uma estrutura determinada pela sua

história coletiva como espécie humana e pela sua história individual no experienciar cotidiano, só consolida o processo de conhecer no nível individual e na sua interação congruente no meio, ou seja, nos vários domínios ou grupos sociais, nos quais transita (família, trabalho, grupos de amigos, etc.). Há que considerar as duas instâncias – individual e social – como uma dinâmica complementar, para se referir ao processo de conhecer. [...] Isso significa que o ser humano não cria representações “do mundo”, mas vive nele numa dinâmica de interações indivíduo e meio. [...] Nesse processo, o conhecer ocorre a todo instante e não somente para se resolver problemas [...]. Portanto, o conhecimento não é obtido. Conhecimento significa conduta adequada em um meio relacional, no qual cada comportamento é um ato cognitivo.

Assim, as tentativas de mapear o cérebro ou reproduzi-lo da mesma forma que o “verdadeiro”, agora dentro de um computador, pela inteligência artificial, ainda é um desafio longe de sua completude e, talvez, algo que sempre será uma incompletude, devido à complexidade envolvida. O fato é que, segundo Regis *et al.* (2012) por enquanto, as TICs precisam ser vistas como meios de ação na busca da atenção, e não, como fontes de informação e comunicação. Elas são ferramentas tecnológicas, e não, a fonte em si. E isso fica mais evidente nos estudos que serão discutidos a seguir, que envolvem, além do contexto social e individual já relatado, outros contextos, como a emoção e a afeição.

Terres (2011), em seus estudos empíricos, constatou que, quando as escolhas e as decisões envolvem situações de consequências mais severas, exige-se que a cognição e as outras variáveis que incluem emoção e/ou afetividade precisam caminhar juntas, para gerar equilíbrio e dar informações relevantes que deem confiança ao indivíduo, em termos de afeição e cognição, de modo que ele possa tomar a melhor decisão. Significa dizer que não se constrói confiança, conseqüentemente, não se produzem resultados, sem ter bom desempenho tanto de aspectos cognitivos quanto afetivos. Logo, o afeto é tão importante quanto a cognição, porquanto exige equilíbrio entre eles, em situações que envolvam escolhas e decisões severas.

## 5. Neuroeconomia, Neuromarketing, Cognição e Informação

A Neuroeconomia, para Moreira *et al.* (2011) é uma área de pesquisa interdisciplinar que alia conhecimentos de economia com psicologia, neurociência e outros e busca analisar a arquitetura interna do cérebro e como ele pode ser

informado para fazer escolhas. Segundo o mesmo autor (ibidem, 2011) já o Neuromarketing também utiliza conhecimentos interdisciplinares, especialmente psicológicos e cognitivos baseados na neurociência para entender, de maneira mais completa, o comportamento de consumo dos indivíduos. Afinal, complementa "o mesmo autor (ibidem, p.102) que, tanto a teoria microeconômica quanto o marketing carecem de maiores detalhes sobre o real comportamento humano. A simples modelagem matemática ou observação passiva nem sempre garantem a compreensão absoluta de um fenômeno".

O final do Século XX e o Século XXI marcam pela grande geração e difusão de informações, gerando inclusive excessos e certa aflição para dar conta de tamanho bombardeio informacional cotidianamente.

Segundo Borges (2005, p.78):

Informação é um construto, ou seja, é a perturbação do meio ou do indivíduo - de qualquer natureza - determinada como “informação”, ou não, pela estrutura do sistema, indivíduo ou meio. Assim, não existe o fenômeno da transmissão da informação, pois ela perde o seu caráter objetivo, de algo que pode ser estocado, codificado e representado. Dessa forma, a determinação do que seja “informação” só pode dar-se nos níveis individual e social ao mesmo tempo e não somente em uma das duas dimensões. De um modo geral, os estudos sobre necessidades, demandas e uso de informação procuram investigar como indivíduos obtêm, processam e usam informação, como buscam informação sobre algo.

Foi nessa perspectiva que estudiosos das teorias cognitivas avançaram seus estudos e buscaram recursos teóricos e metodológicos na Neurociência, conforme aborda Neves (2009, p.41), quando fala que a teoria do processamento da informação

[...] deu origem a grande número de estudos investigando detalhadamente os processos cognitivos envolvidos na resolução de problemas. Tal abordagem vem gradualmente se integrando aos estudos da neurologia, dando origem à neurociência cognitiva. Os teóricos do processamento da informação, de origem mais recente, reúnem diversas abordagens que estudam a mente e a inteligência em termos de representações mentais e seus processos subjacentes ao comportamento observável. Esses pesquisadores consideram o conhecimento como sistema de tratamento da informação.

Vivemos num ambiente onde existe uma acirrada competição de oferta de gamas de informações. Para isso, bombardeia-se o mercado com excessos. Tudo com o intuito de ganhar a preferência do consumidor e vencer a competição dos concorrentes. No entanto, à medida que

todos fazem isso, gera-se uma tendência, "segundo Furlaneto *et al.* (2013, p. 106) de, *comoditização* de produtos, um fenômeno que se caracteriza pela profusão de consumíveis muito parecidos, tanto em termos funcionais, quanto em relação ao seu preço". Isso quer dizer não existe mais uma substancial diferenciação racional para justificar que o indivíduo decida por A ou B, apenas racionalidade. E mesmo que ainda existam diferenciais suficientes (por exemplo, o mesmo produto por preço menor), muitos indivíduos escolhem o que seria considerado racionalmente ilógico (o mais caro, por exemplo).

Baseado em Vera (2013) pode-se dizer que existem equívocos na teoria tradicional e que mostram incompletudes que geram resultados inconsistentes. Por exemplo, supor que as pessoas e as empresas têm fontes ilimitadas de informação e capacidade de analisar todos os cenários possíveis, no mínimo, é ilusório. Afinal, a natureza humana é passível de falhas, de limitações, de crises e de impulsos e se sabe que sempre se estará correndo algum risco por contato das informações assimétricas e pelo fato de que as pessoas tendem a priorizar o futuro imediato em detrimento do futuro de longo prazo.

Kenning e Plassmann (2005 *apud* Moreira *et al.* 2011) afirmam que, nesse contexto de análise, a primeira necessidade de mudança na forma de estudar o indivíduo exige que saia de cena o *homo economicus* e retorne o *homo sapiens*, com natureza social, econômica e comportamental baseada em sua neurobiologia. Assim, para os estudos que fazem um indivíduo escolher e tomar decisões, dentro de um contexto de infinitas opções, exige aprender novos fatores como origem psicológica e biológica, estruturas cognitivas neurais e emoções. Essas últimas variáveis entram em cena com forças poderosas, enquanto a matematização e sua modelagem perdem parte de sua força nesses estudos.

Segundo Terres (2011) quando as decisões envolvem trocas, com consequências mais brandas, facilmente são decididas utilizando-se apenas a cognição, pois envolve menos riscos e repercute menos. Assim, a decisão não envolve tanta emoção, ansiedade ou estresse e será mais racional, mais segura e mais confiante se se empregar mais a cognição como parâmetro de decisão. Já em uma relação com consequências mais severas (por exemplo: decidir sobre um tratamento de câncer ou outro), muitas vezes, vale mais o afeto dispensado ao indivíduo, que gera confiança, do que a orientação puramente técnica.

Para "Ferreira *et al.* (2014, p.9), as emoções ainda criam a 'cognição motivada', onde as pessoas são boas em persuadir 'a si mesmas' de que o que elas gostariam que acontecesse é o que vai acontecer". Ou seja, é pacífico aceitar que o ser humano se autoconvence de algo mesmo não sabendo exatamente como explicar isso.

É preciso considerar, também, o que contextualiza Borges (2005, p.80):

Ao considerarmos os princípios cognitivos contemporâneos, os usuários da informação são seres humanos individuais e sociais, que vivem uma deriva de experiências individuais e intransferíveis, ao mesmo tempo em que vivem em contínua interação com outros indivíduos. Nesse viver cotidiano, os usuários de informação vivem em domínios de ação (empresa, família, lazer, amigos, etc.), fazendo parte de diferentes redes de relações e interações. Cada domínio de ação, pelos quais transitam, possibilita aos indivíduos estabelecerem redes de relações e interações que, embora distintas, influenciam as suas condutas no meio, a sua linguagem e as suas pré-disposições em aceitarem ou não determinada informação. Além disto, ainda que os usuários da informação experienciem relações e interações sociais vivem ao mesmo tempo experiências individuais. [...] somos indivíduos que, apesar de vivermos situações próprias, a nossa formação de sentido é também social e coletiva.

Isso significa que contextos e variáveis muito além do ambiente racional, do conhecimento adquirido e sabido e das experiências vividas do indivíduo o influenciam em suas decisões e escolhas. Há de se considerar o consciente e o subconsciente, o individual e o coletivo, a razão e a emoção, as particularidades e as coletividades em que ele está envolvido.

A Neuroeconomia tem essa perspectiva, já é reconhecida e tem sido premiada por suas contribuições. Como afirma Sandroni (2014, p.15),

a chamada economia experimental, ou "*neuroeconomics*", tem como resultado, até agora, dois prêmios Nobel em Economia: Daniel Kahneman e Vernon Smith. Mas qual é o objetivo desse novo enfoque? Na Era das Incertezas, seu objetivo é oferecer melhores meios para se realizar a espinhosa tarefa de fazer previsões. Se os meandros do cérebro que levam as pessoas a tomar decisões (econômicas) forem conhecidos, as incertezas, acredita-se, poderão ser reduzidas. Basta ligar uns fios [...]

Sobre a importância dos novos estudos aplicados ao marketing, para Furlaneto *et al.* (2013), tem-se que as inclusões de técnicas neurocientíficas permitem moldar melhor o que é ofertado ao consumidor (ator, cidadão), tendo em vista o

que o motive mais, desde a qualidade, a estética e as ideias até o *design*, e outros.

Segundo Moreira *et al.* (2011) isso implica dizer que o Neuromarketing alia as técnicas de marketing às de mapeamento cerebral, de modo a decodificar os processos que ocorrem na mente dos indivíduos e atender ao que eles realmente querem, entendendo não só o seu lado racional, mas também seu lado intuitivo, instintivo, biológico, cultural e emotivo. O risco pode ser a manipulação de tais técnicas de modo a gerar desejos nos cidadãos por coisas que, na verdade, eles não queriam ou não precisavam. Essa técnica já foi muito utilizada na propaganda e na publicidade e até entendida como um dos papéis do marketing para gerar lucro, mas que, nos dias atuais, é eticamente contestável.

Sobre essas técnicas, na perspectiva do Neuromarketing, sempre em consonância com as demandas do mercado, segundo "Furlaneto *et al.* (2013, p. 108) as, "técnicas neurocientíficas vêm despertando atenção das agências e da mídia, tanto por permitirem melhor avaliação de produtos quanto, com menor profusão, por facilitarem desenvolvimentos em *branding* (gerenciamento da marca)."

Essas novas teorias contribuíram com a teoria tradicional, em alguns pontos, de forma incremental, e em outros, de forma radical, inclusive obrigando a descartar certos pressupostos. Passou a ser inegável o papel das outras variáveis que não as racionais nas determinações de escolhas, comportamentos e decisões baseadas nas informações. Graças à neurociência, os aspectos emocionais passaram a assumir função de destaque nesses estudos, denominados de lógica da "revolução cognitiva" em verdadeiras fusões de campos multi e transdisciplinares, nas áreas da Neuroeconomia e da Neuroantropologia, entre outras, especialmente do Neuromarketing, que contribuiu sobremaneira para entender as inter-relações que "Furlaneto *et al.* 2013, p.109 diz serem, funções, como o pensamento, as emoções, o processamento sensorial e os comportamentos".

Para esse avanço acontecer, foi preciso rever teorias tradicionais. Por exemplo, foi revista a máxima da teoria microeconômica do consumidor, ao afirmar que suas decisões eram resultados de processos racionais que ocorrem no córtex cerebral. Moreira *et al.* (2011) levaram em conta estudos de vários autores, como Loewenstein e Prelec (2005); Carmer *et al.* (2005); Platt e Glimcher (1999); Kirouac e Ganguly (1995); Kenning e Plassmann (2005); Bechara e Damásio (2005) e Phelps (2006), e concluíram que, quando o indivíduo é submetido a uma

situação que exija que ele faça escolhas, outras partes do cérebro também são acionadas, e não, apenas, o córtex cerebral. A amígdala, o núcleo accumbens e o córtex insular também foram acionados nessas experiências feitas pelos pesquisadores e sabe-se que essas áreas têm relação direta com a expressão das emoções. Até mesmo as decisões racionais mostraram que, ao ser tomadas ou para ser tomadas, envolviam um processo emocional prévio no cérebro. Ou seja, não é uma questão apenas de viesamento das decisões pela emoção, mas de influência evidente da emoção nas decisões, não apenas em determinado momento, mas desde o início até a execução da decisão tomada.

Ainda Moreira *et al.* (2011) recorreram a outros diversos autores, como Damásio (1995); Bechara e Damásio (2005); Epstein (1990 e 1994), respaldados por diversos estudiosos, como Kahneman e Tversky (1979); Kahneman; Slovic e Tversky (1982) e Sanfey (2007), para dizer que, graças a eles, foi possível afirmar que existe um sistema múltiplo para a tomada de decisões que leva em conta modelos de processamento dual de informações, ou seja, que ocorre no ambiente da razão e da emoção, e não, em apenas um deles. Assim, pode-se dizer que o cérebro tem dois sistemas de tomada de decisão, a saber: o deliberativo ou cognitivo (decisão racional) e o afetivo ou intuitivo (decisão que considera o estado emocional, a intuição e as experiências).

Com isso concorda também Vera (2013), ao afirmar que o comportamento humano, em qualquer que seja a área onde se desenvolve tal decisão, é fortemente influenciado pelos fatores emocionais e afetivos, além dos racionais. Acrescenta que pesquisas feitas demonstram que, no exercício de trabalhos cognitivos, no cérebro, a primeira parte acionada é a afetiva, e a reação do indivíduo dependerá de sua capacidade de controlar suas emoções, pois a parte emocional só entra em segundo momento.

O estado de humor do indivíduo, no momento em que precisar tomar uma decisão, é um exemplo explícito de o quanto a emoção influencia a decisão. Quando está de mau humor, a pessoa tende a ser mais sensível a estímulos negativos e a tomar decisões precipitadas ou radicais estimulada por essa emoção. Já quando está de bom humor, tende a ser influenciada por estímulos positivos e a decidir levando em conta esse clima otimista de influência. Assim, o humor afeta o comportamento, o pensamento e a capacidade de julgar dos indivíduos e impacta diretamente na forma como agem, reagem ou pensam e lidam com informações e, a partir

disso, como fazem escolhas e tomam decisões (Arkes, 2003; Chan, 2005 *apud* Moreira *et al.* 2011). Isso quer dizer que, se uma pessoa está carente, solitária, sensível, aflita, feliz, triste, estressada, raivosa, consolada, cansada, desencansada, aliviada, entusiasmada, informada, desinformada, com alterações hormonais e tantas outras situações possíveis pela condição humana, isso tudo influenciará o seu comportamento de escolha e decisão e não é correto nem justo que esse comportamento seja limitado apenas a questões racionais.

Um exemplo, citado por Moreira *et al.* (2011), de influência em decisões de compra é o hormônio denominado de oxitocina, que é liberado pelo organismo quando vamos fazer negociação e aumenta a confiança de quem está próximo, por exemplo, nesse caso, o vendedor. Tal fato gera um comportamento adequado para a negociação de compra e de venda acontecer. Outro exemplo é o *trade off* de consumo, isto é, comprar ou consumir agora ou postergar o ato. Os estudos procuram saber por que o consumidor tem tanta dificuldade de lidar com isso e descobriram que o sistema cortical é o responsável pelas consequências de longo prazo nas decisões. Mas, como as decisões são instantâneas, de curto prazo, tendem a ser controladas pelo sistema límbico, que é o das emoções, e isso justifica, por exemplo, o comportamento impulsivo.

Segundo Ferreira *et al.* (2014, p.9),

As decisões de um indivíduo impulsivo são excessivamente influenciadas por estímulos, pressões e exigências externos. Essa pessoa pode não conseguir dar uma explicação mais satisfatória de uma ação, exceto que ela automaticamente “achou que deveria agir assim”. O afeto também pode distorcer as avaliações cognitivas. Por exemplo, as emoções influenciam nossas percepções dos riscos – a ira torna as pessoas menos sensíveis aos riscos, e a tristeza, mais.

Assim, num primeiro instante, o mesmo autor (*ibidem*, 2011) diz que a emoção tende a determinar a decisão se não for dado o devido tempo para a região do pós-frontal córtex “raciocinar” as consequências de longo prazo de tal decisão e restabelecer o equilíbrio entre a emoção e a razão. Isso explica, por exemplo, a compra impulsiva e o arrependimento posterior. Um exemplo disso são os testes com crianças que recebem um doce e são desafiadas a aguardar a hora certa para comê-lo. Quem conseguisse “se aguentar” ganharia, como recompensa, mais um doce. A maioria não espera essa hora e come o doce. Porém, depois, demonstra arrependimento por não ter conseguido o segundo doce.

Os autores (*ibidem*) complementam que, a grande contribuição desses estudos práticos que envolvem economia e marketing com a Neurociência e as áreas afins é que foi possível comprovar que os seres humanos não são tão racionais e previsíveis, por modelos e cálculos, como imaginava a teoria tradicional. Já por meio da Neuroeconomia, é possível explicar por que certas escolhas aparentemente irracionais foram feitas. Para confirmarem isso ainda escrevem (*ibidem*, p.12):

Uma escolha poderia ser feita pelo modelo mental do sujeito, contrariando todos os indicadores que se apresentavam, pois a representação mental da escolha preferida condiz com a representação de mundo que o sujeito tem.

Assim, emoções, questões hormonais e mecanismos de captação de informações passaram a desfrutar de *status* destacado para entender as decisões e as escolhas dos indivíduos, sem que seja necessário eliminar a importância da racionalidade, mas dividir espaço com essas novas variáveis.

Enquanto isso Ferreira *et al.* (2014, p.2) citam Kahneman e Hall (1998) para respaldar argumentos nesse sentido:

Assim, métodos da neurociência com o objetivo de observar, mapear e mensurar as “imagens do cérebro” no desempenho de uma tarefa indicaram que existiam diferenças entre o funcionamento comportamental e cognitivo e que imagens mentais (ilusões ou representações mentais) influenciam sobremaneira na escolha do sujeito. [...] uma ilusão cognitiva é a tendência de cometer erro sistemático no processo de tomada de decisão dos humanos. Tais ilusões ocorrem, na visão desses autores, porque os humanos têm dificuldades para julgar subjetivamente quantidades físicas e probabilidades.

Então, o comportamento vai além do que pode ser chamado de uma vontade consciente, racional e sem erros. E tudo o que foi apresentado permite entender por que os indivíduos tomam decisões consideradas não ótimas pelas teorias tradicionais, mas que, para esses cidadãos, são, emocionalmente, as mais adequadas.

#### 5.1. Neurociência: principais técnicas e métodos

No quadro abaixo, apresentam-se as técnicas mais conhecidas utilizadas pelo Neuromarketing e pela Neuroeconomia e buscadas na Neuroeconomia, com base nos apontamentos de Furlaneto *et al.* (2013, p.104-111):

### Ressonância magnética funcional (fMRI)

"A ressonância magnética funcional (fMRI) sobressai-se no campo por permitir a determinação das diversas áreas corticais e subcorticais ativadas pela fruição de imagens estáticas ou dinâmicas (i.e., filmes), em função da distribuição do sangue oxigenado no cérebro, o qual é um indicador poderoso de recrutamento da área cerebral em foco. [...] Trata-se de uma técnica com excelente resolução espacial, que distingue áreas com precisão de 1-10 mm e dá informações valiosas sobre a arquitetura neural relacionada por processos complexos que podem estar potencialmente relacionados à decisão de compra, como a formação de juízos sobre um produto, algo que envolve a associação de memórias afetivas, a formação de conceitos e a experiência sensorial primária. Todavia, é uma técnica que carece de precisão temporal, havendo um atraso de 1-10 segundos entre a exposição ao estímulo e o registro dos dados, uma limitação considerável" (Furlaneto *et al.* 2013, p. 110).

### Eletroencefalografia (EEG)

"A eletroencefalografia é realizada por meio de eletrodos colocados sobre o escalpo do indivíduo, que medem os sinais elétricos emitidos pelo sistema nervoso [...] As técnicas eletrofisiológicas podem dar informações relevantes sobre preferência, como a valência e a saliência da exposição a um objeto ou propaganda. Nesse contexto, a valência é entendida como uma medida de quanto o participante da pesquisa gostou ou não gostou do que viu, e a saliência se refere a quanto o estímulo gera mobilização no participante" (Furlaneto *et al.* 2013, p. 110).

### Rastreamento ocular (eye tracking)

"O rastreamento ocular é uma técnica que pode ser usada para diversos fins, como no auxílio da diagramação de uma propaganda em um site, design de produtos e/ou embalagens, analisar a qualidade de comerciais e/ou imagens fixas, entre outras" (Furlaneto *et al.*, 2013, p.111).

Quadro 1. *Resumo das técnicas de Neurociência adotadas pelo Neuromarketing e pela Neuroeconomia.* (Fonte: Adaptado dos autores citados no Quadro 1.)

Junta-se a essas técnicas da neurociência outra importante técnica da psicologia social: o método do grupo focal e que também auxiliam os estudos do comportamento e reações dos atores pela economia e pelo marketing. A seguir apresenta-se mais detalhes baseados especialmente no mesmo autor (*ibidem*, 2013, p.104-111):

### O método de grupo focal

"A metodologia de grupos focais permanece sendo uma das mais versáteis e indicadas para pesquisas de base qualitativa, recolhendo informações confiáveis sobre a opinião do consumidor em diversos níveis" (Furlaneto *et al.* 2013, p. 111).

#### Afinal

"A partir das proposições apresentadas, faz-se possível identificar sinergias na construção do conhecimento ao se utilizarem métodos distintos e simultâneos ou se-

quenciados, nas pesquisas em marketing. No caso específico dos grupos focais, sugere-se que associações sinérgicas com abordagens capazes de aprofundar o entendimento dos processos mentais que distinguem (e associam) consumidores e shoppers podem levar o campo a um novo estágio de desenvolvimento científico-metodológico. Essas abordagens atingem o seu ápice no campo das neurociências, o que a seção seguinte se propõe a caracterizar" (Furlaneto *et al.* 2013, p. 108).

#### O método recomenda:

"O uso do focus group é particularmente apropriado quando o objetivo é explicar como as pessoas consideram uma experiência, uma ideia ou um evento, visto que a discussão durante as reuniões é efetiva em fornecer informações sobre o que as pessoas pensam ou sentem ou, ainda, sobre a forma como agem" (Oliveira e Freitas, 2010 apud, Furlaneto *et al.*, 2013, p. 104).

#### De onde é originário o método:

"Originárias da Psicologia Social, as pesquisas que empregam grupos indicam que, quando dispostos em grupos (mesmo que entre desconhecidos), a maior parte das pessoas discorre mais livremente e em maior profundidade do que quando estimuladas a fazê-lo a partir de entrevistas e questionários" (Barbour, 2009, apud Furlaneto *et al.*, 2013, p. 104).

#### Forma de aplicação do método:

"Para a condução de grupos focais, geralmente são utilizadas salas monitoradas com auxílio de espelho espião, atrás dos quais pesquisadores ficam ocultos observando as reações das pessoas. Microfones e câmeras de vídeo são colocados de maneira discreta, a fim de evitar a intimidação dos participantes para gravação das discussões, percepções e atitudes dos grupos. Outro fator importante é a informalidade da sala para que não se crie um ambiente muito austero, evitando, assim, o constrangimento das pessoas, o que poderia ocasionar um bloqueio à expressão de ideias e comportamentos espontâneos. A recomendação é de que as pessoas fiquem à vontade" (Oliveira e Freitas, 2010, apud, Furlaneto *et al.* 2013, p. 105).

Por questões de ética e de conduta em pesquisa, antes de o grupo focal passar a ser observado e fazer as discussões, "cabe ao moderador fazer uma breve apresentação dos tópicos a serem abordados, expondo as principais regras, bem como a informação acerca de filmagem ou gravações" (Furlaneto *et al.* 2013, p.105)

Quadro 2. *Resumo de uma das técnicas importadas da Psicologia Social e utilizadas pela Neuromarketing e pela Neuroeconomia.* (Fonte: Adaptado dos autores citados no Quadro 2.)

Como pode ser visto, dos métodos utilizados e recomendados, destacam-se os importados da Neurociência, com o uso de aparelhos para mapear a circulação do sangue no cérebro, a emissão de corrente elétrica pelo cérebro e a questão do rastreamento ocular, ou seja, atividades em laboratório monitorado por máquinas. Somam-se a eles as técnicas de grupo focal, importadas da Psicologia. Portanto, seriam essas as novas "armas" do marketing, da economia e da cognição para desvendar as preferências, as escolhas, os comportamentos e as de-

cisões dos atores, baseados nas informações que lhe são disponíveis, e conseguir alguns resultados mais concretos, embora sem a pretensão de ser uma teoria cabal e autossuficiente.

## 6. Considerações Finais

Há algum tempo, as ciências tem o objetivo de descobrir, mapear, prever e entender melhor o comportamento cognitivo e emocional do ser humano em suas decisões diárias, como comportamento, preferências e escolhas. O surgimento da Neuroeconomia e da Neurociência é um exemplo disso.

A partir dessa nova visão admite-se considerar a complexidade envolvida e a repercussão de curto e em longo prazos, bem como as emoções e afeições, nas escolhas e decisões. E que cada ator se portará de uma forma específica, levando em conta as informações disponíveis, bem como suas condições de controle, sensibilidade e reação às emoções, à afeição, à cognição e à racionalidade.

São novas teorias interdisciplinares que permitem estudar o comportamento dos indivíduos, enriquecendo o entendimento de suas posturas e permitindo novas ferramentas que cercam o estudo desses fenômenos, sem a pretensão de resolver todos os problemas, esgotar as possibilidades ou descartar teorias anteriores.

Entendendo a fragilidade dos atores, quanto a emoção, por exemplo, é possível promover técnicas e ações de marketing que visem educar as pessoas, para que reflitam sobre o que fazer e as consequências do que fazem levando em conta apenas a satisfação impulsiva do curto prazo. Isso inclui, por exemplo, técnicas como ir a um *shopping center* com uma quantidade limitada de capital financeiro e deixando o cartão de crédito em casa. Nesse caso, o ator, ao sentir o impulso da compra, teria que voltar para buscar o cartão. Isso daria tempo suficiente para o cérebro reequilibrar o sistema nervoso entre emoção, sensações e racionalidade, melhorando a qualidade da decisão. Pois, é quando acontece o processamento completo entre o desejo impulsivo e emocional de curto prazo *versus* a reflexão racional das consequências dessa decisão em curto e em longo prazos. Ocorrendo o reequilíbrio é possível que o ator perceba que tal compra é desnecessária, e o retorno ao *shopping* ficará para outra ocasião.

Tratou-se aqui das diversas ciências, de forma interdisciplinar, visando entender melhor do ator nas suas escolhas e decisões baseado em informações. Entendido isso pelas ciências é pos-

sível contribuir, de forma propositiva, com sugestões para que este melhore tais escolhas e decisões, equilibrando emoção, sensações e racionalidade.

## Referencias

- Borges, M. N. (2005). A abordagem contemporânea sobre a cognição humana e as contribuições para os estudos de usuários da informação. *Cadernos BAD*, (2), 74-81.
- Borges, M. N. *et al.* (2004) A Ciência da Informação discutida à luz das teorias cognitivas: estudos atuais e perspectivas para a área. *Cadernos BAD*, (2), 80-91.
- Burguete, J. L. V. (2004). Pasado, presente y futuro de las dimensiones pública y social en el desarrollo conceptual del marketing. *Revista Internacional de Marketing Público y no Lucrativo*, 1 (1), 9-34
- Capurro, R. (2003). Epistemologia e Ciência da Informação. In: V Encontro Nacional de Pesquisa em Ciência da Informação, 5, Belo Horizonte: UFMG, 2003. *Anais...* Recuperado em 10-10-2014 de [http://www.capurro.de/enancib\\_p.htm](http://www.capurro.de/enancib_p.htm)
- Ferreira, D. D. M.; Keller, J.; Spanhol, G. K.; Leal, D. B. (2014). Neuroeconomia: um apoio da neurociência à economia e à gestão do conhecimento. *SEGeT – Simpósio de Excelência em Gestão e Tecnologia*. Recuperado em 14-10-2014 de [http://www.cpge.aedb.br/seget/artigos08/527\\_Neuroeconomia\[SEGeT\\_novo\].pdf](http://www.cpge.aedb.br/seget/artigos08/527_Neuroeconomia[SEGeT_novo].pdf)
- Furlaneto *et al.*, F.; Dias, Á. M.; Souza, C. V.; Akiba, H. T. (2013). Tendências metodológicas em marketing: entre as contribuições dos grupos focais e as abordagens contemporâneas. *Ciências & Cognição*, 18 (2), 102-111, 2013. Recuperado em 02-10-2014 de <http://www.cienciasecognicao.org>
- Hill, R. P.; Martin, K. D. (2014). Broadening the paradigm of marketing as exchange: a public policy and marketing perspective. *Journal of Public Policy and Marketing*, 33(1), 17-33.
- Mankiw, N. G. (2005). *Introdução à Economia*. 3. ed. São Paulo: Pioneira Thomson Learning.
- Moreira, B. C. M.; Pacheco, A. F. A.; Barbato, A. M. (2011). Neuroeconomia e neuromarketing: imagens cerebrais explicando as decisões humanas de consumo. *Ciências & Cognição*, 16(1), 99-111. Recuperado em 02-10-2014 de <http://www.cienciasecognicao.org>
- Neves, Dulce Amélia B. (2012). Representação temática da informação e mapas cognitivos: interações possíveis. *Informação & Sociedade: Estudos*, 22 (especial), 39-47. Recuperado em 15-10-2014 de <http://www.ies.ufpb.br/ojs/index.php/ies/article/view/13300/8208>
- Neves, Dulce Amélia B. (2006). Ciência da Informação e cognição: uma abordagem do processamento da informação. *Ciência da Informação*, 35 (1), 39-44. Recuperado em 13-10-2014. de <http://www.ibict.br/cienciainformacao/viewisue.php?id=39>
- Regis, F.; Timponi, R.; Maia, A. (2012). Cognición integrada, encadenada y distribuida: breve discusión de los modelos cognitivos en la cibercultura. *Comunicação, mídia e consumo*. 9(26), 115-136. Recuperado em 01-10-2014 de [http://revistacmc.espm.br/index.php/revistacmc/article/viewFile/346/pdf\\_1](http://revistacmc.espm.br/index.php/revistacmc/article/viewFile/346/pdf_1)
- Sandroni, P. H. (2014) Admirável neuroeconomia. *Economia: o custo da felicidade RAE executivo*, 16. p.15. Recuperado em 14-10-2104 de <http://rae.fgv.br/sites/rae.fgv.br/files/artigos/2264.pdf>

Vera, J. G. L. (2013). Expectativas racionales versus neuroeconomía: ¿cómo tomamos decisiones los agentes? *Contribuciones a la Economía*. Recuperado em 14-10-2014 de <http://www.eumed.net/ce/2013/expectativas-racionales-versus-neuroeconomia.html>

Terres, M. S. (2011). *A importância do afeto para a confiança do consumidor em trocas de serviços com consequências severas*. Porto Alegre: UFRGS, 2011, (Tese de Doutorado) PPGA, 156f. Recuperado em 01-10-2014 de <http://www.lume.ufrgs.br/handle/10183/36035>

---

# Organização do conhecimento arquivístico: um estudo terminológico comparativo (português, espanhol, francês, inglês) sobre classificação e descrição na *Multilingual Archival Terminology* – ICA.

*Archival knowledge organization: a comparative study (Portuguese, Spanish, French and English) on classification and description in the Multilingual Archival Terminology – ICA.*

---

**Andrieli Pachu da Silva (1), Walter Moreira (2), José Augusto Chaves Guimarães (3), João Batista Ernesto de Moraes (4)**

(1) Universidade Estadual Paulista / UNESP - Marília, andrielipachu@marilia.unesp.br

(2) walter.moreira@marilia.unesp.br

(3) guima@marilia.unesp.br

(4) jota@marilia.unesp.br

## Resumen

A organização do conhecimento arquivístico tem sido tema de discussão na atualidade no que se refere aos seus procedimentos, instrumentos e produtos. No âmbito específico dos procedimentos, a classificação e a descrição arquivística ocupam espaço nuclear na medida em que permitem que um conhecimento produzido pela sociedade e devidamente registrado em documentos possa ser apropriado por essa sociedade para a geração de novos conhecimentos. No entanto, e justamente por relacionar-se a dois termos-chave para a área – a classificação e a descrição - a organização do conhecimento arquivístico defronta-se com um problema de natureza terminológica, na medida em que esses termos passam a ser apropriados pela área em diferentes acepções, evocando diferentes conceitos, o que tem levado a área arquivística, já há quase meio século, ao esforço de sistematizar sua terminologia, em especial em um universo multilíngue, como demonstra o empenho do *International Council on Archives* – ICA, desde 1964. Atualmente o ICA apresenta, em seu site, o *Multilingual Archival Terminology*, com termos disponíveis em dezessete idiomas. Assim, este trabalho busca analisar as conotações conceituais dos termos *classificação* e *descrição* em suas definições e equivalências de tradução nos idiomas: português, espanhol, francês e inglês no *Multilingual Archival Terminology* do ICA. Para tanto foi observado os padrões para verbetes de dicionários terminológicos trabalhados por Krieger e Finatto (2004). A isso se alia a dimensão de análise de conteúdo (Bardin, 2009) de modo a identificar se há diversidade nas conotações conceituais atribuídas ao termo em cada língua. Os resultados sinalizam que quanto ao contexto dos termos nas quatro línguas (originalmente europeias) há casos de informações conceituais ligadas a países específicos,

como o caso de México e Argentina para o espanhol, Brasil, para o português e Canadá para o francês, assim como há termos sem informações contextuais. No âmbito das definições, é possível observar fortes convergências entre as línguas assim como variações consideráveis, como é o caso da questão da classificação. Um problema observado reside nas falhas das remissivas, muitas vezes ocultando termos correspondentes e, em outros casos, remetendo a termos de dimensão conceitual distinta. Tais aspectos reafirmam a diversidade de tradições teóricas e metodológicas da arquivística em distintos países, o que sinaliza para uma diversidade terminológica e conceitual que, se devidamente contemplada por instrumentos terminológicos dessa natureza, pode servir como instrumento de diálogo e de integração nesse importante campo do conhecimento.

**Palabras clave:** Organização do conhecimento arquivístico; Terminologia arquivística; Classificação arquivística; Descrição arquivística

## Abstract

Archival knowledge organization has been discussed in relation to its procedures, tools and products. Regarding the procedures, the aspects of classification and archival description are nuclear since they allow knowledge produced in society and recorded in documents to be appropriated by society to generate new knowledge. However, and just because it is related to the two core elements in the area –classification and description-, archival knowledge organization faces a terminological problem, since those terms are assimilated by the are in their different acceptations and referring to different concepts. This has led the Archival field, for nearly half a century, to efforts on systematization of its terminology, especially in a multilingual environment, as shown by the Interna-

tional Council on Archives – ICA, since 1964. Currently, ICA shows on its website the Multilingual Archival Terminology, with terms available in seventeen languages. This work aims to analyze the conceptual connotation of the terms *classification* and *description* in their definitions and equivalents in other languages, Portuguese, Spanish, French and English, in the ICA's Multilingual Archival Terminology ICA. For this, we studied the standard for dictionary entries by Krieger and Finatto (2004). We also used content analysis (Bardin 2009) to identify the diversity of conceptual connotation of the terms in each language. The results show that in the context of the terms in the four languages (originally European languages) there are cases of conceptual information linked to specific countries, as in the case of Mexico and Argentina for Spanish, Brazil for Portuguese, and Canada for French, and cases of terms without contextual

information. For the definitions, it was possible to observe strong similarities between the languages as well as considerable variations such as in the case of classification. We also observed problems with cross references, often hiding corresponding terms and in other cases referring to terms belonging to a different conceptual dimension. These aspects re-affirm the diversity of theoretical and methodological traditions of Archival Science in different countries, pointing out to a terminological and conceptual diversity that, if properly considered by terminological tools of that nature, can serve as an instrument for dialog and integration in this important field of knowledge organization.

**Keywords:** Archival knowledge organization; Archival terminology; Archival classification; Archival description.

## 1. Introdução

A organização do conhecimento arquivístico tem sido tema de discussão na atualidade no que se refere aos seus procedimentos, instrumentos e produtos (Tognoli; Guimarães, 2011; Tognoli; Guimarães; Tennis 2013), e tal preocupação pode ser observada desde o século XIX, com o manual elaborado pelos arquivistas holandeses, originalmente publicado em 1898 (Muller; Feith; Fruin, 2003).

No âmbito específico dos procedimentos, as atividades de classificação e descrição arquivística ocupam um espaço nuclear ao permitirem que um conhecimento produzido pela sociedade e devidamente registrado em documentos de natureza orgânica possa ser apropriado por essa sociedade para a geração de novos conhecimentos.

Porém, e justamente por relacionar-se a dois termos-chave para a área – a classificação e a descrição – a organização do conhecimento arquivístico defronta-se com um problema de natureza terminológica, na medida em que esses termos passam a ser apropriados pela área em diferentes acepções e evocando diferentes conceitos. Isso tem levado a área arquivística, já há quase meio século, ao esforço de sistematizar sua terminologia, em especial em um universo multilíngue, como demonstra o empenho do *International Council on Archives* – ICA, desde a publicação do *Elsevier's Lexicon of Archive Terminology* em 1964.

Atualmente o ICA apresenta, em seu site, o *Multilingual Archival Terminology*, com termos disponíveis em dezessete idiomas. Isso retoma uma questão terminológica já apontada por Michèle Hudon (1997), relativamente às conotações conceituais que um termo assume a partir

do contexto cultural em que se inserem as diferentes línguas pois, se há arbitrariedade e variação na relação termo-conceito num mesmo idioma, isso se torna ainda mais complexo em idiomas diferentes, considerando-se que cada idioma recorta a realidade de modo diverso.

Desse modo, e adotando como pressuposto que o *status* epistemológico de uma determinada área também se encontra condicionado à qualidade e clareza de sua terminologia, o trabalho tem por objetivo analisar as conotações conceituais dos termos “classificação” e “descrição” em suas definições e equivalências de tradução nos idiomas: português, espanhol, francês e inglês no *Multilingual Archival Terminology do ICA*. Para tanto observaram-se os padrões para verbetes de dicionários terminológicos trabalhados por Krieger e Finatto (2004) como palavras-entrada: registro na forma utilizada, equivalentes em língua estrangeira, informação de categoria gramatical, informação conceitual, fontes contextuais, fontes bibliográficas, gradação sinonímica, remissivas, notas explicativas (linguísticas, técnicas, enciclopédias).

## 2. A arquivística e o estudo terminológico

O estudo terminológico na área da arquivística se faz necessário, tanto no que diz respeito a comunicação científica – condicionado a seu status epistemológico – quanto na organização e representação da informação dos documentos arquivísticos.

A terminologia, segundo Cabré (1993), possui relações com diversas disciplinas, tais como linguística, ciência cognitiva, comunicação, informática e a documentação e, como destaca a referida autora (p. 112, tradução livre) é “medi-

ante a terminologia que os especialistas em documentação descrevem o conteúdo dos documentos que deve formar uma biblioteca especializada”. 112). No entanto, essa relação também pode ser visualizada de forma aplicada nos arquivos quando se refere à organização do conhecimento arquivístico, e em especial nos seus procedimentos de classificação e descrição.

O Manual de arranjo e descrição de arquivos, publicado em 1898, em seu capítulo VI, refere-se ao “uso convencional de certos termos e sinais” e, no tópico 84, destaca que “No interesse da clareza, é proveitoso usar-se sempre a mesma terminologia nos vários inventários”, reportando-se à questão da uniformidade (Arquivo Nacional, 1973, p.147). Essa questão terminológica, como subsídio à organização do conhecimento arquivístico, encontra bases operacionais na análise documental (AD) que, segundo Smit e Guimarães (1999), refere-se ao conjunto de procedimentos que buscam expressar o conteúdo dos documentos a fim de facilitar a recuperação da informação. Tal análise, como destacam os autores, ocorre tanto na fase corrente quanto na fase permanente dos documentos de arquivo, em aspectos ligados à elaboração de planos de classificação e de tabelas de temporalidade assim como na descrição documental. Dessa forma, a definição terminológica torna-se especialmente necessária para designar as funções, as atividades e/ou as temáticas das séries documentais, o que implica controle terminológico e estruturação dos conceitos, bem como normalização dos termos utilizados para descrever o conteúdo dos documentos. Esse controle terminológico influi diretamente na qualidade da representação e da recuperação da informação, pois o vocabulário comum corresponde ao código comunicacional que o usuário deve utilizar para ter acesso à informação. (SMIT; GUIMARÃES, 1999)

A necessidade de estabelecer um código comunicacional é própria dos estudos terminológicos, sendo considerada por Cabré (1993) uma função precípua desses estudos. Assim, para a autora (p.37) “a terminologia só tem sentido em relação às linguagens de especialidade e de comunicação, podendo responder e atingir diferentes finalidades, e todas elas relacionadas com o mundo da comunicação e da informação”, o que revela sua dimensão linguística e comunicativa e reitera aquilo que considera como as três grandes tendências de estudos terminológicos: uma orientada ao sistema linguístico, outra à tradução e uma de planificação linguística (normalização).

Os termos, segundo Krieger e Finatto (2004), têm duas funções essenciais: a de representar e transmitir o conhecimento especializado. Ao longo dos anos, a arquivística vem se dedicando a essa questão na medida em que eu, como destaca Daniels (1999), desde a Segunda Guerra Mundial os arquivistas do mundo todo vêm dando atenção ao estudo terminológico de tal forma que “a terminologia arquivística é um grupo flexível de palavras em comum que tem um significado especializado para os arquivistas” (Daniels, 1999).

Bellotto (2007) discute a terminologia como meio de expressão e comunicação técnica, ressaltando a função de um dicionário terminológico bem como as vantagens e justificativas para um controle terminológico na área dos arquivos. Destaca a autora que “os dicionários, ou pelo menos, as suas edições, devem ser renovadas, refletindo os avanços e as obsolescências da área, evidentes no surgimento de novos termos e no desaparecimento ou metamorfoses de outros” (Bellotto, 2007, p. 52). Esse instrumento facilita e melhora o entendimento e a apropriação terminológica tanto por estudiosos avançados quanto por iniciantes na área, na medida em que:

Uma terminologia própria da arquivística é um dos elementos essenciais para a definitiva consolidação, não só da profissão do arquivista, como da própria área, contribuindo para uma maior nitidez dos seus contornos, de modo a distingui-la das outras profissões e áreas do conhecimento (Bellotto, 2007, p.53)

Dryden (2005), ao abordar a importância dos estudos terminológicos no âmbito da arquivística, destaca o manual dos arquivistas holandeses como o marco inicial da padronização terminológica na área. Nesse contexto, a referida autora resgata exemplo como a dificuldade de definição do “*respect des fonds*” em outras línguas e relata exemplos de tentativas de padronização em âmbito nacional e internacional, tais como a *Standardized Terminology List*, de 1986, dos arquivistas de Ontario, e o primeiro dicionário internacional da área, o *Elsevier's Lexicon of Archive Terminology*, de 1964 que, embora buscando uma abrangência internacional, partia de definições em francês para chegar a termos equivalentes nas outras cinco línguas: alemão, inglês, espanhol, italiano e holandês (Dryden, 2005, p.7).

Em 1977, o ICA formou um comitê para revisar o léxico de 1964 e, em 1984, publicou o *Dictionary of Archival Terminology*, que foi estruturado em ordem alfabética dos termos em inglês e francês, bem como suas respectivas definições.

Não são apresentadas definições para os termos em holandês, alemão, italiano, russo e espanhol, para esses aparecem apenas os termos correspondentes.

Dryden (2005) ainda apresenta as dificuldades de padronizar os termos da área. No entanto, para superar tais dificuldades é preciso trazer novas abordagens, tais como o conceito de garantia cultural (Beghtol, 2002, 2005), que dialoga com a ideia de contexto terminológico.

Duchein (2007), ao referir-se ao *Dictionary of Archival Terminology*, retrata três dificuldades relacionadas à tradução: a imprecisão de definições e de usos nacionais, a divergência de vocabulário de uma mesma língua entre países homófonos, e o fato de a arquivologia possuir uma relação intrínseca com os sistemas jurídicos e administrativos de cada país, aspecto que se reflete no vocabulário, em que todo um conjunto de conceitos passa a ser dificilmente transportável de um país a outro. No entanto, apesar de tais dificuldades, Duchein (2007, p.21-22) destaca a importância da arquivística internacional comparada, na medida em que, a partir dela, pode-se observar não apenas as diferenças contextuais mas, e principalmente, quais são as bases teóricas da arquivologia que transcendem os limites territoriais, como é o caso do *respect des fonds*, que subsidia teoricamente as diferentes práticas profissionais na área, ainda que sua tradução seja sempre uma aproximação, como é o caso do termo “princípio da proveniência”.

O ICA, no ano de 2010, com Secção de Educação Arquivística e Treinamento (ICA-SAE), abraçou o desafio de criar um banco de dados de terminologia arquivística. O projeto foi liderado por Luciana Duranti, da Universidade de British Columbia, e consolidado pelo ICA e pelo *InterPARES*. No estágio inicial do projeto, um grupo formado por membros da ICA-SAE, escolheu 300 conceitos da área e os identificou nos países de língua inglesa, o que resultou em um conjunto de 320 termos em inglês.

Num segundo momento, um grupo composto por acadêmicos e profissionais da área identificou as definições dos termos utilizados em países de língua inglesa assim como os termos e definições correspondentes, caso existentes, em outras quinze línguas. O projeto resultou na publicação, pelo ICA, em 2013, da ferramenta *Multilingual Archival Terminology*, disponível, a arquivistas e pesquisadores, em catalão, chinês, holandês, inglês, francês, alemão, grego, italiano, japonês, polonês, português, punjabi, russo, espanhol, sueco e croata, com uma função interativa, em que a base oferece aos membros do

ICA a oportunidade de adicionar línguas, termos e definições.

### 3. Classificação e descrição: atividades-núcleo para a organização e representação do conhecimento arquivístico

Considerando que a classificação e a descrição constituem o cerne dos procedimentos de organização e representação do conhecimento arquivístico, em um ambiente que produz, recebe, envia, gere e utiliza documentos como fonte de informação e conhecimento.

A classificação, segundo Sousa (2007a) é uma atividade matricial, que precede principalmente as atividades de avaliação e descrição, sendo “[...] uma tarefa fundamental, um momento especial no processo de organização, primeiro passo para permitir o acesso à informação” (Sousa, 2007a, p.6).

No entanto, apesar de existir um consenso sobre a atividade de classificação, ainda há muito o que avançar sobre o assunto pois, como aponta Sousa (2007b), a arquivística ainda não utiliza, na classificação, as contribuições da filosofia, da teoria da classificação e da teoria do conceito, e nem tampouco efetiva uma comunicação mais intensa com outras disciplinas da própria área. Tal aspecto se faz especialmente importante em um momento em que a arquivística contemporânea se depara com grandes volumes de documentos acumulados diariamente, exigindo uma maior sofisticação dos esquemas de classificação e o desenvolvimento teórico da organização de documentos arquivísticos.

Na literatura brasileira, por algum tempo, predominou o pensamento da distinção da classificação entre os arquivos correntes e permanentes, utilizando o termo classificação para o arquivo corrente e arranjo para o arquivo permanente, fundamentados na literatura de Schellenberg, que atribuiu aos *records management* a responsabilidade pelo arquivo corrente. No entanto, como destaca Sousa (2007b), iniciaram-se no Brasil e no Canadá, pensamentos que “[...] perceberam a classificação como uma função a ser desenvolvida por um método que independe da fase ou idade dos documentos ou dos seus vários usos” e que o “[...] uso que se faz dos conjuntos documentais altera-se com as idades, ou melhor, novos usos vão sendo agregados, mas essa é uma questão a ser resolvida por uma outra função arquivística: a descrição” (Sousa, 2007b, p.85).

Surge então, a necessidade de estabelecer alguns significados sobre os conceitos que são

utilizados no processo classificatório, que são: classificação, ordenação, arquivamento, codificação e instrumentos de classificação. Atenta-se aqui para os três primeiros, sendo que o termo classificação é utilizado para “identificar a ação intelectual de construir esquemas para agrupar documentos e partir de princípios estabelecidos”; a ordenação é “a forma de disposição dos tipos documentais dentro das divisões estabelecidas no esquema de classificação” e o arquivamento é “a ação física de colocar os documentos em pastas ou caixas orientadas pelo esquema de classificação e pela ordenação definida” (Sousa, 2007b, p.85). Tem-se, então, a classificação como o princípio para organizar o conhecimento arquivístico, uma ação intelectual que elabora e constrói categorias a partir do princípio da proveniência.

Segundo Lopez (2002, p. 12), as “atividades de classificação só conseguem ter seus objetivos plenamente atingidos mediante a descrição documental”. Para a construção de uma descrição que atenda e reflita, da melhor forma possível, o fundo documental, é necessário que essa se inicie simultaneamente com a classificação abrangendo o documento tanto em sua fase corrente como em sua fase permanente, e os instrumentos gerados a partir da descrição (guias, inventários, catálogos, índices) precisam ser sempre revistos, mantendo-se atualizados.

Bellotto (1991), ao tratar sobre a descrição nos arquivos permanentes, aponta que “o processo da descrição consiste na elaboração de instrumentos de pesquisa que possibilitem identificar, rastrear, localizar dados, seja pela via sumária, seja pela analítica” (Bellotto, 1991, p.108). A informação sumária está ligada à representação do conjunto documental, enquanto a analítica consiste na representação da unidade documental.

A Norma Geral Internacional de Descrição Arquivística – *ISAD (G)*, elaborada pelo *ICA*, apresenta como objetivo para a descrição: “identificar e explicar o contexto e o conteúdo de documentos de arquivo a fim de promover o acesso aos mesmos”. Nesse contexto, a referida norma destaca que os processos relacionados à descrição “podem começar na ou antes da produção dos documentos e continuam durante sua vida”, permitindo “instituir controles intelectuais necessários para tornar confiáveis, autênticas, significativas e acessíveis descrições que serão mantidas ao longo do tempo” (Conselho Internacional de Arquivos, 2000, p. 11).

Heredia Herrera (1995), define a descrição como “[...] a ponte que comunica o documento com os usuários” e o arquivista é o profissional

que está na “cabeceira da ponte”, realizando uma tarefa de “[...] análise que supõe identificação, leitura, resumo e indicação que transmite ao usuário [...]”. (Heredia Herrera, 1995, p.300, tradução livre).

As atividades de classificação e a descrição devem ser observadas como atividades que se originam de um mesmo núcleo – o arquivo – de forma simultânea, tendo como mesmo objetivo dar acesso à informação registrada. Por ocorrerem de forma simultânea, pressupõe-se que seja elaborada uma linguagem normalizada dos termos utilizados para organizar e representar essa informação, levando em conta o fundo ao qual esta pertence.

Assim, a necessidade de estudos terminológicos para a arquivística, tanto para organizar e representar o conhecimento no ambiente dos arquivos, quanto para a delimitação e consolidação da área e sua comunicação científica, torna-se cada vez mais premente.

#### 4. Metodologia

A partir da identificação dos padrões prototípicos dos dicionários terminológicos propostos por Krieger e Finatto (2004), elaborou-se, para a coleta de dados junto ao *Multilingual Archival Terminology* do *ICA*, uma ficha terminológica que permitiu observar os verbetes nos seguintes aspectos: palavras-entrada (verificando o registro na forma utilizada), equivalentes em língua estrangeira, informação de categoria gramatical, fontes contextuais, fontes bibliográficas, gradação sinonímica, remissivas, notas explicativas (linguísticas, técnicas, enciclopédicas). Além dessas informações, sentiu-se a necessidade de criar o campo “observações”, para realizar alguns comentários durante a realização da coleta dos dados, como se verifica a seguir (Figura 1).

| Verbetes 00   |  |
|---|--|
| Palavra-entrada: registro na forma utilizada                |  |
| Equivalentes em língua estrangeira                          |  |
| Informação de categoria gramatical                          |  |
| Informação conceitual                                       |  |
| Fontes contextuais  |  |
| Fontes bibliográficas                                       |  |
| Gradação sinonímica   |  |
| Remissivas  |  |
| Notas explicativas (linguísticas, técnicas, enciclopédicas) |  |
| Observações   |  |

Figura 1. Ficha de coleta de dados. Fonte: adaptado pelos autores a partir de Krieger e Finatto (2004).

A realização da coleta teve início a partir do termo *classificação* no idioma português, efetuando-se busca por meio do radical “clas\*” (que é comum às quatro línguas analisadas). Desse modo cobriram-se 6 termos relacionados a classificação, sendo: classificação, classificar, desclassificar, documento classificado, plano de classificação e sistema de classificação. A partir desses 6 termos e das equivalências apresentadas em cada um deles, foram preenchidas as demais fichas em espanhol, francês e inglês.

O mesmo procedimento foi realizado para o termo *descrição*, e a busca foi realizada pelo radical “descr\*”, permitindo a localização dos seguintes termos: descrição, descrição arquivística, nível de descrição e unidade de descrição.

A análise de conteúdo de Bardin (2009) foi utilizada no quesito *informação conceitual* de modo a identificar se há diversidade nas conotações conceituais atribuídas ao termo em cada língua. Para tanto foram estabelecidas duas categorias de análise, uma buscando o conceito em comum e, outra, as diferenças dos verbetes analisados.

## 5. Apresentação dos resultados

No tocante à descrição, e tendo por base os dez campos do formulário de coleta de dados, obtiveram-se os seguintes resultados:

Quanto à Palavra-entrada, foi identificado um total de 19 verbetes tendo-se, em português, *classificação*, *classificar*, *desclassificar*, *documento classificado*, *plano de classificação* e *sistema de classificação*, em espanhol, *clasificar*, *documento de archivo clasificado*, *cuadro de clasificación*, em francês, *classification*, *classifié* e *plan de classement* e, em inglês, *classification*, *classify*, *declassify*, *classified records*, *file plan*, *classification scheme* e *classification system*.

Quanto ao segundo quesito Equivalentes em língua estrangeira, foram encontradas, em cada verbete, as seguintes informações:

**Classificação:** *clasificar* (espanhol), *classification* (francês) e *classification* e *classify* (inglês).

**Classificar:** *clasificar* (espanhol), *classification* (francês) e *classification* e *classify* (inglês). É necessário ressaltar que as informações conceituais tanto que classificação como classificar são semelhantes, vindo da mesma fonte bibliográfica, no entanto o verbete classificar apresenta uma definição a mais.

**Desclassificar:** apenas *declassify* (inglês). Ressalta-se aqui que apesar de não apresentar

equivalência em espanhol e francês, não quer dizer que nesses idiomas não há esse termo. No espanhol, foi encontrado o verbete *desclasificar*, e no francês, encontrou-se *déclassification*, ambos não apresentam equivalência no idioma português.

**Documento classificado:** *documento de archivo clasificado* (espanhol), *classifié* e *classification* (francês) e *classified records* (inglês).

**Plano de classificação**, semelhante ao termo **desclassificar**, apresenta sua equivalência somente no inglês, *file plan*. No entanto, no espanhol foi encontrado o verbete *esquema de clasificación*, com a remissiva ver *cuadro de clasificación*, também foi encontrado o verbete *sistema de clasificación*, relacionados ao *filig system* (inglês) e ao *sistema de arquivamento* em português.

Os termos relacionados ao verbete *file plan* (inglês) são: *cuadro de clasificación* (espanhol), *plan de classement* (francês) e *plano de classificação*, *quadro de arranjo* e *sistema de classificação* (português).

**Sistema de classificação:** *cuadro de clasificación* (espanhol), *plan de classement* (francês) e *classification scheme*, *classification system* e *file plan* (inglês).

No tocante à categoria gramatical, observa-se que, do total de verbetes analisados 31% são verbos e 68% são substantivos, sendo que, no Francês, o verbete *classification* apresenta-se tanto como substantivo quanto como verbo e, o verbete *classifié*, como adjetivo.

Quanto a informação contextual tem-se os seis verbetes em português relacionados ao Brasil. No caso do espanhol, tem-se dois verbetes relacionados ao México e um à Argentina. No francês, tem-se um verbete relacionado ao Canadá e os outros dois sem informação contextual. No caso do inglês, nenhum dos verbetes apresenta informações contextuais.

A partir da análise de conteúdo (Bardin, 2009), foi possível verificar que, nos termos relacionados à Classificação, o seguinte:

Relativamente ao termos *Classificação*, as definições nos quatro idiomas analisados convergem ao referirem-se à classificação como a organização dos documentos de acordo com o plano de classificação, assim como os processos de análise e identificação dos documentos, e elaboração de categorias e grupos. No entanto, as definições em francês e em inglês acrescentam a possibilidade de a classificação servir também para limitar o acesso a determinados

documentos, de acordo com a necessidade administrativa.

No que se refere ao termo *Classificar*, representado em português e em inglês, fica patente a ideia de organização dos documentos e de análise e identificação do conteúdo para elaboração de categorias para que o documento possa ser recuperado.

O termo *Desclassificar*, representado em português e em inglês, corresponde a liberação à consulta de um documento que anteriormente foi sujeito à um grau de sigilo

Para *Documento classificado*, em português, espanhol e inglês tem-se o documento que foi submetido a um código ou sistema de classificação. Nesse contexto, é interessante notar também a relação desse termo ao grau de sigilo do documento, aspecto que verificado em português, francês e inglês, sem correspondência no espanhol.

Já no que se refere ao *Plano de Classificação*, em português e em inglês tem-se a distribuição dos documentos a partir da sua produção, visando à sua recuperação. No caso do português, observa-se uma forte referência a arquivos correntes, sendo importante destacar que aspecto que essa ideia de distinção entre plano de classificação para arquivo corrente e quadro de arranjo para arquivo permanente já foi superado na literatura arquivística brasileira, que atualmente utiliza o termo plano de classificação tanto para documentos correntes como permanentes. Na versão em espanhol, tem-se a menção a um instrumento que decorre do ato de classificar que formaliza e representa a sistematização do conteúdo documental de um fundo, de um arquivo ou de uma coleção. Mas é em francês que esse termo apresenta maior complexidade conceitual, como atividade de categorização e ordenação de atividades, organização hierárquica de das séries, dossiês, sub-dossiês, volumes e documentos, como representação conceitual e taxonômica das atividades funcionais de uma organização ou como instrumento hierárquico e lógico que subsidia a descrição e a recuperação. Tais aspectos se reiteram na versão em inglês, que se refere a planos, diagrama, mapa ou padrão, segundo categorias previamente estabelecidas e logicamente estruturadas segundo regras procedimentais utilizados para descrever e organizar materiais com características similares, organizando e relacionando os conceitos neles contidos.

Por fim, a questão do Sistema de Classificação, apresenta-se, em português e em inglês, com uma abordagem mais geral, com referência a

um sistema no qual o material relacionado é arquivado sob um tema mais amplo e seus subtemas.

As informações bibliográficas relacionadas aos verbetes foram, na ordem dos mais citados: *A Glossary Archival and Records Terminology* (citados 8 vezes), *Dicionário de Terminologia Arquivística* (citado 5 vezes em), *Glossary of Records and Information Management Terms* (citado 5 vezes), *Dictionary of Archival Terminology DAT III* (citado 3 vezes), *Glosario InterPARES de Preservación Digital: InterPARES 3* (citado 2 vezes), *ICA-REQ* (citado 2 vezes), *InterPARES 2 Terminology Database* (citado 2 vezes), e o *Aprendix 1: Dictionary of technical terms, Art and Architecture Thesaurus, Glossary of Archival and Recordkeeping Terms, La grande dictionnaire terminologique, Lenguaje y vocabulario arquivístico algo más que um dicionario, Les fonctions de l'archivistique contemporaine, MoReq2, Normes et procédures archivistiques des Archives nationales du Québec, Nouveau glossaire de l'archivage e Select List of Archival Terminology*, foram citados uma vez.

Quanto aos quesitos gradação sinonímica, remissivas e notas explicativas (linguísticas, técnicas, enciclopédicas), nada foi encontrado.

Com relação à Descrição, a partir da Palavra-entrada em português por meio dos verbetes *descrição, descrição arquivística, nível de descrição e unidade de descrição*, chegou-se, em espanhol, aos verbetes *descripción, descripción archivística e unidad de descripción*, em francês, a *description, description archivistique e unité archivistique à décrire* e, em inglês, a *description, archival description, levels of description e unit of description*, totalizando 14 verbetes relacionados à descrição.

No tocante aos *Equivalentes em língua estrangeira*, tem-se:

**Descrição:** *descripción e descripción archivística* (espanhol), *description e description archivistique* (francês) e *archival description e description* (inglês), sendo que estes últimos apresentam os links duplicados, mas encaminhando para o mesmo conteúdo.

**Descrição arquivística**, apresentanda as mesmas informações do verbete **descrição**.

**Nível de descrição**, apresenta equivalência somente no inglês *levels description*. No entanto, ao analisar a listagem em espanhol, há o verbete *nível de descripción*. O mesmo ocorre com na listagem em francês, onde aparece o verbete *niveau de description*. Porém, nenhum

destes últimos aparece relacionado ao português.

**Unidade de descrição:** *unidad de descripción* (espanhol), *unité archivistique à décrire* (francês) e *unit of description* (inglês).

Na categoria gramatical, dos idiomas analisados, 100% dos verbetes são substantivos, de acordo com as informações apresentadas nos verbetes.

Na informação contextual tem-se, no português, os quatro verbetes relacionados ao Brasil, no espanhol, dois relacionados a Argentina e dois ao México, no francês, um termo relacionado ao Canadá e dois sem fontes contextuais e, no inglês, nenhum deles apresenta relações contextuais.

No âmbito da Descrição, observa-se, de pronto, a inexistência de divergências conceituais nos idiomas analisados. Nesse sentido tem-se que termo *Descrição* apresenta convergência conceitual nos quatro idiomas analisados, enquanto um conjunto de procedimentos que analisa elementos formais e de conteúdo dos documentos, afim de recupera-los. Essa convergência se reitera quando se verticaliza a abordagem para *Descrição arquivística*, entendida, nos quatro idiomas, como uma representação precisa de uma unidade de descrição levando em conta seu contexto de produção, para uma futura recuperação.

No que se refere ao *Nível de descrição*, previsto apenas em português e inglês, tem-se a posição em que a unidade de descrição (o documento), se encontra na hierarquia estabelecida no fundo ou na coleção analisada enquanto a Unidade de descrição, prevista nos quatro idiomas, compreende os documentos (o fundo, grupo, subgrupo, séries e item documental) que são objetos de análise para o processo de descrição.

As informações bibliográficas relacionadas aos verbetes foram: Dicionário Brasileiro de Terminologia Arquivística (citado 5 vezes), ISAD (G): Norma geral internacional de descrição arquivística (citada 5 vezes), *Dictionary of Archival Terminology DAT III* (citado 4 vezes), *A Glossary of Archival and Records Terminology* (citado 3 vezes), Bureau canadiens de archivistes, Règles pour la description des documents d'archives (citado 2 vezes), *Glosario InterPARES de Preservación Digital: Parte InterPARES 3* (citado 2 vezes), e o *Dictionnaire de terminologie, ISAAR (CPF): International Standard Archival Authority Record for Corporate Bodies, Lenguaje y vocabulário arquivísticos* algo mais que um dicionário, foram citados uma vez.

Quanto à gradação sinonímica e notas explicativas (linguísticas, técnicas, enciclopédicas), nada foi encontrado.

O quesito remissiva revelou, para o verbete *descripción* (espanhol), a remissiva *descripción arquivística*, acompanhada de informações conceituais e para o verbete *description* (inglês), a remissiva *archival description*, sem informações conceituais.

## 6. Considerações Finais

Como se pode observar, a classificação e a descrição ocupam espaço nuclear no que se refere aos procedimentos envolvidos na organização do conhecimento arquivístico e se regem, por um lado, por princípios internacionalmente consolidados, mas, por outro lado, ficam à mercê dos aspectos contextuais inerente a cada regime jurídico-administrativo do país em que se situam os arquivos.

Isso traz uma maior complexidade à questão terminológica na área, em especial no que se refere aos léxicos especializados, uma vez que a diversidade de línguas envolvidas exige que se vá além da dimensão da tradução para incorporar os aspectos idiossincráticos inerentes a cada cultural envolvida, confirmando questões anteriormente discutidas por Hudon (1997) e Beghtol (2002, 2005), e reiterando que as atividades de classificação e de descrição arquivística trazem, consigo, um “poder de nomear” (Olson, 2002) concedido pela sociedade ao arquivista, dele exigindo, portanto, uma “ética transcultural de mediação” (García Gutiérrez, 2002).

Foi possível também observar, no instrumento terminológico analisado, que aspectos importantes, como os de gradação sinonímica e notas explicativas não se fizeram presentes, o que leva a uma maior dificuldade de aplicação desse léxico. Alia-se a isso uma heterogeneidade de tratamento dos termos nas diferentes línguas, na medida em que aspectos como fontes contextuais e remissivas, fazem-se presentes ora sim ora não. Observa-se, ainda, casos de duplicação de informação, tanto nas definições, quanto termos de equivalências de outro idioma.

As informações de categorias gramaticais, variaram entre verbos e substantivos, o que se revelou consistente pois a classificação e a descrição são, por um lado, processos, e, por outro, institutos do campo arquivístico.

No que tange às fontes utilizadas para a configuração do instrumento terminológico utilizado destacam-se duas fontes altamente utilizada

para orientar a configuração de definições em apenas uma língua, como o *Glossary of Archival and Records Terminology* (2005), para o inglês, e o *Glosario InterPARES de Preservación Digital: Parte InterPARES 3* (2012), para o espanhol.

Por outro lado, observaram-se fontes que ofereceram subsídios para a configuração de definições em mais de um idioma, como é o caso do Dicionário Brasileiro de Terminologia Arquivística (2005), para português e espanhol, do Dictionary of Archival Terminology DAT III (1999; 2004) e do Bureau canadiens des archivistes, Règles pour la description des documents d'archives (1990, 2008), e do Glossary of Records and Information Management Terms (2007), para inglês e português. Tais aspectos revelam que a área já dispõe de um conjunto respeitável de fontes terminológicas que, por sua vez, fornecem importantes subsídios para um instrumento terminológico multilíngue, como o ora analisado.

O instrumento analisado revelou forte preocupação com as variações terminológicas das quatro línguas (originalmente europeias) em seus usos no continente americano (com referências, em ordem decrescente, a Brasil, México, Argentina e Canadá), o que reitera, por um lado, a importância desses encontros no cenário da ciência Arquivística na atualidade e, por outro, o reflexo da especificidade de cada um desses universos contextuais para a configuração terminológica da área.

Retomando as discussões de Michèle Hudon (1997), para quem o conceito de um termo se relaciona ao contexto cultural que está inserido. Assim, para a autora, o grau de arbitrariedade inerente à relação termo-conceito em um mesmo idioma torna-se ainda mais complexa quando de têm diferentes idiomas envolvidos – como é o caso do instrumento analisado na medida em que cada idioma recorta a realidade de modo diverso. A vista disso, a falta de informações contextuais, como as observadas na análise, pode gerar confusão quando se trata da comunicação científica em uma área, bem como para o seu avanço profissional.

Espera-se, com este estudo, contribuir para o aperfeiçoamento do *Multilingual Archival Terminology* do ICA, em especial no que se refere à inserção, configuração e relacionamento entre seus termos, haja vista a inegável importância e utilidade do referido instrumento terminológico, para a área arquivística, sempre tão permeada por nuances culturais, jurídicas, administrativas políticas, e em constante evolução.

## Referencias

- Beghtol, C. (2002) A proposed ethical warrant for global knowledge representation and organization systems. *Journal of Documentation*, London, v. 58, n. 5, p. 507-532.
- Beghtol, C. (2005) Ethical decision-making for knowledge representation and organization systems for global use. *Journal of the American Society for Information Science and Technology*, New York, v. 56, n. 9, p. 903-912.
- Bellotto, H.L. (2007). A Terminologia das Áreas do Saber e do Fazer O caso da arquivística. *Acervo*, Rio de Janeiro, v.20, nº1-2, p.47-56.
- Cabré, M. T. (1993). La terminología: teoría, metodología, aplicaciones. Barcelona: Antártida/Empúries, 1993.
- Conselho Internacional de Arquivos. (2000) ISAD(G). Norma Geral Internacional de Descrição Arquivística. 2. ed. Rio de Janeiro: Arquivo Nacional, 2000. Recuperado 13-01-2015 de [http://www.conarq.arquivonacional.gov.br/Media/publicacoes/isad\\_g\\_2001.pdf](http://www.conarq.arquivonacional.gov.br/Media/publicacoes/isad_g_2001.pdf)
- Daniels, M.F. (1999). Introduction to Archival Terminology. Recuperado 14-01-2015 de <http://www.archives.gov/research/alic/reference/archives-resources/terminology.html#note>
- Dryden, J. (2005), A Tower of Babel: Standardizing Archival Terminology. *Archival Science* v.5, p.1-16. Recuperado 14-01-2015 de [http://download.springer.com/static/pdf/875/art%253A10.1007%252Fs10502-005-9001-3.pdf?auth66=1421838555\\_7eb97fda7158f6f0475e1a42a0531acd&ext=.pdf](http://download.springer.com/static/pdf/875/art%253A10.1007%252Fs10502-005-9001-3.pdf?auth66=1421838555_7eb97fda7158f6f0475e1a42a0531acd&ext=.pdf)
- Duchain, M. (2007). Os Arquivos na Torre de Babel Problemas de terminologia arquivística internacional. *Acervo*, Rio de Janeiro, v.20, n.1-2, p.13-22.
- Herredia Herrera, A. (1995). *Arquivística general: teoría y práctica*. 7. ed. Sevilla: Diputación Provincial de Sevilla.
- Hudon, M. (1997). Multilingual thesaurus construction: integrating the views of different cultures in one gateway to knowledge and concepts. *Knowledge Organization*, Würzburg, v. 24, n. 2, p. 84-91.
- Krieger, M. G.; Finatto, M. J. B. (2004). *Introdução à terminologia: teoria e prática*. São Paulo: Contexto.
- Muller, S.; Feith, J.A.; Fruin, R. (2003). *Manual for the arrangement and description of archives*. Chicago: Society of American Archivists.
- Smit, J. W.; Guimarães, J.A.C. (1999). Análise documental. In: Johanna W. Smit; Yedda Dias Lima; Heloísa Liberalli Bellotto. (Org.). *Organização de arquivos*. São Paulo: IEB-USP, v. 3, p. 63-77.
- Sousa, R. T.B. de. (2007b). A classificação como função material do que-fazer arquivístico. In: SANTOS, Vanderlei Batista dos (org.). *Arquivística. Temas contemporâneos: classificação. Preservação Digital. Gestão do Conhecimento*. Distrito Federal: SENAC.
- Sousa, R.T. B de. (2007a). Os princípios da teoria da classificação e o processo de organização de documentos de arquivo. *Arquivo&Administração*. Rio de Janeiro, v.6, n.1, p. 5-26.
- Tognoli, N.B. T.; Guimarães, J. A. C. (2011). A organização do conhecimento arquivístico: perspectivas e renovação a partir das abordagens científicas canadenses. *Ciência da Informação*, v.16, n.1, p.21-44.
- Tognoli, N.B.; Guimarães, J.A.C.; TENNIS, J.T. (2013). Diplomats as a methodological perspective for archival knowledge organization. In: *The Fourth North American Symposium on Knowledge Organization - NASKO, 2013, Milwaukee*. NASKO 2013- Transition Cultures, Transition KO: Evolving Exploration, Critical Reflection, and Practical Work, v. 1. p. 216-227.

---

# Políticas de indización en bibliotecas escolares de Brasil y Portugal: análisis comparativo

*Indexing policies in Brazilian and Portuguese school libraries: a comparative analysis*

---

Ana Lúcia Terra (1), Mariângela Fujita (2), M<sup>a</sup> Carmen Agustín Lacruz (3)

(1) Instituto Politécnico do Porto (Portugal)

(2) Universidade Estadual Paulista Júlio Mesquita Filho – Campus Marília (Brasil)

(3) Universidad de Zaragoza (España)

## Resumen

Considerando un marco teórico que incluye la indización dentro del Análisis Documental y de la Organización y Representación del Conocimiento, se describen los resultados obtenidos de una investigación realizada sobre las políticas de indización existentes en las bibliotecas escolares brasileñas y portuguesas.

Se describen de forma concisa los contextos legales y organizativos en los que se encuadran estas bibliotecas. Se detalla la metodología y el cuestionario desarrollado para la recogida y el proceso de los datos obtenidos de 150 bibliotecas escolares brasileñas y 406 portuguesas.

Se identifican los niveles educativos, la adscripción institucional y ubicación geográfica de las bibliotecas escolares de la muestra. Se determina la existencia y uso de catálogos. Se describen los lenguajes de indización utilizados así como la existencia de manuales para guiar la indización. Se detallan aspectos relacionados con el número de descriptores recomendados, las recomendaciones y normas utilizadas, comparando la realidad brasileña y portuguesa.

**Palabras clave:** Políticas de indización. Bibliotecas escolares. Brasil. Portugal.

## Abstract

The aim of this investigation is to study the indexing policies of Brazilian and Portuguese school libraries.

Legal and organizational contexts of these libraries are described. The methodology and the questionnaire create to gather and analyse data from 150 Brazilian and 406 Portuguese school libraries are explained.

Educational levels, institutional context and geographical location of the school libraries sample are identified. The existence and use of catalogues is identified. Used indexing languages are described as well as the existence of manuals to guide the indexing procedures. The number of selected descriptors is detailed, as well as the use of manual procedure and indexing standards, comparing the Brazilian and Portuguese reality.

**Keywords:** Indexing policy. Schools library. Brazil. Portugal.

## 1. Introducción y objetivos

Las bibliotecas escolares son centros de información que proporcionan recursos educativos impulsores de las innovaciones pedagógicas y lugares que democratizan el acceso al conocimiento. Sobre estos presupuestos, Portugal, desde finales del siglo XX y Brasil, a partir de 2010, impulsan la creación de bibliotecas escolares capaces de responder a estos desafíos.

Con la intención de crear servicios alineados con las necesidades de sus usuarios, la investigación científica sobre estas unidades de información puede impulsar el conocimiento de la

realidad y ayudar a desarrollar mejores bibliotecas escolares.

En este sentido, este texto muestra los resultados de una investigación sobre las políticas de indización que se desarrollan en las bibliotecas escolares de Brasil y de Portugal. Se explicitan los fundamentos teóricos que orientan el estudio, en particular en lo que concierne al concepto de política de indización y a sus concreciones en las funciones de organización y de recuperación de la información en el ámbito de las bibliotecas escolares. También se tratarán los contextos específicos, definidos en términos legales y

organizativos en los que se encuadran las bibliotecas escolares de Brasil y de Portugal.

Los datos sobre las políticas de indexación en estas unidades de información específicas se recogieron mediante una encuesta disponible en línea, distribuida a través de correo electrónico a las direcciones de las bibliotecas escolares en Brasil y Portugal. La encuesta fue igual en ambos países, salvo pequeñas adaptaciones puntuales, tenía 35 preguntas, la mayoría de las cuales eran de respuesta cerrada. El período de recogida de respuestas se desarrolló entre noviembre de 2012 y mayo de 2013. En Brasil se recogieron 150 encuestas y 406 en Portugal.

En este texto se analizan algunos de los principales datos recogidos en esta encuesta, para alcanzar los siguientes objetivos:

- 1) identificar los lenguajes de indexación utilizados en las bibliotecas escolares estudiadas.
- 2) estudiar las prácticas de cooperación entre bibliotecas escolares en materia de indexación, incluyendo el intercambio de registros y la participación en catálogos colectivos.
- 3) analizar aspectos relacionados con la especificidad de la indexación realizada en las bibliotecas escolares.

## 2. Marco conceptual: Indexación y políticas de indexación

Antes de adentrarnos en las políticas de indexación de las bibliotecas escolares, vamos a examinar el concepto de indexación desde la perspectiva del Análisis Documental y de la Organización y Representación del Conocimiento.

Dahlberg –fundadora de la Sociedad Internacional para la Organización del Conocimiento (ISKO) y creadora del esquema de clasificación de esta misma área– definió en 2006 la Organización del Conocimiento como "[...] los objetos y las actividades de la teoría del concepto, clasificación e indexación y representación del conocimiento [...]".

El esquema de clasificación concebido por Dahlberg, denominado "Systematifier", incluye las subdivisiones del área de Organización del Conocimiento clasificadas en 10 grupos, numerados de 0 a 9. La explicación de Systematifier, según Fujita (2008), es proporcionada por Dahlberg (1993) a partir de la secuencia de sistematización de los grupos de 1 a 9, sin considerar el Grupo 0, en tres subdivisiones:

*Grupos 1-3:* representan las divisiones componentes/constituyentes del área, caracterizadas por 1) Fundamentos teóricos, 2) Estructura y

construcción de sistemas de clasificación y tesauros, e 3) Clasificación e Indexación;

*Grupos 4-6:* representan las aplicaciones de las divisiones componentes/constituyentes del área: 4) Sistemas Universales, 5) Sistemas de clasificación orientados a objeto y tesauros, 6) Sistemas orientados a materias específicas;

*Grupos 7-9:* representan la influencia, aplicación y ambientes del área, 7) influencias externas: problemas de la representación del conocimiento por el lenguaje y la terminología, 8) aplicación de la clasificación e indexación en diferentes tipos de datos, enunciados, documentos, y 9) niveles nacionales e internacionales de la organización del conocimiento, en lo relativo a educación y formación, aspectos jurídicos y económicos, estudios de usuarios y normalización.

El Grupo 3, de forma específica, considera la indexación en el ámbito de la Organización del Conocimiento a partir de sus aspectos teóricos y metodológicos relativos a los fundamentos conceptuales, proceso y técnica, indexación automática y evaluación y calidad de la indexación.

*Grupo 3:* Metodología de clasificación e indexación. Contenidos: Teoría de Clasificación e Indexación, Materias, Técnicas de Clasificación e Indexación, Clasificación e Indexación automática, Ordenación manual y automática, Generación de índices, reclasificación, evaluación de la clasificación e indexación, Codificación y decodificación (FUJITA, 2008).

Con la perspectiva del ambiente institucional y de la concepción del conocimiento registrado, Hjørland (2008) considera que, en sentido específico, la Biblioteconomía y las Ciencias de la Documentación son las disciplinas centrales de la Organización del Conocimiento y comprende actividades tales como "[...] descripción documental, indexación y clasificaciones bibliotecarias, bases de datos bibliográficas, archivos y otros tipos de "instituciones de la memoria" en las que trabajan bibliotecarios, archiveros, analistas del contenido" (HJØRLAND, 2008, p. 86).

De forma específica, la Organización del Conocimiento se institucionalizó en ISKO. No obstante, tanto la indexación como la clasificación –entendidas como áreas de actividad de la Biblioteconomía y las Ciencias de la Documentación– tienen orígenes más antiguos y están vinculadas, en la concepción de Guimarães (2008), al desarrollo del Tratamiento Temático de la Información (TTI):

[...] a lo largo de la historia fue, poco a poco, siendo impregnado por distintas concepciones, que propiciaron la delimitación de tres corrientes teóricas –*catalogación temática, indexación y análisis documental*– que hoy encuentran espacio de interlocución en el marco de la *International Society for Knowledge Organization* (GUIMARÃES, 2008, p. 79).

El Tratamiento Temático de la Información incluye el conocimiento teórico y metodológico disponible sobre el tratamiento descriptivo del soporte material de la información así como el tratamiento temático del contenido de la información. Esta dicotomía se explica, por un lado por el distinto desarrollo teórico y metodológico alcanzado por cada una de las dos áreas, y por otro por el tratamiento diferenciado que exigen los aspectos materiales y temáticos.

De esta forma, el tratamiento del soporte material de los documentos es denominado Representación Descriptiva, según Kobashi (1994), porque se ocupa de la descripción normalizada de los aspectos físicos de los registros (título, autoría, edición, publicación, fecha y lugar de publicación) para, justamente, propiciar la identificación y accesibilidad física. Estas tareas son desarrolladas de forma específica, según Guinchat y Menou (1994) por la descripción bibliográfica o por la catalogación, en el contexto de la representación descriptiva.

El tratamiento del contenido recoge una descripción de los temas o materias de un documento para permitir una representación condensada orientada a la recuperación temática.

La indización como proceso de análisis temático tiene raíces teóricas y metodológicas de origen inglés y norteamericano. Tuvo desarrollos diferenciados dependiendo de los entornos institucionales y de las tipologías documentales tratadas. Así la indización se realiza en los servicios de indización y resumen sobre artículos de publicaciones periódicas y documentación científica en general, mientras que la catalogación temática se realiza en las bibliotecas, sobre monografías y publicaciones convencionales.

Debido a la necesidad creciente que tienen los servicios y sistemas de recuperar la información de forma rápida, precisa y especializada, su tratamiento cuenta con un aparato metodológico e instrumental más diversificado y orientado al contexto cada documento.

La indización, entendida como construcción de índices, es una práctica de tratamiento documental muy antigua. En las bibliotecas de la Antigüedad ya existían listas de los documentos custodiados. Cuando la ordenación de esos listados requirió una organización temática, se produjeron cambios importantes en la construcción de los índices. Es decir, se introdujo el análisis del contenido para crear representaciones documentales secundarias.

La indización, vista por Chaumier (1980, p. 42) como la "[...] parte más importante del Análisis Documental", es una combinación metodológica

estratégica entre el tratamiento del contenido de los documentos y su recuperación por un usuario, demostrando una relación estrecha entre el proceso y la finalidad de la indización.

La literatura sobre la indización y lenguajes documentales (tesauros) es muy importante en cantidad y calidad gracias a las aportaciones de Foskett, Lancaster, Austin, Farradane, Metcalfe, Aitchinson, Gilchrist, entre otros estudiosos, además de normas como ISO-5693 y ABNT-12.676.

Por otro lado, la catalogación temática propia de los entornos bibliotecarios evolucionó ligada a las prácticas profesionales de los catalogadores a partir de las normas establecidas por Charles Ammi Cutter. Según Guimarães (2008), tanto la catalogación temática como la indización son líneas teóricas incluidas dentro del tratamiento temático de la información (TTI).

La primera de las dos corrientes procede de la actividad profesional desarrollada en las bibliotecas, orientadas al desarrollo del catálogo como producto representativo del tratamiento de la información. Esta línea recibió una fuerte influencia de la Escuela de Chicago, de los principios establecidos por Cutter y de la tradición sobre encabezamientos de materia de la *Library of Congress*. En este contexto, la indización es todavía entendida por los profesionales de la información y por los bibliotecarios, de modo general, como operación que se realiza solamente en los servicios de información dedicados a producir bases de datos.

La evolución de las tecnologías de la información ha cambiado el modo en que se almacena, trata y recupera el conocimiento. Y las bibliotecas, que durante mucho tiempo han elaborado sus catálogos para una comunidad local que acudía a consultar sus fondos *in situ*, los disponen en la web ahora para su consulta en línea. Además, no deja de crecer la inmensa cantidad de documentos y recursos de información que están disponibles en la web sin que sea necesario acceder a bases de datos o catálogos bibliotecarios.

Las dos vías diferenciadas, entre los servicios de indización y resumen que producen bases de datos y las bibliotecas, que producen catálogos, ya fue puesta de manifiesto por Lancaister en el prefacio de su obra:

Aunque me concentré en la indización y redacción de resúmenes, como son practicadas por los servicios de indización y resúmenes en forma de publicación (en papel o soporte electrónico), y no en catálogos de bibliotecas, los principios son los mismos, y espero que este libro sea de interés y utilidad a aquellos preocupados con el mejoramiento del acceso temático en los catálogos en línea (LANCASTER, 1993, p. VII).

Compartiendo el enfoque convergente de Lancaster, consideramos necesario que las bibliotecas dispongan de propuestas de políticas de indización que aborden los parámetros, normas y factores que deben considerarse en la catalogación temática para mejorar el acceso a los catálogos en línea.

A pesar de que existe una considerable literatura científica dedicada a la indización, son pocos los autores que se han ocupado de su sistematización teórica y metodológica. Necesitamos retrotraernos a la publicación en 1991 de *Indización y resúmenes: teoría y práctica*, la obra clásica de Lancaster. También es muy destacable la aportación del profesor e investigador español Isidoro Gil Leiva que, en su *Manual de Indización: teoría y práctica*, se inserta en la estela anglosajona de la indización y no en la franco-española del análisis documental de contenido.

Aunque Lancaster no dedica un capítulo específico a las políticas de indización, sí que se encuentran en su monografía referencias a ella:

[...] Son muchos los factores que determinan si una búsqueda en una base de datos es o no exitosa. Estos comprenden la cobertura de la base de datos, su política de indización, sus reglas de indización [...] (LANCASTER, 1993, p. 4).

La principal decisión política es respecto a la exhaustividad de la indización [...]. Las decisiones políticas sobre exhaustividad no deben tomar la forma de límites absolutos sobre el número de términos a ser asignados. Más bien, la política podría sugerir una franja de términos [...] (LANCASTER, 1993, p. 23).

Por su parte Gil Leiva (2008) dedicó un epígrafe a las políticas de indización dentro del capítulo "Práctica de la indización" circunscrito a las bases de datos documentales producidas por servicios de indización y resúmenes AGRICOLA, INSPEC, CURRENT CONTENTS, ERIC y MEDLINE. Así mismo, desarrolla un ejemplo de políticas de indización de la biblioteca de la Universidad Nacional Autónoma de México.

A partir de las aportaciones de estos dos reconocidos autores –recogidas sintéticamente en la Figura 1)– podemos inferir que la aplicación teórica y metodológica de la indización al tratamiento temático de los documentos está vinculada a un conjunto de decisiones que se articulan, por un lado, al contexto de gestión de la información orientada a la recuperación y por otro, a los procedimientos, normativas e instrumentos de la indización. En este sentido, Fujita (2012, p. 22) define política de indización como "un conjunto de procedimientos, materiales, normas y técnicas orientadas por decisiones que reflejan la práctica y los principios teóricos de la cultura organizacional de un sistema de información."

| Lancaster (1993)  | Gil Leiva (2008)  |
|---|---|
| - Práctica de indización: exhaustividad y especificidad.  | a) Indización:<br>- Cualidades de la indización (exhaustividad y especificidad, corrección y consistencia);<br>- Normas sobre indización;<br>- Relación entre indización y recuperación;<br>- Indización en Internet; |
| - Principios de la indización: etapas, tematicidad, vocabularios controlados (especificidad del vocabulario, traducción). | b) Herramientas para la indización:<br>- listas de encabezamientos de materias y tesauros;<br>- interoperabilidad entre vocabularios controlados.   |
| - Métodos adoptados en los servicios de indización y resúmenes.   | c) Práctica de indización:<br>- proceso de indización;<br>- indización de documentos audiovisuales, sonoros, gráficos e textuales;<br>- política de indización.   |
| - Indización automática.  | d) Indización automática.   |
| - Evaluación de la indización: calidad y consistencia.  | e) Evaluación de la indización: intrínseca e extrínseca.  |

Figura 1: Aspectos de la indización según Lancaster y Gil Leiva

Considerando los diferentes aspectos abordados por Lancaster y Gil Leiva es posible que las bibliotecas inicien la elaboración de su política de indización estableciendo una secuencia de acciones:

1. Práctica de indización: proceso de indización, tematicidad, normas, manuales de indización, i indización automática;
2. Cualidades de la indización: exhaustividad, especificidad, corrección y consistencia;
3. Herramientas para la indización: listas de encabezamientos de materia y tesauros;
4. Evaluación de la indización.

En la formulación de la política de indización, cada uno de estos aspectos será evaluado en función de las características estructurales y de gestión de cada biblioteca o sistema bibliotecario, de tal forma que la política de indización de cada centro será específica, según establece la propuesta de Fujita (2012).

De esta manera, la política de indización determina procedimientos de análisis y representación de los temas; aspectos cualitativos de la indización como la precisión, la especificidad, la exhaustividad y la revocación de los términos; instrumentos de control de vocabulario, además de la evaluación de la indización.

Para Gil Leiva (2008, p. 297) esta concepción de la que política de indización incluye "[...] tanto

la forma de realizar la indización de una determinada institución como el esfuerzo por concretar, sistematizar y plasmar en guías los procesos seguidos en la indización”.

Garantizar una indización uniforme y de calidad es la premisa básica de la política de indización, pues habitualmente varios profesionales se dedican a estas tareas en lugares y tiempos distintos. Así, las bibliotecas o los sistemas bibliotecarios necesitan disponer de un manual para asegurar tanto la calidad como la uniformidad del trabajo. Gil Leiva (2008, p. 297) considera que “Esta plasmación de los procedimientos en guías suele estar en relación directa con el número de profesionales que se dedican a esta tarea o el grado de calidad perseguido por el sistema.” El manual de indización, por tanto, constituye una necesidad en estos servicios.

Para que las bibliotecas puedan cumplir las funciones que la sociedad les demanda es necesario que la política de indización de cada centro y todo aquello que la integra – metodologías para la representación descriptiva y temática, lenguajes documentales, estudios de usuarios, entre otros elementos– contemple de forma adecuada la elaboración de normas y procedimientos así como las técnicas de implantación, implementación y evaluación de la indización efectuada.

En síntesis, las bibliotecas necesitan de forma imperiosa sistematizar los procesos y las formas de indizar y para eso dependen de las políticas de indización.

### 3. Bibliotecas escolares

La biblioteca escolar es un componente fundamental de cualquier sistema educativo y su utilización es un elemento nuclear dentro de los procesos de enseñanza-aprendizaje orientados a la formación integral de niños y jóvenes capaces de convertirse en ciudadanos comprometidos con la mejora de su comunidad.

Actualmente, las bibliotecas escolares disponen de acceso a colecciones de recursos de información muy diversas, tanto temáticamente, como en cuanto a soportes y canales de comunicación.

De acuerdo con la *International Association of School Librarianship* (IASL, 1993), la biblioteca escolar debe estar directamente involucrada en todas las actividades escolares, cumpliendo cuatro funciones esenciales:

a) Informativa, proporcionando acceso fácil y rápido a información de calidad, integrada dentro de las redes regionales y nacionales de información;

b) Educacional, promoviendo la educación a lo largo de la vida a través de la creación de ambientes propicios para el aprendizaje; de la orientación en la búsqueda de información; de la integración dentro del currículo escolar y con apoyo de la libertad intelectual;

c) Cultural, sustentando la mejora de la calidad de vida a través del contacto con las artes, incentivando la creatividad y las relaciones humanas positivas;

d) Recreativa, fomentando la ocupación del tiempo libre mediante el acceso a la lectura literaria y a actividades enriquecedoras.

Según IASL (1993) "la sociedad que invierte en las bibliotecas escolares para sus niños, invierte en su propio futuro", y por ello consideramos que corresponde a los gobiernos, a través de políticas y programas adecuados, la creación y mantenimiento de bibliotecas escolares capaces de afrontar los retos de la Sociedad del Conocimiento.

En este contexto se enmarca el estudio de las iniciativas brasileñas y portuguesas que se describen en los siguientes apartados, y que prosigue una investigación anterior sobre herramientas de indización (AGUSTÍN; FUJITA y TERRA, 2014).

#### 3.1. Bibliotecas escolares en Brasil

En Brasil, el sistema de enseñanza público y privado, conforme a la Ley nº 9.394 de 20 de diciembre de 1996 que establece las directrices y bases de la educación nacional, establece la educación básica obligatoria y gratuita de los 4 a los 17 años de edad y la educación infantil gratuita para los niños hasta los 5 años de edad. La educación básica está organizada en preescolar, enseñanza básica o fundamental y enseñanza media.

Las escuelas de las redes municipales y los estados federados son públicas y gratuitas y se mantienen con presupuestos gubernamentales.

En la red de enseñanza municipal, el presupuesto procede del Gobierno Federal a través del Ministerio de Educación de Brasil, mientras que la red de enseñanza de los estados federados es financiada por cada Estado de la república federal de Brasil. Las escuelas privadas dependen de su propio presupuesto. Por tanto, el público que usa las bibliotecas escolares en Brasil está formado por niños y jóvenes de 4 a 17 años.

Pieruccini (2008) define esta biblioteca escolar como:

[...] un dispositivo complejo, compuesto de elementos heterogéneos: arquitectura y entorno, técnicas y tecnologías, procesos y productos, normas y reglamentos, contenidos materiales e inmateriales, responsable de la superposición de significados a los significados guardados por ella, convirtiéndose en elementos de su naturaleza.

Aunque cada escuela debería tener una biblioteca y por lo menos, un bibliotecario, es hecho que esto no es una realidad en todos los estados y municipios de Brasil. Sin embargo la regulación que establece las directrices y los planes de las políticas públicas, como el *Plan Nacional de Biblioteca en la Escuela* (PNBE) del Ministerio de Educación, recogen que la biblioteca escolar es una condición importante para la enseñanza en la escuela.

Campello (2011, p. 106) afirma que "a pesar de toda esta valorización en el plano del discurso oficial, la situación real de la biblioteca escolar en Brasil continua siendo precaria". Cita el Censo Escolar de 2005 realizado por el Instituto de Pesquisas Econômicas Aplicadas (IPEA) en el que se muestra que el 63.72% de las escuelas tenía una biblioteca o sala de lectura. Por el contrario, 108.500 escuelas con 13.165.377 alumnos de enseñanza fundamental carecían de biblioteca escolar. Este mismo estudio reveló que en la mayoría de las escuelas la biblioteca era tan solo una sala de lectura sin planificación arquitectónica como espacio bibliotecario y sin ningún tipo de organización del fondo bibliográfico para poder ser utilizado. En contrapartida, el Censo Escolar de 2004 realizado por INEP (CONSELHO FEDERAL DE BIBLIOTECONOMIA, 2008) apuntó que el 51% de los alumnos de enseñanza básica y profesional disponía de bibliotecas escolares. 10 años más tarde, en 2013, el Censo Escolar (INEP, 2014) revela que 24.694.440 alumnos (75%) de enseñanza básica de la red pública están escolarizados en centros que disponen de biblioteca o sala de lectura. El crecimiento de las bibliotecas escolares ha sido muy importante.

Sin embargo, al carecer la mayoría de las bibliotecas de profesionales cualificados, los libros enviados por los programas gubernamentales no siempre son puestos a disposición de los lectores, o los espacios destinados a la lectura son pequeños, inadecuados, o son reaprovechados para otros fines distintos de los bibliotecarios. Dos importantes acciones de los bibliotecarios brasileños han sido decisivas para la movilización política a favor de la biblioteca en la escuela: la primera fue el desarrollo y la publicación en 2008 de "*Projeto Mobilizador: biblioteca escolar construção de uma rede de informação para o ensino público*" por el Consejo Federal de Biblioteconomía junto con los Conse-

jos Regionales de Biblioteconomía (CFB/CRB). La segunda fue la búsqueda de alianzas para lograr los objetivos propuestos en el proyecto con el objetivo final de caracterizar y definir lo que es una biblioteca escolar mediante el estudio de modelos.

Este estudio se realizó en colaboración con el Grupo de Estudios sobre la biblioteca escolar (GEBE) de la Escola de Ciência da Informação da Universidade Federal de Minas Gerais (UFMG) y dio lugar a la publicación *Biblioteca escolar como espaço de produção do conhecimento: parâmetros para bibliotecas escolares brasileiras* (CAMPELLO, 2010). En ella se indican los siguientes parámetros para la organización del fondo: en el *nivel básico*, el catálogo de la biblioteca incluye al menos los libros de la colección, lo que permite la recuperación por autor, título y tema; en el *nivel ejemplar*, el catálogo de la biblioteca está informatizado y permite el acceso remoto a todos los elementos de la colección; permite –además de por autor, título y tema– la recuperación por otros puntos de acceso.

El *Projeto Mobilizador* es un proyecto llamado así por su intención de mover políticamente a los órganos profesionales de la Biblioteconomía brasileña, y también es un plan estratégico con objetivos, metas y responsabilidades claras, cuyo propósito es conseguir "la universalización de la biblioteca en la escuela pública situada en un sistema de información que atienda de forma consistente y coherente, no sólo al alumnado, sino también a todos los actores involucrados en la educación" (CONSELHO FEDERAL DE BIBLIOTECONOMIA, 2008, p. 24). Para alcanzar este propósito, el proyecto propone "[...] la adopción e integración dentro del sistema propuesto, que requerirá la existencia previa o la implantación de bibliotecas en las escuelas públicas, dentro de cada dirección de enseñanza, de cada municipio o región [...]" (CONSEJO FEDERAL DE BIBLIOTECA, 2008, p. 25).

Este concepto de universalización de la biblioteca escolar propone la creación de una red en la enseñanza pública que incluya a todos los agentes involucrados, especialmente a los gestores de las políticas educativas públicas, pues se trata de que la escuela se apropie de los beneficios proporcionados por la universalización de la biblioteca, dentro de los principios que rigen la educación.

En esta misma línea y fruto de la movilización, se promulgó por la Presidencia de la República de Brasil la Ley nº 12.224 de 24 de mayo de 2010 que instituye la universalización de las bibliotecas en las instituciones de enseñanza

brasileñas. En su Artículo 3 se establece que “los sistemas de enseñanza del país deberán desarrollar esfuerzos progresivos para conseguir que la universalización de las bibliotecas escolares, en los términos previstos en esta, sea efectiva en un plazo máximo de diez años, respetada la profesión de bibliotecario”.

Debido a esta ley, se ha desplegado mucho trabajo y esfuerzo para aumentar la concienciación de los órganos y los gestores del sistema educativo municipal y estatal. Se observa en la actualidad la existencia de redes de bibliotecas escolares subordinadas a las Secretarías Municipales y Estatales de Educación que han conseguido implantar y gestionar bibliotecas escolares, a partir de equipos de bibliotecarios profesionales que actúan de forma centralizada desde estas Secretarías.

También ha sido muy importante la movilización conseguida, antes de la promulgación de la Ley, mediante la celebración de *Forums* de discusión sobre bibliotecas escolares. Un caso ejemplar es el de la Red de bibliotecas escolares del municipio de Vitória, capital del Estado de Espírito Santo, impulsada por la Secretaría Municipal de Educação de Vitória. Según el relato de Silva (2012), dicha red existía desde 1999 y, anticipándose a la Ley, se contrató a 47 bibliotecarios en 2005, alcanzado el número de 54 en la Secretaría Municipal de Educación.

### 3.2. Bibliotecas escolares en Portugal

En Portugal, las bibliotecas escolares comenzaron a desarrollarse de forma notable a partir de la última década del siglo XX, con la creación de la *Rede de Bibliotecas Escolares* (RBE). La RBE es un proyecto del Ministerio de Educación portugués iniciado en 1996, que cuenta con la participación de las Câmaras Municipais, de las bibliotecas públicas municipales y de las direcciones regionales de educación, bajo la coordinación del *Gabinete da Rede Bibliotecas Escolares*. Con esta iniciativa, el gobierno pretendía dotar a todas las escuelas públicas, independientemente de su nivel de enseñanza, de una biblioteca escolar, creando en los alumnos el hábito de la lectura y del acceso a la información, a partir del 1º ciclo de escolaridad. Se trata de un proyecto concebido en fases sucesivas, con el propósito final de llegar a todas las escuelas de todos los ayuntamientos del país.

Actualmente, la RBE cubre cerca de 2500 bibliotecas escolares (PORTUGAL, 2014), ya que desde finales de 2008 todas las escuelas de 2º y 3º ciclo de educación básica, escuelas básicas integradas y agrupamientos escolares disponen ya de biblioteca integrada.

En julio de 2009, fue creada legalmente la función de profesor bibliotecario, a través de la Portaria 756/2009, sucesivamente actualizada en 2010 e 2011. Su objetivo fue dotar a las bibliotecas escolares de los recursos humanos necesarios para garantizar su papel innovador en el modelo organizacional de las escuelas, promoviendo en el conjunto de la comunidad educativa, competencias de acceso y uso de la información.

Generalmente, en los protocolos de cooperación establecidos entre las Câmaras Municipais y la RBE, queda establecido que las primeras prestan apoyo técnico a las escuelas secundarias y a las escuelas básicas, poniendo a su disposición los programas de software y los registros bibliográficos.

En el nivel de 1º Ciclo, la Câmara Municipal asume la construcción o adaptación del espacio que acoge a la biblioteca, bien mediante la adquisición de mobiliario específico, equipamientos y fondos bibliográficos y multimedia, con la colaboración de las Direcciones regionales de Educación.

## 4. Políticas de indización en las bibliotecas escolares brasileñas y portuguesas: estudio exploratorio

Partiendo de la revisión de la literatura sobre indización y sobre políticas de indización de las bibliotecas escolares brasileñas y portuguesas, en el último trimestre de 2012, se diseñó una investigación para recoger datos sobre las prácticas de indización de las bibliotecas escolares en el contexto lusófono.

### 4.1. Material y método de la investigación

Para recoger los datos del estudio se diseñó un cuestionario con 35 preguntas, que fueron adaptadas para adecuarse a las diferencias entre el contexto bibliotecario brasileño y el portugués. La adaptación implicó de facto la creación de dos cuestionarios: uno dirigido al contexto brasileño (QB) y el otro para el contexto portugués (QP). La encuesta se distribuyó en línea a través de *SurveyMonkey* ([www.surveymonkey.com](http://www.surveymonkey.com)) herramienta especializada en la gestión de cuestionarios Web.

Las preguntas están organizadas en diez grupos. El primero está dirigido a la identificación general de las bibliotecas escolares, teniendo en cuenta su adscripción institucional (QB1) o geográfica (QP1), los niveles de educación que cubre (QB2, QP2), el tamaño de la colección (QB3, QP3), el número de profesionales que intervienen en la biblioteca (QB4, QP4) y la in-

tegración en una red (QB5, QP5). En el segundo grupo de preguntas, se pregunta por la existencia del catálogo en la biblioteca escolar (QB6, QP6), por el tipo de catálogo (manual o electrónico) (QB7), la existencia de catálogo individual o colectivo (QB8, QP7), y en este último caso, por las entidades cooperantes (QB9, QP8).

Estas preguntas iniciales, aunque no son específicas sobre la indexación, se refieren a aspectos que pueden influir en ella y, por ello, se consideraron pertinentes. La identificación de los lenguajes de indización utilizados en la biblioteca escolar (QB10, QP9) es la primera pregunta directamente relacionada con el tema central de la investigación. Se identifica a los profesionales involucrados en la indización (QB11, QP10, QP11). La formación de los indizadores se formula en las siguientes cuatro preguntas (QB12, QB13, QB14, QB15, QP12, QP13, QP14, QP15).

También se interroga por la importación de registros (QB16, QB17, QP16, QP17). Las dos preguntas siguientes tratan de determinar la existencia de un manual de procedimientos para la indexación (QB18, QB19, QP18, QP19).

El bloque siguiente se centra en aspectos concretos del proceso de indización, incluyendo el uso de herramientas automáticas o semiautomática (QB20, QB21, QP20, QP21), la determinación de la especificidad (QB22, QB23, QP22, QP23), el existencia de directrices sobre el número de términos seleccionados para cada documento (QB24, QB25, QP24, QP25), la determinación del tiempo medio dedicado a la indización de cada documento (QB26, QB27, QP26, QP27). En el noveno grupo de cuestiones se pregunta por la normalización, identificando el uso de normas nacionales e internacionales (QB28, QB29, QP28, QP29), el uso del sistema de validación/corrección automática de términos/temas (QB30, QB31, QP30, QP31) o el uso de términos libres sin control del vocabulario (QB32, QP32). Por último, se abordan aspectos relacionados con la evaluación de la indización (QB33, QB34, QP33, QP34). La última pregunta permitía a los encuestados dejar sus comentarios libremente.

La recolección de respuestas se desarrolló desde noviembre de 2012 hasta mayo de 2013. La respuesta más antigua recogida en Brasil es de 20 de noviembre de 2012 y la más reciente de 30 de marzo de 2013. En Portugal, la más antigua es de 28 de noviembre de 2012 y la más reciente de 23 de mayo de 2013.

La estrategia de difusión de la encuesta fue distinta en cada uno de los dos países. En Por-

tugal se recogieron los datos de contactos de las bibliotecas escolares a partir de la información disponible en el sitio web de la Red de Bibliotecas Escolares en el último trimestre de 2012. Se obtuvieron unas 2.000 direcciones de correo electrónico a la que se envió un mensaje con el enlace a la encuesta y un breve texto explicativo sobre la investigación.

En Brasil, no existe ningún directorio de bibliotecas escolares. El Censo escolar realizado por el INEP con datos cuantitativos, incluye la biblioteca escolar o sala de lectura. De acuerdo con el censo escolar de 2013, publicado en 2014, hay alrededor de 120.500 escuelas públicas de educación básica y el 47% de ellas tienen una biblioteca o sala de lectura. Por ello, la encuesta se envió a las bibliotecas de las escuelas públicas dependientes de las Secretarías Municipales y Estatales seleccionadas a partir del registro de participantes en el *1º Forum Brasileiro de Biblioteconomia Escolar* promovido por el Grupo de Estudios de Bibliotecas Escolares de la UFMG y especialmente gracias a la colaboración del coordinador de la red de bibliotecas de la Secretaría Municipal de Espírito Santo. También se adhirieron las redes municipales de Victoria, Curitiba, Belo Horizonte, Porto Alegre y Sao Paulo. En Portugal, 406 personas comenzaron el cuestionario, pero como no todas las preguntas eran obligatorias, algunas quedaron sin respuesta. En Brasil, el número de personas que iniciaron la encuesta fueron 150. En el análisis de los datos, el cálculo de los porcentajes se basará en el número total de respuestas obtenidas por cada pregunta.

#### 4.2. Identificación genérica de las bibliotecas escolares

En Brasil, respecto a la dependencia institucional, las bibliotecas escolares de la muestra son un 54,7% (n = 82) municipales, un 7,3% (n = 11) estatales y un 38% (n = 57) federales. En Portugal, la encuesta recoge datos de 162 municipios, en los que hay 398 escuelas. Hay dos respuestas no válidas y seis encuestados que no respondieron.

Uno de los elementos caracterizadores de las bibliotecas escolares se refiere a la identificación de los niveles educativos cubiertos, lo que determina las características de su público, sus necesidades de información y, por tanto, la formación y organización de la colección, incluyendo aquí todos los aspectos relacionados con la indexación.

En Portugal, había 394 bibliotecas para identificar su ciclo educativo. En este conjunto, el 44,41% (n = 175) se encuadran en un solo ciclo

de los estudios: 13,96% (n = 55) en el 1º ciclo, 19,80% (n = 78) en el 2º ciclo y 10,66% (n = 42) en secundaria. Sin embargo, el grupo más frecuente (55,83%) está constituido por bibliotecas que pertenecen a centros con más de un ciclo de estudios. De hecho, el 10,41% (n = 41) de bibliotecas escolares cubren los cuatro ciclos de estudio, el 22,59% (n = 89) cubren el 1º y 2º/ 3º ciclo, el 22,59% (n = 89) se incluyen en el 2º/3º ciclo y en secundaria. En esta pregunta hubo 12 encuestados que no proporcionaron ninguna respuesta.

En Brasil, de las 150 bibliotecas cuyo nivel de estudios había que identificar, se encontró que el 0,67% (n = 1) son de nivel infantil, el 21,33% (n = 32) son de nivel fundamental o básico y 54% son de enseñanza media. El 58% (n = 87) es de un único ciclo, el 42% restante (n = 63) cubren diversos ciclos de estudio: el 18% (n = 27) se suman infantil, primaria y secundaria, el 14,67% (n = 22) incluye infantil y básica y el 9,33% (n = 14) básica y enseñanza media.

#### 4.3. Lenguajes de indización

Analizando los datos relativos a los lenguajes de indización utilizados en las bibliotecas públicas de Brasil y Portugal [Anexo 1, Anexo 2] se constata una diferencia significativa entre los dos países, lo que refleja sus respectivas áreas de influencia y las opciones reflejadas también en otros tipos de unidades de información. De hecho, en Brasil hay un uso predominante de la Clasificación Decimal Dewey (DDC) con el 72,4%, mientras que en Portugal la práctica totalidad de la muestra (97,4%) declaró usar la Clasificación Decimal Universal (CDU). En Portugal, predomina de forma abrumadora el uso de la CDU, un lenguaje documental de estructura jerárquica, que es complementado por un 21,3% de las bibliotecas que utilizan también un lenguaje de tipo combinatorio como la *Lista de Cabeçalhos de Assunto para Bibliotecas*. Tan solo un 11,1% de las bibliotecas consultadas utiliza una lista controlada de términos elaborada por la propia biblioteca, lo que supone que en estos centros hay profesionales con capacidad para crear un lenguaje de indización.

Además de las opciones mencionadas, también se daba la posibilidad de identificar otros lenguajes de indización. Hubo seis encuestados que informaron de la utilización de la Tabla de Clasificación propuesta por la Federación Internacional de Archivos Filmicos (FIAF) que distingue documentos de ficción y documentos de no ficción, clasificándolo según géneros, temas y tópicos. También hubo encuestas en las que se identificó la "Tabla de París", posiblemente refiriéndose a la "Principios de Clasificación de los Documentos musicales aplicables a las colecciones de préstamo", desarrollado en los años ochenta por un conjunto de bibliotecas y fonotecas de París y adaptada en Portugal en los años noventa (Portilheiro, Rodrigues, 1994). Una escuela informó de que utilizaba el tesoro Artes, explicando que impartía educación artística. En otros casos, se indicó que se empleaban listas de términos controlados desarrolladas por la propia biblioteca escolar o por la biblioteca municipal, o bien que empleaban adaptaciones simplificadas de la CDU, desarrolladas internamente.

En Brasil, con un claro predominio de la DDC, no hay una uniformidad tan marcada como en el contexto portugués. La CDU es la segunda clasificación más usada, presente en el 25,4% de las bibliotecas escolares. Por lo tanto, existe una preferencia clara por los lenguajes documentales de estructura jerárquica para la representación del conocimiento. Los lenguajes combinatorios son adoptados por un número insignificante de bibliotecas, un 17,2% de centros usan listas propias de términos controlados y un 15,7% usan la *Lista de Cabeçalhos de Assuntos para Bibliotecas*. En Brasil, se observa una inversión más significativa en la construcción de lenguajes de indización que en Portugal.

En Brasil, con un claro predominio de la DDC, no hay una uniformidad tan marcada como en el contexto portugués. La CDU es la segunda clasificación más usada, presente en el 25,4% de las bibliotecas escolares. Por lo tanto, existe una preferencia clara por los lenguajes documentales de estructura jerárquica para la representación del conocimiento. Los lenguajes combinatorios son adoptados por un número insignificante de bibliotecas, un 17,2% de centros usan listas propias de términos controlados y un 15,7% usan la *Lista de Cabeçalhos de Assuntos para Bibliotecas*. En Brasil, se observa una inversión más significativa en la construcción de lenguajes de indización que en Portugal.

#### 4.4. Manual de procedimientos para la indización

De acuerdo con los datos recogidos en el Anexo 3, se constata que en más de la mitad (64,9%) de las bibliotecas escolares brasileñas no existe manual de procedimientos para la indización. Por el contrario, en Portugal un 41,5% afirma disponer de un manual de indización. Un 30,1%, reconoce que no existe ningún documento específico de orientación de la indización, aunque los procedimientos de esta tarea están incluidos en un manual general aplicable a toda la biblioteca escolar. En este sentido, es necesario aclarar que un número elevado de bibliotecas escolares portuguesas disponen de un manual de procedimientos que recoge todas las fases del tratamiento de la información, en el que se recogen aspectos que guían la indización.

Los datos señalan que en Portugal existe un mayor grado de formalización de las prácticas de indización. Este aspecto es relevante para la consistencia de la indización, sobre todo en relación con la movilidad de los profesores-bibliotecarios y otros funcionarios. Por otro lado, refleja también que los servicios de información están ya bien implementados y cuentan con tiempo de existencia.

Dado que algunas de las bibliotecas escolares estaban integradas en redes y otras participaban en catálogos colectivos, resultaba interesante descubrir si las orientaciones en materia de indización se aplicaban a todas las bibliotecas de la red, o bien estaban circunscritas al ámbito individual de cada biblioteca escolar. Así, la pregunta 19 de los cuestionarios aplicados en Brasil y Portugal se planteó en estos términos "Si la biblioteca escolar usa un manual de procedimientos para la indización, refiera si este documento está disponible para todas las entidades que integran el catálogo colectivo". De acuerdo con el Anexo 4, se verifica que esta pregunta fue respondida por un número bastante bajo de centros (44), equivalente a menos de un tercio de la muestra brasileña. En Portugal, se obtuvieron 233 respuestas, lo que corresponde a un poco más de la mitad de la muestra. En el conjunto de los que respondieron a esta pregunta, compartir un manual de indexación o un manual de procedimientos generales no es una práctica dominante, verificándose en el 39,5% de las bibliotecas portuguesas y en el 29,5% de las brasileñas.

#### 4.5. Catálogo

Se comprueba la existencia de catálogos en las bibliotecas escolares [Anexo 5] y que este instrumento de información y acceso a la información existe en el 71,3% de las bibliotecas escolares brasileñas y en el 86,3% de las portuguesas.

En Brasil, se preguntó si, comprobada la existencia del catálogo, este era manual o electrónico. El 80,3% (n=86) de los casos es electrónico, y tan solo en el 11,2% (n=12) de los casos es un catálogo manual, coexistiendo ambas modalidades en un 8,4% (n=9) de las bibliotecas escolares. 43 encuestados no respondieron a esta pregunta.

Atendiendo a la importancia del trabajo cooperativo y a la integración de las bibliotecas escolares en redes de colaboración, es relevante conocer si el catálogo es una herramienta de esta cooperación. Sobre este aspecto indaga la pregunta 8 del cuestionario brasileño y la pregunta 7 del cuestionario portugués. En Brasil, predomina el catálogo independiente, en un 67% (n=67), mientras que en el 33% (n=33) de las bibliotecas escolares existe un catálogo colectivo. 50 encuestados no respondieron. Esta tendencia también se observa en Portugal, con un 58% (n=195) de bibliotecas que disponen de un catálogo individual frente a un 42% (n=141) que tienen un catálogo colectivo. 70 encuestados no responden.

En cuanto a la alimentación del catálogo, se consideró que esto se podría hacer bien incorporando registros nuevos o bien descargándolos de otro catálogo o base de datos. Como demuestran los anexos 6 y 7, el procedimiento más común consiste en copiar registros, así se confirma en el 54,3% de las bibliotecas escolares brasileñas y en el 68,6% de las portuguesas. En Brasil, esta tendencia no es tan fuerte como en Portugal pues casi la mitad (45,7%) de los que respondieron a la pregunta dijeron que no utilizan este procedimiento.

Recurrir a copiar registros de otros catálogos es una práctica común en un gran número de bibliotecas y puede contribuir a uniformizar las descripciones, aunque puede ser problemática cuando incide sobre la descripción de contenido, atendiendo a la necesidad de adaptación al perfil del usuario. Esta práctica refuerza la colaboración entre las bibliotecas escolares y su trabajo en red, mediante el aprovechamiento de tareas comunes y la racionalización de inversiones en personal y tiempo. Simultáneamente, para que se mantengan los criterios de calidad, la captura de registros debe basarse en principios y reglas conocidos y aceptados por todos, procurando asegurar la uniformidad y la coherencia de las prácticas de indización y sus productos.

En este sentido, las bibliotecas escolares fueron preguntadas acerca de la existencia de orientaciones relativas a la especificidad y al número de materias/términos a asignar y al uso de las normas.

#### 4.6. Orientaciones para la práctica de indización

Respecto a la especificidad de la indización, importaba saber si se aplicaban orientaciones para definir el grado de exactitud de los descriptores. Analizados los datos del Anexo 8, se constata que en el 45,9% de las bibliotecas escolares brasileñas que responden a esta pregunta afirman que sí y que las cifras que responden lo contrario son equiparables (44,3%). Un 9,8% afirma que no sabe responder a esa pregunta.

Los resultados obtenidos en Portugal son muy distintos, ya que la especificidad de la indización solo es reconocida por el 21,3% de las bibliotecas escolares que responden a esa pregunta. Y esa situación no se produce en la mayoría de las bibliotecas (52,8%). Una cuarta parte de los centros (25,9%) contesta no saber.

La pregunta 23 del cuestionario brasileño y portugués solicitaba que se especificase el grado de especificidad, en el caso de que estuviese

establecido. Se trataba de una pregunta abierta a la que respondieron 30 encuestados de cada uno de los dos países. Entre las respuestas destacan las siguientes: "Es específico en la medida en que se adapte a las necesidades de los niños", "si el documento trata de perros, no se indexará en animales o mascotas, sino en 'perros'", "Literatura infantil y juvenil – Cuentos; Literatura infantil y juvenil – Poesía" "Si el libro habla de árboles, el término de indexación tiene que ser árboles, o el nombre específico de cada tipo de árboles si el libro trata de árboles específicos" y "Flores Use nombres específicos. Ejemplo: Margaritas, violetas, rosas".

Respecto al número de términos seleccionados para representar el contenido de cada documento, en la mayoría de las bibliotecas encuestadas predomina la inexistencia de indicaciones en este sentido [Anexo 9]. En Brasil, esta situación se da en el 72,6% (n. 85) y en Portugal en 60,7% de los casos. En este último país, afirman desconocer si había indicaciones sobre esta cuestión el 15,6% de los centros.

El uso de normas en el proceso de indexación ofrece un marco metodológico y conceptual para las operaciones de análisis y síntesis del contenido de los *items* informativos. Las normas cumplen una función de guía y promueven la normalización de las prácticas, aspectos especialmente relevantes cuando las bibliotecas escolares trabajan en red BE y / o cuando se realiza la indexación por varias personas.

Al analizar los resultados obtenidos, se constata que en Brasil el uso de normas no está extendido, pues solo el 42,2% de los que responden a la pregunta, declaran realizarlo. Por el contrario, el 58,5% de las bibliotecas portuguesas que respondieron a esta pregunta declaran seguir normas nacionales o internacionales en el proceso de indexación

Es destacable que casi una cuarta parte del total de encuestados no respondieron (2,1%) y el 19% dice que lo desconoce.

Por lo tanto, parece que en las bibliotecas escolares es necesario mejorar la definición de las orientaciones para la indexación, tanto con respecto a la especificidad, como sobre el número de palabras o temas seleccionados e incluso sobre el uso de las normas nacionales o internacionales.

## 5. Conclusiones

Considerando que la indexación está condicionada, entre otros aspectos, por los procedimientos, las normas, las técnicas y los lenguajes documentales utilizados, este trabajo ha anali-

zando las políticas de indexación practicadas en las bibliotecas escolares brasileñas y portuguesas. La revisión bibliográfica realizada constató que en el análisis de la política de indexación de un centro o sistema, es necesario tener en cuenta aspectos como el uso de las normas, la existencia de manuales de indexación, las orientaciones sobre la exhaustividad y la especificidad, o el tipo de herramientas (listas de encabezamientos de materia, tesauros, clasificaciones, etc.). Desde estos supuestos teóricos, podemos hacer algunas observaciones sobre los resultados obtenidos en nuestra muestra.

Las bibliotecas escolares brasileñas y portuguesas se encuentran en estadios de desarrollo distintos, pero ambos países coinciden en enfatizar el valor que asignan a estas bibliotecas dentro del proceso de enseñanza-aprendizaje.

Los lenguajes de indexación más usados en ambos países son clasificaciones jerárquicas, que sirven tanto para representar y organizar los contenidos de la colección como para ordenar físicamente el fondo. Por el contrario, son menos usados los lenguajes asociativos o combinatorios y tampoco es frecuente en ninguno de los dos países las creaciones o adaptaciones de lenguajes existentes.

El hecho de que las bibliotecas escolares portuguesas comenzasen a finales del siglo pasado a articularse dentro de redes bibliotecarias, explica que sea más frecuente en ellas la existencia de manuales para orientar la indexación, pues a medida que se consolida la experiencia de la biblioteca escolar se incrementa la tendencia a formalizar los procedimientos. En Brasil, todavía son minoría las bibliotecas escolares que disponen de manuales de procedimientos.

A pesar de todo, se observa que en Portugal, solo una quinta parte de las bibliotecas estudiadas afirman disponer orientaciones relacionadas con la especificidad de la indexación. Y quizá podría analizarse en trabajos futuros la falta de correlación entre la constatación de esta falta de indicaciones, cuando existen tantos manuales de procedimientos. Por su parte, en Brasil, los manuales de procedimientos nos son tan frecuentes y casi la mitad de los encuestados afirman disponer de indicaciones en materia de especificidad en la indexación.

El conocimiento detallado de estas prácticas de tratamiento intelectual es muy relevante, ya que la tendencia a establecer redes colaborativas de trabajo es cada vez mayor. Además, esta comprensión va a posibilitar la creación de instrumentos de acceso a la información de calidad, especialmente a través de las búsquedas por materias, para apoyar los procesos de ense-

ñanza y aprendizaje, función primordial de todas las bibliotecas escolares.

## Referencias

- AGUSTÍN LACRUZ, M. C.; FUJITA, M.S.L y TERRA, A. L. *S. Linguagens documentais para as bibliotecas escolares: o caso da Espanha, Portugal e Brasil, Informação & Sociedade: Estudos*, João Pessoa, v. 24, n. 3, p. 83-97, 2014.
- ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. *NBR 12676: Métodos para análise de documentos - determinação de seus assuntos e seleção de termos de indexação*. Rio de Janeiro, 1992.
- BRASIL. Presidência da República. *Lei nº 12.244 de 24 de maio de 2010* – Dispõe sobre a universalização das bibliotecas nas instituições de ensino do país. Brasília, DOU de 25 de maio de 2010.
- CAMPELLO, B. (Org.) *Biblioteca escolar como espaço de produção do conhecimento: parâmetros para bibliotecas escolares*. Belo Horizonte: UFMG, Sistema CFB, 2010. 36p.
- CAMPELLO, B. et al. Parâmetros para bibliotecas escolares brasileiras. *Informação & Sociedade: Estudos*, João Pessoa, v.21, n.2, p.105-120, maio/ago. 2011.
- CONSELHO FEDERAL DE BIBLIOTECONOMIA. *Projeto mobilizador: biblioteca escolar construção de uma rede de informação para o ensino público*. Brasília: Sistema CFB/CRB, 2008.
- CHAUMIER, J. *Travail et methodes du/de la documentaliste: connaissance du problème*. Paris : ESF/Libraries Techniques. Exposé 3, Chap.3: L'indexation, p. 42-7, 1980.
- DAHLBERG, Ingetraut. Knowledge organization: a new science? *Knowledge Organization*, Frankfurt, v.33, n.1, p. 11-19, 2006.
- DAHLBERG, Ingetraut. Knowledge organization: its scope and possibilities. In: *Knowledge Organization*, Frankfurt, v.20, n.4, p.211-222, 1993.
- FUJITA, M. S. L. A política de indexação para representação e recuperação da informação. In: GIL LEIVA, I., FUJITA, M. S. L. (Orgs.) *Política de indexação*. São Paulo: Cultura Acadêmica; Marília: Oficina Universitária, 2012. p. 17-28
- FUJITA, M. S. L. Organização e representação do conhecimento no Brasil: análise de aspectos conceituais e da produção científica do ENANCIB no período de 2005 a 2007. *Tendências da Pesquisa Brasileira em Ciência da Informação*, v.1, p.1 - 32, 2008.
- GIL LEIVA, I. *Manual de indización*. Teoría y práctica. Gijón: Trea, 2008.
- GUIMARÃES, J. A. C. A dimensão teórica do tratamento temático da informação e suas interlocuções com o universo científico da International Society for Knowledge Organization (ISKO). *Revista Ciência da Informação*, Brasília, v. 1, p. 77-99, jan./abr. 2008.
- GUINCHAT, C.; MENO, M. *Introdução geral às ciências e técnicas da informação e documentação*. 2.ed. rev. aum. Brasília: IBICT, 1994.
- HJORLAND, Birger. What is knowledge organization? *Knowledge Organization*, v.35, n.2/3, p.86-101, 2008.
- IASL IASL Policy Statement on School Libraries. Recuperado 26.06.2015, de [http://www.iasl-online.org/about/organization/sl\\_policy.html](http://www.iasl-online.org/about/organization/sl_policy.html)
- INEP. *Censo escolar da educação básica 2013: resumo técnico*. Brasília: O Instituto, 2014.
- INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. ISO 5693:1985: documentation: methods for examining documents, determining their subjects, and selecting indexing terms. Geneve, 1985.
- KOBASHI, N. Y. A elaboração de informações documentárias: em busca de uma metodologia. Tese (Doutorado em Ciências) – Escola de Comunicação e Arte, Universidade de São Paulo, São Paulo, 1994.
- LANCASTER, F. W. *Indexação e resumos: teoria e prática*. Tradução de A. A. Briquet de Lemos. Brasília: Briquet de Lemos/Livros, 1993.
- PIERUCCINI, Ivete. *A ordem informacional dialógica: estudo sobre a busca de informação em Educação*. Tese, Doutorado em Ciência da Informação e Documentação, Universidade de São Paulo, Escola de Comunicação e Artes, 2004.
- PORTILHEIRO, Joaquim; RODRIGUES, Júlio Vaz – “Classificação e Cotação de Documentos Audiovisuais em Bibliotecas de Leitura Pública”. In Congresso Nacional de Bibliotecários Arquivistas e Documentalistas, 5, Lisboa 1994 – Multiculturalismo: actas. Lisboa: BAD, 1994, vol. I, p. 221-259.
- PORTUGAL. Ministério da Educação e Ciência. Gabinete da Rede Bibliotecas Escolares. Portal RBE: Lista de escolas integradas na RBE por direção de serviços regional [Em linha]. Lisboa: RBE, atual. 17-11-2014. [Consult. 26-06-2015] Disponível em WWW: <URL: <http://www.rbe.mec.pt/np4/578.html>>
- SILVA, E. V. da Relato de experiência da implantação e fortalecimento da rede de bibliotecas escolares do município de Vitória-ES. In: FORUM DE PESQUISA EM BIBLIOTECA ESCOLAR, 1, 24 e 25 de maio de 2012. *Anais...* Belo Horizonte: UFMG/GEPE, 2012.

**Anexo 1**

## Lenguajes de indización utilizados en la biblioteca escolar | Brasil (Q10)

|  | %     | n.  |
|--|-------|-----|
| Clasificación Decimal Universal (CDU)                                  | 25,4% | 34  |
| Clasificación Decimal Dewey (CCD)                                      | 72,4% | 97  |
| Lista de Encabezamientos de Materia para Bibliotecas                   | 15,7% | 21  |
| Lista Controlada de Términos (elaborada por las bibliotecas escolares) | 17,2% | 23  |
| Tesoro Brasileño de Educación  | 1,5%  | 2   |
| Tesoro Europeo da Educación  | 0,0%  | 0   |
| Tesoro da Educación de la Unesco                                       | 0,7%  | 1   |
| No sabe  | 1,5%  | 2   |
| Otras(s)   | 14,2% | 19  |
| Pregunta respondida  |       | 134 |
| No responden   |       | 16  |

**Anexo 2**

## Lenguajes de indización utilizados en la biblioteca escolar | Portugal (Q9)

|  | %     | n.  |
|--|-------|-----|
| Clasificación Decimal Universal (CDU)                                  | 97,4% | 370 |
| Clasificación Decimal Dewey (CCD)                                      | 0,8%  | 3   |
| Lista de Encabezamientos de Materia para Bibliotecas                   | 21,3% | 81  |
| Lista Controlada de Términos (elaborada por las bibliotecas escolares) | 11,1% | 42  |
| Tesoro Brasileño de Educación  | 0,0%  | 0   |
| Tesoro Europeo da Educación  | 3,2%  | 12  |
| Tesoro da Educación de la Unesco                                       | 3,4%  | 13  |
| No sabe  | 1,8%  | 7   |
| Otras(s)   | 4,7%  | 18  |
| Pregunta respondida  |       | 380 |
| No responden   |       | 26  |

**Anexo 3**

## Existencia de manual de procedimientos para la indización | Brasil (Q18) Portugal (Q18)

|   | Brasil |     | Portugal |     |
|---|--------|-----|----------|-----|
|   | %      | n.  | %        | n.  |
| Si  | 35,1%  | 46  | 41,5%    | 145 |
| No  | 64,9%  | 85  | 28,4%    | 99  |
| Los procedimientos para la indización están incluidos en un manual general de los procedimientos de la biblioteca escolar | --     | --  | 30,1%    | 105 |
| Pregunta respondida   |        | 131 |          | 349 |
| No responden  |        | 19  |          | 57  |

**Anexo 4**Disponibilidad de manual de procedimientos para las entidades participantes en el catálogo colectivo |  
Brasil (Q19) Portugal (Q19)

|  | Brasil |     | Portugal |     |
|--|--------|-----|----------|-----|
|  | %      | n.  | %        | n.  |
| Si   | 29,5%  | 13  | 39,5%    | 92  |
| No   | 15,9%  | 7   | 20,6%    | 48  |
| No se aplica (biblioteca no integradas en ningún catálogo colectivo) | 54,5%  | 24  | 39,9%    | 93  |
| Pregunta respondida  |        | 44  |          | 233 |
| No responden   |        | 106 |          | 173 |

**Anexo 5**

Existencia de catálogo en la biblioteca escolar | Brasil (Q6) Portugal (Q6)

|                     | Brasil |     | Portugal |     |
|---------------------|--------|-----|----------|-----|
|                     | %      | n.  | %        | n.  |
| Si                  | 71,3%  | 107 | 86,3%    | 340 |
| No                  | 28,7%  | 43  | 13,7%    | 54  |
| Pregunta respondida |        | 150 |          | 394 |
| No responden        |        | 0   |          | 12  |

**Anexo 6**

Indización de documentos realizada a partir de registros copiados de algún catálogo o base de datos |

Brasil (Q16)

|                     | %     | n.  |
|---------------------|-------|-----|
| Si                  | 54,3% | 70  |
| No                  | 45,7% | 59  |
| Pregunta respondida |       | 129 |
| No responden        |       | 21  |

**Anexo 7**

Indización de documentos realizada a partir de registros copiados de algún catálogo o base de datos |

Portugal (Q16)

|                     | %     | n.  |
|---------------------|-------|-----|
| Si (todos)          | 8,7%  | 29  |
| Si (algunos)        | 59,9% | 199 |
| No                  | 31,3% | 104 |
| Pregunta respondida |       | 332 |
| No responden        |       | 74  |

**Anexo 8**

Establecimiento del grado de especificidad de la indización | Brasil (Q22) Portugal (Q22)

|                     | Brasil |     | Portugal |     |
|---------------------|--------|-----|----------|-----|
|                     | %      | n.  | %        | n.  |
| Si                  | 45,9%  | 56  | 21,3%    | 69  |
| No                  | 44,3%  | 54  | 52,8%    | 171 |
| No sabe             | 9,8%   | 12  | 25,9%    | 84  |
| Pregunta respondida |        | 122 |          | 324 |
| No responden        |        | 28  |          | 82  |

**Anexo 9**

Indicaciones sobre el número de términos/materias a seleccionar por documento | Brasil (Q24) Portugal

(Q24)

|                     | Brasil |     | Portugal |     |
|---------------------|--------|-----|----------|-----|
|                     | %      | n.  | %        | n.  |
| Si                  | 22,2%  | 26  | 23,7%    | 76  |
| No                  | 72,6%  | 85  | 60,7%    | 195 |
| No sabe             | 5,1%   | 6   | 15,6%    | 50  |
| Pregunta respondida |        | 117 |          | 321 |
| No responden        |        | 33  |          | 85  |

## Anexo 10

Uso de normas nacionales o internacionales para la indización

|                     | Brasil |     | Portugal |     |
|---------------------|--------|-----|----------|-----|
|                     | %      | n.  | %        | n.  |
| Si                  | 42,2%  | 49  | 58,5%    | 185 |
| No                  | 49,1%  | 57  | 22,5%    | 71  |
| No sabe             | 8,6%   | 10  | 19,0%    | 60  |
| Pregunta respondida |        | 116 |          | 316 |
| No responden        |        | 34  |          | 90  |

---

# Categorización de revistas científicas españolas en las bases de datos ISOC

*Categorization of spanish academic journals in ISOC Databases*

---

José Ignacio Vidal Liy (1), Luis Rodríguez Yunta (1)

(1) Centro de Ciencias Humanas y Sociales, C./ Albasanz 26-28, Madrid, nacho.vidal@cchs.csic.es, lryunta@cchs.csic.es

## Resumen

Esta comunicación presenta el nuevo sistema de categorización de revistas científicas españolas aplicado en las bases de datos ISOC a partir de 2014. Es un sistema basado en la explotación de los datos contenidos en la propia base. Hasta la fecha, los criterios que se venían aplicando eran los utilizados por la red Latindex centrados en las características que de cada revista declaraban sus respectivos editores, las citas recibidas, encuestas respondidas por la comunidad científica o presencia en bases de datos internacionales. Consideramos que la evaluación, caracterización y categorización de las revistas científicas españolas cubre una necesidad tanto para la gestión interna de la base de datos como para la comunidad de investigadores nacionales, al disponer de un núcleo de publicaciones consolidadas, abiertas y que puedan competir en el terreno internacional. Con esta nueva propuesta estas bases de datos pasan pues a ser fuente para la caracterización y categorización de las revistas científicas españolas, aprovechando la ingente cantidad de datos que contiene.

**Palabras clave:** Sistemas de Información, Evaluación de revistas, Indicadores bibliométricos, Bases de datos bibliográficas, Criterios de selección, Categorización.

## 1. La necesidad de un sistema de categorización de revistas en la base de datos ISOC.

De sobra es conocida en el ámbito académico nacional la relevancia de las bases de datos ISOC (CCHS-CSIC). Sus cuarenta años de vida hacen de ellas uno de los mejores sistemas de información bibliográfica sobre documentación científica española, pues analizan toda la producción científica española en Humanidades y Ciencias Sociales desde 1975, indizando en la actualidad registros de más de 1.400 revistas vivas de Ciencias Sociales y Humanidades (Abejón (Abejón Peña, Maldonado Martínez ,

## Abstract

This paper presents the new system of categorization of Spanish scientific journals in ISOC data bases. It is based on the exploitation of the data contained in the database system itself. To date, the criteria had been applied were those used by the network Latindex focused on the characteristics that of each journal declared their publishers, citations, surveys answered by the scientific community or presence in international databases. We think that the testing, characterization and classification of Spanish scientific journals fills a need for both internal management of the database and the community of national researchers, to have a consolidated core, open publications that can compete in the international arena. With this new proposal these databases to be passed as source for characterization and categorization of the Spanish scientific journals, taking advantage of the huge amount of data it contains.

**Keywords:** Information systems, evaluation of journals, bibliometrics, bibliographic databases, selection's criteria, categorisation.

Rodríguez Yunta, & Rubio Liniers, 2009) (Rodríguez Yunta, 2015).

Las revistas científicas constituyen en la actualidad el medio idóneo en los que la comunidad científica presenta, visibiliza y difunde sus trabajos (Torres Salinas & Cabezas-Clavijo, 2013). La reciente eclosión de las Nuevas Tecnologías, la difusión de Internet, los weblogs, la multiplicación de repositorios y plataformas abiertas para el depósito de archivos ha disparado la expansión de las publicaciones a escala mundial, sin que ello esté respaldado por una calidad o un prestigio necesariamente.

Con el propósito de establecer filtros que beneficiaran y aseguraran una calidad mínima de los contenidos de las bases de datos ISOC, en el

año 2006 se estableció una política de selección de revistas que pivotaba en torno al cumplimiento de los criterios editoriales establecidos por el sistema Latindex. Éste incluye 3.449 registros de publicaciones españolas vivas (1), pero no es ni una categorización ni un ranking de revistas, sino una plataforma que establece unos criterios amplios para filtrar publicaciones que cumplen los parámetros de calidad mínima: Antigüedad mínima de un año, especificación de los miembros del consejo de redacción o comité editorial, nombre completo del director/a de la revista, identificación completa de los autores, afiliación institucional de los mismos (lugar de trabajo), mención de la entidad editora de la revista, dirección postal o electrónica de la Administración de la revista, lugar de edición, ISSN, sumario o tabla de contenidos, datos de identificación de la revista en las páginas de portada y cubierta, presentación de resumen de los artículos, al menos el 40% de los contenidos deben ser trabajos de interés científico.

La aplicación de este sistema supuso la jerarquización de las revistas analizadas en las bases ISOC en tres grupos –A, B y C– en función de su calidad editorial, siendo A el que mayor número de criterios cumplía y C integraba a aquellas publicaciones que, sin reunir los requisitos exigidos por los grupos A y B, cumplían el mínimo de ellos exigidos por las bases de datos ISOC. Así, para formar parte del grupo de revistas de mayor calidad (A), éstas debían contar con al menos 25 criterios Latindex, disponer de un sistema de revisores externos para la selección de artículos originales, estar presente en al menos una base de datos de prestigio internacional, cumplir con la periodicidad declarada y haber obtenido una valoración de al menos 70 puntos sobre 100 en la encuesta realizada a la comunidad científica. Las revistas B son aquellas que cumplen con 18 de los 33 criterios Latindex y al menos 40 puntos sobre 100 en la encuesta al profesorado universitario y personal científico. Finalmente, las revistas clasificadas en la categoría C, si bien preciso tenían menos de 18 criterios Latindex, si reunían al menos los requerimientos mínimos exigidos por las bases de datos ISOC.

La metodología aplicada en Latindex no es del todo suficiente en algunos aspectos de la medición, dada la poca exigencia de algunos de sus criterios. Así, para establecer la apertura de la autoría se fija un baremo por el que al menos el 50% de los trabajos publicados sean de autores externos a la entidad editora. Además es un criterio que beneficia a las entidades privadas o asociaciones, pues en estos casos sólo son autores pertenecientes a la entidad editora los

que forman parte del equipo de la revista. Otro caso similar es la apertura del consejo editorial de la misma. (Rodríguez Yunta, 2015)

Cualquier revista puede alcanzar con cierta facilidad una alta puntuación en el cumplimiento de los criterios Latindex. De hecho, en esta comunicación ejemplificaremos como en el todavía vigente sistema de evaluación, son numerosos los casos en los que revistas que no disponen de una categorización alta, cumplen sin embargo con la mayoría de los criterios editoriales exigidos por Latindex.

## 2. Una nueva propuesta para la categorización de revistas.

La insuficiencia del sistema Latindex para medir la calidad de las revistas ha supuesto la generación de otros sistemas que establecen categorizaciones más graduadas (DICE, RESH, INRECS, CIRC). Lamentablemente, éstos se han visto afectados por los recortes en la financiación de sus respectivos proyectos. Como consecuencia del XI Seminario Hispano-Mexicano de Investigación en Bibliotecología y Documentación, en las bases ISOC se han tomado medidas para establecer un nuevo sistema de categorización de revistas fundamentado en torno a la explotación de los datos bibliográficos contenidos en la propia base de datos, tratados éstos con criterios más precisos y graduales que superasen los hasta ahora meramente formales o editoriales. Con carácter previo, se han tenido en cuenta una serie de principios metodológicos en los que ha de inspirarse cualquier sistema de evaluación de revistas. Éstos han quedado recogidos en la “Metodología aplicada en la categorización de las revistas españolas en la base de datos ISOC” (Rodríguez Yunta, 2015)

- a) La evaluación tiene que ser equilibrada y multifacetada, esto es, basada en la combinación de diferentes indicadores. No debe depender de una única variable. (Alcaín Partearroyo, Román Román, & Giménez Toledo, 2008)
- b) La forma de medida debe aplicarse de forma idéntica a cualquier modelo de publicación, de tal manera que las condiciones sean independientes del tipo de entidad editora, sea ésta una universidad, un organismo público, una asociación cultural o una empresa privada.
- c) Aplicar el sistema a un período determinado. Se propone un período de análisis de 10 años como intervalo temporal representativo para medir la trayectoria de una revista y valorar su situación actual.

- d) Independencia de la evaluación de la clasificación temática de las revistas. Se debe de poder aplicar un sistema que valora la revista por sí misma.
- e) Es recomendable que el sistema de puntuación sea gradual y sencillo, esto es, con franjas para una puntuación graduada que evite marcos rígidos basados en respuestas sí/no o valoraciones 0/1.

Cada indicador debe poder puntuarse un sistema de baremación con franjas graduales. Se establecen filtros arbitrarios de los valores que pueden considerarse mínimos y de suficiente excelencia en cada uno de los indicadores.

De acuerdo a estos principios, el planteamiento del nuevo sistema de evaluación se basa en establecer 9 indicadores agrupados en tres bloques. Éstos son:

1. Trayectoria de la revista, que valora como factor de calidad de la misma el comportamiento temporal de la revista por un período de 10 años, su esfuerzo editorial y el cumplimiento de la regularidad en la edición.

Indicadores: Años cubiertos en el período, mediana artículos año, prácticas editoriales.

2. Apertura de la revista, que tiene en cuenta la diversidad de la autoría y de la variedad institucional. Se tiene en cuenta si existen colegios invisibles o si una única persona pública de forma reiterada.

Indicadores: Diversidad de la autoría, Índice de concentración de autoría principal, concentración institucional.

3. La internacionalidad de la revista, medida por el idioma, la procedencia institucional de los autores y por su presencia en bases de datos internacionales reconocidas.

Indicadores: Índice compensado de internacionalidad lingüística, índice compensado de internacionalidad en la autoría, difusión en BBDD internacionales.

En el nuevo sistema de categorización de revistas se valoran las prácticas editoriales en función de su antigüedad, del cumplimiento de su periodicidad y de su presencia tanto en el Sistema Latindex (considerándose, como veremos más adelante, en el baremo como valor óptimo cumplir sus 33 criterios) como en el listado de revistas con el sello de calidad FECYT. Sin embargo, el método que se propone en la base ISOC permite que cualquier tipo de revista alcance la categoría máxima aunque no tenga puntuación en alguno de los indicadores. Para ello se han desarrollado los indicadores referi-

dos a la apertura y a la internacionalidad de las revistas.

Se tiene en cuenta por tanto la posible existencia de 'colegios invisibles' en las revistas, la publicación reiterada de un mismo autor o el grado de endogamia institucional en las mismas. La internacionalidad también es un pilar importante de la evaluación. En cuanto a la difusión, la base de datos ISOC analiza y valora la difusión de las revistas en bases de datos internacionales, en especial en los índices de citas Web of Science y Scopus, si bien su puntuación no es un requisito de obligado cumplimiento. Hay que tener en cuenta, que los índices de citan miden la calidad a partir del impacto, lo cual condiciona sus limitaciones, sus insuficiencias: sus resultados dependen de la cobertura de las fuentes de datos, tanto para los títulos de revistas como por el período de años seleccionado. Además, las prácticas de citación son variables según las disciplinas, especialmente en las pertenecientes a las Ciencias Sociales y Humanidades, cuyas publicaciones tienen a dispersarse en diferentes lenguas y comunidades científicas.

Este nuevo sistema de categorización está todavía en "fase beta" y, por tanto, sujeto a revisiones futuras. Se ha aplicado a las siguientes áreas (entre paréntesis, el vínculo electrónico al informe):

- Historia (<http://hdl.handle.net/10261/103402>).
- Antropología (<http://hdl.handle.net/10261/104217>).
- Geografía (<http://hdl.handle.net/10261/108674>).
- Biblioteconomía y Documentación (<http://hdl.handle.net/10261/109523>).
- Arqueología y Prehistoria (<http://hdl.handle.net/10261/111062>).
- Lingüística (<http://hdl.handle.net/10261/113785>).
- Revistas españolas de ciencias humanas y sociales con sello de calidad FECYT (<http://hdl.handle.net/10261/116883>).
- Revistas españolas de ciencias humanas y sociales en Web of Science y/o Scopus (<http://hdl.handle.net/10261/116893>).

Los informes dedicados a disciplinas concretas se han completado con dos análisis dedicados a conjuntos multidisciplinares de revistas españolas de ciencias humanas y sociales: uno para los títulos que contaban con el Sello de Calidad

FECYT en 2014 y un segundo estudio para las publicaciones españolas de estas áreas que son utilizadas como fuente en los índices de citas Web of Science y Scopus. A partir de los resultados obtenidos en estos casos puede concluirse que estas características no constituyen por sí solas una garantía de calidad de las publicaciones. Si bien la mayoría de las revistas que cumplían estas condiciones sí fueron incluidas en las categorías A o B, también se localizaron casos de publicaciones que quedaron en C o incluso en la categoría D o que no pudieron analizarse por no cumplir los requisitos mínimos aplicados.

### 3. Criterios de selección y análisis de las revistas.

#### 3.1. Criterios de selección de revistas en la base de datos ISOC

- Estar presentes en el Directorio de Revistas de Ciencias Sociales y Humanidades de las bases de datos bibliográficas del CSIC, y por tanto, estar sujetas al Manual de normas para el análisis documental en la base de datos ISOC (Documento de trabajo ISOC 2015/07).
- Publicaciones impresas o electrónicas españolas de contenido científico. Pueden incluirse versiones españolas de revistas editadas por organismos internacionales.
- Criterios normalizados (superar cierto umbral de calidad editorial e interés científico-técnico).

#### 3.2. Criterios de análisis de contenido aplicados en la base de datos ISOC

- Artículos que aportan contenido científico (artículos originales, excluyendo editoriales, crónicas, necrológicas, artículos breves, etc).
- Se reflejan todos los autores de un documento y todas sus afiliaciones institucionales cuando hay variaciones en los datos. Cuando todos los autores se adscriben a una misma institución, ésta sólo se recoge una vez.
- De la misma manera, sólo se recoge una afiliación institucional en caso de que un autor tenga varias. En estos casos, sólo se recoge la primera de ellas.
- No se considera afiliación institucional cuando lo que se indica en los artículos fuente es el lugar de estudio o la universidad donde se obtuvo la titulación.
- Se refleja las versiones en diferentes idiomas que pueda tener un documento.

- Se considera versión en otro idioma el caso de los textos abreviados, con extensión superior a una página y que no son meros resúmenes.

#### 3.3. Criterios de selección de revistas para la categorización

1. Estar presentes en el Directorio de Revistas de Ciencias Sociales y Humanidades de las bases de datos bibliográficas del CSIC, con la consideración de estar viva o vigente.
2. Disponer de un vaciado completo de sus artículos científicos en la base ISOC. Se excluyen en este sentido, las revistas no calificadas en el sistema Latindex que han dejado de indizarse en la base de datos, así como algunas otras publicaciones no vaciadas, aquellas multidisciplinares que se analizan selectivamente o los casos en los que un título se ha incorporado solo recientemente sin que aún se hayan introducido aún todos los registros de años anteriores.
3. Tener en la base de datos ISOC registros de artículos de un mínimo de 3 años e incluyendo algún número en el periodo 2011-2013.
4. Tener completados en la base de datos ISOC información de la afiliación institucional de los autores, al menos en el 40% de los artículos del periodo 2004-2013.

#### 3.4. Categorización de las revistas

Se ha diseñado un sistema de categorización para la evaluación de las publicaciones aplicable a cualquier área del conocimiento, de acuerdo a la trayectoria de cada revista, su grado de apertura en la autoría y su internacionalidad.

El sistema de baremación se basa en valores arbitrarios preestablecidos, en vez de aplicar una medida relativa en función de la media alcanzada por el conjunto de publicaciones analizadas. Con ello se pretende que se pueda aplicar la evaluación tanto a áreas del conocimiento, como a estudios comparativos entre disciplinas, evaluación de revistas de forma individual o en conjuntos heterogéneos (por ejemplo las revistas de un mismo editor o las publicaciones seleccionadas por un producto determinado). Además, cuando se aplica sobre un área de conocimiento, la aplicación de valores arbitrarios preestablecidos permite que los resultados obtenidos por cada título sean independientes de los criterios utilizados para seleccionar las revistas objeto de análisis.

De acuerdo con estos parámetros, se ha establecido un sistema de 9 indicadores. Cada uno

de estos índices se cuantifica por tramos equivalentes por un sistema de puntuación de siete valores posibles entre 0 y 3. En cada caso se ha baremado como puntuación máxima (3) el rango que se ha juzgado a priori como indicador óptimo de suficiente cumplimiento y como valor mínimo (0) el rango considerado como claramente insuficiente, otorgando los valores intermedios 0,5, 1, 1,5, 2 y 2,5 a los rangos intermedios con una distribución equivalente o aproximada. Por tanto, cuando se valora un indicador con 3 puntos a partir de cierto valor, por ejemplo a partir de un 30% en la internacionalidad de los autores, significa que se considera este valor como indicador suficiente de excelencia, sin que se otorgue una mayor valoración a los casos que obtengan valores superiores (un caso de 31% puntúa igual que uno de 80 o 100%).

En la mayor parte de los indicadores los cálculos se realizan sobre el conjunto de artículos publicados en el periodo analizado. Se calculan porcentajes, sin hacer distinciones entre las publicaciones por su mayor número de artículos, excepto en el indicador específico que valora este aspecto.

A partir del baremo de cada indicador, se obtiene, para cada publicación, una puntuación general que puede oscilar entre un mínimo de 0 y un máximo de 27 puntos. Este resultado final muestra una valoración equilibrada sobre los diferentes factores evaluados, y permite establecer un ranking de las revistas del área.

Además se incluye en los informes la cuantificación del número de indicadores en los que cada publicación mejora la media del conjunto de su disciplina, con un valor general que oscila entre un mínimo de 0 y un máximo de 9. Este dato tiene un valor meramente informativo, pero no influye sobre la asignación de categoría a cada revista. En definitiva, la distribución de las revistas de una disciplina por categorías no se fija por cuartiles, los resultados de la evaluación pueden establecer varias

#### 4. Sistema de indicadores. Baremación.

Como se mencionó en el segundo apartado de esta comunicación, se computan 9 indicadores de calidad, tres sobre trayectoria de la revista (años cubiertos en el periodo, mediana artículos por año y prácticas editoriales), tres sobre apertura de la autoría (índice diversidad de la autoría, índice de concentración de la autoría e índice de concentración institucional) y otros tres últimos sobre internacionalidad (índice de internacionalidad lingüística, índice de internacionalidad en la autoría y difusión en bases de datos internacionales).

#### 4.1. Indicadores sobre trayectoria

4.1.1. Años cubiertos en el periodo. Número de años cumplimentados en el periodo estudiado, sobre 10 años.

Baremo:

| Rango de valores numéricos | Puntuación |
|----------------------------|------------|
| Menos de 5                 | 0          |
| 5 años                     | 0.5        |
| 6 años                     | 1          |
| 7 años                     | 1.5        |
| 8 años                     | 2          |
| 9 años                     | 2.5        |
| 10 años completados        | 3          |

Tabla I. Puntuación de las revistas según los años cubiertos

Se ha considerado como valor óptimo haber editado números en todos los años del periodo, y como valor insuficiente tener completados menos de la mitad. No se hace distinción a si se trata de revistas nacidas en los últimos años o si se debe a que se han producido lagunas o números dobles, pues se pretende puntuar positivamente a aquellas revistas que ofrecen una trayectoria que las convierte en publicaciones consolidadas y fiables. Se considera que una publicación mejora la media, cuando su valor numérico es superior a ésta.

#### 4.1.2. Mediana artículos/año

Mediana anual de los artículos publicados respecto a los años completados.

Baremo:

| Rango de valores numéricos | Puntuación |
|----------------------------|------------|
| Menos de 5                 | 0          |
| Entre 5 y 9.5              | 0.5        |
| Entre 10 y 14.5            | 1          |
| Entre 15 y 19.5            | 1.5        |
| Entre 20 y 24.5            | 2          |
| Entre 25 y 29.5            | 2.5        |
| 30 o más                   | 3          |

Tabla II. Puntuación de la mediana anual de los artículos publicados.

Se ha considerado como valor óptimo haber editado al menos 30 artículos anuales de forma regular, y como valor insuficiente menos de 5 contribuciones al año. Se aplica la mediana en lugar de la media, para reflejar la regularidad en la trayectoria de la publicación. Se considera que una publicación mejora la media, cuando su valor numérico es superior a ésta.

#### 4.1.3. Prácticas editoriales

Se valora este parámetro de acuerdo a las valoraciones recibidas por la publicación en el Sistema Latindex3 y a su presencia en la lista de revistas que han recibido el sello de calidad de la FECYT. En los casos en los que el Sistema Latindex aporta dos fichas para una publicación (versión impresa y versión electrónica) se tiene en cuenta la que sea más favorable.

| Rango de valores numéricos  | Puntuación |
|---|------------|
| Menos de 25 criterios Latindex  | 0          |
| Entre 25 y 27 para revistas impresas, 25-28 en electrónicas                             | 0.5        |
| Entre 28 y 30 para revistas impresas, 29-32 en electrónicas                             | 1          |
| Entre 31 y 32 para revistas impresas, 33-35 en electrónicas                             | 1.5        |
| 33 para revistas impresas, 36 en electrónicas   | 2          |
| Puntuación adicional para las revistas que han recibido el sello de calidad de la FECYT | +1         |

Tabla III. Puntuación de las prácticas editoriales de las revistas.

Se ha considerado como valor óptimo tener el sello de calidad FECYT y cumplir los 33 criterios Latindex para revistas impresas o los 36 para las electrónicas. Se considera que una publicación mejora la media, cuando su valor numérico es superior a ésta.

#### 4.2. Apertura de la autoría

##### 4.2.1. Diversidad de la autoría

Media entre número de autores diferentes y el número total de autorías o firmas presentes en los artículos durante el período analizado.

Baremo:

| Rango de valores numéricos | Puntuación |
|----------------------------|------------|
| Menos de 0.65              | 0          |
| Entre 0.65 y 0.69          | 0.5        |
| Entre 0.70 y 0.74          | 1          |
| Entre 0.75 y 0.79          | 1.5        |
| Entre 0.80 y 0.84          | 2          |
| Entre 0.85 y 0.89          | 2.5        |
| 0.90 o más                 | 3          |

Tabla IV. Puntuación de diversidad de la autoría.

Se ha considerado como valor óptimo cuando el número de autores diferentes en el periodo, dividido por el número total de autorías en los

artículos, es igual o mayor a 0,90, y como valor insuficiente cuando esta relación se sitúa por debajo del 0,65. Se considera que una publicación mejora la media, cuando su valor numérico es superior a ésta.

##### 4.2.2. Índice de concentración de autoría principal

Porcentaje que representa el autor principal sobre el número de artículos publicados en el período analizado.

| Rango de valores numéricos | Puntuación |
|----------------------------|------------|
| 10 o más                   | 0          |
| Entre 8.50 y 9.99          | 0.5        |
| Entre 7 y 8.49             | 1          |
| Entre 5.50 y 6.99          | 1.5        |
| Entre 4 y 5.49             | 2          |
| Entre 2.50 y 3.99          | 2.5        |
| Menos de 2.50              | 3          |

Tabla V. Puntuación del porcentaje de artículos del autor principal respecto al total de artículos publicados.

Se ha considerado como valor óptimo cuando el autor de mayor frecuencia de publicación en el periodo representa menos de un 2'5% del total de artículos publicados en el periodo por la revista, y como valor insuficiente cuando la frecuencia es de un 10% de los artículos o superior. Se considera que una publicación mejora la media, cuando su valor numérico es inferior a ésta.

##### 4.2.3. Concentración institucional

Porcentaje que representa la afiliación de los autores más frecuentes en el período analizado, sobre el número de registros que contienen información sobre el lugar de trabajo de los autores. Se considera el primer nivel institucional (Universidad/CSIC), no cada uno de los centros que forman parte de las mismas.

| Rango de valores numéricos | Puntuación |
|----------------------------|------------|
| 35 o más                   | 0          |
| Entre 30 y 34.99           | 0.5        |
| Entre 25 y 29.99           | 1          |
| Entre 20 y 24.99           | 1.5        |
| Entre 15 y 19.99           | 2          |
| Entre 10 y 14.99           | 2.5        |
| Menos de 10                | 3          |

Tabla VI. Puntuación de la concentración institucional.

Se ha considerado como valor óptimo cuando la afiliación institucional de los autores, con mayor frecuencia en el periodo, representa menos de un 10% del total de artículos publicados en el periodo por la revista sobre los que existen datos de afiliación, y como valor insuficiente cuando la frecuencia es de un 35% de los artículos o superior. Se considera que una publicación mejora la media, cuando su valor numérico es inferior a ésta.

#### 4.3. Internacionalidad

##### 4.3.1. Índice compensado de internacionalidad lingüística

Porcentaje de artículos en idiomas no iberoamericanos, con especial peso para el inglés. Se contabiliza como 1 los artículos editados en inglés y como 0,5 los editados en otros idiomas excepto los de la Comunidad Iberoamericana (español, portugués, catalán, gallego o euskera). Se considera como editado en inglés o en otros idiomas para los artículos bilingües o que incluyen una versión reducida de más de una página.

Baremo:

| Rango de valores numéricos | Puntuación |
|----------------------------|------------|
| 0                          | 0          |
| Entre 0.1 y 6              | 0.5        |
| Entre 6.1 y 12             | 1          |
| Entre 12.1 y 18            | 1.5        |
| Entre 18.1 y 24            | 2          |
| Entre 24.1 y 30            | 2.5        |
| Más de .0                  | 3          |

Tabla VII. *Puntuación de la internacionalidad de las revistas.*

Se ha considerado como valor óptimo cuando la suma de artículos en idioma inglés y el valor dividido por dos de otros idiomas no iberoamericanos, supone más del 30% del total de artículos publicados en el periodo por la revista, y como valor insuficiente cuando su porcentaje es nulo. Se aplica en este indicador un mayor peso al idioma inglés, al considerarlo especialmente relevante para el objetivo de internacionalidad de las revistas. El idioma portugués no se valora al considerar que su grado de internacionalidad en revistas españolas es similar al del español pues pueden representar en ambos casos artículos de autores latinoamericanos. Se considera que una publicación mejora la media, cuando su valor numérico es superior a ésta.

##### 4.3.2. Índice compensado de internacionalidad en la autoría

Porcentaje de autores con afiliación ajena a los dos países de mayor frecuencia, en los artículos publicados en el periodo analizado.

Baremo:

| Rango de valores numéricos | Puntuación |
|----------------------------|------------|
| 0                          | 0          |
| Entre 0.1 y 6              | 0.5        |
| Entre 6.1 y 12             | 1          |
| Entre 12.1 y 18            | 1.5        |
| Entre 18.1 y 24            | 2          |
| Entre 24.1 y 30            | 2.5        |
| Más de .0                  | 3          |

Tabla VIII. *Puntuación del índice compensado de internacionalidad.*

Se ha considerado como valor óptimo cuando los autores que indican una afiliación institucional no ligada a los dos países de mayor presencia, suponen más del 30% del total de artículos con información del lugar de trabajo de los autores, publicados en el periodo analizado, y como valor insuficiente cuando su porcentaje es nulo. Se aplica este indicador en base a los valores ajenos a los dos países de mayor frecuencia, en lugar del país editor como un factor de compensación, para valorar positivamente los casos en los que las contribuciones externas a la revista se distribuyen de forma amplia y no provienen casi exclusivamente de la colaboración con un único país predominante. Se considera que una publicación mejora la media, cuando su valor numérico es superior a ésta.

##### 4.3.3. Difusión en BBDD Internacionales

Valoración de la presencia en bases de datos internacionales, en especial en los índices de citas en WoS y Scopus.

Baremo:

- 1 punto por la inclusión entre las revistas fuente de las bases de datos de Web of Science Core Collection4: Arts & Humanities Citation Index y/o Social Sciences Citation Index.
- 1 punto por la inclusión entre las revistas fuente en Scopus.
- 0,5 por la inclusión en otras bases de datos documentales internacionales recogidas en el Sistema Latindex6, con un máximo de 1 punto en total para la suma en diferentes productos.

Se considera valor óptimo cuando las publicaciones están recogidas tanto en Web Science Core Collection, como en Scopus y en al menos otras dos bases de datos internacionales, y como valor insuficiente cuando no está presente

en ninguno de estos recursos. Se considera que una publicación mejora la media, cuando su valor numérico es superior a ésta.

#### 4.4. Sistema de categorización. Baremación.

El objetivo de la categorización es establecer agrupaciones representativas de la calidad de las revistas. Los datos pueden presentarse en forma de ranking exclusivamente con valor informativo, para indicar si la categoría de una publicación está justificada suficientemente o puede depender de variaciones leves en una única variable.

Sobre la puntuación total obtenida por las publicaciones, entre 0 y 27 puntos, se establece una distribución en cuatro categorías:

- Categoría A: entre 18 y 27 puntos, considerando que es suficiente obtener una media de 2 puntos por indicador para entrar en la máxima categoría.
- Categoría B: entre 13,5 y 17,5 puntos, para las revistas que superan la mitad de la puntuación posible, sin llegar al filtro establecido para la categoría A.
- Categoría C: entre 9 y 13 puntos, asignada a las publicaciones que obtienen una media de 1 punto por indicador y que no alcanzan el filtro establecido para la categoría B.
- Categoría D: menos de 9 puntos, considerando como insuficiente el caso de aquellas revistas que no alcanzan una media de 1 punto por indicador.

La puntuación final queda recogida en unos informes parciales. En cada uno de ellos, además de los resultados de categorización de las revistas, se ofrecen diferentes rankings con la finalidad de facilitar la interpretación de los datos para áreas temáticas o conjuntos determinados de publicaciones. Aunque el sistema determina una puntuación única para cada revista de forma independiente a su adscripción temática, la interpretación de que publicaciones constituyen un núcleo central o preferente sí puede realizarse por disciplinas concretas o conjuntos determinados. Los rankings complementarios que se ofrecen en cada caso son los siguientes:

1. Ranking de las revistas del grupo analizado ordenado por el total de puntuaciones obtenidas (de 0 a 27 puntos).
2. Ranking de las revistas del grupo analizado ordenado por el número de indicadores en los que la revista mejora la media del grupo (de 0 a 9 puntos).

3. Ranking de las revistas del grupo analizado ordenado por las puntuaciones obtenidas en apertura de la autoría (de 0 a 9 puntos).

4. Ranking de las revistas del grupo analizado ordenado por las puntuaciones obtenidas en internacionalidad (de 0 a 9 puntos).

### 5. Análisis de caso. Las revistas de Arqueología y Prehistoria

Con intención de ejemplificar los resultados de este nuevo sistema de categorización de revistas, se ha procedido a escoger una muestra pequeña de las revistas mejor y peor valoradas del área de Prehistoria y Arqueología de acuerdo al antiguo sistema, y compararlas con su nueva valoración. Las revistas escogidas son Complutum (Universidad Complutense de Madrid), Trabajos de Prehistoria (CSIC), Archivo de Prehistoria Levantina (Univ. Valencia), Arqueología de la Arquitectura (CSIC) y Mainake (Universidad de Málaga).

Ateniéndonos al sistema de evaluación de 2006, todas las revistas superan el cumplimiento de más de 30 criterios Latindex, sin que ello haya revertido en la mejor clasificación. Sin embargo, al aplicar el resto de criterios de evaluación, tal y como se muestra en la tabla siguiente, sólo dos revistas se integran en la categoría A (Complutum y Trabajos de Prehistoria), una en la B (Archivo de Prehistoria Levantina) y dos en la categoría C (Arqueología de la Arquitectura y Mainake).

| Título                           | Crits<br>Latindex | Cat<br>(2006) | Cat.<br>(2014) |
|----------------------------------|-------------------|---------------|----------------|
| Complutum                        | 33                | A             | A (19,5)       |
| Trabajos de Prehistoria          | 33                | A             | B (17,5)       |
| Archivo de Prehistoria Levantina | 31                | B             | C (7)          |
| Arqueología de la Arquitectura   | 33                | C             | C (14)         |
| Mainake                          | 32                | C             | C (12,5)       |

Tabla IX. Ejemplo de la recategorización de revistas de acuerdo a la aplicación de los nuevos criterios.

\* Entre paréntesis se indica la puntuación obtenida en el sistema de categorización que se presenta.

Se muestra así que el cumplimiento exclusivo de los 33 criterios Latindex no son garantía suficiente para asegurar una buena valoración de las revistas. Desde que este sistema entrara en vigencia en 1997, los editores han ido volcando sus esfuerzos en cumplir con todos los criterios editoriales, sin que ello haya revertido necesariamente en una mejora de su calidad editorial.

El nuevo sistema de evaluación, como ya se ha explicado en las páginas anteriores, utiliza criterios nuevos que redefinen la nueva categorización de las revistas con criterios más precisos y que analizan el comportamiento de cada revista en sí misma independientemente de la disciplina a la que se adscriba.

De acuerdo a los indicadores que se han descrito en los apartados anteriores, la evaluación de las revistas mencionadas quedaría expuesta como se indica en el gráfico 1 (Anexo1).

Complutum y Trabajos de Prehistoria -que formaban parte del grupo A-, tienen buena puntuación en los indicadores de trayectoria (especialmente en los de periodicidad y requerimientos Latindex/sello FECYT), pero en cuanto a los indicadores de apertura o de internacionalidad, están superadas por revistas de categoría inferior como Mainake o Arqueología de la Arquitectura (C). Todavía más llamativo es el ejemplo de Archivo de Prehistoria Levantina (B) que presenta los índices más bajos de la muestra, sobre todo en los índices referidos a apertura de la autoría e internacionalidad, con independencia de la composición de sus comités científicos o de redacción.

La comparación para el resto de las revistas de las áreas analizadas ha supuesto cambios importantes. El caso del área de Arqueología y Prehistoria en donde los cambios son más agudos, se debe a la propia naturaleza del área que tiende a la endogamia en la autoría propiciada por el predominio de los estudios de carácter local. Así, el área pasa de tener 4 revistas "A" a sólo dos, a la par que crece mucho el número de revistas "D", esto es, revistas que no pasan el umbral mínimo de calidad. De hecho, es un comportamiento muy parecido al que muestran las revistas de estudios locales de Historia. La reducción de las revistas "A" en esta área también guarda relación con las dificultades propias del área de difundir las revistas en BBDD internacionales. La recategorización general para todas las áreas se muestra en los gráficos contenidos en los anexos 2 y 3.

Salvo en el caso del área de Prehistoria y Arqueología, el resto de las áreas ha incrementado la calidad de sus revistas. Significativamente además en algunos casos. Es importante indicar que las revistas "C" son las más numerosas debido a que ha sido el grupo en donde se han ido incorporando todas aquellas revistas que fueron añadiéndose en la base de datos desde 2006. El área de Lingüística, el más llamativo quizás, pasa de tener 8 revistas A a 36, a la par que Historia lo aumenta en 10 (de 16 a 26) y Antropología incorpora 7 en esta categoría

en la que anteriormente no tenía ningún ejemplar. Las revistas B también disfrutaban de un aumento generalizado, lo cual redundaba en la reducción de las categorías menos valoradas, las C y D.

## 6. Conclusiones

El sistema de categorización aplicado en las bases de datos ISOC desde 2006 no es suficiente en la medida en que no recoge el comportamiento de las revistas en su trayectoria temporal, no analiza autores ni instituciones, así como que los indicadores de internacionalidad son muy limitados.

A partir de este nuevo sistema de categorización, la base de datos ISOC amplía su utilidad como herramienta para la comunidad investigadora. Al dar un cuadro de resultados más ajustado, minucioso y preciso con el comportamiento de las revistas, permite a éstas mejorar en aquellas cuestiones en donde muestran unos indicadores bajos.

No obstante, quedan cuestiones pendientes de mejora. ¿Cómo integramos la multi o interdisciplinariedad? Sabemos que la evaluación de las revistas académicas son inseparables de su clasificación temática, pero también es una realidad que los estudios en Ciencias Sociales y Humanidades tienden cada vez más a ser multidisciplinarios. Aunque está pendiente de estudio, sería deseable añadir indicadores que midieran la interdisciplinariedad de las revistas desde el análisis de las clasificaciones temáticas a nivel de revista y a nivel de artículo. De esta manera, podrían valorarse revistas que, como "Documentación de la Salud", que aúnan dos áreas -Documentación y Medicina- con diferentes comportamientos editoriales.

En este sentido también se puede indicar que el sistema que se propone tenga capacidad para valorar el comportamiento de una revista en sí misma, pero que al mismo tiempo pueda analizarlas comparativamente con publicaciones de su misma naturaleza. En este sentido, los rankings incluidos en los informes parciales pueden constituir una herramienta muy útil, pues permite valorar cada revista en el conjunto de su disciplina.

## Notas

- (1) Esta cifra es relativa en cualquier caso, pues las publicaciones con doble versión tienen dos registros en Latindex

## Referencias

- Abejón Peña, T., Maldonado Martínez, Á., Rodríguez Yunta, L., & Rubio Liniers, M. (2009). La base de datos ISOC como sistema de información y fuente para el análisis de las ciencias humanas y sociales en España. *El Profesional de la Información*, 18(5), 521-528.
- Alcaín Partearroyo, M., Román Román, A., & Giménez Toledo, E. (2008). Categorización de las revistas españolas de Ciencias Sociales y Humanas en RESH. *Revista Española de Documentación Científica*, 31(1), 85-95.
- Rodríguez Yunta, L. (2015). Evaluación de las revistas españolas: pasos hacia un sistema que mida la apertura y la internacionalidad. *Anuario ThinkEPI*, 9, 189-195.
- Rodríguez Yunta, L. (2015). *Metodología aplicada en la categorización de las revistas españolas en la base de datos ISOC: indicadores sobre trayectoria, apertura de la autoría e internacionalidad en el periodo 2004-2013*.
- Torres Salinas, D., & Cabezas-Clavijo, Á. (2013). *Cómo publicar en revistas científicas de impacto: consejos y reglas sobre publicación científica*.

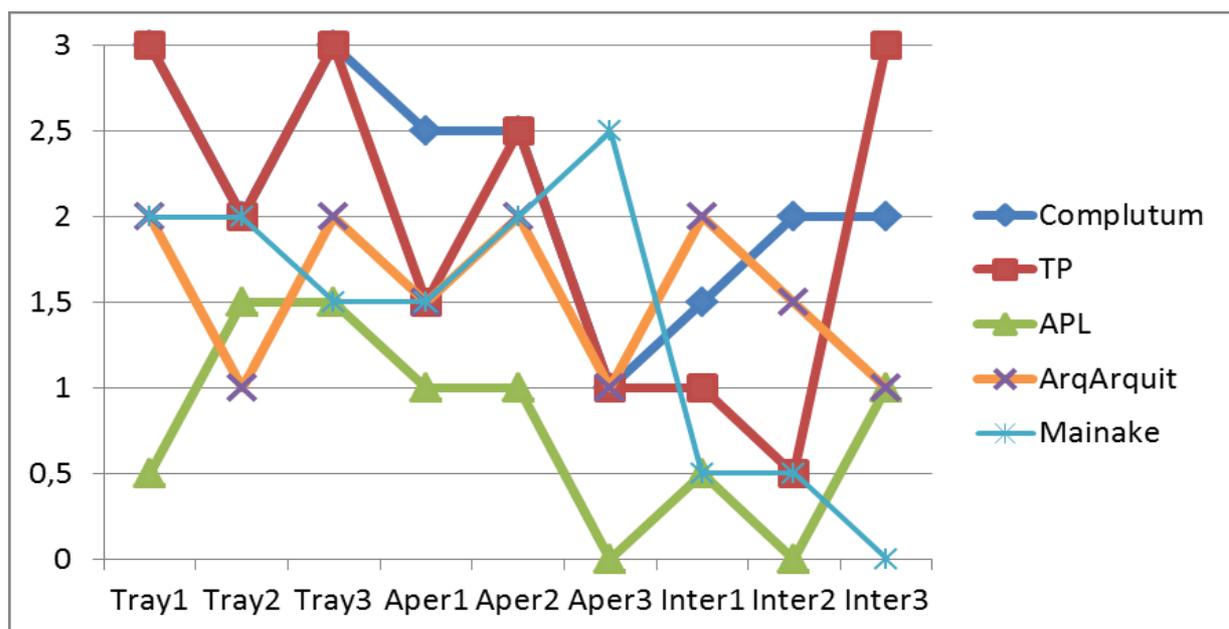
## Informes

- Andrés Verdú, Charo; Ortega Priego, José Luis; Rodríguez Yunta, Luis; Vidal Liy, Nacho; Abejón Peña, Teresa. Caracterización de las revistas españolas de Historia: Producción científica y perfil temático en el periodo 2004-2013. Documento de trabajo ISOC 2014/01. <http://hdl.handle.net/10261/103401>
- Ortega Priego, José Luis; Rodríguez Yunta, Luis. Caracterización de las revistas españolas de Antropología: Producción científica y perfil temático en el periodo 2004-2013. Documento de trabajo ISOC 2014/03. <http://hdl.handle.net/10261/104214>
- Barreales Tomé, Mar; Ortega Priego, José Luis; Rodríguez Yunta, Luis. Caracterización de las revistas españolas de Geografía: Producción científica y perfil temático en el periodo 2004-2013. Documento de trabajo ISOC 2014/05. <http://hdl.handle.net/10261/108672>
- Cerdán Medina, José Carlos; Ortega Priego, José Luis; Rodríguez Yunta, Luis. Caracterización de las revistas españolas de Biblioteconomía y Documentación: Producción científica y perfil temático en el periodo 2004-2013. Documento de trabajo ISOC 2015/01. <http://hdl.handle.net/10261/109522>
- Vidal Liy, Nacho; Ortega Priego, José Luis; Rodríguez Yunta, Luis. Caracterización de las revistas españolas de Arqueología y Prehistoria: Producción científica y perfil temático en el periodo 2004-2013. Documento de trabajo ISOC 2015/03. <http://hdl.handle.net/10261/111061>
- Campos Alberca, Julia; Ortega Priego, José Luis; Rodríguez Yunta, Luis. Caracterización de las revistas españolas de Lingüística: Producción científica y perfil temático en el periodo 2004-2013. Documento de trabajo ISOC 2015/05. <http://hdl.handle.net/10261/113784>
- Caracterización de las revistas españolas de Ciencias Sociales y Humanas con sello de calidad FECYT: Producción científica y perfil temático en el periodo 2004-2013. Documento de trabajo ISOC 2015/09. <http://hdl.handle.net/10261/116883>
- Caracterización de las revistas españolas de ciencias humanas y sociales en Web of Science y/o Scopus: Producción científica y perfil temático en el periodo 2004-2013. Documento de trabajo ISOC 2015/11. <http://hdl.handle.net/10261/116893>
- vectoria, apertura de la autoría e internacionalidad en el periodo 2004-2013. Documento de trabajo ISOC 2014/02. <http://hdl.handle.net/10261/103402>
- Rodríguez Yunta, Luis. Categorización de las revistas españolas de Antropología: Propuesta de indicadores sobre trayectoria, apertura de la autoría e internacionalidad en el periodo 2004-2013. Documento de trabajo ISOC 2014/04. <http://hdl.handle.net/10261/104217>
- Rodríguez Yunta, Luis. Categorización de las revistas españolas de Geografía: Propuesta de indicadores sobre trayectoria, apertura de la autoría e internacionalidad en el periodo 2004-2013. Documento de trabajo ISOC 2014/06. <http://hdl.handle.net/10261/108674>
- Rodríguez Yunta, Luis. Categorización de las revistas españolas de Biblioteconomía y Documentación: Propuesta de indicadores sobre trayectoria, apertura de la autoría e internacionalidad en el periodo 2004-2013. Documento de trabajo ISOC 2015/02. <http://hdl.handle.net/10261/109523>
- Vidal Liy, Nacho; Rodríguez Yunta, Luis. Categorización de las revistas españolas de Arqueología y Prehistoria: indicadores sobre trayectoria, apertura de la autoría e internacionalidad en el periodo 2004-2013. Documento de trabajo ISOC 2015/04. <http://hdl.handle.net/10261/111062>
- Rodríguez Yunta, Luis. Categorización de las revistas españolas de Lingüística: indicadores sobre trayectoria, apertura de la autoría e internacionalidad en el periodo 2004-2013. Documento de trabajo ISOC 2015/06. <http://hdl.handle.net/10261/113785>
- Rodríguez Yunta, Luis. Categorización de las revistas españolas de Ciencias Sociales y Humanas con sello de calidad FECYT: indicadores sobre trayectoria, apertura de la autoría e internacionalidad en el periodo 2004-2013. Documento de trabajo ISOC 2015/10. <http://hdl.handle.net/10261/116886>
- Rodríguez Yunta, Luis. Categorización de las revistas españolas de ciencias humanas y sociales en Web of Science y/o Scopus: Indicadores sobre trayectoria, apertura de la autoría e internacionalidad en el periodo 2004-2013. Documento de trabajo ISOC 2015/12. <http://hdl.handle.net/10261/116894>

## Informes de categorización

- Rodríguez Yunta, Luis. Categorización de las revistas españolas de Historia: Propuesta de indicadores sobre tra-

## Anexo 1



TRAY1: Años cubiertos en el período; TRAY2: Mediana arts./año; TRAY3: Prácticas editoriales (Sello Fecyt/Latindex); APER1: Índice diversidad de la autoría; APER2: Índice de concentración de la autoría; APER3: Índice de concentración institucional; INTER1: Índice de internacionalidad lingüística; INTER2: Índice de internacionalidad en la autoría; INTER3: Difusión en Bases de Datos Internacionales.

## Anexo 2

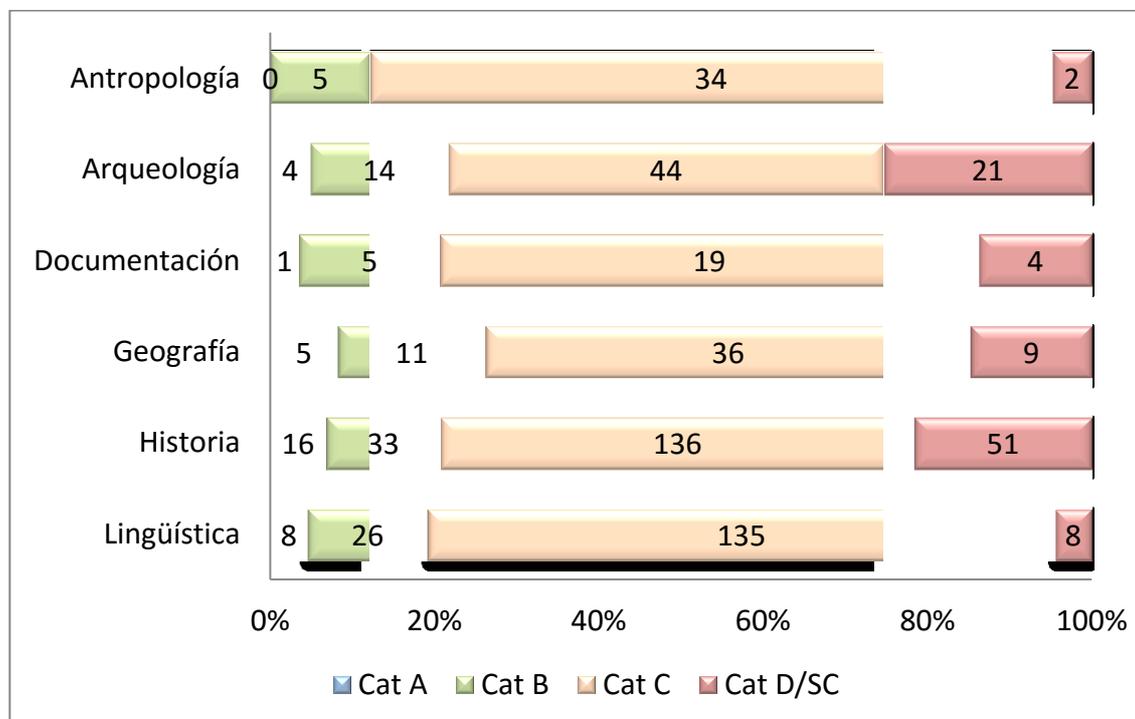


ILUSTRACIÓN 1. CATEGORIZACIÓN DE REVISTAS DE ACUERDO AL SISTEMA DE EVALUACIÓN APLICADO DESDE 2006

## Anexo 3

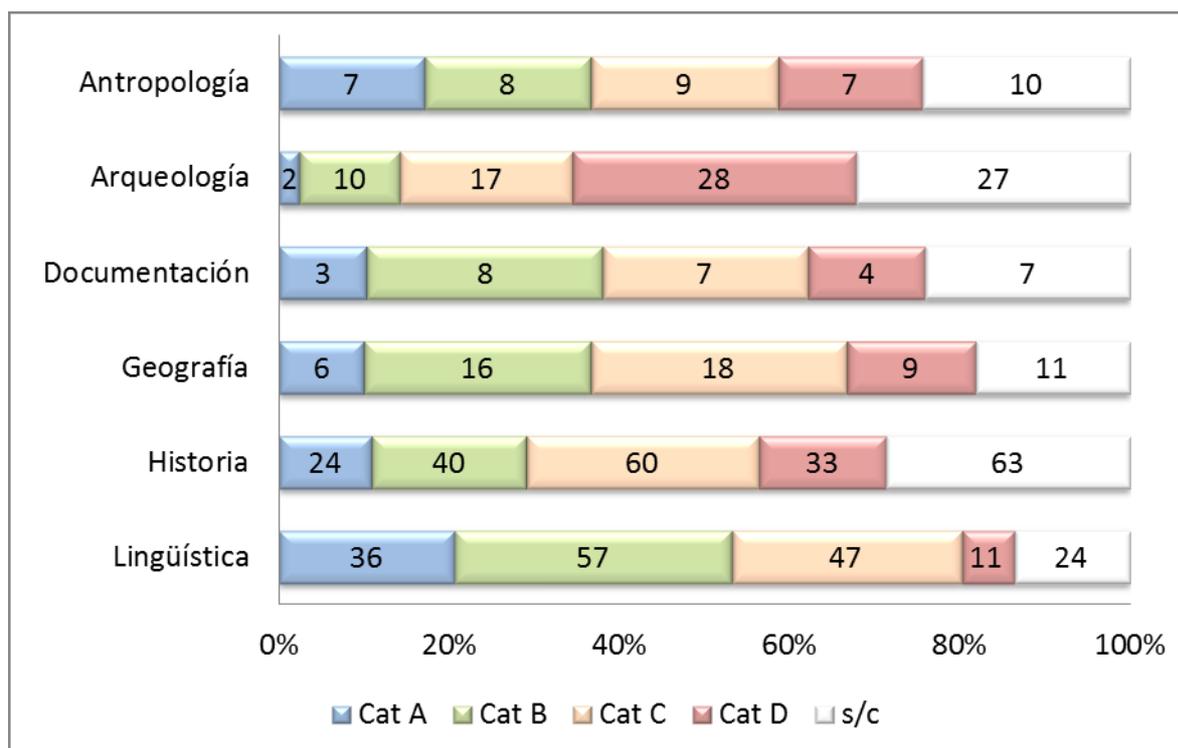


ILUSTRACIÓN 2. CATEGORIZACIÓN DE REVISTAS DE ACUERDO A LA NUEVA PROPUESTA

**Hito 1: Modelos de organización del conocimiento. Modelaje de sistemas y ciclo de vida de la información**

**Hito 2: Entornos y sistemas para la organización del conocimiento. Diseño, normalización e interoperabilidad de sistemas de información abiertos.**

**Hito 3: Instrumentos para el acceso y transparencia de la información: gestión de documentos, plataformas digitales y archivos**

**Hito 4: Usuarios y usos de la información. Patrones vinculados a la transparencia y el gobierno abierto**

**Hito 5: Gobernanza y Auditoría de información en la organización del conocimiento.**

**Hito 6: Representación del conocimiento: procesos y herramientas en sistemas abiertos y multilingües**

**Hito 7: Epistemología en la organización del conocimiento**

**Hito 8: Igualdad, género y diversidad en la organización del conocimiento**

---

# Aportaciones Semióticas de la Documentación en España a la Organización Del Conocimiento: un análisis preliminar

*Semiotics Contributions in Spain Documentation to knowledge organization: a preliminary analysis*

---

**Carlos Cândido de Almeida (1), Francisco Javier García Marco (2)**

(1) Estancia de investigación FAPESP. Departamento de Ciencias de la Documentación y Historia de la Ciencia - Universidad de Zaragoza y Departamento de Ciência da Informação - Universidade Estadual Paulista. Correo eletrônico: carlosalmeida@marilia.unesp.br.

(2) Profesor Doctor Catedrático de la Universidad, Departamento de Ciencias de la Documentación y Historia de la Ciencia - Universidad de Zaragoza. Correo eletrônico: jgarcia@unizar.es

## Resumen

Las teorías semióticas en las líneas lógicas y lingüísticas han sido objeto de estudio de diversas disciplinas pues tienen como reto principal conocer los problemas asociados al significado de palabras, textos, imágenes y sonidos. De hecho, ese ha sido también el caso del campo de la Información y Documentación en España, en la que se han investigado los temas semióticos, aunque no hayan estado entre los que más atención han recibido en los últimos años. Las contribuciones de la Documentación española en esta materia aún no han sido investigadas detenidamente. Así que, el objetivo general de esta investigación, aún en desarrollo, ha sido conocer las aportaciones semióticas de la Documentación en España. El análisis de los datos fue cualitativo y reunió los discursos de expertos en documentación sobre el tema. Respecto a los resultados, se ha notado que las propuestas semióticas a la organización del conocimiento, generadas en los años ochenta, han sido fundamentales, incluso para constituir una de las aportaciones más originales a los estudios de Documentación.

**Palabras clave:** Semiótica. Semiótica Documental. España. Documentación. Organización del Conocimiento.

## 1. Introducción

Las teorías semióticas, en sus aportaciones lógicas y lingüísticas, han sido objeto de interés de diversas ciencias a lo largo de las últimas décadas, pues tienen como reto principal conocer los problemas asociados al significado de palabras, textos, imágenes y sonidos. La Semiótica investiga una multitud de códigos y ha contribuido con aportes teóricos y aplicados a los métodos del tratamiento de dichos códigos.

Desde luego, los procesos documentales de lectura documental, indización, clasificación de materias, etc., son, en concreto, actividades

## Abstract

The semiotic theories on logical and linguistic lines have been studied in several disciplines whose main challenge know the problems associated with the meanings of words, texts, images and sounds. In fact, that was also the case in the Information and Documentation field in Spain, which have been investigated semiotic issues, but have not been received more attention in recent years. The contributions of the Spanish documentation on this matter have not been thoroughly investigated. So, the overall goal of this research was to determine the semiotic contributions of Spanish documentation. The data analysis was qualitative and analyzed researches speeches about the subject. Regarding the results, it has been noticed that the semiotic proposals to the knowledge organization generated in the 1980's have been critical, even to be one of the most original contributions in Documentation studies.

**Keywords:** Semiotics. Documental semiotics. Spain. Documentation. Knowledge Organization.

semióticas por su propia naturaleza y, por lo tanto, suponen el manejo, la traducción, la transmisión y la interpretación de los signos. Los signos son los elementos básicos de la decodificación y de la representación de información.

En el caso del área información y documentación en España, los temas semióticos y lingüísticos no solo han sido investigados, sino que se les han propuesto nuevas conexiones teóricas cuyos problemas documentales siempre estuvieron considerados. No obstante, las aportaciones a éstos temas no han sido totalmente detalladas, tampoco han logrado la atención merecida en los últimos años en las revistas

científicas en España. Por lo tanto, habría que revisarlos, puesto que constituye una vía para dar a conocer la historia de la teoría de la documentación en España. Las aportaciones semióticas que han dado lugar en la Documentación en España estén entre los principales planteamientos teóricos reconocidos en otros países. Consecuentemente, hay que tener en cuenta que muchos autores de lengua castellana y portuguesa suelen utilizar las aportaciones teóricas creadas en España, por lo cual no se las debería olvidar.

Así que merece la pena escribir algunas líneas más sobre esta historia, pues las contribuciones de la documentación española en este tema han sido potentes y hay que explicitarlas detenidamente. Tal vez sus logros sólo serán reconocidos en el futuro, si se supiera recuperarlos. Con el reto de contribuir al debate se ha subrayado este capítulo de la historia de la teoría de la documentación en España, el proyecto lingüístico-semiótico. El objetivo general de la investigación ha sido conocer las aportaciones semióticas de la documentación en España, enmarcadas en el contexto más amplio de los estudios semióticos y documentales. En particular, se ha intentado, a lo largo de la investigación: identificar los marcos teóricos de la semiótica en la documentación en España; caracterizar las líneas semióticas adoptadas por la documentación en España, y evaluar el nivel de contacto disciplinario entre documentación y semiótica. En este trabajo ponen de relieve los resultados de la investigación respecto a las líneas semióticas, a la concepción de la semiótica como ciencia y el nivel de interdisciplinariedad comentado por un grupo de investigadores en documentación con reconocido prestigio en el tema.

Con respecto a las teorías semióticas, estas han sido divididas en dos frentes: uno formado por las teorías peirceanas, las cuales recurren a la propuesta semiótica de Charles S. Peirce (1834-1914), y el otro titulado teorías no peirceanas, que no están vinculadas a los estudios peirceanos y suelen vincularse a las propuestas lingüístico-semiológicas. Como consecuencia, se ha sido posible entender mejor las distintas posiciones teóricas asumidas por los expertos en documentación. En la investigación también se ha adoptado la idea de otro frente de teorías semióticas híbridas que recoge elementos conceptuales de distintas formaciones de la semiótica, tales como la semiótica de las imágenes, semiótica de la cultura, etc., que suelen adoptarse para hacer frente a demandas específicas y, en concreto, solucionar problemas de tratamiento de contenido documental de códigos no verbales.

En efecto, ha sido posible apuntar las diferencias entre los planteamientos y entender la vinculación de sus conceptos. A continuación, se comentarán los procedimientos metodológicos que se han adoptado en esta investigación.

## 2. Aspectos metodológicos

En términos metodológicos, por una parte, se han recolectado los datos bibliográficos para conocer las principales contribuciones teóricas en este sentido y, por otra parte, se han manejado los datos discursivos de investigadores del área para componer una síntesis de las líneas semióticas de la documentación en España, las cuales han contribuido a suministrar, teórica y metodológicamente, elementos a la organización del conocimiento.

En total, fueron consultados 20 expertos (18 investigadores del área de documentación, un experto de semiótica y uno dedicado a los temas de lingüística), 16 mediante entrevistas grabadas, 2 contestaron vía cuestionario, y 2 fueron entrevistados sin el registro del diálogo. La recolecta de datos ocurrió en el período de noviembre de 2014 a marzo de 2015, y los temas de la entrevista versaron sobre sus temas de investigación, su concepción sobre la ciencia semiótica, sus aplicaciones en documentación y el nivel de interdisciplinariedad.

El análisis de los datos fue cualitativo, pues se buscó entender los sucesos que han influido en el desarrollo de la semiótica en la documentación. La técnica que se utilizó la llamada análisis del discurso del sujeto colectivo, muy utilizada en el tratamiento de datos en las ciencias humanas y ciencias de la salud. La técnica fue creada por Lèfevre y Lèfevre (2003) y se fundamenta en la concepción de que el "habla social" de una colectividad puede ser representada por el conjunto de "hablas individuales", con sus similitudes y contradicciones, puesto que la presencia de categorías de análisis (como del análisis de contenido) no sustituyen en absoluto el discurso propio del grupo de personas. Las categorías se convierten en las ideas centrales y sólo se las utiliza para la organización del discurso colectivo, visto que el habla debe representarse a sí misma.

Resulta del tratamiento colectivo de los discursos un bloque de ideas asociado al pensamiento social, formado a través del habla de los entrevistados. Por lo tanto, el guión de entrevista tenía que respetar la dinámica de la técnica y promover la fluidez de la conversación. El procedimiento obedece a reglas de semejanza entre los temas hablados (Lèfevre y Lèfevre recurrieron al concepto peirceano de iconicidad

porque este pone de relieve la conexión entre los discursos), de reducción de la redundancia (pues muchas ideas se repiten en más de una entrevista), y de la estructura textual con el uso de expresiones que garanticen la unión de oraciones (se debe notar que lo que se ha subrayado es necesario para la cohesión del discurso).

Lo que llama la atención es que el habla resultante puede ser leída como un fragmento del pensamiento social del grupo. Lèfevre y Lèfevre (2003) se basaron en la teoría de las representaciones sociales y en la noción durkheimiana de representaciones colectivas, pues al fin y al cabo, éstas se refieren a una idea existente del colectivo investigado. No se trata de una suma simple de frases, sino una reconstrucción del discurso del colectivo en una forma coherente. En el caso estricto de este trabajo, el pensamiento colectivo que será expuesto a través de la técnica tratará de la concepción de semiótica y del nivel de interdisciplinariedad, que han manifestado el grupo de investigadores de la documentación en España entrevistados.

### 3. Aportaciones Semióticas

Respecto a los resultados de la investigación, se ha notado que las propuestas semióticas en organización del conocimiento se concentran en las reflexiones conceptuales de la Lingüística Documental y las asociadas a la epistemología de una gran disciplina que ha de ser llamada Semiótica Documental. Hay que explicar que otras líneas teóricas han sido identificadas, como las aportaciones de la lingüística computacional, del análisis de imágenes, etc., sin embargo, no serán examinadas en este trabajo. A continuación, se expondrán las ideas centrales de las dos aportaciones semióticas principales y sus conexiones teóricas.

#### 3.1. Lingüística documental

La Lingüística documental es una disciplina fundamental para entender la influencia de la semiótica en el discurso de la documentación en España. Hay que tener un poco de cuidado cuando se emplea la palabra disciplina, visto que no parece consensual esta condición ante la Documentación, incluso en Lingüística. La obra más influyente sobre la materia fue escrita por García Gutiérrez y fue publicada en 1984, en la cual se encuentra propuesta la Lingüística documental como una "interdisciplina de lenguajes documentales", o mejor dicho:

[...] una disciplina impregnada de otros campos científicos, como son básicamente la Lógica, la Estadística, y la Informática y más concretamente,

la Lexicología, la Archivística, la Biblioteconomía y la Telemática, bajo los imperativos específicos de ámbito del saber que va a ser controlado por un lenguaje documental. (García Gutiérrez, 1984, p.138).

Esta disciplina adopta un concepto de lenguaje documental que, según el autor, es su principal objeto de interés. Por lenguaje documental se entiende un sistema de signos naturales o artificiales que tiene como objetivo la identificación de documentos que son utilizados en el análisis y la recuperación de la información. El lenguaje funciona como medio de expresión creado, por una parte, para controlar y, por otra, para realizar la comunicación, es decir, una función intermediaria.

Merece la pena llamar la atención sobre la acepción de lenguaje. Un lenguaje es un sistema de signos inscrito en un documento.

El sistema de signos escritos en el documento es lenguaje humano pero el sistema de signos naturales o artificiales utilizados para identificar ese documento, entre un conjunto de ellos, es lenguaje documental, aunque éste no pueda existir si el documento carece del contenido que le otorga la acción intelectual humana. (García Gutiérrez, 1984, p. 136).

Desde luego, el lenguaje es un sistema de signos, hecho que no está de acuerdo con la definición estructuralista de lenguaje en Barthes, tampoco en Saussure, puesto que el dispositivo de ejecución del lenguaje está ausente en la definición. Como se puede suponer, el énfasis en la noción del lenguaje que ha sido ampliamente aceptada en Documentación está en el sistema y en el código, no en el habla. En igual forma, el concepto de Lingüística empleado por el autor está resumido en la fórmula "ciencia que estudia el lenguaje humano". Desde un punto de vista aplicado, García Gutiérrez (1984, p.139) enumeró los conceptos derivados del cruce disciplinar con la Lingüística: el descriptor, o unidad significativa mínima del mensaje documental; la frase documental, que se entiende por el conjunto de descriptores vinculados o no vinculados a las relaciones sintácticas artificiales; y el resumen documental, un producto resultante de la aglutinación de mensajes de un documento que pueden ser expresadas en un lenguaje controlado.

La aportación de García Gutiérrez se basa concretamente en una línea semiótica no peirceana, aunque sus contribuciones alejadas de un acercamiento superficial a Lingüística. Además, hay que subrayar el debate sobre el lenguaje científico y técnico, un léxico especializado adoptado por una comunidad científica, pues la ciencia depende de la comunicación entre co-

munidades científicas, y el documentalista, así como los lingüistas y terminólogos, han de ser guardianes de la lengua (García Gutiérrez, 1984, p.141).

En efecto, el autor defendió que los lenguajes documentales no pueden ser sólo vocabularios técnicos, como asevera Trujillo (*apud* García Gutiérrez, 1984, p. 144), puesto que

Este razonamiento no es válido con el de los lenguajes documentales, en los que, si bien a niveles simples, se dan relaciones semánticas entre los conceptos e incluso sintácticos, por tanto hay que hablar de lenguajes dinámicos frente al estatismo interno de un vocabulario. (García Gutiérrez, 1984, p.144).

El autor defendió también la tríade documentación-terminología-normalización y propuso que el documentalista debe tener conocimientos de terminología y normalización para realizar su labor (García Gutiérrez, 1984, p. 147). Luego, las tesis generales de la Lingüística documental le ponen en un sitio de cruce entre disciplinas, cuyo problema no ha sido planteado por los abordajes tradicionales.

De acuerdo con García Gutiérrez (1990), la Lingüística documental no debe ser vista como una lingüística aplicada, sino como una genuina teoría de la documentación cuyo método de abordaje se centra en la noción de estructura y sistema relacional. La noción de estructura de la Documentación es muy importante en Lingüística documental. Según el autor:

La estructura de la Documentación, como organización de contenidos codificables y decodificables, es el objeto de la Lingüística documental en un doble sentido: 1) La estructura de la producción de información, la formación, organización y presentación de las ideas por parte del productor, como corpus de observación y descripción; 2) La estructura de la representación del discurso del productor, a la que se accede mediante formulaciones metodológicas y modelos de síntesis y traducción. (García Gutiérrez, 1990, p. 24).

En lo tocante a las bases teóricas generales, en Lingüística documental son fundamentales los conocimientos provenientes de la Lingüística general, la Semántica, la Lingüística textual, la Sociolingüística, el Estructuralismo, el Generativismo, la Teoría de la comunicación y de la información, la Teoría del contexto, la Teoría de la representación y la Teoría de la traducción. Sin embargo, no se puede perder la individualidad y supuesta autonomía ante éstos campos; la Lingüística documental se acerca a estas aportaciones para responder a las demandas de interés de la Documentación.

Siguiendo las ideas de García Gutiérrez, López Yepes y Ros García (1993) establecen que existe una concepción lingüística de la Documentación. Comentan ellos que hay una influencia de la Lingüística documental como un área que suministra conceptos y teorías a la Documentación:

Como se sabe, en el análisis documental en sentido amplio, cobra especial relevancia el uso de los lenguajes de clasificación y de indización. Su estudio profundo. Riguroso y eficaz ha dado lugar a la formación de una asignatura denominada Lingüística Documental o Lenguajes documentales. (López Yepes y Ros García, 1993, p. 121)

Lo que se ha notado en el tema de la Lingüística documental es su estatuto epistemológico, así como su implantación y desarrollo institucional en documentación, pues al fin y al cabo habría que lograr un mejor espacio en los departamentos universitarios y en las diplomaturas/licenciaturas/grados en Documentación. En ese mismo sentido, se corrobora la afirmación de Tálamo y Lara (2006, p. 210, traducción libre),

El cuerpo de conceptos de la Lingüística Documental no goza, hasta la fecha, de estabilidad. Hay todavía una variación de su nombre muy grande que ha expresado el estadio embrionario de los contornos del dominio. Esa es, no obstante, sólo una característica de la estructuración del campo.

Las autoras han aportado una excelente contribución en la explicación del concepto de Lingüística documental a comunidad brasileña de expertos y profesionales en Documentación, hasta el punto de establecer espacios epistemológicos de acción. La Lingüística documental, según han comentado Tálamo y Lara, ha sido considerada un subdominio de la Documentación y sostiene características de interdisciplina, puesto que se apropia de conceptos de la Lingüística estructural, Semiótica (de matiz morrisiana), Terminología y Lógica formal. El objeto principal han de ser los productos documentales: "cabe a Lingüística Documental componer los cuadros de referencia para el análisis, evaluación y construcción del lenguaje documental, que se entiende como lenguaje de información, que asocia los niveles sintáctico-semántico-pragmático [...]" (Tálamo y Lara, 2006, p. 206, traducción libre).

Entre sus temas de análisis están las estructuras simbólicas de la documentación, las cuestiones lingüísticas de la relación productores-consumidores de información, la elaboración de lenguajes documentales, el análisis de las operaciones con los lenguajes documentales. Lara y Tálamo (2007) entienden que el trabajo de

García Gutiérrez ha sido pionero y remonta a las iniciativas de Jean-Claude Gardin, así como a de muchos autores brasileños, según sostienen las autoras.

Es posible establecer una conexión teórica entre los trabajos de García Gutiérrez y las contribuciones lingüísticas a la documentación que provienen de diferentes autores, con mayor énfasis, con las aportaciones de Moreiro González (1993; 2011). Pero, no se deben olvidar las aportaciones de la Lingüística computacional, la Lingüística cognitiva y Terminología a una Lingüística documental. Se ha notado, pues, un proceso de establecimiento de contribuciones semiótico-lingüísticas, en las cuales sobresale el matiz estructural.

Se supone que no solo el hecho de que la Lingüística documental esté en un nivel intermedio de desarrollo describe su situación ante la Documentación, sino que los problemas que ha procurado solucionar y las herramientas teóricas de las que dispone, tal vez no sean los más relevantes actualmente. Hay que contrastar la Lingüística Documental con una Semiótica Documental, y poner de relieve la tesis de desplazamiento de los temas, según la cual unos abordajes han que seguir siendo lingüísticos y los demás deben recurrir a una teoría semiótica.

### 3.2. Semiótica documental

Se pasa ahora a presentar las ideas asociadas a Semiótica documental que se han producido en la Documentación en España. Aquí se debe contestar otra vez a los que piensan que la propuesta “semiótica documental” es solo un análisis semántico aplicado a la representación en vocabularios. Esta visión es muy limitada, ignora lo que se entiende por Semiótica, y desconoce las tesis planteadas en Semiótica documental. La Semántica, especialmente, la Semántica lingüística que suele utilizarse en Documentación, no comprende todos los temas tratados por la Semiótica, aunque esté volcada al significado; y sus fundamentos no permiten explicar el proceso de interpretación fuera del signo lingüístico y del registro verbal. Por lo tanto, se está totalmente en desacuerdo con lo que se ha sugerido de cambiar la expresión “semiótica documental” por “aplicaciones semánticas a la Documentación”. La epistemología de la Semiótica documental, si ha sido bien entendida, defiende que sus problemas están más allá de lo lingüístico.

Desde luego, la propuesta de mayor proyección en ese sentido ha sido la de Izquierdo Arroyo que acuñó la expresión “Semiótica Documental”, utilizada desde finales de los años 1980. La idea

de una Semiótica documental surgió por primera vez, según se entiende, en el proyecto docente escrito por Izquierdo Arroyo, redactado para el proceso de admisión como profesor para la Universidad de Murcia. En entrevista concedida en 2014, Izquierdo Arroyo explicó que estaba escribiendo los “Esquemas de Lingüística Documental”, entre los 1989 y 1990, “Pero el hilo conductor de mi propuesta era ya la Semiótica peirceana. Para la elaboración del trabajo Sobre la transducción, había reunido en Burgos bastante material de Semiótica y Semiología.” (Izquierdo Alonso y Izquierdo Arroyo, 2014, p. 111). No obstante, se puede inferir que su contacto con el pensamiento de Peirce más detenido fue a finales de los años 1970.

Instalado en la Filosofía del Lenguaje (de corte más bien analítico) y en la Semiótica – acababa de leer a Peirce –, consideré que la aproximación más adecuada sería la lógico-lingüística (término este que ya empezaba a aparecer en algunas publicaciones recientes de entonces). A ello obedeció el título y orientación del libro *Concepción lógico-lingüística de la Documentación*. (Izquierdo Alonso y Izquierdo Arroyo, 2014, p. 111)

El acercamiento a la obra del filósofo Peirce ocurrió muy temprano, concretamente en una asignatura mientras estudiaba la licenciatura en Filosofía en la Universidad Pontificia de Salamanca, entre 1962 y 1965. Según Izquierdo Arroyo (2014) ha relatado “[...] Delgado, nos daba un curso de Peirce en el año 63, o sea, ya te lo digo todo, o sea, era un adelantado, era un adelantado.” Su tesis de fin de curso en Filosofía versó sobre la ontología y el sistema categorial de Nikolai Hartmann (Izquierdo Alonso y Izquierdo Arroyo, 2014, p. 112). Ya en 1981, leía su tesis de doctorado en la Universidad Complutense de Madrid sobre historia de la lógica medieval, titulada “Lógica proposicional sumulista: bases histórico-textuales” que todavía sigue inédita.

Antes de incorporarse a Universidad de Murcia y efectivamente a la docencia en Documentación, Izquierdo Arroyo dictó las asignaturas de Literatura y Crítica Literaria, en Burgos, entre 1975 y 1985, después pasó a Málaga, donde impartió clases de Lógica e Filosofía de la Ciencia, de 1985 a 1989. En las asignaturas Lingüística documental I, Lingüística documental II, Seminario de Estudios I, Seminario de Estudio II y Trabajo de Fin de Curso, en la diplomatura de Biblioteconomía y Documentación de la Universidad de Murcia, se impartieron temas semióticos, los cuales conllevaban reflexiones originales sobre sus aplicaciones en el campo de la Documentación (Izquierdo Arroyo, 1990).

Según Izquierdo Arroyo (1993, p. 200), la Semiótica Documental es el marco acogedor de las denominadas Ciencias del Texto en la aplicación al tratamiento documental. La definición de Semiótica posee la intención de reconstituir el papel de la tríade expresión/contenido/referente. Para constituir la Semiótica documental Izquierdo Arroyo ha propuesto sustituir la Lingüística textual, puesto que esta no cubre los tres planos: el plano de expresión (las palabras), el plano del referente (las cosas), y la base mediadora entre lengua y mundo, el plano de contenido (Izquierdo Arroyo, 1993, p. 202). Consecuentemente, la propuesta Semiótica Documental pone de relieve la importancia de los tres planos, por lo tanto, supone una teoría triádica del signo.

En fin, la sustitución de "lingüística" por "semiótica" se hace asumiendo la Semiótica peirceana como telón de fondo. El análisis de los conceptos peirceanos ha estado presente en la teoría de Izquierdo Arroyo. Lo que se nota por el continuo uso de los conceptos nucleares de la lógica, signo, tricotomía, interpretante, realidad, cosa, etc. (Izquierdo Arroyo, 1993). Según esta línea, hay que sustituir la Lingüística documental por Semiótica documental, visto que la primera no cubre todos los tipos de signos en la representación del discurso y la multitud de símbolos, íconos e índices.

El argumento hace creer que se debe reemplazar los objetos de la Lingüística documental y adoptar una concepción más corriente y ajustada a los problemas de la Documentación que, hoy por hoy, exigen planeamientos conceptuales y soluciones aplicadas a los más distintos códigos no verbales. Esta visión de unidad y amplitud de la Semiótica documental es la que ha enunciado Izquierdo Arroyo hace dos décadas.

Bajo estas breves consideraciones, hay que consultarles a los expertos para saber cómo entienden ellos las aportaciones semióticas y su nivel de interdisciplinariedad con la Documentación y, consecuentemente, su relación con los temas de la organización del conocimiento.

#### **4. Los discursos colectivos sobre la concepción semiótica y la interdisciplinariedad**

Bajo la exposición de las dos aportaciones semióticas principales y sus conexiones teóricas, una marcadamente no peirceana y otra de matiz peirceano, hay que referirse a la concepción semiótica y al nivel de interdisciplinariedad que se han encontrado en el análisis de las entrevistas.

Por una parte, en el pensamiento colectivo de la documentación obtenido a través de la construcción discursiva, es notoria la claridad del papel de la semiótica y de sus temas respecto a la Documentación. Según el discurso colectivo, la Semiótica es la ciencia de los signos, de la semiosis, del significado, del lenguaje y de la representación, tal como se ha constado en los manuales de teoría semiótica. A continuación, se presenta como resultado el discurso colectivo de las personas consultadas:

Por Semiótica entiendo que es la ciencia de los signos, es decir, sería la ciencia de representar mediante el lenguaje conceptos o realidades. Entiendo que el objeto de estudio de la Semiótica son los signos, es la representación de los signos. Una ciencia que explica el significado de los sistemas de signos, no exclusivamente lingüísticos. Resulta una disciplina, o sea, un marco disciplinar muy eficaz para comprender de forma integrada la producción de los significados. También, se trata de una ciencia interdisciplinar con conexiones que irían desde la Filosofía a la Teoría de la Comunicación y la Lingüística. Es el análisis de los signos en procesos de semiosis social, si fuera Semiótica, [...] yo asocio Semiótica con Comunicación, si yo entiendo que hay una comunicación documental y es verdad porque igual que hay un hecho semiótico hay un objeto documental que es lo que nos configura como disciplina diferente, tiene que haber una Semiótica Documental también. Es el análisis de los signos, de los símbolos, entendiendo el documento como símbolo, como signo y dentro del documento y del contexto los lenguajes como otro signo, de su naturaleza, de su función y todo en un proceso de semiosis comunicativa, de semiosis documental. Tiene que ver con la comunicación, con los signos, con todo ese tipo de cuestiones y ahí los sistemas tienen una parte importante porque realmente están hechos para eso, están formados por signos, símbolos, sirven para comunicarse y claro que tiene importancia desde ese punto de vista, por supuesto que sí. La semiótica es la explicación de la comunicación, de la interpretación del signo. La interpretación de la realidad a través del lenguaje (y no tanto su significado como sería el objetivo de la Semántica). Es la representación mediante signos de los documentos, bueno de los documentos, digamos del mundo, del entorno en el que nos movemos. Es una definición operativa, pero que explica los códigos de signos, las formas de producción de los sistemas de significación cuando no son lingüísticos. Y que resulta muy operativa para comprender otro tipo de realidades, otro tipo de sistemas de comunicación, de modos, sistemas de comunicación que operan en paralelo a lo lingüístico en todo tipo de situaciones. En nuestro caso, en lo que tiene que ver con la Documentación nos hemos dedicado a la Semiótica en el sentido de cómo representamos la información de una determinada forma. Sin embargo, yo casi prefiero hablar más de Semántica, pero por el texto, porque claro, si vamos a la imagen, el análisis de imagen,

¿cómo se hace? En primer lugar tienes que comprender cómo están estructurados determinados mensajes para que sean entendidos por los usuarios, por los destinatarios de esos mensajes. En la corriente saussuriana el texto constituye una manifestación del lenguaje y por tanto es fundamental para la Semiótica y la Lingüística. Y posteriormente, las posibilidades que aporta después para la recuperación. Yo te hablo desde el punto de vista práctico, de para qué me sirve a mí la semiótica. Yo gracias a la semiótica, puedo emplear los signos para representar los documentos, el contenido de los documentos.

Se ha comprobado que hay un entendimiento del significado de la Semiótica, y que está más allá de la definición denotativa, puesto que se han presentado las cuestiones epistemológicas que le pertenecen. En definitiva, la Semiótica ha sido considerada una disciplina, pero con un matiz de campo interdisciplinar, y entre sus corrientes más importantes se encuentra la saussureana. Adicionalmente, se ha destacado la conexión de la Semiótica con las disciplinas Lingüística y Comunicación, y lo fundamental, el uso que se ha hecho en Documentación de sus conceptos y teorías. Hay que subrayar también el matiz de orientación semántica, puesto que la Semántica debería ser una disciplina que iba concentrar en los problemas del significado.

Por las consideraciones anteriores, el dicho discurso colectivo ha revelado que la comunidad de científicos en Documentación en España que se ha investigado está concienciada sobre el concepto general de semiótica. Probablemente esto se debe al hecho de que poseen un recorrido histórico de análisis teórico y de contribuciones en esa misma línea, lo que convierta la Semiótica en una importante base teórica, aunque no esté claramente presente en las publicaciones recientes y en las asignaturas de muchos cursos de licenciatura/grado en documentación. El discurso obtenido es una pequeña imagen de como han pensado los investigadores en Documentación sobre el concepto de semiótica.

Por otra parte, se ha identificado un discurso que también se hace presente y que ha demostrado el desconocimiento de lo que significa Semiótica, ni de lo que puede aportar a la Documentación, hasta el punto de afirmarse que no lo tiene claro o que no es un asunto de la Documentación en España.

Nosotros aquí no se da semiótica. Es que el tema de documentación, aquí, desde luego, nunca se ha dado por esa vertiente. Solo sé el nombre, no conozco realmente las teorías ni he leído nada de Semiótica. Entonces, no me atrevo a opinar, sinceramente. No soy experta en la materia [...]. No lo tengo nada claro. La semiótica yo no la trabajo. Pues mira, hay profesores muy buenos de se-

miótica, en Ciencias de la Información [...], pero no en Documentación.

Este segundo discurso colectivo denota la inseguridad con relación al tema y aún más, se cuestiona la pertinencia de sus teorías para la Documentación. Se comenta claramente que no es un asunto muy propio a la Documentación, sino más de las Ciencias de la Información, es decir, periodismo, comunicación audiovisual, publicidad y relaciones públicas. Éstos dos discursos colectivos que están ubicados en parte de la comunidad, han impactado en la manera de actuar de los expertos frente al uso de las teorías semióticas en la solución de problemas relacionados con el tratamiento del lenguaje. Los expertos en Documentación habrían de tomar conciencia de los discursos conflictivos que han tenido lugar en la Documentación antes de poner en marcha estrategias de acercamiento a las teorías semióticas, sean peirceanas o no peirceanas.

Con respecto al análisis de los contactos disciplinares, se ha cuestionado al grupo sobre el nivel de interdisciplinariedad entre Semiótica y Documentación. Se ha presentado un discurso colectivo mayoritario y potente que destaca que no han ocurrido contactos interdisciplinarios fuertes, y al mismo tiempo que, cuando han sucedido, se han producido en un nivel muy bajo o elemental. La propia composición del discurso se hace esclarecedora con respecto a eso:

No hay buenos vínculos, está muy dissociado [...]. Yo creo que en general, unos están en una teoría y otros están en lo otro. Es difícil. Yo creo que estamos en un momento, no sé si fuera, pero estamos en un momento de crisis porque va más rápido la aplicación práctica y tecnológica que la cuestión teórica, va muchísimo más rápido. Creo que vivimos de espaldas unos a otros, esa es mi sensación. No sé si en España estamos trabajando de forma interdisciplinar en ese sentido, la verdad es que desconozco si hay algún equipo de trabajo en el que participen lingüistas e investigadores del ámbito de la Documentación [...]. Fíjate que aquí nosotros convivimos [...] en una facultad de letras y que hay filólogos y lingüistas y nunca o prácticamente nunca hemos trabajado en conjunto. Ellos piensan que lo suyo tiene una categoría superior, que esto nuestro es algo aplicado que se resuelve muy fácilmente y no. Los lingüistas, que encima se defienden, no quieren saber nada con la Documentación, y dentro del campo de la Documentación pues hay líneas diferentes. El contacto disciplinario está dirigido de la Documentación hacia la Semiótica y la Lingüística (fundamentalmente por encontrar respuestas a los problemas que plantea la representación lingüística del conocimiento) y no al contrario. Yo no lo veo muy estrecho sinceramente. Yo creo que son, vamos, no te digo divergentes, pero sí paralelos.

Yo no creo que exista, salvo Izquierdo que trabajó más en ese ámbito, en el de la semiótica, intentando aplicarlo a este campo, que yo no sé hasta qué punto lo habrá conseguido. Comunicación, poca, desde mi experiencia poca, porque hay un cierto miedo por parte de las personas de Documentación o una sensación de que se están metiendo en nuestro terreno. Hay muy poca relación interdisciplinar. Quiero decir, muy poca en el sentido de que las personas, los investigadores de un campo trabajen con investigadores de otro campo, o incluso, en el sentido de que investigadores que proceden de un campo determinado, se muevan y acaben aterrizando en otro campo de este tipo. Es decir, yo creo que aquí, las disciplinas están muy parceladas y además en general, muy poco interconectadas entre ellas ¿eh? Hay también razones de tipo burocrático o administrativo ¿no? para eso. Desde mi punto de vista y desde mi experiencia, no solamente en España, sino también en el extranjero, la interdisciplinariedad es siempre un aspecto muy defendido, muy reivindicado por algunas áreas como la terminología y la documentación, que son de naturaleza transdisciplinaria. Teóricamente, una disciplina que se aplica a todas es más fácil que reivindicar también su interdisciplinariedad, esta necesidad de coger elementos teóricos también de otras disciplinas y esto se reconoce y se vive, yo creo, como una cosa de esencia que caracteriza precisamente este tipo de disciplinas. Pero en la práctica, esto es muy difícil. Las relaciones desde mi punto de vista, muy flojas. No, ni más ni menos, ni mejor ni peor, está estancado por completo, o sea, no está. Pero ya digo, yo creo que los autores que se han apartado más a esta parte lingüística y semiótica están muy divorciados de la parte teórica y de la parte práctica. No, cada vez más.

Hay gente que está sin modelos, que se preocupa solo de productos, de procesos. Son más generalidades que otra cosa. Creo que a lo mejor es más desde el punto de vista general de la manera en que otros tratan la semántica también, más generalista. Yo creo que lo que están haciendo es excesivamente básico. El trabajo es muy, muy básico, y centrado en algunos aspectos yo creo que muy menores. Yo creo que muy menores. Yo encuentro que en nuestro campo en España ha habido más teóricos, más profesores con una base fuerte lingüística que han introducido una parte lingüística importante, pero no tanto. Tampoco encontraba aquí a mis teóricos de referencia y creo que eso ha sido muy perjudicial para la práctica porque no se han desarrollado métodos de acuerdo con las nuevas teorías, por lo menos para reconocer hasta dónde llega el método que tú estás enseñando.

Aquí es difícilísimo hacer equipos individuales, es decir, equipos de una sola persona, no vas a ninguna parte. Sí, haces investigación, pero no vas a aplicarla a una auténtica solución, a una auténtica mejora de los procesos y procedimientos en nuestra área. Entonces, está muy claro que tienes que formar un equipo. Pero el corporativismo yo creo que impide que exista una interrelación entre es-

tos colectivos. De hecho yo concretamente, si te soy sincera, no he colaborado con ningún lingüista ni con nadie que se dedique sólo a la semántica. Yo conozco poca gente que esté trabajando, por ejemplo, en elaboración teórica sobre semiótica, sobre representación de conocimiento, etc., etc. Es un trabajo más de tipo práctico, de probar sistemas, adaptar sistemas más que elaboración teórica. O sea, es verdad que yo tengo esa formación y a lo mejor no he necesitado ayuda, pero que no se ha planteado ningún proyecto de investigación ni un trabajo, pero no me suena a mí que haya mucha interacción de personas en documentación que trabajen con personas en los campos de la lingüística en general o la semántica, no lo conozco, no lo sé.

Mi sensación sobre esto es que estamos siempre un poco dándole la vuelta a las mismas cosas. Yo treinta años casi ya, siempre veo que la gente se pregunta lo mismo y no acabo de entenderlo, porque no... El momento, el paradigma es diferente y esto es otra profesión que tiene que tener otros conocimientos y que tiene que saber hacer equis cosas, que la titulación se llamará como se llamará, pero al final la gente tiene que saber lo que tiene que saber. Están muy impulsados por las modas, las coyunturas académicas, las evaluaciones, esto y lo otro, los investigadores vamos pensando más bien en lo que puedo publicar que se ajuste a lo que ahora te están pidiendo y quedarme descansando una vez que lo saque y luego ya veré por dónde tiro, ¿entiendes?

Mejor que cuantificar el discurso es ponerlo de relieve como lo es. Por esta razón, se ha optado por presentarlo como obtenido después del análisis y construcción, sin resumirlo en pocas palabras, y traerlo a los lectores como ha sido presentado de manera viva. El discurso produce un impacto muy distinto si no se adopta una mera representación en una tabla de datos. Es lo que propone la técnica de Lèfevre y Lèfevre (2003), que las representaciones colectivas, las que pueden regular la acción de unos, puedan ser presentadas y decir a los lectores lo que está por detrás de la inteligencia acumulada y de las formas de razonar sobre los problemas de los encuestados.

Como puede observarse, la interdisciplinariedad no sólo con la Semiótica, sino también con respecto a otras disciplinas, no se ha llevado a cabo o se encuentra en una situación demasiado complicada de suceder. No sólo demuestra que hay que ponerse al día con las nuevas teorías, sino que se ha notado que se están arriesgando los contactos disciplinares en general, sea por la estructura académica disponible o por la propia concepción de investigación en Documentación que ha guiado a mucha gente hacia las modas o tendencias temáticas ocasionales.

Esa notoria preocupación con las modas, por ejemplo en organización del conocimiento, no siempre ha revelado el estadio de investigación, ni tampoco las reales necesidades de investigación en un país. De hecho, este fenómeno se agudiza, como se ha comentado en el discurso colectivo, en una situación de crisis económica y falta de inversión en las universidades e investigadores.

Debido a esto, los investigadores habrían sido llevados a publicar, lo que, en teoría, tendría más aceptación en las revistas de impacto y no han invertido tiempo y esfuerzo en lo que, probablemente, le importaría a la realidad social a largo plazo. Esto pone de relieve que los expertos están concienciados de los problemas reales de la interdisciplinariedad en sus universidades y departamentos. También han manifestado que el nivel de contactos personales necesario para garantizar el intercambio de conocimientos es muy bajo, a pesar de que el nivel de los estudios y aplicaciones han logrado un nivel destacable. En el caso de la Semiótica, como le suele pasar, no siempre hay una predilección por sus temas de estudio, así que se quedaría fuertemente fuera del eje de las modas.

## 5. Consideraciones finales

A manera de resumen final, se debe subrayar que las aportaciones antes presentadas, generadas en el contexto de los años ochenta y noventa, en pleno proceso de expansión de los cursos de licenciatura en Documentación en España, han sido fundamentales para la organización del conocimiento en los años que se siguieron, incluso - o tal vez principalmente -, en países latinoamericanos.

De los resultados de la investigación presentados aquí, se deducen las siguientes conclusiones principales. En primer lugar, las aportaciones semióticas que han dado lugar en España tal vez constituyan las contribuciones más originales a los estudios de Documentación en la organización del conocimiento. Desde luego, si no se puede llamarlas de escuela o corriente de la Documentación española, por lo menos hay que pensarlas como una referencia fundamental que ha aportado cuestiones y soluciones conceptuales.

En segundo lugar, se ha notado en el análisis empleado que la concepción semiótica de los expertos demuestra que ellos son conscientes de la noción general de Semiótica y su importancia a la Documentación. Por otra parte, en tercer lugar, el discurso colectivo del "desconocimiento" que también está presente en el pen-

samiento social del grupo investigado, puede servir de impedimento frente a la introducción de los temas y teorías semióticas en la Documentación, especialmente en la organización del conocimiento, que nos parece la principal puerta de entrada de estos abordajes.

En cuarto lugar, hay que tener en cuenta el hecho del problema interdisciplinario. La interdisciplinariedad con la Semiótica es un reto que todavía no se ha superado, y por lo demostrado en el discurso, está en un estadio extremadamente bajo y de difícil superación. Se infiere de lo que se ha analizado la dificultad de establecer contactos colaborativos en el ámbito científico que faciliten el desarrollo del campo. En ese mismo sentido, se está de acuerdo con lo que comentó García Marco (1997, p. 11; 1999), en la segunda mitad de la década de 1990, sobre una clara dificultad de formar grupos de investigación en organización del conocimiento en España.

Por otra parte, el análisis sugiere que tal vez una forma alternativa de entender el impacto de las teorías semióticas en la Documentación en España sea vía un examen cualitativo y diacrónico, más allá de las lecturas sincrónicas que no han explicado la disciplina en el movimiento histórico, y consecuentemente, sus condicionantes políticos, sociales y académicos y, por supuesto, sus contradicciones.

En definitiva, el protagonismo de la documentación en España en el tema de la Semiótica puede que haya sido sustituido por un alineamiento hacia las tendencias temáticas internacionales que genera dudas sobre si están radicadas en las propuestas teóricas autóctonas de la Documentación, y desde luego, de la organización del conocimiento.

Por fin, parece esclarecedora la reflexión de Guy sobre la filosofía española, cuando se refiere a la cita de Castillejo. Lo que comenta el autor presenta una cierta analogía con el caso de la Documentación. Según Guy:

[...] hay que reconocer con Castillejo que, a consecuencia de una acumulación de desgraciados avatares circunstanciales, a menudo contrarios a la expansión del pensamiento libre, un gran número de instituciones originales y de innovaciones mentales surgidas en España no se han podido desarrollar normalmente y con continuidad; así, éstas han pasado al extranjero, de donde con frecuencia han vuelto posteriormente a sus orígenes hispánicos primeros, pero bajo una forma más acabada. (Guy, 1985, p. 10).

Con efecto, se puede afirmar, respetando y haciendo los matices necesarios, que lo mismo podría llegar a pasar con el tema de los estudios

semióticos y lingüísticos en la Documentación en España. Guy (1985) hizo el comentario con respecto a las diversas corrientes innovadoras de la filosofía occidental que tuvieron su origen distante en España, tales como: el socialismo cristiano de Arnau de Vilanova, la lógica simbólica de Llull, el pacifismo de Vives, la ética internacional de Vitoria y Suárez, el existencialismo trágico de Unamuno, y muchos otros.

Tales corrientes filosóficas volvieron a España, bajo una forma más sistematizada, construida fundamentalmente en el extranjero. En concreto, se espera que no pase así con las aportaciones semióticas en Documentación que, venidas desde el extranjero, se presentarán como más teóricas y más sistematizadas, cuando, en realidad, han tenido lugar en la Documentación española.

## 6. Referencias

- García Gutiérrez, A. L. (1984). *Lingüística documental: aplicación a la documentación de la comunicación social*. Barcelona: Mitre D.L.
- García Gutiérrez, A. L. (1990). *Estructura lingüística de la documentación*. Murcia: Universidad de Murcia.
- García Marco, F. J. (1997). Avances en organización del conocimiento en España: los II encuentros sobre organización del conocimiento en los sistemas de información y documentación. En: García Marco, F. J. Coord. (1997). *Organización del conocimiento en sistemas de información y documentación: actas del II encuentro de ISKO-España, Getafe (Madrid), 16 y 17 de noviembre de 1996*. Zaragoza: Universidad de Zaragoza. 7-12.
- García Marco, F. J. (1999). Casi una década de la Sociedad Internacional para la Organización del Conocimiento en España. En: García Marco, F. J. (1999). *Organización del conocimiento en sistemas de información y documentación: actas del III encuentro de ISKO-España, Getafe (Madrid), 19 a 21 de noviembre de 1997*. Zaragoza: Universidad de Zaragoza. 7-21.
- Guy, A. (1985). *Historia de la filosofía española*. Barcelona: Anthropos Editorial del Hombre.
- Izquierdo Alonso, M. (2000). Nuevos enfoques en el estudio del tratamiento documental de contenido desde los presupuestos de las ciencias del lenguaje. *Scire*. 6 (1), 143-163.
- Izquierdo Alonso, M. (2004). Nuevos retos en el análisis documental de contenido: gestión de la forma documental del contenido. *Scire*. 10 (1), 31-50.
- Izquierdo Alonso, M.; Izquierdo Arroyo, J. M. (2014). Entrevista a José María Izquierdo Arroyo, realizada por Mónica Izquierdo Alonso em Alcalá de Henares (Madrid), junio de 2014. *Scire*. 20 (1), 109-116.
- Izquierdo Arroyo, J. M. (1990). *Esquemas de lingüística documental*. Barcelona: Promociones y Publicaciones Universitarias. 3 Tomos.
- Izquierdo Arroyo, J. M. (1993). De la semiótica del discurso a la semiótica documental. En: Moreno González, J. A. (1993). *Aplicación de las ciencias del texto al resumen documental*. Madrid: Universidad Carlos III. 199-216.
- Izquierdo Arroyo, J. M. (1990). La ciencia de la búsqueda documental secundaria. *Documentación de las Ciencias de la Información*. 13, 87-111.
- Izquierdo Arroyo, J. M. (1992). Cuatro trabajos en curso. *Documentación de las Ciencias de la Información*. 15, 35-65.
- Izquierdo Arroyo, J. M. (1995). *La organización documental del conocimiento*. Madrid: Tecnidoc.
- Izquierdo Arroyo, J. M. (2014). *Entrevista a José María Izquierdo Arroyo*. Realizada en Alcalá de Henares (Madrid), noviembre de 2014.
- Lara, M. L. G. ; Tálamo, M. F. G. M. (2007). Uma experiência na interface Linguística Documentária e Terminologia. *DataGramaZero*. 8 (5). Recuperado 03-04-2013 de [http://www.dgz.org.br/out07/Art\\_01.htm](http://www.dgz.org.br/out07/Art_01.htm)
- Lefèvre, F.; Lefèvre, A. M. C. (2003). *O discurso do sujeito coletivo: um novo enfoque em pesquisa qualitativa (desdobramentos)*. Caxias do Sul: UDUCS.
- Lopez Yepes, J. (1995). *La documentación como disciplina: teoría e historia*. Pamplona: EUNSA.
- López Yepes, J.; Ros García, J. (1993). *¿Qué es documentación?: teoría e historia del concepto en España*. Madrid: Editorial Síntesis, S. A.
- Moreiro González, J. A. (1993). *Aplicación de las ciencias del texto al resumen documental*. Madrid: Universidad Carlos III De Madrid; Boletín Oficial Del Estado.
- Moreiro González, J. A. (2011). *Linguagens documentárias e vocabulários semânticos para a Web: elementos conceituais*. Salvador: UFBA.
- Moreno Fernández, L.M.; Izquierdo Alonso, M. (2014). El pensamiento de José María Izquierdo Arroyo em la organización y representación del conocimiento: una sistematización desde fundamentos filosóficos y semióticos. *Scire*. 20 (1), 21-32.
- Tálamo, M.F.G.M.; Lara, M.L.G. (2006). O campo da Linguística Documentária. *Transinformação*. 18 (3), 203-211.

---

# Referentes Teóricos em Organização do Conhecimento: uma Análise de Domínio na ISKO Espanha (2005-2013)

*Theoretical references in Knowledge Organization: a Domain Analysis in ISKO Spain (2005-2013)*

---

**Renata Cristina Gutierrez Castanha (1), Maria Cláudia Cabrini Grácio (2), Suellen Oliveira Milani (3)**

(1) Universidade Estadual Paulista “Júlio de Mesquita Filho”, Av. Hygino Muzzi Filho, 737 – 17525-000 – Marília, São Paulo, Brasil, regutierrez@gmail.com

(2) cabrini@marilia.unesp.br (3) suellenmilani@marilia.unesp.br

## Resumo

A Organização do Conhecimento (OC) configura-se como um campo interdisciplinar que oferece aportes teóricos e metodológicos para a Biblioteconomia e Ciência da Informação, mas seu conceito e escopo ainda encontram-se em construção. Considerando a contribuição da caracterização epistemológica para se evidenciar as interlocuções teóricas do domínio da OC, este artigo tem como objetivo analisar os últimos cinco Congressos da ISKO Espanha, ocorridos entre 2005 e 2013, por meio da Análise de Domínio. Para tanto, propõem-se identificar os autores com a presença mais consolidada na ISKO Espanha e identificar as relações entre os autores citantes e seus referentes teóricos, assim como as cocitações estabelecidas por esta comunidade. Finalizando, objetiva-se identificar os temas mais recorrentes, por meio da análise das palavras-chave. Identificaram-se 16 investigadores mais produtivos, dos quais destacam-se: J. A. C. Guimarães, M. C. A. Lacruz, L. S. Bufrem e M. S. L. Fujita. A análise dos termos mais recorrentes no conjunto de 67 trabalhos publicados pelos autores mais citados demonstra que 33% dos termos dizem respeito aos processos de representação e recuperação da informação, ao passo que 25% das palavras-chave são compostas pelos termos gerais organização do conhecimento e organização da informação. Evidencia-se que estudos sobre os instrumentos de representação da informação aparecem em 17% das palavras-chave analisadas, reafirmando a própria trajetória do evento. Como resultados da rede de citação, destaca-se o conjunto dos autores mais citados pelos autores mais produtivos: Hjørland, citado por cinco investigadores diferentes, Dahlberg, citada por quatro investigadores, e Guimarães, López-Huertas, Mai e Beghtol, citados por dois pesquisadores diferentes. A rede de cocitação evidenciou que a maior frequência de cocitação ocorreu entre Dahlberg, Hjørland e Olson, seguida de Beghtol, Guimarães e López-Huertas. Conclui-se que há um consistente núcleo de referentes teóricos que frequentam os eventos do capítulo espanhol da ISKO, assim como percebe-se uma interlocução entre os autores citados por esse núcleo.

**Palavras-chave:** Organização do Conhecimento, Análise de Domínio, Análise de citação, Análise de cocitação, Isko Espanha.

## Abstract

Knowledge Organization (KO) is configured as an interdisciplinary field that offers theoretical and methodological contributions to the Library and Information Science, but its concept and scope are still under construction. Regarding the contribution of epistemological characterization to evidence the theoretical dialogues in KO domain, this article aims to analyze the last five ISKO Spain Conferences, that occurred between 2005 and 2013, through Domain Analysis. Therefore, we aim to identify the authors with more consolidated presence in ISKO Spain and identify the relationships between citing authors and their related theoretical as well as established co-citations by this community. Finally, we aim to identify the most recurrent themes through keywords analysis. We identified 16 most productive researchers, which includes J. A. C. Guimarães, M. C. A. Lacruz, L. S. Bufrem and Fujita M. S. L. The analysis of the most frequent terms in the set of 67 papers published by most cited authors shows that 33% of the terms relate to the processes of information representation and retrieval, while 25% of the keywords are composed of the general knowledge organization and information organization. It is evident that studies on the process of information representation appear in 17% of the keywords analyzed, reaffirming the trajectory of the Conferences. As the citation network results, there is the set of authors most cited by the most productive authors: Hjørland, cited by five different researchers, Dahlberg, cited by four researchers, and Guimarães, López-Huertas, Mai and Beghtol cited by two different researchers. The co-citation network showed that most frequently co-citation occurred between Dahlberg, Hjørland and Olson, followed by Beghtol, Guimarães and López-Huertas. We conclude that there is a consistent core of theoretical referenced who attend the events of the Spanish chapter of ISKO, as we can see a dialogue between the authors cited by this core.

**Keywords:** Knowledge Organization, Domain Analysis, Citation Analysis, Cocitation Analysis, ISKO Spain.

## 1. Introdução

Considerando que a Biblioteconomia e Ciência da Informação atuam no intuito de promover o acesso à informação construindo caminhos para que as pessoas encontrem essa informação registrada, destaca-se o aporte teórico oferecido pela organização do conhecimento.

Na prática profissional, catalogadores, classificadores e indexadores lidam com esquemas de seleção de discursos, análises, processos de decodificação, representação e acesso à informação registrada e as decisões que eles tomarem “darão forma” à biblioteca que representam e serão bem recebidas ou não pelas comunidades discursivas que são genericamente denominadas usuários.

As comunidades usuárias interagem no contexto de domínios do conhecimento. Um domínio de conhecimento é “[...] um espaço social de produção e circulação de conhecimento, uma cultura particular, dotada de lógica própria de ordenação dos discursos [...]” (Guimarães et al., 2009, p. 109). Um domínio é composto por instituições que dão a ele a sua forma como: organizações profissionais, eventos, revistas científicas, frentes de pesquisa e todas as manifestações de poder, políticas ou culturais que essas instituições possuem. Essas instituições influenciarão diretamente a forma, as ações, e os fluxos desse domínio. O domínio atua, então, não apenas oferecendo instrumentalidade para o mapeamento de um campo, suas disciplinas e subdisciplinas, mas, também, como reflexo de uma comunidade discursiva.

Assim, os domínios envolvem diretamente a epistemologia elaborada, inclusive, a partir da literatura gerada por seus especialistas. Na organização do conhecimento, estudos nesse sentido foram realizados por Travieso-Rodríguez, C. et al. (2011), Smiraglia (2011). Arboit et al (2012), Guimarães, Oliveira e Grácio (2012), Souza (2013), entre outros.

O termo Organização do Conhecimento (OC) foi estabelecido por volta de 1900 por autores como Charles Ami Cutter e Ernest Cushingon Richardson e consolidado por meio de autores como W. C. Berwick Sayers e Henry Bliss (Hjørland, 2008).

O livro *The organization of knowledge and the system of the sciences*, de Henry Bliss (1929) “[...] representa uma das principais contribuições intelectuais no campo” (Hjørland, 2008, p. 97).

Quando se pensa na organização do conhecimento no contexto da Biblioteconomia e Ciência da Informação em um âmbito teórico, “[...] tem-

se a necessidade de sistematização e consolidação de um conhecimento (conjunto de saberes) verificável em uma dada sociedade em um dado momento histórico, com o objetivo de *transmissão*” (Guimarães, 2000, p. 208). Nesse sentido, o autor destaca os trabalhos dos filósofos ligados à Teoria do Conhecimento, tais como Aristóteles, Platão, Porfírio, Bacon e Harris, e educadores, tal como Comênio.

Por outro lado, tem-se, em um âmbito mais pragmático, “[...] a necessidade de resgate do conhecimento registrado em documentos, visando ao seu acesso, com um objetivo de recuperação” (Guimarães, 2000, p. 208), destacando os trabalhos de documentalistas, tais como Calímaco, Otlet e La Fontaine.

Quando se discute a Organização do Conhecimento no contexto da Biblioteconomia e Ciência da Informação, geralmente, busca-se um aporte teórico que o ampare as questões relacionadas ao controle bibliográfico, que seria esse contexto mais pragmático observado por Guimarães (2000).

Howarth (2005, p. 39) ensina que:

Dedicadas à criação de substitutos – ou registros bibliográficos – para representar itens e objetos reais – e, mais recentemente, itens ou objetos virtuais, a teoria e a prática do controle bibliográfico têm adotado abordagens sistemáticas, uniformes e consistentes para descrever o conteúdo intelectual ou artístico e as características físicas desses itens e objetos.

Em outras palavras, o controle bibliográfico é “[...] o processo sistemático de criação de substitutos que ocupem o lugar dos itens que de fato contêm informação” (Olson, 2002, p. 100).

Os bibliotecários realizam o controle bibliográfico por meio dos processos de catalogação e representação de assunto e utilizam instrumentos, tais como códigos de catalogação, esquemas de classificação, lista de cabeçalhos de assunto e tesouros, para construir os registros catalográficos, índices, resumos e notações de classificação.

A organização do conhecimento enquanto disciplina trata-se de um universo que

[...] procura, então, apresentar um modelo conceitual adequado às diversas práticas e atividades sociais vinculadas com o acesso do conhecimento, e pretende operar como instrumento de tratamento da informação, e da gestão de uso da informação, abarcador e integrador dos fenômenos e das aplicações vinculados à estruturação, à disposição, ao acesso e à difusão do conhecimento socializado (Barité, 2001, p. 39-40, tradução nossa).

Esse universo é amparado pela ISKO, primeira sociedade acadêmica internacional dedicada à teoria e à prática da Organização do Conhecimento.

Considerando que a caracterização epistemológica contribui para se evidenciar as interlocuções teóricas do domínio estudado, este artigo tem como objetivo analisar os últimos cinco Congressos da ISKO Espanha, mais especificamente os eventos que ocorreram entre 2005 e 2013 por meio de uma Análise de Domínio (Hjørland, 2002) ressaltando as relações entre os autores citantes e os autores citados e as citações estabelecidas por esta comunidade.

Considera-se que o estudo proposto traz elementos para a consolidação do próprio campo de organização do conhecimento, considerado aqui como um domínio do conhecimento. Para tanto, propõem-se identificar os investigadores com a presença mais consolidada no capítulo espanhol da International Society for Knowledge Organization (ISKO) nos últimos cinco anos.

## 2. Metodologia

Com o objetivo de identificar os referentes teóricos presentes nos Anais dos últimos cinco eventos do capítulo espanhol da ISKO, esse estudo configura-se como exploratório e bibliográfico. Os resultados serão analisados sob a ótica da Análise de Domínio.

Entre as abordagens para a análise de domínio, optou-se pelos estudos bibliométricos, em especial pelos estudos de citação e cocitação, os quais destacam-se como procedimentos adequados para a visualização da estrutura cognitiva e social de um domínio (Hjørland, 2002).

A frequência de citações de autores ou documentos e a coocorrência destes viabilizará, nesse contexto, que o domínio seja investigado. A análise de citações objetiva a visualização de um campo científico enquanto um domínio, por meio da forma principal com a qual a comunidade desse domínio se legitima – as citações. (Piovezan e Fujita, 2014).

A frequência com que dois autores são citados simultaneamente em produções científicas posteriores é captada pela análise de cocitação de autores, e, neste contexto, a citação consiste medida da efetivação da comunicação científica (Oliveira e Grácio, 2013).

### 2.1. Corpus da Análise

O universo escolhido para esta análise foi o capítulo espanhol da ISKO uma vez que essa é a primeira sociedade acadêmica internacional

dedicada à teoria e à prática da organização do conhecimento.

McIlwaine e Mitchell (2008) sinalizam que apesar de a maioria dos membros da ISKO advirem da área da Biblioteconomia e Ciência da Informação, a organização do conhecimento não se limita a esse campo e destacam as contribuições interdisciplinares e a confusão sobre o significado de organização do conhecimento e sua relação com outros campos como gestão do conhecimento, por exemplo.

A ISKO foi criada no dia 22 de julho de 1989, em Frankfurt, Alemanha, com a missão de unir nacional e internacionalmente membros pessoais e institucionais interessados na pesquisa, desenvolvimento e aplicação de todos os métodos de organização do conhecimento em campos gerais ou particulares, integrando, especialmente, a aproximação das pesquisas em classificação e inteligência artificial.

Visando cumprir o proposto, a ISKO tem o objetivo de: a) promover a pesquisa, desenvolvimento e aplicações de sistemas de organização do conhecimento no que diz respeito ao avanço de abordagens filosóficas, psicológicas e semânticas de ordenação do conhecimento; b) fornecer aos seus membros meios e redes de comunicação acerca da organização do conhecimento; c) funcionar como um elo entre as instituições e sociedades nacionais, trabalhando com problemas relacionados à organização conceitual e ao processamento do conhecimento.

Desse modo, a ISKO tem atuado como um fórum internacional em um âmbito amplo e interdisciplinar, contribuindo para o avanço da produção literária em OC. Como atividades mais importantes, destacam-se: conferências internacionais a cada dois anos; conferências nacionais e regionais em tópicos especiais; publicações de caráter científico na *Knowledge Organization*; um boletim informativo incorporado a *Knowledge Organization* intitulado *ISKO News*; e as séries *Advances in Knowledge Organization* e *Knowledge Organization in Subject Areas*. Entre os capítulos regionais, destaca-se a ISKO Espanha, com seu primeiro congresso realizado em 1993 em Madri. A partir de 2013, a ISKO Espanha inclui Portugal a esse capítulo e torna-se, a partir daí, ISKO Espanha-Portugal, ou ISKO Ibérico. Nesse mesmo ano, a conferência foi realizada na cidade do Porto, em Portugal, consolidando a parceria dos países.

### 2.2. Métodos

O foco desta análise parte do investigador para o universo da ISKO Espanha por meio de suas

citações. Para tanto, foram elencadas todas as referências utilizadas nos 229 trabalhos publicados nos Anais dos seguintes eventos: Barcelona (2005), León (2007), Valencia (2009), Ferrol (2011) e Oporto (2013).

Sabe-se dos estudos e controvérsias acerca do ato não neutro de citar, conforme expõem Vanz e Caregnato (2003), mas a presente análise considera a lista de referências de um trabalho científico como uma descrição explícita do conjunto de trabalhos, autores, instituições e eventos que influenciaram o autor, ou seja, uma descrição de um domínio do conhecimento delimitado por um autor especificamente.

Para identificar os autores mais produtivos no universo analisado, valeu-se da lei do elitismo de Price.

Para Price, a distribuição da produtividade dos autores em uma coordenada cartesiana é tão inclinada que o inspirou a propor a Lei do Elitismo. Esta lei consiste na raiz quadrada do total de autores, a qual representa a elite do domínio em estudo. A Lei do Elitismo apresenta aplicações e repercussões imediatamente eficazes para a política científica de um país, sendo responsável por grande parte da quantificação da produtividade científica, que é normalmente mensurada em termos de trabalhos publicados, dos quais se originam os cálculos necessários para a medição desta produtividade dos autores (Alvarado, 2009).

Uma vez que a elite científica do domínio analisado foi identificada como composta por 16 investigadores, recuperaram-se 67 artigos proeminentes dessa elite e analisaram-se suas principais referências bibliográficas, por meio da análise de redes sociais, construindo as seguintes redes: autores citantes-citados e cocitação.

No âmbito dos estudos métricos, as redes sociais constituem elementos de visualização gráfica para que se possa melhor compreender a configuração de um dado domínio científico (Otte e Rousseau, 2002).

Em relação à análise de citação, representada aqui pela rede citante-citado, esta tem por finalidade identificar os grupos de cientistas, suas publicações, e evidenciar os pesquisadores de maior impacto de uma determinada área; indica seus paradigmas, procedimentos metodológicos pertinentes, bem como sua frente de pesquisa (Glänzel, 2003). Ainda, as citações definem o domínio (Smiraglia, 2011), e a análise de cocitação, produz representações válidas da estrutura intelectual de um domínio científico (Miguel, Moya-Anegón e Herrero-Solana, 2008).

Para compor a rede citante-citado, foram considerados os investigadores mais produtivos e os autores citados pelo menos três vezes pelo investigador (citante). Salienta-se que foram eliminadas as autocitações.

A rede de cocitação foi construída a partir da ocorrência simultânea de citação a dois autores, dentro da produção científica de cada investigador mais produtivo. No que se refere à cocitação, esta tem a seguinte premissa: quando dois documentos ou autores são citados juntos em uma literatura posterior, há uma proximidade de assunto entre os citados, na perspectiva dos citantes. Assim, quanto maior a incidência de cocitação, maior a similaridade entre os citados.

Os estudos de cocitação avaliam a interlocução entre pesquisadores, constituindo instrumento significativo para a visualização do processo de comunicação científica, a fim de explicar a estrutura social de um domínio do conhecimento. Ainda, a análise do conjunto de referências de trabalhos científicos, representa reflexo de uma comunidade discursiva, destacando os autores reconhecidos por tal comunidade, retratando características próprias desse domínio (Grácio e Oliveira, 2014).

Para a construção da rede citante-citado foi utilizado o software Pajek, pela maior facilidade para construção de redes *2-mode* (citante, citado) quando alguns autores são simultaneamente citante e citados, sem duplicá-los na figura da rede, e para a rede de cocitação foi usado o software Ucinet, pela simplicidade operacional quando se trabalha com redes *1-mode*, como as redes de cocitação, com universos pequenos.

Em um segundo momento, realizou-se uma análise das palavras-chave em inglês presentes nos 67 trabalhos selecionados a fim de identificar as temáticas mais recorrentes neste domínio. É importante dizer que alguns trabalhos não continham as palavras-chave em inglês, as *keywords*, então realizaram-se traduções para que o controle terminológico e as aproximações temáticas pudessem ser realizadas.

### 3. Resultados

Dos 229 trabalhos publicados nos Anais dos últimos cinco eventos da ISKO Espanha, 67 (29,3%) trabalhos deles foram publicados pelos 16 investigadores mais produtivos, excluídas as duplicações de contagem decorrentes das co-autoria entre eles, evidenciando a forte presença desses investigadores neste evento nas últimas edições. A Tabela 1 apresenta os 16 investigadores mais produtivos, com respectivos

totais de trabalhos publicados no período analisado.

| <i>Autores</i>                  | <i>Trabalhos</i> |
|---------------------------------|------------------|
| José Augusto Chaves Guimarães   | 10               |
| María del Carmen Agustín Lacruz | 8                |
| Leilah Santiago Bufrem          | 6                |
| Mariângela Spotti Lopes Fujita  | 6                |
| Carlos Cândido de Almeida       | 5                |
| Evelyn Goyannes Dill Orrico     | 5                |
| Jesús Tramullas Saz             | 5                |
| José Antonio Moreira González   | 5                |
| Marilda Lopes Ginez de Lara     | 5                |
| Aida Varela                     | 4                |
| Dulce Amelia de Brito Neves     | 4                |
| Juan Carlos Fernández Molina    | 4                |
| Luis Rodríguez Yunta            | 4                |
| Piedad Garrido Picazo           | 4                |
| Rosa San Segundo Manuel         | 4                |
| Thiago Henrique Bragato Barros  | 4                |

Tabela I. *Autores mais produtivos*

Destaca-se o brasileiro J. A. C. Guimarães como o autor mais produtivo nas últimas cinco edições do ISKO Espanha, com 10 trabalhos publicados. Em seguida, aparecem a espanhola M. C. A. Lacruz, com 8 trabalhos publicados, e as brasileiras L. S. Bufrem e M. S. L. Fujita, com 6 trabalhos cada uma.

Ao realizarem uma análise bibliométrica sobre toda a trajetória do ISKO Espanha (1993-2009), Travieso-Rodríguez, C. et al. (2011) já sinalizavam para a forte presença das instituições brasileiras nas edições deste evento, uma vez que entre os 17 mais produtivos, 4 eram brasileiros.

No período analisado (2003 a 2012), observa-se, entre os 16 mais produtivos no evento, a presença de 9 brasileiros, indicando uma tendência de crescimento da participação brasileira no evento.

É importante destacar que os referidos autores são docentes de instituições brasileiras e espanholas e investigam, principalmente, temáticas relacionadas ao Tratamento Temático da Informação e a Organização do Conhecimento como disciplina, perfil também percebido por Travieso-Rodríguez, C. et al. (2011). A professora e pesquisadora L. S. Bufrem aprofunda-se, ainda, em questões métricas nesse contexto.

Os autores C. C. Almeida (Brasil), E. G. D. Orrico (Brasil), J. Tramullas Saz (Espanha), J. A. Moreira González (Espanha) e M. L. G. Lara (Brasil) apresentaram cinco trabalhos cada um.

Por fim, os pesquisadores A. Varela (Brasil), D. A. B. Neves (Brasil), J. C. Fernández Molina (Espanha), L. Rodríguez Yunta (Espanha), P. G. Picazo (Espanha), R. San Segundo Manuel (Espanha) e T. H. B. Barros (Brasil), totalizaram quatro trabalhos cada.

Ao elencar as palavras-chave utilizadas nos 67 trabalhos analisados, chegou-se a um conjunto de 303 palavras-chave. Em um segundo momento, realizou-se um exercício de aproximação temática que resultou em um conjunto de 16 palavras-chave que apareceram três ou mais vezes. O conjunto apresentado no gráfico a seguir (Figura 1, Anexo 1) expõe as 16 palavras-chaves predominantes e as suas 82 incidências.

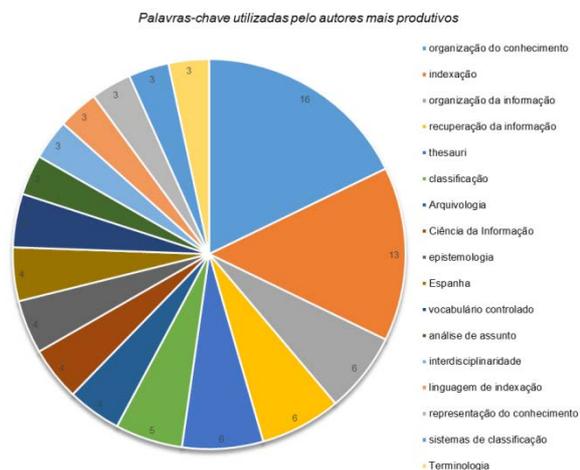


Figura 1. *Palavras-chave utilizadas pelos autores mais produtivos*

O conjunto de palavras-chave elencadas possui um alto grau de especificidade, mas ao analisarmos esse conjunto no contexto no qual ele foi publicado, isto é, os Anais do ISKO Espanha, conclui-se que as palavras-chave poderiam apresentar maior exaustividade. Isso fica evidente quando se verifica que 25% das palavras-chave são compostas pelos termos *organização do conhecimento* (18%) e *organização da informação* (7%).

Com 33% das incidências, destaca-se a presença dos processos meio e fim da Organização do Conhecimento, respectivamente: a *representação do conhecimento* (3%), a *indexação* (14%) e a *classificação* (6%) e a *recuperação da informação* (7%). Não identificaram-se palavras-chave que mencionassem o processo inicial de produção da informação.

Os instrumentos de representação temática da informação sob o temo geral *vocabulário controlado* (4%) e *linguagens de indexação* (3%) e sob os termos específicos *tesauros* (7%) e *sistemas de classificação* (3%) integram 17% das palavras-chave identificadas.

Por fim, destaca-se a presença de reflexões acerca da *Arquivologia* que aqui está representada por apenas 4% dos termos, mas que no conjunto das 303 palavras-chave aparece em outras situações como, por exemplo: arquivos, estudos arquivísticos, manuais de Arquivologia etc.

A palavra-chave *Espanha* advém dos trabalhos que utilizam a abordagem dos estudos métricos presentes nos eventos e o termo *epistemologia* salienta o caráter mais teórico dos eventos do capítulo espanhol da ISKO.

O termo geral *Ciência da Informação* aparece em 4% dos trabalhos e o termo *interdisciplinaridade* em 3%. A área interdisciplinar *Terminologia* também apareceu três vezes no conjunto de palavras-chave analisado.

Com o objetivo de identificar as interlocuções teóricas dos autores mais produtivos nas temáticas mencionadas, extraíram-se as referências dos 67 trabalhos analisados e construiu-se a rede de citação a seguir (Figura 2, Anexo 2).

Os autores mais produtivos (citantes) são apresentados com círculos vermelhos, enquanto os círculos azuis apresenta os autores que compõem as referências (citados) e os autores destacados com círculos verdes são, simultaneamente, citantes e citados.

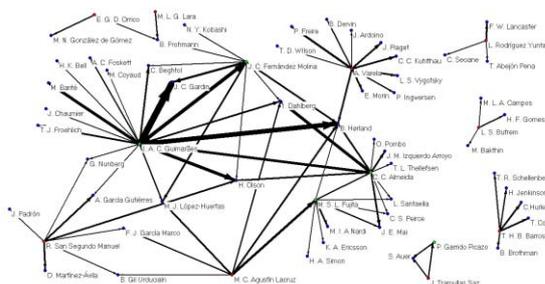


Figura 2. Rede citante-citado entre os autores mais produtivos e seus referentes.

No conjunto dos autores mais citados pelos investigadores mais produtivos, destacam-se: Hjørland, citado por cinco investigadores diferentes, Dahlberg, citada por quatro investigadores, e Guimarães, López-Huertas, Mai e Beghtol, citados por dois pesquisadores diferentes. Os autores mais citados advém de diferentes países, a saber: Dinamarca, Alemanha, Brasil,

Espanha e Canadá, o que caracteriza uma internacionalização dos referentes teóricos.

Os cinco autores que são simultaneamente citantes e citados são: J.A.C. Guimarães, J.C. Fernández Molina, M.S. L. Fujita, C.C. Almeida e P. Garrido Picazo.

Observa-se, na Figura 2 (Anexo 2), a formação de cinco componentes, sendo quatro menores, centrados nos seguintes investigadores citantes: Rodríguez Yunta, o qual trabalha, principalmente na Recuperação da Informação, tema amparado na literatura por seus citados F. W. Lancaster, C. Seoane e T. Abejón Pena.

O próximo componente centra-se em Bufrem, que utiliza, majoritariamente, referenciais brasileiros, como Gomes e Campos, as quais são, inclusive, coautoras em vários artigos da área da Ciência da Informação. Bufrem tem como foco de sua pesquisa os seguintes temas: Organização e Representação do Conhecimento, mais especificamente a Análise de Domínio, Metodologia Científica e Epistemologia, temas que são amparados na literatura por seus citados Gomes e Campos. A autora também trabalha com a teoria do Círculo de Bakhtin aplicada a conceitos relacionados à área de Organização do Conhecimento, daí sua citação específica para esse autor.

Barros se encontra distante do núcleo do componente principal, fato que poderia ser explicado por sua temática de trabalho, a qual foca na área da Arquivística e Análise do Discurso, temas que são amparados na literatura por seus citados, B. Brothman, T. Cook, C. Hurley, H. Jenkinson e T. R. Schellenberg, os quais são pesquisadores atuantes na área do autor.

Por fim, Tramullas Saz e Garrido Picazo, ambos citando o autor Auer. Os autores tem como foco de pesquisa a Recuperação e Sistemas da informação. Destaca-se que Garrido Picazo (círculo verde) é tanto citado quanto coautor de Tramullas Saz.

O quinto e maior componente ampara o restante dos autores mais produtivos, e apresenta ao mesmo tempo, autores citantes e citados.

Destaca-se o forte apoio para fundamentação teórica de Guimarães, com intensa citação a Gardin, Hjørland e também a Olson, Barité e Fernández- Molina. Ainda, Guimarães e Fernández-Molina compartilham de múltiplos autores referenciais, como Beghtol, Olson, Gardin, Dahlberg e Hjørland devido à parceria acadêmica destes; inclusive, citam um ao outro.

San Segundo Manuel e Guimarães compartilham citações a três autores: Nunberg, Garcia-Gutierrez e Olson.

Ainda, Agustín Lacruz e Guimarães amparam suas pesquisas em três autores em comum, a saber: López-Huertas, Dahlberg e Hjørland.

Guimarães possui também a maior amplitude de referencial, composto por 13 autores citados, no mínimo, três vezes.

Observa-se ainda que Varela tem uma significativa amplitude de referencial não compartilhada pelos demais investigadores citantes desse componente e possui um único autor citado em comum com os demais citantes da rede, a saber, Hjørland.

Destaca-se também o compartilhamento da fundamentação teórica de Fujita e Almeida, por meio de coautoría e citação dos seguintes autores: Santaella, Peirce e Mai.

A Figura 3 (Anexo 3) apresenta a rede de cocitação entre os autores mais citados pelos investigadores mais produtivos das últimas cinco edições da ISKO Espanha.

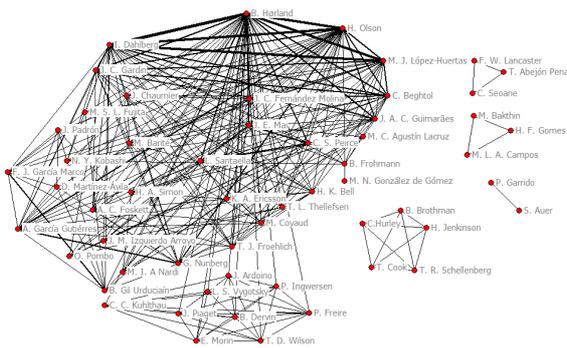


Figura 3. Rede de cocitação entre os referentes mais citados pelos autores mais produtivos.

Em consequência da rede de citação (Figura 2, Anexo 2), a rede de cocitação também é formada por cinco componentes, sendo o maior deles composto por 47 autores cocitados.

Identificam-se ainda, três cliques. A primeira clique é composta por cinco citados: Cook, Jenkinson, Brothman, Hurley e Schellenberg, citados pelo investigador arquivista Barros. A clique seguinte é composta por três pesquisadores, a saber: Seoane, Lancaster e Abejón Pena, citados por Rodriguez Yunta.

Por fim, tem-se a clique formada por Campos, Gomes e Bakhtin, recorrente de Bufrem.

Observa-se a ligação de Auer e Garrido, utilizados por Tramullas Saz.

O maior componente demonstra que a maior frequência de cocitação ocorreu entre Dahlberg, Hjørland e Olson, seguida de Beghtol, Guimarães e López-Huertas, comprovando a importância teórica destes autores para a área de organização do conhecimento no universo da ISKO Espanha.

#### 4. Considerações Finais

Há um consistente núcleo de referentes teóricos que frequenta os eventos do capítulo espanhol da ISKO, assim como observa-se uma intensa interlocução entre os autores citados por esse núcleo, o que indica o efetivo papel realizado pela ISKO Espanha como um centro de construção teórica no campo da organização do conhecimento.

Os autores mais produtivos na ISKO Espanha são docentes e pesquisadores, entre os quais destacam-se o brasileiro J. A. C. Guimarães, a espanhola M. C. A. Lacruz e as brasileiras L. S. Bufrem e M. S. L. Fujita. É interessante observar a forte presença de autores brasileiros no evento, o que demonstra uma consolidada interlocução internacional entre pesquisadores e instituições dos países em questão.

Os autores mais cocitados são o trio Dahlberg, Hjørland e Olson, seguido pelo trio Beghtol, Guimarães e López-Huertas, os quais representam importantes referentes teóricos para a área de organização do conhecimento, especificamente no que se diz respeito ao universo da ISKO Espanha.

Consideram-se as análises das palavras-chave, de citação e de cocitação, conjuntamente, abordagens eficientes para a visualização e compreensão do domínio científico analisado, especificamente no que se diz respeito à dinâmica social do processo de comunicação científica.

Ressalta-se que, em continuidade a esta pesquisa, análises mais aprofundadas poderão ser feitas por meio de métodos de análises de conteúdo dos artigos.

#### Referências

- Alvarado, R. U. (2009). Elitismo na literatura sobre a produtividade dos autores. *Ciência da Informação*, 38:2 (2009) 69-79.
- Arboit, A. et al (2012). The relationship between authors and main thematic categories in the field of knowledge organization: a bibliometric approach. In: Neelamenghan, A.; Raghavan, K. S. (Org.). *Categories, contexts and relations in knowledge organization*. Würzburg: Ergon Verlag, 1 (2012) 44-50.

- Barité, M. (2001). Organización del conocimiento: un nuevo marco teórico-conceptual en Bibliotecología y Documentación. En Carrara, K. (org.). (2001) *Educação, universidade e pesquisa*. São Paulo: FAPESP, 35-60.
- Esteban Navarro, M. A.; García Marco, F. J. (1995). Las "primeras jornadas sobre organización del conocimiento: organización del conocimiento e información científica". *Scire*. 4 (1995) 149-157.
- Guimarães, J. A. C. (2000). Perspectivas de ensino e pesquisa em organização do conhecimento em cursos de biblioteconomia do Mercosul: uma reflexão. En Encuentro de EDIBCIC, 5., Granada. *La formación de profesionales y investigadores de la formación para la sociedad del conocimiento*: Actas. Granada: Universidad de Granada (2000) 206-216.
- Guimarães, J. A. C. et al. (2009). Aspectos éticos da organização da informação: abordagens teóricas acerca da questão dos valores. En Gomes, H. F.; Bottentuit, A. M.; Oliveira, M. O. E. de (org.). (2009) *A ética na sociedade, na área da informação e da atuação profissional: o olhar da Filosofia, da Sociologia, da Ciência da Informação e da formação e do exercício profissional do bibliotecário no Brasil*. Brasília: Conselho Federal de Biblioteconomia, 94-129.
- Guimarães, J.A.C.; Oliveira, E.F.T.; Grácio, M. C. C. (2012). Theoretical referents in knowledge organization: a domain-analysis of Knowledge Organization journal. In: Neelamenghan, A.; Raghavan, K. S. (Org.). *Categories, contexts and relations in knowledge organization*. Würzburg: Ergon Verlag, 1 (2012) 31-38.
- Hjørland, B. (2002). Domain analysis in information science: eleven approaches-traditional as well as innovative. *Journal of Documentation*. 58:4 (2002) 422-462.
- Hjørland, B. (2008). What is knowledge organization (KO)? *Knowledge Organization*. 35:2/3 (2008) 86-101.
- Howarth, L. C. (2005). Metadata and bibliographic control: soul-mates or two solitudes? *Cataloging & Classification Quarterly*. 40: ¾ (2005) 37-56.
- Mcllwaine, I. C.; Mitchell, J. S. (2008). Preface to especial issue. "What is knowledge organization". *Knowledge Organization*. 35:2/3 (2008) 79-81.
- Oliveira, E. F. T.; Grácio, M. C. C. (2013). Studies of author cocitation analysis: a bibliometric approach for domain analysis. *Iris*. 2:1 (2014) 12-23.
- Olson, H. A. (2002). *The power to name: locating the limits or subject representation in libraries*. Dordrecht: Kluwer Academic Publisher.
- Otte, E.; Rousseau, R. (2002) Social network analysis: a powerful strategy, also for the information sciences. *Journal of Information Science*, Cambridge, 28:6, (2002) 441-453.
- Piovezan, L. B.; Fujita, M. S. L. (2015). Análise de cocitação de autores: uma aplicação em estudos de indexação. *Em Questão*. 21:1 (2015) 110-129.
- Smiraglia, R.P. (2011). ISKO 11's Diverse Bookshelf: an editorial. *Knowledge Organization*. 38:3 (2011) 179-186.
- Souza, C.D. (2013). A organização do conhecimento: Estudo bibliométrico na base de dados ISI Web of Knowledge. *Biblios: Revista de Bibliotecología y Ciencias de la Información*. 51 (2013) 20-32.
- Travieso-Rodríguez, C. et al. (2011). La Organización del Conocimiento en España a partir del análisis bibliométrico de los Congresos ISKO-Capítulo español. En Pérez Pais, M. C.; Bonome, M. G. (org.). *20 Años del Capítulo Español de ISKO: Actas del X Congreso ISKO Capítulo Español* (Ferro, 2011). Coruña: Universidade da Coruña. 127-141.
- Vanz, S. A. de S.; Caregnato, S. E. (2003). Estudos de citação: uma ferramenta para entender a comunicação científica. *Em questão*. 9:2 (2003) 295-307.

## Anexo 1

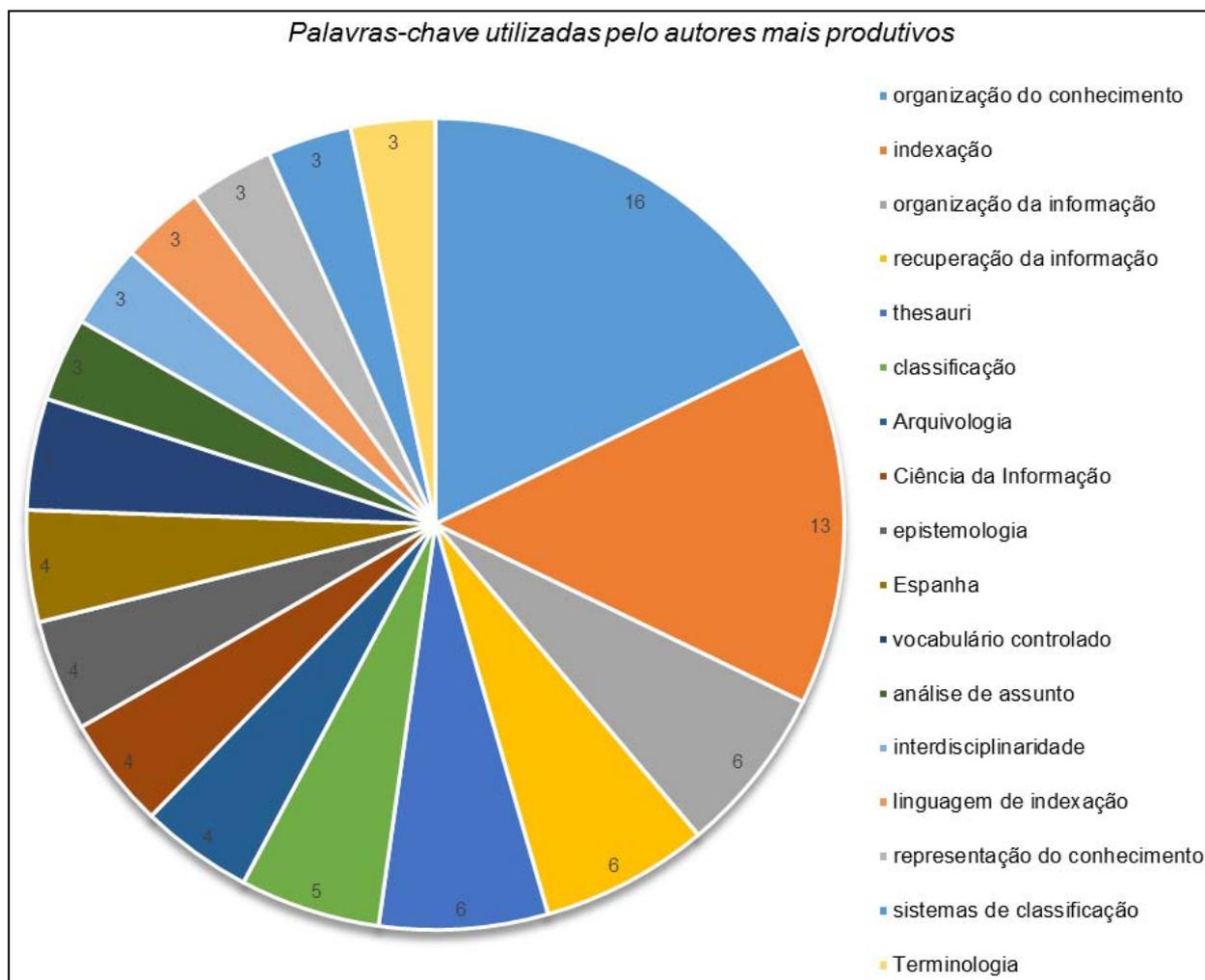


Figura 1. *Palavras-chave utilizadas pelos autores mais produtivos*

## 5. Anexo 2

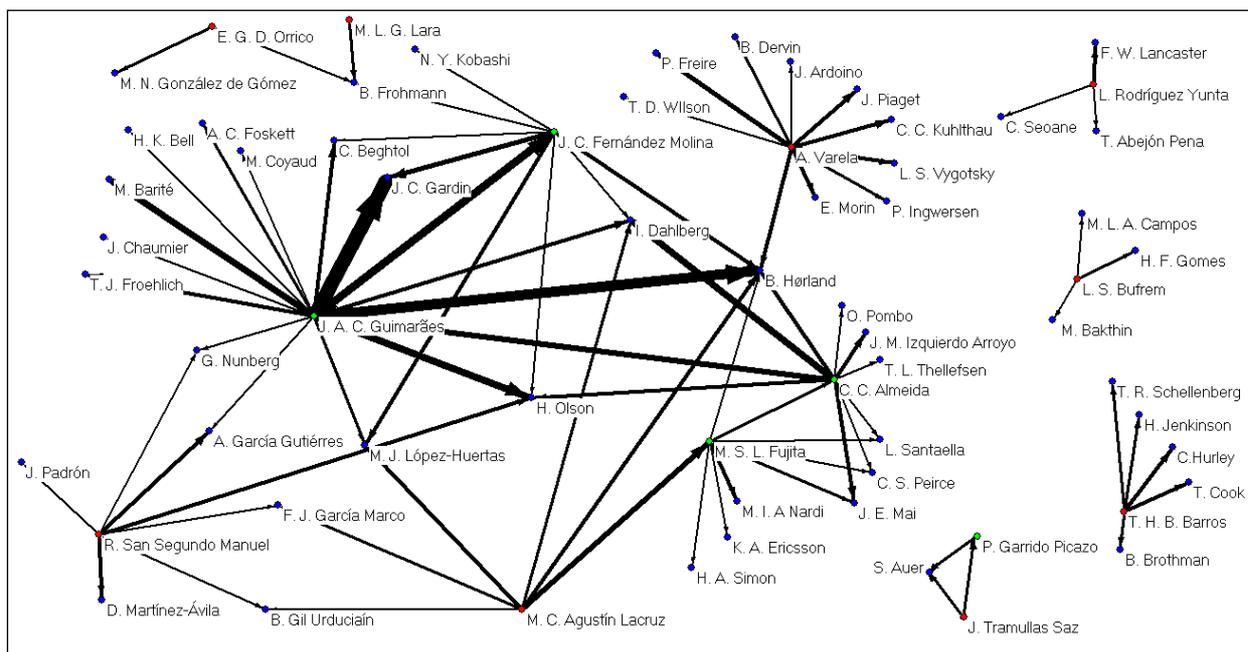


Figura 2. Rede citante-citado entre os autores mais produtivos e seus referentes.

## Anexo 3

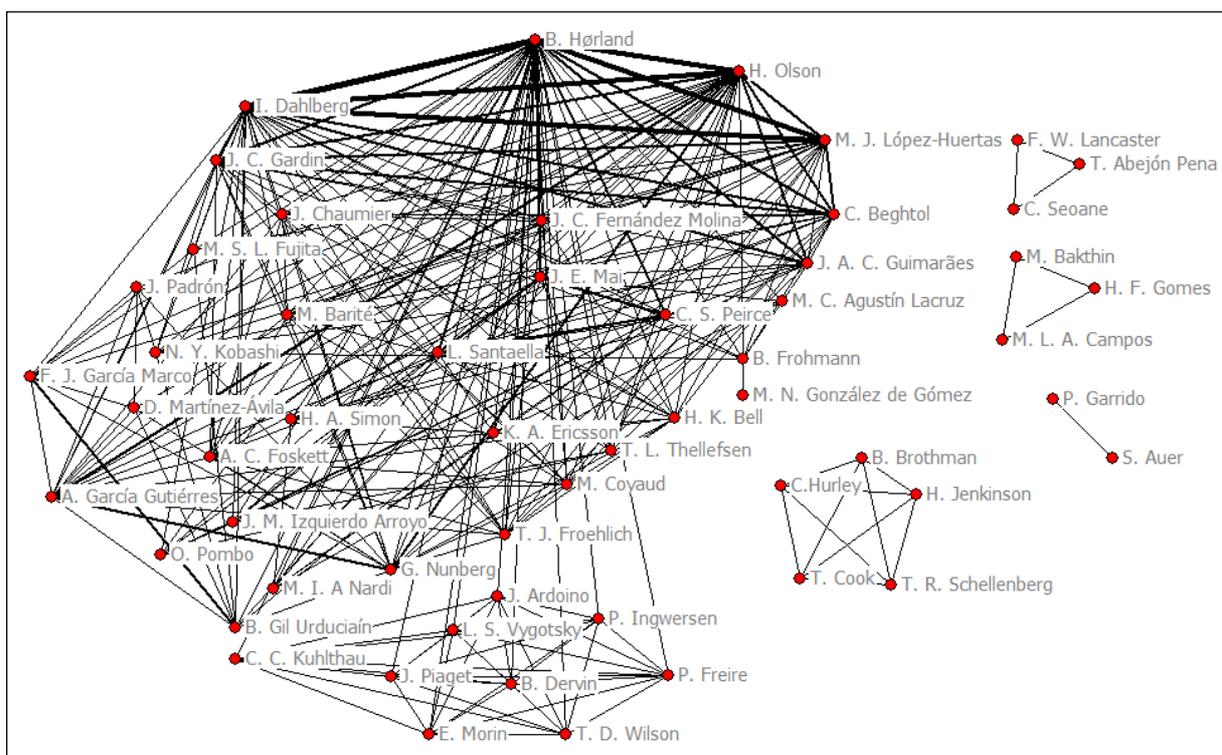


Figura 3. Rede de cocitação

---

# Referentes e interlocuções teóricas em análise documental no contexto brasileiro de Ciência da Informação

*Theoretical influences and scientific dialogues in subject analysis in Brazil*

---

**José Augusto Chaves Guimarães (1), Rafael Cacciolari Dalessandro (2), Daniel Martínez-Ávila (3)**

(1) Departamento de Ciência da Informação – Universidade Estadual Paulista, Av. Hygino Muzi Filho, 737 – Marília – SP – 17525-900 – BRASIL, guima@marilia.unesp.br

(2) rafaelcacciolari@marilia.unesp.br

(3) dmartinezavila@marilia.unesp.br

## Resumo

Considerando a análise documental – AD - como dimensão teórica do tratamento temático da informação - TTI (Guimarães, 2008, 2009), com forte desenvolvimento no Brasil, objetivou-se identificar e analisar, na produção científica brasileira em AD, quais as relações dialógicas estabelecidas entre os autores assim como os referentes teóricos norteadores dessa produção. Para tanto, desenvolveu-se um resgate histórico e teórico-conceitual de AD e uma análise de domínio, nas dimensões epistemológicas e bibliométricas (Hjørland, 2002). Tendo como corpus a produção científica periódica dos pesquisadores cujo currículo Lattes apresenta incidência superior a 20% dos termos análise documental / análise documentária, leitura documental / leitura documentária, condensação documental / condensação documentária, representação documental / representação documentária, linguagem documental / linguagem documentária e linguagens documentais / linguagens documentárias, analisam-se os autores e instituições mais produtivos, as coautorias, os periódicos mais utilizados para publicação e elaboram-se redes de coautoria e de citação, para verificar os referentes teóricos que norteiam essa produção científica.

**Palabras clave:** Análise documental. Análise de domínio. Análise de citações.

## 1. Introdução

No âmbito das discussões sobre as dimensões teóricas do tratamento temático da informação (Guimarães, 2008, 2009; Guimarães; Pinho; Ferreira, 2012), observa-se, especificamente no caso da análise documental (AD), uma configuração distinta das demais, pois, a catalogação de assunto e a indexação apresentam matriz anglo-americana e de utilização predominante na América do Norte e na Europa anglo-saxônica e, a análise documental, matriz francesa, com influência grande na Espanha e no

## Abstract

Considering subject analysis as one of the theoretical dimensions of the subject approach to information (Guimarães, 2008, 2009), we aim to identify, in the Brazilian subject analysis scientific production, the dialogical relations established between these researchers and the theoretical influences that guide the scientific production of these authors. We conducted a historical and theoretical-conceptual analysis of subject analysis and a domain analysis, in its epistemological and bibliometric approaches (Hjørland, 2002). We worked with the corpus of journal publications by researchers whose lattes curricula present an incidence higher than 20% of the terms document analysis, documental reading, documentary condensation, documentary representation, document language, document languages (in Portuguese); then we analyzed the most productive authors and institutions, co-authorships, the most used journals for the dissemination of this production and, and using the software Pajek, co-authorship and citation (citing / cited) network to check the theoretical influences that guide this scientific production.

**Keywords:** Citation analysis. Domain analysis. Subject analysis.

Brasil, com forte presença lógico-linguística a partir, principalmente, das concepções de Jean-Claude Gardin (Guimarães; Sales, 2010).

No entanto, um problema parece subsistir: a terminologia da área se disseminou amplamente, na literatura científica brasileira, sem necessariamente refletir uma concepção teórica específica. Assim, não é raro encontrar-se a utilização da referida terminologia para designar situações claramente voltadas a um universo de conceitos e de referentes ligados à indexação ou a catalogação de assunto, o que muitas ve-

zes dificulta a própria caracterização da construção epistemológica dessa área.

Desse modo, teve-se como objetivo geral, identificar as dimensões teóricas da análise documental no Brasil, a partir da produção científica dos pesquisadores da área. Decorrendo desse objetivo geral, apresentaram-se, como objetivos específicos: a) identificar, a partir da produção científica, a comunidade de pesquisadores brasileiros que integram esse universo de AD; b) identificar as relações dialógicas que se estabelecem entre esses pesquisadores a partir de suas coautorias; e c) analisar os referentes teóricos que norteiam a produção científica desses autores, no domínio e as relações que se estabelecem entre eles.

Para tanto, parte-se de uma revisão teórica sobre análise documental para, em seguida, analisar a literatura sobre análise documental publicada em periódicos por autores brasileiros, tendo como base de pesquisa o Currículo Lattes do CNPq – Conselho Nacional de Desenvolvimento Científico e Tecnológico. Com base em tais aspectos, analisa-se a referida produção em termos de sua configuração temática, relações bibliométricas e referentes teóricos.

## 2. Análise documental - AD: considerações teóricas

O Tratamento Temático da Informação – TTI - ocupa espaço nuclear na Ciência da Informação pois propicia a mediação entre a produção e o uso da informação, refletindo a necessidade de se “reunir e organizar para achar” (SMIT, 1986, p.12), atuando no ciclo de operações documentárias (Kobashi, 1994).

Assim, e tendo por objeto os “aspectos vinculados à análise, descrição e representação do conteúdo dos documentos, bem como suas inevitáveis interfaces com as teorias e sistemas de armazenamento e recuperação da informação.” (BARITÉ, 1998, P. 124), o TTI divide sua função nuclear com o Tratamento Descritivo da Informação, voltado aos aspectos externos do documento para contribuir para que se efetive a recuperação da informação.”

No âmbito do TTI, merecem destaque os trabalhos de Jean-Claude Gardin e de Coyaud, que, como destacam Guimarães & Sales (2010), desenvolveram, na França, nas décadas de 60 e 70 do século XX, estudos de forte matriz linguística para análise do conteúdo de documentos, ao que denominaram *análise documental*, voltada para a identificação e a representação do conteúdo intelectual do documento para facilitar a consulta ou a posterior recuperação.

Chaumier (1982, p.27) explica que “a análise documental abrange dois tipos de tratamentos diferentes: a condensação, que se vale de uma redução do texto para fins de difusão da informação, e a indexação, que se vale da extração de conceitos para servir de apoio à recuperação”.

Essa operação intelectual consiste, para Fox (2005, p.22), em um trabalho de “distinção e decomposição das partes de um todo (...) que, em um universo informacional, realiza-se por meio de uma operação de “reconhecimento e de estudo de um documento”.

Em seu *Diccionario de Bibliología y ciencias afines*, José Martínez de Sousa define a análise documental como um “conjunto de operações realizadas para representar o conteúdo de um documento de forma distinta da original, com o fim de facilitar a consulta ou a recuperação da informação” e explica que, em termos gerais, a AD abrange, por um lado, a descrição bibliográfica e a catalogação e, por outro, a indexação e a elaboração de resumos. (MARTÍNEZ DE SOUSA, 1989, P.19).

Observa-se, assim, que a AD pressupõe a questão dos procedimentos, de como se efetiva esse *conjunto de operações*, o que pressupõe uma lógica interna, uma sequência coerente de etapas, cada uma com seus precípuos. O objeto desses procedimentos, como destaca está na expressão de um conteúdo documental, ou seja, dos “...elementos informativos de um documento original a fim de expressar seu conteúdo de forma abreviada, resultando na conversão de um documento primário em documento secundário” (Coll Vinent & Bernal Cruz (1990, p.108). Essa expressão do conteúdo se faz por meio de uma linguagem, denominada genericamente linguagem documental (classificações, listas de cabeçalhos de assunto, tesouros) de forma a permitir a recuperação da informação que se busca em meio a um dado conjunto documental (LÓPEZ YEPES, 2004, v.1, v.51).

Como sintetiza Fox (2005, p. 25, 28), tem-se, na AD, um processo comunicativo em que um conjunto de operações cognitivas de natureza analítico-sintética transforma, por meio do reconhecimento e da representação do conteúdo (ou, como prefere Chaumier, 1982, p.27, do *reconhecimento* e da *classificação* de conteúdos), o documento original em outro mais exequível, fornecendo especial apoio à pesquisa científica. Tem-se, assim, como resultado, a produção de um novo documento, denominado por Fox (2005, p.25) “documento secundário” ou, como prefere Kobashi (1994), “informação documental”, que, conforme a natureza da recuperação

da informação almejada reveste-se da forma de indexação (aqui compreendidas tanto as notações classificatórias como os termos descritores) ou de resumo.

Chaumier (1982, p.13, 37-38), ao defender a existência de operações de duas ordens - uma de ordem eminentemente intelectual, voltada para o reconhecimento e a extração dos conceitos informativos e outra, posterior e de natureza mais aplicada, voltada para a tradução desses conceitos em uma linguagem documental, destaca que a AD “condiciona o valor do sistema documental” uma vez que, do conjunto de operações de tratamento da informação utilizado na cadeia documental é que dependerão os resultados obtidos na fase de recuperação.

Pode-se observar, portanto, que a AD, como uma operação – ou conjunto de operações - de decomposição (análise) e representação do conteúdo informacional dos documentos, pressupõe um conjunto sistemático e sequencial de procedimentos que possam ser explicitados, com respaldo em aportes interdisciplinares (com os da Lingüística, da Lógica e da Terminologia) e necessita de ferramentas, denominadas linguagens documentais. Desse modo, a explicitação dos processos desenvolvidos torna-se fundamental para que os produtos daí gerados possam ter confiabilidade.

A partir da literatura estudada, pode-se dizer que a AD centra-se na explicitação de um conjunto de processos ou operações de teor analítico, voltados para a análise, síntese e representação do conteúdo documental que se valem de instrumentos normalizados, também denominados linguagens documentais para gerar produtos simbólicos ou textuais que subsidiam a recuperação da informação. Esses processos se pautam por princípios de pertinência, precisão, relevância, univocidade, coerência, desambiguidade e seletividade e contribuem para viabilizar um processo comunicativo de tal modo que a informação possa ser efetivamente apropriada, agregando valor ao sistema documental.

No universo estudado, os estudos bibliométricos, entendidos como a aplicação de métodos quantitativos para a análise de processos da comunicação escrita, revelam-se especialmente importantes pois, a partir de análises de citações, por exemplo, pode-se melhor visualizar e compreender como se comporta um campo científico e quais os referentes teóricos que lhe são mais importantes. Isso contribui para que melhor se compreenda o universo epistemológico de uma área e como seus atores interagem (Vanz; Caregnato, 2003).

A partir desses estudos torna-se possível à construção e visualização gráfica de redes sociais de citação e de cocitação, especialmente útil na visualização do comportamento de um dado domínio de conhecimento.

### 3. Metodologia

A seleção do corpus de pesquisa se constituiu a partir de uma busca dos termos *análise documental/análise documentária, leitura documental/leitura documentária, condensação documental/condensação documentária, representação documental/representação documentária, linguagem documental/linguagem documentária, linguagens documentais/linguagens documentárias* em currículos de pesquisadores doutores brasileiros, da área de Ciência da Informação, constantes da Plataforma Lattes, do CNPq, tendo sido realizada no dia 7 de Março de 2013.

Em seguida, levantou-se a produção científica desses autores selecionados, especificamente em artigos de periódicos, por ser a produção cientificamente mais visível e que possui critérios objetivos de avaliação em nível internacional (ISI) e nacional (Qualis).

Dessa forma, considerou-se esse conjunto documental como representante de um domínio a ser analisado, lançando-se mão da análise de domínio.

A vista do exposto e considerando os objetivos da presente pesquisa, valeu-se da análise de domínio – no caso, a comunidade de pesquisadores brasileiros que tem se dedicado a questões de análise documental – mais especificamente na intersecção de duas abordagens: os estudos epistemológicos (na medida em que o cerne reside na dimensão do conceitual da organização do conhecimento e em suas articulações com os conceitos de organização da informação, de conhecimento e de informação) e os estudos bibliométricos, entendidos como a aplicação de métodos quantitativos para a análise de processos da comunicação escrita e, indo além, como subsídio para a análise da natureza e do comportamento de campos científicos. Para tanto, recorreu-se ao aporte teórico-metodológico da análise de citações como subsídio à identificação, visualização e maior compreensão do universo de marcos teóricos que permeiam a área.

Desse modo, e uma vez identificada a produção científica periódica total dos autores selecionados (para permitir uma análise diacrônica), foram coletados, em cada artigo, os seguintes dados: autor(es), filiação institucional, e referências.

Uma primeira análise fez-se a partir do conjunto de artigos, no sentido de verificar quais os periódicos em que foram publicados e o status dado a esses periódicos (Qualis/ISIS) pela comunidade científica.

Em seguida, foram analisados os períodos de maior produção, para que se possam aquilatar as tendências de desenvolvimento e/ou decréscimo dessa corrente teórica.

As afiliações institucionais permitiram identificar as instituições que mais têm se dedicado a essa temática, como subsídio à identificação de núcleos investigativos e a análise de coautorias, por sua vez, permitiu analisar as interlocuções ocorridas para a produção desses artigos. Já as referências (excluindo autocitações e autoria de entidades coletivas), por sua vez, foram analisadas a partir de redes, elaboradas com a utilização do software PAJEK, na seguinte conformidade: a) rede de autores citantes e citados: para identificar núcleos teóricos a partir de escolas de referentes, e b) rede de coautoria.

Com base nas referidas informações tornou-se possível traçar um panorama da pesquisa brasileira nessa área temática, relativamente ao seu núcleo de autores e instituições, as influências teóricas por eles sofridas, aos veículos por eles utilizados para socializar o conhecimento por eles produzido e, em uma perspectiva diacrônica, como se deu a evolução dessa linha teórica.

Vale destacar que foram desconsideradas as autocitações e, no caso de um artigo figurar mais de uma vez, em virtude de coautoria, esse artigo foi contado para o nome do autor que figurou em primeira posição.

#### 4. Apresentação, análise e discussão dos resultados

Do processo de análise dos currículos Lattes, obteve-se como resultado um conjunto de 375 autores sob o assunto análise documental / análise documentária, 2 em condensação documentária, 19 em leitura documental / leitura documentária, 25 em representação documental / representação documentária, e 103 em linguagem(ns) documental(ais) / linguagem(n)s documentária(s).

Em seguida, fez-se a consolidação dos dados, uma vez que muitos autores figuravam em várias listas. Desse universo, foram selecionados os autores brasileiros que pertencem à área de Ciência da Informação, chegando-se a um grupo de 67 autores com produção efetiva sobre a temática analisada, veiculada em artigos de periódicos, aos quais se aplicou um corte correspondente a uma produção científica na temá-

tica analisada que fosse igual ou superior a 20% de sua produção total. Chegou-se, assim, a um conjunto de 360 artigos, a cargo de 45 autores, nomeadamente: Silva, I.S. (100%); Lima, V.M.A.(100%); João Batista Ernesto de Moraes (89%); Maria Luiza de Almeida Campos (83%); Milena Polsinelli Rubi (83%); Marcos Luiz Cavalcanti de Miranda (80%); Mariângela Spotti Lopes Fujita (78%); Marilda Lopes Ginez de Lara (78%); Rodrigo de Sales (75%); Marisa Brascher Basílio Medeiros (71%); Dalgiza Andrade Oliveira (67%); Fabio Assis Pinho (60%); Maria de Fátima Gonçalves Moreira Tálamo (60%); Regina Helena van der Laan (60%); Vera Regina Casari Boccato (60%); Cibele Araújo Camargo Marques dos Santos (57%); Nair Yumiko Kobashi (55%); Lourival Pereira Pinto (50%); Simone Bastos Vieira (50%); Wilmara Rodrigues Calderon (50%); Ligia Maria Arruda Café (48%); Anna Maria Marques Cintra (47%); Hagar Espanha Gomes (47%); Cristina Dotta Ortega (43%); José Augusto Chaves Guimarães (43%); Maria Cristiane Barbosa Galvão (43%); Cláudio Omar Iahnke Nunes (42%); Regina Keiko Obata Ferreira Amaro (40%); Ulf Gregor Baranow (40%); Lídia Alvarenga (38%); Maria Aparecida Moura (33%); Miriam Paula Manini (30%); Luciana de Souza Gracioso (28%); Maria Odaisa Espinheiro de Oliveira (28%); Virgínia Bentes Pinto (28%); Clarinda Rodríguez Lucas (27%); Rosali Fernandez de Souza (27%); Gustavo Silva Saldanha (25%); Johanna Wilhelmina Smit (25%); Wanda Maria Maia da Rocha Paranhos (25%); Eliane Serrão Alves Mey (22%); Ilza Leite de Azevedo Santos Lopes (22%); Walter Moreira (21%); Maria Isabel Asperti Nardi (20%); Rosa Inês de Novais Cordeiro (20%). Esse conjunto documental, desconsiderando-se as incidências múltiplas em função de coautorias entre autores analisados, levou a um total de 280 artigos diferentes dos quais 21 (8%) não puderam ser recuperados, recaindo a análise sobre um total de 259 artigos.

Reorganizando esse universo pela quantidade de artigos sobre a temática efetivamente analisada, chegou-se a um conjunto de 45 autores e 360 artigos, com a seguinte produtividade:

Fujita (47 artigos); Lara (28 artigos); Tálamo (27 artigos); Guimarães (24 artigos); Boccato (21 artigos); Campos (20 artigos); Kobashi (16 artigos); Café e Moura (12 artigos cada); Rubi e Brascher (10 artigos cada); Sales (9 artigos); Moraes e Alvarenga (8 artigos cada); Cintra; Gomes e Smit (7 artigos cada); Ortega; Pinho; Galvão e Pinto, V. (6 artigos cada); Nunes e Moreira (5 artigos cada); Santos; Gracioso; Miranda e Lima (4 artigos cada); Lucas; Manini; van der Laan; Cordeiro e Souza (3 artigos ca-

da); Oliveira, D.; Mey; Saldanha; Lopes; Nardi; Oliveira, M.O.; Amaro; Vieira e Baranow (2 artigos cada); Silva; Pinto, L.; Paranhos e Calderon (1 artigo cada).

Especificamente no que se refere à temática analisada, foi possível observar que 11 dos 45 autores – Fujita; Lara; Tálamo; Guimarães; Boccato; Campos, M.L.A.; Kobashi; Café; Moura; Rubi e Brascher respondem por 63% da produção, o que revela sua centralidade na área estudada. Nesse universo, Fujita se destaca, pois apenas sua produção responde por 13% do total. Desse modo, em termos de produção geral, tem-se a Universidade Estadual Paulista - UNESP respondendo por 38% do total, com 137 artigos, seguida da Universidade de São Paulo - USP, com 94 artigos, correspondendo a 26% do total, o que leva a produção da UNESP somada à da USP corresponder à ordem de 645 do total.

Em termos de autores, em uma dimensão institucional, USP e UNESP, ambas no estado de São Paulo, lideram, respondendo por respectivamente 9 e 7 pesquisadores cada (correspondendo, juntos a um terço do total), seguidas por Universidade Federal de Minas Gerais - UFMG e Universidade Federal Fluminense - UFF (com 4 pesquisadores cada), por Universidade de Brasília - UnB e Universidade Federal de São Carlos - UFScar (3 pesquisadores cada), ao que se soma um conjunto de 9 outras universidades que congregam um total de 15 pesquisadores na área. É interessante observar que 4 das 17 instituições apresentadas – USP, UNESP, UFMG e UFF - respondem, juntas, por metade da produção, o que revela serem espaços de excelência científica nessa temática, aspecto que se justifica, ainda, por serem temáticas que formalmente integram linha de pesquisa de seus cursos de pós-graduação em Ciência da Informação. Essa liderança da USP e da UNESP se explica historicamente, pois essas instituições foram mais nitidamente influenciadas pelas ideias de Jean-Claude Gardin no âmbito da análise documental.

Relativamente aos veículos de divulgação (periódico científico), tem-se o seguinte resultado: Datagramazero (46 artigos); Scire (36 artigos); Perspectivas em Ciência da Informação (34 artigos); Ciência da Informação; Transinformação (28 artigos cada); Encontro Bibli (19 artigos); Informação e informação (18 artigos); Informação e Sociedade (13 artigos); Revista de Comunicações e Artes (10 artigos); Revista Brasileira de Biblioteconomia e Documentação (9 artigos); Brazilian Journal of Information Science (7 artigos); Biblos; Cuadernos de ADAB; Informare; Knowledge Organization; Revista de Iniciação Científica da F.F.C.- UNESP e Revista

Digital de Biblioteconomia e Ciência da Informação (5 artigos cada); Biblos; Comunicação e Educação; InCID: revista de Ciência da Informação e Documentação; Revista EDICIC e Sciences de la Société (4 artigos cada); Anales de Documentación; Cuadernos de Terminología; Ponto de acesso e Revista Eletrônica de Comunicação, Informação & Inovação em Saúde – RECIIS (3 artigos cada); Arquivo e Administração; Cuadernos da F.F.C.; Cuadernos BAD; Cenário Arquivístico; Cibertextualidades; História, Ciências, Saúde – Manguinhos; International Journal of Library and Information Science; META; Revista Brasileira de Educação Especial; Revista CRB-8 digital; Revista do Centro Interdepartamental de Tradução e Terminologia da USP; Revista Ibero-Americana de Ciência da Informação; Rumores e Tendências da Pesquisa Brasileira em Ciência da Informação (2 artigos cada); África e Africanidades; Boletim da ABDF; Boletim Informativo da ANPOLL; Cuadernos da Pós-Graduação do Instituto de Artes da UNICAMP; Conexões; Cuadernos de la Universidad de Panama; Delta - Documentação de estudos em linguística teórica e aplicada; Escrita; Information Research; Interatividade: revista eletrônica de pesquisas e estudos acadêmicos; Journal of Information and Data Management – JIDM; Liinc em Revista; Nuovi Annali della Scuola Speciale per Archivisti e Bibliotecari; Páginas A&B; Revista Acervo; Revista de Educação, Cultura e Comunicação da FATEA; Revista de Letras; Revista do GELNE; Revista do Museu de Arqueologia e Etnologia; Revista Latino Americana de Documentación; São Paulo em perspectiva e Significação (1 artigo cada)

Trata-se de um conjunto de 62 periódicos, sendo 48 publicadas no Brasil (77% do total) e 14 no exterior (23% do total) em que se observa uma forte concentração de publicações em um grupo de onze revistas – Datagramazero, Scire, Perspectivas em Ciência da Informação, Ciência da Informação, Transinformação, Encontros Bibli, Informação & Informação, Informação e Sociedade, Revista de Comunicações e Artes da USP, Revista Brasileira de Biblioteconomia e Documentação e Brazilian Journal of Information Science – que, juntas, respondem por 49% de toda a produção. Em termos de produção total, observa-se que 19% foi publicada no exterior, o que revela uma inserção internacional ainda tímida na área.

Uma análise dessa produção tendo em vista o conceito Qualis na área de Ciências Sociais Aplicadas, revela que os três extratos considerados de excelência – A1, A2 e B1 perfazem, juntos, 76% do total da produção. Essa produção, por sua vez, se deu prioritariamente em periódicos

B1 (53% do total), seguida de periódicos A1 (22,50% do total) e de periódicos A2 (0,5% do total), o que revela ser uma produção não apenas expressiva em termos quantitativos mas, e principalmente, em termos qualitativos.

Uma análise cronológica dessa produção, ocorrida entre os anos de 1974 e 2012, revelando uma produção de 6 artigos (2%) na década de 70, 16 artigos (4%) na década de 80, 61 artigos na década de 90 (17%), 211 artigos na década de 2000 (59%) e 66 artigos nos anos de 2011 e 2012 (18%). Nesse período, observa-se um forte crescimento de produção científica notadamente a partir de 2006, sendo que a produção de 2006 a 2012, de 224 artigos, responde por 62% de toda a produção analisada.

No tocante às citações ocorridas nesses artigos, obteve-se um conjunto de 2234 autores citados em um total de 5769 citações. Desse universo, observou-se que 1440 autores (64% do total) foram citados apenas uma vez. Aplicando-se a Teoria de elitismo de Price aos 2234 autores citados, chegou-se a 47 autores, que, figurando na categoria daqueles que receberam 14 citações, elevou o número para os 49 autores (22% do total), que receberam cinco ou mais citações (correspondendo a 50% das citações), a saber: Dahlberg, I. (98 citações); Gardin, J.-C. (79 citações); Hjørland, B. (77 citações); Lancaster, F.W. e Lara, M.L.G. (64 citações cada); Campos, M.L.de.A. (60 citações); Smit, J.W. (56 citações); Cintra, A.M.M. e Kobashi, N.Y. (54 citações cada); Guarino, N. (51 citações); Talamo, M.de.F.G.M. (48 citações); González De Gómez, M.N. e Guimarães, J.A.C. (47 citações cada); Fujita, M.S.L. (45 citações); Gomes, H.E. (44 citações); Ranganathan, S.R. (43 citações); Cabré, M.T. e Foskett, A.C. (38 citações cada); Beghtol, C. (29 citações); Eco, U. e García Gutiérrez, A. (28 citações cada); Barité, M. (27 citações); Vickery, B.C. (25 citações); Gruber, T.R. (24 citações); Albrechtsen, H.; Chaumier, J. e Dias, E.W. (22 citações cada); Koch, I.G.V. e Robredo, J. (21 citações cada); Olson, H.A. (20 citações); Austin, D. e Peirce, C.S. (19 citações cada); Cesarino, M.A.N. e Naves, M.M.L. (18 citações cada); Capurro, R.; Piedade, M.A.R. e Uschold, M. (17 citações cada); Barreto, A.A.; Pinto Molina, M.; Svenonius, E. e Wüster, E. (16 citações cada); Fernández-molina, J.C.; Hutchins, W.J. e Sowa, J.F. (15 citações cada); Miranda, M.L.C.; Cunha, I.M.R.F.; Foucault, M.; Nardi, M.I.A. e Pinto, M.C.M.F. (14 citações cada)

Esse grupo de 49 autores, perfazendo um total de 1600 citações recebidas revela certo equilíbrio entre autores estrangeiros (57%) e autores nacionais (43%), aspecto que se repete nas

citações recebidas, sendo 56% de citações de autores estrangeiros e 44% de autores nacionais. Tal aspecto revela-se extremamente positivo, pois demonstra, por um lado, uma preocupação em coadunar-se com aquilo que se discute cientificamente no mundo e, por outro, revela que o país já dispõe de considerável massa crítica nessa área.

Do total de 49 autores mais citados que compõem essa elite, observa-se que 16 deles (33%) respondem por 58% das citações recebidas no universo pesquisador, sendo seis autores estrangeiros – Dahlberg, Gardin, Hjørland, Lancaster, Guarino e Ranganathan) e dez autores brasileiros (Lara, Campos, Smit, Cintra, Kobashi, Talamo, González de Gómez, Guimarães, Fujita e Gomes). No grupo brasileiro dos mais citados, três universidades se destacam, nomeadamente a USP (com 5 autores e 276 citações), a UFF (com 2 autores e 104 citações) e a UNESP (com 2 autores e 92 citações). Esse cenário confirma os dados anteriores sobre as instituições mais produtivas, na medida em que isso se reitera nas universidades mais citadas uma vez que USP, UFF e UNESP respondem, juntas, por 30% das citações recebidas no âmbito desse grupo de elite e 8% do total de citações ocorridas nos artigos analisados. Considerando, como já explicado anteriormente, que foram eliminadas previamente as autocitações, tal cenário revela-se significativo.

Dos autores estrangeiros mais citados, observa-se, por um lado, a questão da teoria da classificação, representada por Dahlberg e Ranganathan (141 citações), da análise documental propriamente dita (Gardin, com 79 citações), da abordagem sócio-cognitiva em tratamento temático da informação (Hjørland, com 77 citações) da teoria da indexação (Lancaster, com 64 citações) e das ontologias (Guarino, com 51 citações). Ainda sob uma outra ótica, observa-se que os estudos de organização do conhecimento (Dahlberg e Hjørland com um total de 175 citações, representando 11% das citações no âmbito dessa elite) oferecem base teórica para o campo estudado.

Sobre esse contexto elaborou-se, por meio do software Pajek, a rede de citações, como se pode observar na Figura 1, ao final deste trabalho.

Uma análise da rede apresentada revela uma diversidade de autores citados, o que pode apontar para um aspecto altamente positivo, qual seja, a diversidade de influências teóricas no campo estudado. Dessa forma, enquanto Campos se baseia fortemente nas teorias de organização do conhecimento de Dahlberg as-

sim como nos estudos de ontologia de Guarino, Fujita se aproxima de uma tradição anglo-saxônica do campo, representada por Foskett e Lancaster, aliada aos estudos sócio-cognitivos de Lancaster. Já Kobashi, Guimarães e Lara revelam uma concepção mais francesa, com forte orientação gardiniana. Destaca-se, ainda, entre os referentes nacionais a presença de González de Górmex, fortemente citada por Moura, o que evidencia uma preocupação de caráter epistemológica no campo estudado.

Tendo por base os autores mais produtivos, construiu-se, também com base no software PAJEK, a rede de colaboração, como se pode observar na Figura 2, ao final este trabalho.

A análise dessa rede evidencia parcerias significativas, como as de Fujita com Boccato e com Rubi, a de Campos com Gomes e um interessante núcleo que revela o trabalho do Grupo Temma, da USP: Lara, Kobashi, Tálamo, Cintra e Smit que guardam forte relação entre si. Nesse âmbito, forte é a colaboração entre Lara e Tálamo, evidenciando uma preocupação com a linguística documental. Guimarães atua como ponto de enlace entre o grupo da USP (com Smit), a UNESP (com Fujita e Moraes), a UFF (Sales) e a UFPE (Pinho).

## Conclusão

A partir da literatura internacional, observa-se que a análise documental revela-se como um conjunto de processos de análise e de síntese de conteúdo documental, que se vale tanto da linguagem natural quanto de linguagens documentais para fins de representação para posterior recuperação da informação por meio de produtos simbólicos ou textuais.

Objetivando viabilizar o processo comunicativo, a análise documental tem por objeto o conteúdo documental e caracteriza-se por uma forte ligação com práticas profissionais específicas (catalogação de assunto, indexação, etc) fornecendo subsídios a avaliação da recuperação da informação (*pertinência, precisão, relevância temática*) e da composição textual (*univocidade, coerência, ausência de ambiguidade, etc.*).

Analisando-se a produção científica periódica a partir de palavras-chave dos currículos Lattes de pesquisadores brasileiros, recuperaram-se 259 artigos em um grupo final de 45 autores que provém em grande parte (cerca de um terço) de USP e UNESP, revelando serem espaços naturais de pesquisa na temática estudada.

Essa produção foi veiculada em 62 periódicos, predominantemente publicados no Brasil (77% do total), em que se observa uma forte concen-

tração de publicações em um grupo de onze revistas – Datagramazero, Scire, Perspectivas em Ciência da Informação, Ciência da Informação, Transinformação, Encontros Bibli, Informação & Informação, Informação e sociedade, Revista de Comunicações e Artes da USP, Revista Brasileira de Biblioteconomia e Documentação e Brazilian Journal of Information Science – que, juntas, respondem por 49% de toda a produção.

Uma análise dessa produção, tendo em vista o conceito Qualis na área de Ciências Sociais Aplicadas, que avalia, no Brasil, as revistas da área de Ciência da Informação revela que os 76% do total da produção situa-se em revistas que possuem categorização e excelência (A1, A2 e B1), assim como se revela uma produção atual, pois o período de 2006 a 2012 responde por 62% de toda a produção analisada.

No tocante às citações, e aplicada a Teoria de Elitismo de Price, chegou-se a um conjunto de 49 referentes teóricos, em equilíbrio entre autores brasileiros e autores estrangeiros, o que se revela positivo pois demonstra, por um lado, uma preocupação em coadunar-se com aquilo que se discute cientificamente no mundo e, por outro, revela que o país já dispõe de considerável massa crítica nessa área. Nesse universo, Dahlberg, Gardin, Hjørland, Lancaster, Guarino, Ranganathan, Lara, Campos, Smit, Cintra, Kobashi, Tálamo, González de Górmex, Guimarães, Fujita e Gomes respondem por mais da metade das citações, evidenciando uma posição nuclear.

Uma análise da rede de citações aponta uma diversidade de influências teóricas no campo estudado em que enquanto Campos se baseia fortemente nas teorias de organização do conhecimento de Dahlberg assim como nos estudos de ontologia de Guarino, Fujita se aproxima de uma tradição anglo-saxônica do campo, representada por Foskett e Lancaster, aliada aos estudos sócio-cognitivos de Lancaster. Já Kobashi, Guimarães e Lara revelam uma concepção mais francesa, com forte orientação gardiniana. Destaca-se, ainda, entre os referentes nacionais a presença de González de Górmex, fortemente citada por Moura, o que evidencia uma preocupação de caráter epistemológica no campo estudado.

A rede de colaboração evidencia parcerias significativas, como a de Fujita com Boccato e com Rubi, de Campos com Gomes, e um interessante núcleo que revela o trabalho do Grupo Temma, da USP: Lara, Kobashi, Tálamo, Cintra e Smit que possuem forte relação entre si. Especialmente forte é a colaboração entre Lara e

Tálamo, evidenciando uma preocupação com a linguística documental. Guimarães atua como ponto de enlace entre o grupo da USP (com Smit), a UNESP (com Fujita e Moraes), a UFF (Sales) e a UFPE (Pinho).

Por fim, observa-se ser a análise documental um campo de estudo efetivamente consolidado no Brasil, com referencial teórico estabelecido, uma massa crítica considerável e uma produção atualizada e divulgada em veículos abalizados, aspectos que atestam a maturidade desse domínio. Desse modo, e considerando esse cenário, atenta-se para a necessidade de uma maior permeabilidade internacional desse grupo de pesquisadores, propiciando uma interlocução mais diversificada

## Referencias

- Barité, M. (1998). Referenciales teóricos vigentes en el área de tratamiento temático de la información y su expresión metodológica. Porto Alegre: ABEBD.
- Chaumier, J. (1982). Analyse et langages documentaires: le traitement linguistique de l'information documentaire. Paris : Entreprise Moderne d'Édition.
- Coll Vinent, R. ; Bernal Cruz, F.J. (1990). Curso de documentación. Madrid : Dossat.
- Fox, V. (2005). Análisis documental de contenido: principios y prácticas. Buenos Aires : Alfagrama.
- Guimarães, J.A.C. (2008). Ciência da Informação, arquivologia e biblioteconomia: em busca do necessário diálogo entre o universo teórico e os fazeres profissionais. En: Guimarães, J.A.C.; Fujita, M.S.L. (org.). Ensino e pesquisa em Biblioteconomia no Brasil: a emergência de um novo olhar. São Paulo: Cultura Acadêmica, 2008, p. 33-44.
- Guimarães, J.A.C. Abordagens teóricas em tratamento temático da informação: catalogação de assunto, indexação e análise documental. en: GARCIA MARCO, F.J. (org.). (1990). Avances y perspectivas en sistemas de información y de documentación. Zaragoza: Prensas Universitarias. p. 105-117.
- Guimarães, J.A.C.; PINHO, F.A.; FERREIRA, G.M. (2012). Relações teóricas da organização do conhecimento com as abordagens de catalogação de assunto, indexação e análise documental: uma análise de domínio da revista Scire (1995-2010). Scire, 18(2), 31-41.
- Guimarães, J.A.C.; SALES, R. (2010). Análise documental: concepções do universo acadêmico brasileiro em Ciência da Informação. Datagramazero, 11(1) artigo 2. Recuperado em 12.02.2015, de [http://www.dgz.org.br/fev10/F\\_I\\_art.htm](http://www.dgz.org.br/fev10/F_I_art.htm)
- Hjørland, B. (2002). Epistemology and the socio-cognitive perspective in Information Science. Journal of the American Society for Information Science and Technology, 53(4), 257-270.
- Kobashi, N. Y. (1994). A elaboração de informações documentárias: em busca de uma metodologia. Tese (Doutorado) – Escola de Comunicação e Arte, Universidade de São Paulo, São Paulo.
- López Yepes, J. (2004). Diccionario enciclopédico de Ciencias de la Documentación. Madrid : Síntesis. 2v..
- Martínez de Sousa, J. (1989). Diccionario de Bibliología y ciencias afines. Madrid : Fund. Germán Sánchez Ruipérez.
- Smit, J.W. (1986). O que é documentação. São Paulo : Brasiliense.
- Vanz, S.A.S.; Caregnato, S.E. (2003). Estudos de citação: uma ferramenta para entender a comunicação científica. Em questão: revista da Faculdade de Biblioteconomia e Comunicação da UFRGS, 9(2), 295-307.

## Apêndice: Redes de citação e de colaboração

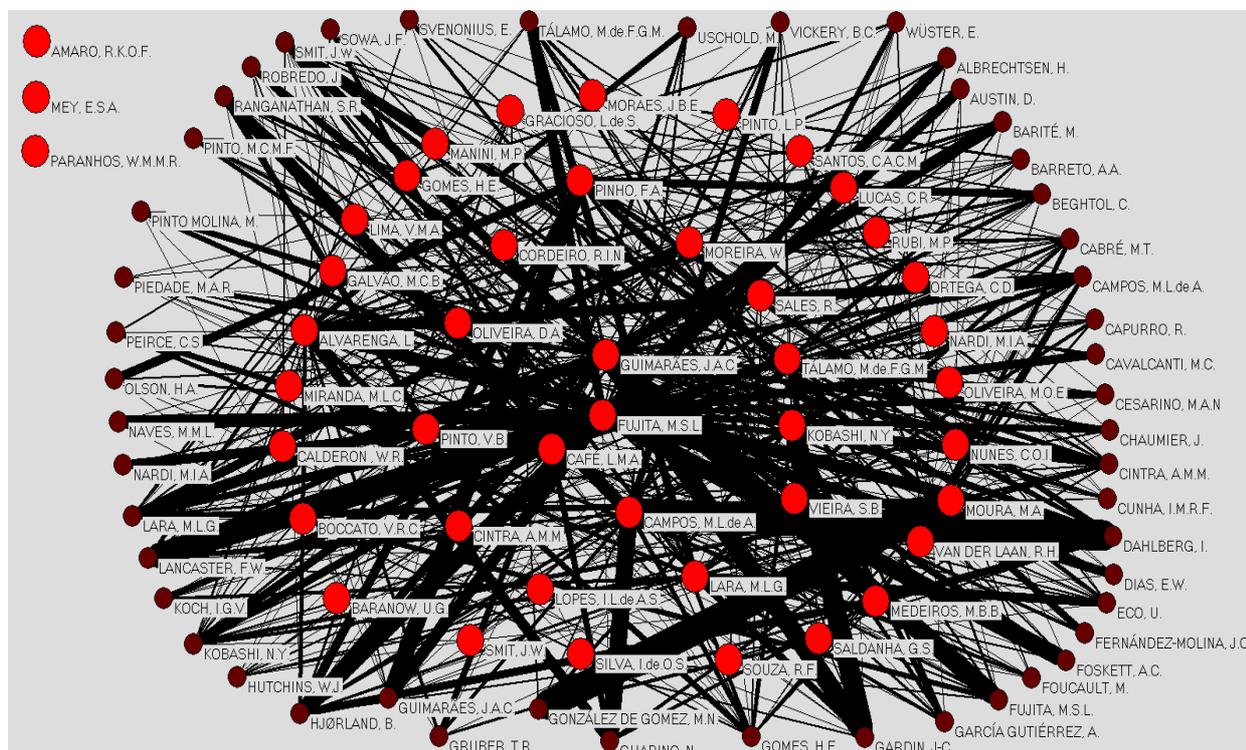


Figura 1: Rede de citação (vermelho = citantes; marrom = citados). Fonte: Elaborado pelos autores.

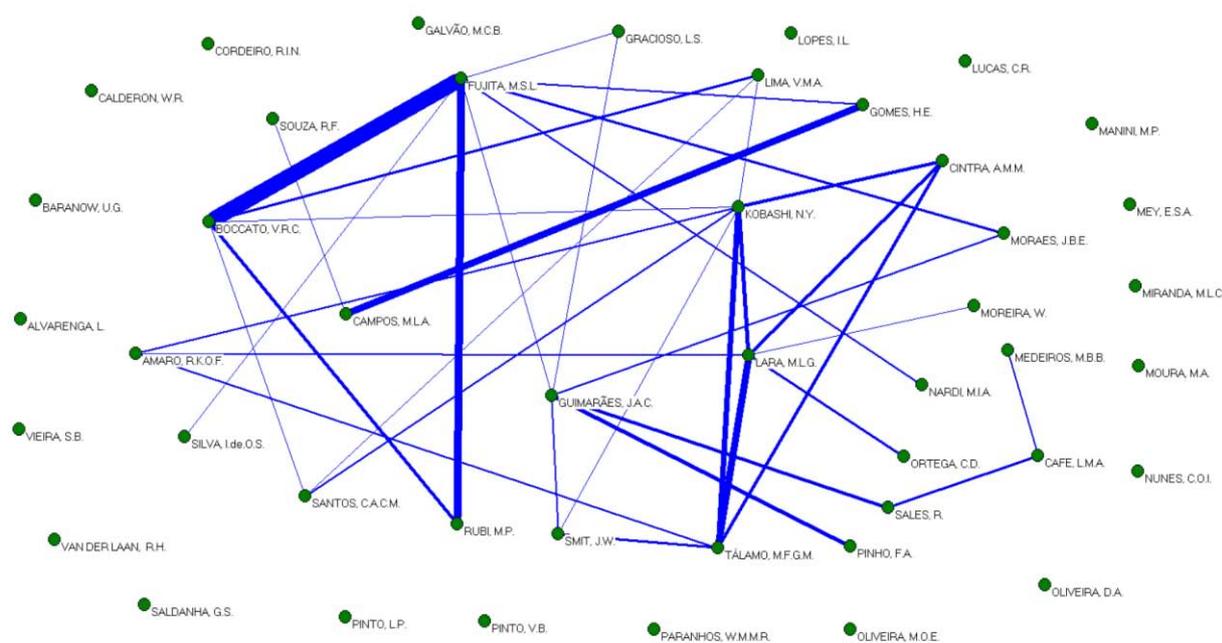


Figura 2: Rede de colaboração dos autores mais produtivos. Fonte: Elaborado pelos autores.

---

# Exaustividade e especificidade na indexação: uma análise de conteúdo como perspectiva de investigação sobre o tema

*Exhaustivity and specificity in indexing: a content analysis as a research perspective on the topic.*

---

**José Augusto Chaves Guimarães (1), Isadora Victorino Evangelista (2)**

(1) Universidade Estadual Paulista, Av. Hygino Muzzi Filho, 737 – Marília/SP, guima@marilia.unesp.br.

(2) isadora.biblio@marilia.unesp.br

## Resumen

A indexação, enquanto processo da representação temática na organização do conhecimento, possui outros dois micro-processos, responsáveis para o aumento de sua eficiência – a exaustividade e a especificidade. O presente artigo tem por objetivo identificar as abordagens conceituais acerca de exaustividade e especificidade na literatura internacional da Ciência da Informação. Para tanto, utilizou-se como método a análise de conteúdo de artigos recuperados pela base de dados SCOPUS, SCIELO, Web of Science e LISA, que continham os termos em suas áreas de maior conteúdo informacional (título, resumo e palavras-chave) publicados em periódicos científicos do campo da Ciência da Informação, sem limite temporal. Foram recuperados um total de 51 artigos, os quais 10 não foram encontrados para leitura na íntegra e 14 foram excluídos por não conterem as variáveis de inferência. Foram estabelecidas como categorias de inferência Conceitos, Características e Objetivo. Os resultados revelam uma preponderância de abordagens conceituais genéricas e "matemáticas" na medida em que se referem a questões de mensuração, sem que se revelem preocupações mais efetivas com aspectos ligados às necessidades dos usuários, objetivos da instituição e a questões éticas. Conclui-se, portanto, que existe a necessidade de maiores estudos na área, para que o tema seja visto de maneira mais ampla

**Palabras clave:** 1. Exaustividade. 2. Especificidade. 3. Indexação. 4. Ética Informacional. 5. Análise de Conteúdo.

## 1. Introdução

O campo da organização e representação do conhecimento materializa-se em duas dimensões complementares – descritiva e temática, a primeira tendo como resultado os catálogos que serão representados em bases de dados, de maneira que a representação chegue o mais próximo possível do documento físico e, a segunda, voltada ao conteúdo do recurso, envolvendo processos como a classificação e a inde-

## Abstract

Indexing, as a process of subject representation in knowledge organization, presents two more micro-processes that are responsible for the improvement of its efficiency –exhaustivity and specificity. This paper aims to identify the conceptual approaches to exhaustivity and specificity in the international literature on Information Science. For this, we used the method of content analysis for the articles retrieved from the databases SCOPUS, Scielo, Web of Science and LISA, that include the terms in their main areas of information content (title, abstract and keywords), published in scientific journals in the area of Information Science without time limit. A total of 51 articles were retrieved, where 10 of them could not be accessed and 14 were excluded for not including the inference variables. These variables included: Concepts, Characteristics and Objective. Results show a predominance of generic and "mathematical" conceptual approaches that refer to questions of measuring, without revealing more effective concerns on aspects related to users' needs, goals of the institution and ethical aspects. We conclude that there is a need of more studies in the area.

**Keywords:** Exhaustivity. Specificity. Indexing. Information ethics. Content analysis.

cação, que visam a resgatar os principais assuntos do documento, traduzindo-os por meio de termos que serão utilizados para futura busca na base de dados.

No caso da indexação, dois procedimentos encontram-se diretamente voltados à questão da eficiência: exaustividade, que diz respeito à amplitude da cobertura da temática do documento e a especificidade, entendida como a precisão e particularidade que será dada à re-

apresentação. Esses dois procedimentos possuem poucos estudos na área, entendidos muitas vezes apenas como questões matemáticas relacionadas à indexação, embora estudos recentes (Milani, 2007; Guimarães et al., 2008; Evangelista, 2013) demonstraram ser entendidos também como valores éticos.

O presente artigo tem por objetivo demonstrar como a literatura internacional da área de Ciência da Informação vem abordando conceitualmente a exaustividade e a especificidade, procedimentos intrínsecos à indexação valendo-se, para tanto, da metodologia de análise de conteúdo (Bardin, 2008).

Desse modo, parte-se de uma abordagem conceitual sobre indexação para se chegar especificamente à exaustividade e à especificidade. Em seguida, adentra-se nas questões éticas da organização do conhecimento como subsídio para a análise que se realiza no material explorado, de modo a identificar e discutir as dimensões conceituais almeçadas.

## 2. Indexação

A organização e representação do conhecimento é uma área de atuação da Ciência da Informação, centrada em tratar o documento para posterior recuperação e disseminação. Para que essa recuperação ocorra, têm-se a necessidade de alguns aportes teóricos que contribuam para tanto, um deles é a indexação.

A principal tarefa da indexação é a descrição temática do documento, empregando termos representativos geralmente selecionados a partir de um vocabulário controlado. De acordo com Lancaster (2004), a indexação é composta por duas etapas: análise conceitual, onde o assunto do documento é definido e a tradução, onde os assuntos são traduzidos para termos, de acordo com uma linguagem documentária. Ainda segundo o autor "(...) os termos atribuídos pelo indexador servem como pontos de acesso mediante os quais um item é localizado e recuperado, durante uma busca por assunto." (Lancaster, 2004, p. 06).

Hjørland define assunto como sendo "potenciais epistemológicos ou informativos dos documentos" e ainda ressalta que "(...) a melhor análise temática é a que faz o melhor prognóstico quanto ao uso futuro do documento" (Hjørland, 2011, p. 776).

Para Leiva (2012), "(...) a indexação gera palavras-chave, índices ou os cabeçalhos de assunto de um documento" (Leiva, 2012, p. 31). O autor ainda sistematiza algumas definições da

literatura para o termo indexação, como pode ser visto a seguir.

Analisar o conteúdo informacional dos registros do conhecimento e expressar o conteúdo informacional na linguagem do sistema de indexação. (Borko; Bernier, 1978, p. 8).

Representação pelos elementos de uma linguagem documentária ou natural, das noções resultantes da análise do conteúdo de um documento para facilitar sua localização. (NF Z 47 – 102 1978)

Descrever o conteúdo de documentos ou demandas documentárias para possibilitar a elaboração de estratégias de recuperação mediante conceitos ou assuntos. (García Gutiérrez (1984, p. 105)

Ação de descrever ou identificar um documento em relação ao conteúdo. (ISO 5963 – 1985)

Identificar informação numa entidade de conhecimento (que seja texto ou não) e organizá-la para que esteja disponível num sistema de recuperação. (Cleveland, D. B.; Cleveland, A. D., 2001, p. 97)

Determinar o assunto temático dos documentos e expressar em índices (por exemplo, descritores, cabeçalhos de assunto, números de chamada, códigos de classificação ou índice) para tornar possível a recuperação temática. (Mai, 2005, p. 599)." (Leiva, 2012, p.68).

Ou seja, a indexação visa a reduzir o documento a um grupo de termos, para que o armazenamento e a busca em bases de dados sejam facilitados e deve satisfazer as necessidades informacionais dos usuários, o que é buscado deve corresponder com o que for encontrado.

Considerando que os documentos irão responder a questões internas dos usuários, é essencial que a análise de assunto esteja preparada para as perguntas da comunidade em que a unidade informacional está inserida. Ou seja, o assunto do documento deve estar diretamente ligado ao objetivo daquela unidade (Hjørland, 2011).

Para Lancaster (2004), não há uma indexação "correta", uma indexação eficiente é aquela que além de levar em conta os assuntos de determinado documento, também reflete sobre o porquê dele ser importante para determinada comunidade. Comunidades com interesses diferentes buscarão através de termos diferentes e, portanto, não há apenas um conjunto de termos para aquele documento.

A indexação utiliza-se de processos que contribuem para melhorar sua eficácia, no presente artigo, busca-se focar em dois deles: a exaustividade e a especificidade.

### 3. Exaustividade e especificidade

A exaustividade caracteriza-se, basicamente, pelo nível de abrangência que a indexação vai optar por atribuir. A exaustividade concerne ao número de termos que serão utilizados para representar um documento e qual a extensão do conteúdo que esses termos irão cobrir.

Para Lancaster (2004), a quantidade de termos escolhidos para serem representativos daquele documento influencia o quanto o documento tornar-se-á acessível, e como resultado, recuperado, considerando que no momento da busca, a proporção de termos indexados/termos buscados será proporcionalmente maior. O autor ainda ressalta que ainda irá ocorrer o aumento de características que distingam aquele documento de outros.

Fosket (1973) divide a exaustividade em duas partes, a primeira dizendo respeito ao conteúdo global de um documento e a segunda, aos subtemas que correspondem a ele.

De acordo com Fujita (2004), quando optar pela exaustividade, o indexador deverá atentar-se ao número de termos que irá extrair de cada recurso, os selecionando baseados sempre no critério da abrangência.

Leiva (2012) diz que, embora não se tenha claro qual a quantidade exata de termos a serem atribuídos aos documentos, parece lógico que documentos como resumos gerem menos termos que textos completos, por exemplo.

Já a especificidade, consolida-se por ser o grau de particularidade dado à representação ou o nível de precisão que iremos utilizar na extensão de cobertura dos documentos. Decide-se quão específico serão os termos atribuídos no momento da representação.

A especificidade é atribuída prezando pelo interesse dos grupos que utilizam daquele sistema de busca e depende do nível de especificidade dos usuários (Lancaster, 2004). Ainda de acordo com o autor, enquanto a exaustividade diz respeito ao âmbito de cobertura, a especificidade refere-se à profundidade de tratamento do conteúdo. Estabelece-se a exaustividade como uma decisão da política de indexação, enquanto a especificidade é uma propriedade do vocabulário adotado na indexação (Lancaster, 2004, p. 203).

Foskett (1973), define a especificidade como sendo também a extensão dada ao documento, mas nesse caso, sendo relativa à extensão da precisão ao se especificar um assunto. O autor ainda ressalta que, quanto maior a especificidade, maior será a possibilidade de documentos

relevantes serem encontrados. Com uma precisão limitada, consequentemente, a relevância dos documentos será baixa.

Segundo Fujita (2004), a especificidade depende de uma decisão estabelecida pelo sistema: se essa será ou não uma recomendação dada ao indexador; e caso seja, o indexador deve realizar a leitura documental tendo em mente o nível de especificidade exigida pelo sistema. Ainda segundo a autora, “a especificidade está relacionada ao nível de abrangência que a biblioteca e a linguagem documentária permitem especificar os conceitos identificados documento.” (Fujita, 2004, p. 85).

Fujita (2004) ainda demonstra, com o seguinte exemplo, como os dois processos – exaustividade e especificidade – podem trabalhar de maneira conjunta adequadamente: caso um livro trate especificamente do tema “tilápia, pacus e lambaris”, e o indexador optar também pela exaustividade, o usuário que buscar pelo termo “peixes” também encontrará o mesmo documento.

Tem-se, portanto, que esses são dois processos complementares, que dizem respeito diretamente à abrangência e à profundidade. Estudos realizados na área (Milani, 2007; Guimarães et al, 2008; Evangelista, 2013), demonstram que esses dois aspectos da indexação, podem ser vistos como valores éticos.

### 4. Ética em representação do conhecimento

Essa dimensão das atividades de indexação faz com que se remonte à Ética, enquanto ramo da Filosofia, nascida na Grécia Antiga, cuja reflexão está centrada no estudo das normas, princípios e condutas que orientam o homem na vida em sociedade, a partir da busca por um bem-estar individual e social. Dela decorre a Moral, que situa as questões éticas a uma dada realidade de costumes e valores, no tempo e no espaço (Guimarães, 2007).

A questão ética, apenas mais recentemente vem sendo objeto de pesquisa em organização do conhecimento, no campo da Ciência da Informação, ainda que, há mais de duas décadas, Ingetraut Dalghberg (1992), fundadora da ISKO, já enunciava fatores alarmantes que levavam a tal necessidade. Passada mais de uma década, López-Huertas (2008), ao analisar as tendências de pesquisa em organização do conhecimento na atualidade, ainda apontava a falta de suficiente zelo com o assunto.

Guimarães (2000), referindo-se ao universo da organização e representação do conhecimento,

mais especificamente no que se refere ao aspecto temático destaca cinco níveis de compromissos éticos: com as necessidades informacionais do usuário, com os objetivos da organização, com a especificidade e a natureza da própria informação disponibilizada, com o bom nome da profissão e com o próprio universo de valores do indexador. Observam-se, assim, cinco níveis de valores a serem considerados: os do usuário, os da organização, os inerentes ao conteúdo informacional, os da profissão coletivamente considerada e os do profissional indexador em sua individualidade. Assim, a questão dos valores é central na discussão das questões éticas, pois revela um universo de normas de conduta que uma dada sociedade considera importantes para atingir o bem-comum. Da negativa desses valores nascem os denominados problemas éticos, que a sociedade considera como impeditivos para que se atinja o bem-comum.

Ao analisarem a literatura internacional da área no período de 1995-2004, relativamente às questões éticas em organização e representação do conhecimento, Guimarães et al. (2008) depararam-se com um núcleo axiológico composto por três níveis de valores: a) valores maiores (ou supervalores) que permeiam toda a atividade informacional, b) valores antes havidos como requisitos profissionais, na medida em que integram a essência do fazer profissional na área; c) valores antes havidos como meras medidas de ORC, mas que hoje se integram ao universo axiológico da área, como Precisão (também entendido como Especificidade). Decorrendo da negativa dos valores, surgem os problemas que, de acordo com Guimarães et al. (2008), podem ser categorizados em dois contextos: os Problemas que permeiam o mundo atual; e os Problemas que afetam diretamente as atividades de ORC, quando do exercício profissional.

Milani (2007) e Evangelista (2013) analisaram, respectivamente aos períodos de 1995- 2004, e de 1990-1994; 2005-2010, a presença de valores e problemas éticos na literatura internacional de organização e representação do conhecimento, identificando um conjunto de 54 valores éticos, com a preponderância de Privacidade, Garantia cultural, Precisão (relativo à especificidade), Autoria e Exaustividade. Relativamente aos problemas, obteve-se um conjunto de 44, sendo predominantes: Vigilância, Negligência, Censura Direcionamento informacional, Falta de garantia cultural e Má representação.

Viu-se, portanto, a necessidade de realizar a análise de conteúdo de maneira que se demonstre o que a literatura internacional da área

vem dizendo a respeito desses dois processos e em que âmbito eles são entendidos.

## 5. Material e métodos

A pesquisa demonstra-se ser qualitativa, uma vez que se utiliza inferências e interpretações para demonstrar como se estabelecem dois conceitos, de acordo com a literatura.

A análise de conteúdo caracteriza-se por possibilitar “uma análise com base em inferências extraídas de conteúdos de documentos – a partir de uma interpretação controlada por meio de variáveis ou indicadores” partindo de “definições concebidas por pesquisadores, cujas características relativas à forma são homogêneas: argumentos descritivos registrados textualmente” (Guimarães & Sales, 2010, p. 05).

O método foi aperfeiçoado por Bardin (2003), que o descreve como um conjunto de métodos aplicáveis a discursos, qualquer que seja sua natureza, possuindo um ponto principal: a inferência. A autora ainda ressalta, que o objetivo dessa análise são dois: a ultrapassagem da incerteza e o enriquecimento da leitura, e ainda possui duas funções: a função heurística e a administração da prova (hipóteses).

Primeiramente, é necessário que se realize uma fase de pré-análise, que de acordo com Salles (2008) caracteriza-se por ser “a escolha dos documentos que serão submetidos à análise, (...) construção de um corpus de análise, a formulação de hipóteses e objetivos, e a elaboração de indicadores fundamentais à interpretação dos resultados.” (SALLES, 2008, p. 100).

Foi definido como material para posterior aplicação da análise de conteúdo, artigos que contenham os termos exaustividade e especificidade em inglês – *exhaustivity* e *specificity*, respectivamente -, recuperados pelas bases de dados Scopus, SCIELO, Web of Science e LISA, através do campo de busca que compreende o título do artigo, o resumo e as palavras-chave. Os termos foram procurados separadamente e não se estabeleceu limite de tempo de publicação, mas foram selecionados apenas os artigos publicados em periódicos da área de Ciência da Informação.

Para criação do corpus de pesquisa, Bardin (2003) sugere quatro regras principais: regra de exaustividade, regra de representividade, regra da homogeneidade e regra da pertinência. Decidiu-se não utilizar-se da regra da exaustividade, visto que essa define-se como a não-seleção dos documentos encontrados.

O processo de coleta de dados gerou dois corpora de pesquisa, a saber:

Corpus 1: conjunto de todos os artigos recuperados pela base de dados. Foram recuperados 51 artigos, os quais 10 não foram encontrados na íntegra. Realizou-se a leitura dos resumos dos 41 recuperados, para saber se constituiriam o Corpus 2.

Corpus 2: textos selecionados por tratarem do assunto no âmbito da Ciência da Informação e ainda, conterem as variáveis de inferência definidas, resultando em um corpus de 27 artigos. Para tanto, leu-se os textos na íntegra.

Conforme recomendado por Bardin (2003), passa-se para a fase de elaboração de índices e indicadores. Para Salles (2008) essa é fase tem “a função de indicar quais são esses índices, e organizá-los sistematicamente com indicadores (frequência, em caso de análise quantitativa, e, presença, em caso de análise qualitativa) relativos a eles” (Salles, 2008, p. 106). Foi definido como variáveis de inferência três concepções: conceitos (as definições ou as funções dos termos), características (quais os atributos que caracterizam cada um dos termos) e objetivos (qual a intenção de cada um dos termos).

Posteriormente, passa-se a fase das inferências, onde se procura sintetizar em sentenças, o conteúdo dos documentos. A análise gerou dois quadros – um correspondente ao termo exaustividade e outro, correspondente ao segundo termo, especificidade – que serão explorados na seção a seguir.

## 6. Análise dos resultados

A seguir, procurou-se definir de forma resumida, o que os autores dos artigos do corpus de pesquisa, interpretam sobre esses dois processos da indexação. Em um primeiro momento, procura-se demonstrar as principais inferências dadas ao termo exaustividade:

No que se refere a categoria *Conceito*, duas dimensões se colocam: uma ligada à análise (matemática) de distribuição de ocorrência de dados e outra ligada à cobertura de termos utilizados na indexação.

Relativamente às *Catacterísticas*, foram identificados os seguintes atributos: Seletividade, Representatividade, Abrangência, Relação direta com a revocação, Relação inversa com a precisão, Gradatividade, Consistência, e Quantificação.

Em termos de *Objetivos*, destacam-se o estabelecimento de relações entre os documentos, a redução da densidade dos documentos encon-

trados, o fornecimento de subsídios à distinção entre documentos na recuperação e a promoção da relevância.

O que pode se perceber, dessa análise, é que, de maneira geral, a exaustividade é definida como a abrangência que o sistema permite indexar um documento e que está diretamente ligada à revocação (maior recuperação de documentos) do sistema. Também devemos chamar a atenção que, muitas vezes, ela é vista apenas como uma expressão matemática, não tomando consideração por elementos como as necessidades dos usuários, os objetivos da instituição e questões éticas. Outro fator importante é a constatação por alguns atores de que ela está inversamente relacionada à precisão e especificidade do sistema.

O resultado da análise de conteúdo da literatura sobre especificidade revela, em termos de *Conceito*, uma referência aos documentos diretamente ligados àquela temática e ao nível de minúcia dada a um documento, quando representado por meio de termos.

As *Características* da especificidade se revelam por meio do detalhamento, da profundidade, e de sua relação direta com a precisão e inversa à revocação.

No que se refere aos *Objetivos*, a especificidade se revela como uma estratégia para uma organização e representação do conhecimento efetiva por meio da análise e seleção de sentenças para posterior categorização.

De maneira geral, podemos identificar, na especificidade, a particularidade que será dada à indexação, na hora da escolha dos termos.

Podemos estabelecer algumas semelhanças e algumas diferenças entre os termos. Por semelhanças, podemos perceber que fica claro que os dois tratam de características do sistema. A questão matemática também está sempre presente, visto que muitas vezes, os processos são vistos apenas pensando na quantidade de termos que eles irão gerar. Os autores também evidenciam que os dois métodos ainda causam potenciais problemas em sua aplicação, fator que pode se dever aos poucos estudos na área.

Como diferenças, o principal aspecto é a desigualdade entre um termo estar totalmente relacionado à alta revocação enquanto outro, completamente relacionado à alta precisão.

## 7. Conclusões

A abordagem do processo de indexação a partir de dois dos micro-processos que a compõem -

a exaustividade e a especificidade – e a análise dessa temática a partir do aspecto conceitual na literatura especializada, com aporte da análise de conteúdo evidenciou que, exaustividade e especificidade são entendidos como atributos do sistema, no mais das vezes ligados à questão quantitativa (quantidade de termos que deve ser atribuído ao documento) sem maiores reflexões sobre as consequências dessas ações ou, em última instância, os valores envolvidos que integram a sua dimensão ética.

Isso leva à necessidade de uma ampliação de horizontes rumo a uma abordagem mais social (Hjørland, 2002) que vá além dos estritos limites do sistema para ir ao encontro do papel mediador que a indexação deve realizar entre o contexto de produção da informação (o conteúdo do documento), e o contexto de uso dessa informação (necessidades do usuário), aproximando-se do que Antonio García Gutiérrez (2002) denomina como “ética transcultural de mediação”.

**Agradecimentos:** Agradece-se o apoio da Fundação de Amparo à Pesquisa do Estado de São Paulo, FAPESP (proc. 2015/06278-6).

## Referências

- Bardin, L. (2009) *Análise de conteúdo*. Tradução de Luiz Antero Reto e Augusto Pinheiro. Lisboa: Edições 70..
- Dahlberg, I. (1992). Ethics and knowledge organization: in memory of Dr. S. R. Ranganathan in his centenary year. *International Classification*, 19:1, p. 1-2.
- Foskett, A. C. (1973) *A abordagem temática da informação*. Tradução de Antonio Agenor Briquet de Lemos. São Paulo: Polígono.
- Evangelista, I. V. (2013) *Ética em representação do conhecimento: uma análise de valores e problemas a partir da literatura internacional da área (1991-2010)*. Trabalho de Conclusão de Curso (Bacharelado em Biblioteconomia) - Universidade Estadual Paulista Júlio de Mesquita Filho, Marília.
- García Gutiérrez, A. (2002) Knowledge organization from a “culture of the border” towards a transcultural ethics of mediation. En: López-Huertas, M. J. (ed.). (2002) *Challenges in knowledge representation and organization for the 21st century: integration of knowledge across boundaries*. Würzburg: ERGON, p. 516-522.
- Guimarães, J. A. C. et al. (2008) Ethics in the Knowledge Organization Environment: an overview of values and problems in the LIS literature. En: Arsenault, C.; Tennis, J. T. (eds.). (2008) *Cultural and Identity in Knowledge Organization*. Würzburg: ERGON, p. 340-346
- Guimarães, J. A. C. (2000) O profissional da informação sob o prisma de sua formação. En: VALENTIM, M. L. P. (ed.). *Profissionais da Informação: formação, perfil e atuação profissional*. São Paulo: Polis, p. 53-70.
- Guimarães, J. A. C.; Sales, R. (2010) Análise documental: concepções do universo acadêmico brasileiro em Ciência da Informação. *Datagramazero*, 11:1.
- Hjørland, B. (2002) Domain analysis in information science. Eleven approaches - traditional as well as innovative. *Journal of the American Society for Documentation*, 58:4, p. 422-462.
- Hjørland, B. (2001) Toward a theory of aboutness, subject, topicality, theme, domain, field, content...and relevance. *Journal of the American Society for Information Science and Technology*, 52: 09, p. 724 – 778.
- Fujita, M. S. L. (2009). *A indexação de livros a partir da percepção de catalogadores e usuários de bibliotecas universitárias: um estudo de observação do contexto sociocognitivo com protocolos verbais*. São Paulo : Cultura Acadêmica.
- Lancaster, F. W. (2004) *Indexação e resumos: teoria e prática*. Tradução de Antonio Agenor Briquet de Lemos. Brasília: Briquet de Lemos.
- Leiva, I. G. (2012) Aspectos conceituais da indexação. In: LEIVA, I. G.; FUJITA, M. S. L. *Política de indexação*. São Paulo: Cultura Acadêmica, p. 31 – 106.
- López-huertas, M. J. (2008) Some current research questions in the field of knowledge organization. *Knowledge Organization*, 35: 2/3, p. 113-136.
- Sales, R. (2008) *Tesauros e ontologias sob a luz da teoria comunicativa da terminologia*. Dissertação Mestrado em Ciência da Informação. Forianópolis: PGCIN/ UFSC.

---

# La construcción de la Biblioteconomía como ciencia y su relación con la clasificación

*The Construction of Library Science and its Relationship with Classification*

---

Daniel Martínez-Ávila (1), José Augusto Guimarães (2)

(1) Department of Information Science, São Paulo State University – UNESP, Av. Higyno Muzzi Filho, 737, Marília (17525-900) – São Paulo – Brazil, dmartinezavila@marilia.unesp.br

(2) guima@marilia.unesp.br

## Resumen

La construcción histórica de la Biblioteconomía como ciencia presenta una serie de problemas terminológicos y discursivos que han ido conformando y transformando diversos conceptos. La narrativa oficial sobre el área presenta la clasificación como una de las prácticas discursivas clave en la emergencia de la Biblioteconomía, la cual también parece haber sido relegada por otras prácticas en la construcción de otros conceptos como el de la Ciencia de la Información. En el presente trabajo se utiliza un análisis del discurso de Foucault combinado con una investigación bibliográfica para analizar los principales discursos y narrativas relacionadas con el campo de la Biblioteconomía y la clasificación como práctica en el sentido de Foucault. El objetivo del estudio es revelar las discontinuidades discursivas entre diferentes áreas relacionados con la Biblioteconomía como ciencia en su relación con la influencia de la clasificación para la investigación y la enseñanza en el área.

**Palabras clave:** Clasificación, Análisis de discurso, Biblioteconomía, Documentación, Ciencia de la Información.

## 1. Introducción

La construcción histórica de la Biblioteconomía como ciencia presenta una serie de problemas terminológicos y discursivos, tanto en los contextos hispano-hablantes como anglosajones, que han ido conformando y transformando diversos conceptos. La narrativa oficial sobre el área presenta la clasificación como una de las prácticas discursivas clave en la emergencia de la Biblioteconomía, la cual también parece haber sido relegada por otras prácticas en la construcción de otros conceptos como el de la Ciencia de la Información. Estas prácticas discursivas y de control influyen no solo en la delimitación histórica y actual de las áreas sino también en la configuración de los planes de investigación y estudios en relación a la clasificación. Por otra parte, desde la fundación de la ISKO, siendo la revista *Knowledge Organization* uno de sus principales canales para la difusión de su

## Abstract

The construction of Library Science presents several terminological and discursive problems that have shaped and transformed several related concepts. The official narrative of the area presents classification as one of the key discursive practices in the emergence of Library Science, while also being relegated by other practices in the construction of other concepts such as Information Science. In this paper we use a combination of Foucauldian discourse analysis and bibliographic research to study the main discourses and narratives related to the field of Library Science, and classification as a practice in the sense of Foucault. The aim of the study is to reveal the discursive discontinuities between different areas related to Library Science regarding the influence of classification on research and teaching.

**Keywords:** Classification, Discourse analysis, Library science, Documentation, Information science.

investigación, la visión armonizada de la clasificación ha sido contemplada no solo en sus discursos seminales, sino también ratificada tras los principales cambios editoriales e incluso nominales de la revista. Todos estos discursos afectan la Organización del Conocimiento y su estudio, tanto en los procesos identitarios del área como en las posibilidades de diálogo e investigación.

La metodología del presente trabajo es considerada un análisis del discurso de Foucault (1972) que consiste en la utilización de técnicas historiográficas para la revelación de continuidades y discontinuidades discursivas en los diferentes momentos de la construcción de un concepto. Para ello, se combina con una investigación bibliográfica en la que se analizan los principales discursos y narrativas relacionadas con el campo de la Biblioteconomía y la clasificación, como práctica en el sentido de Foucault.

## 2. El problema de los aspectos terminológicos en la Biblioteconomía como ciencia

Birger Hjørland (2003) afirma que la organización del conocimiento (KO) es un concepto más amplio que la Biblioteconomía y Ciencia de la Información (pp. 87-88). La mera traducción de esta afirmación, que no idea, ya supone un problema para aquellos hispano-hablantes que han visto como la construcción histórica del área y sus estudios, y especialmente en España, han utilizado diferentes términos para traducir el concepto. Específicamente, Hjørland utiliza la expresión (L)IS, por *Library and Information Science* (Biblioteconomía y Ciencia de la Información), omitiendo como alternativa la palabra Biblioteconomía ("*Library*-" pero no "*Science*", que va junto a "*Information*"), para simplificar al denotar el área.

En el caso de España, ese *Information Science*, fue trasladado al área académica, al menos en los contenidos de su enseñanza universitaria, en un primer momento, a las Facultades de Ciencias de la Información (hoy Facultades de Comunicación). Ya desde esos orígenes en España, la articulación entre los estudios de la información/documentación y comunicación ha sido tan estrecha que desde los inicios de la formación universitaria en ambos campos (1971), se produjo una identificación, todavía no suficientemente y, a nuestro juicio, no necesariamente anulada: denominar los estudios de comunicación "Ciencias de la Información", justamente abordándose la investigación y la formación en Documentación (*Information Science*) desde las Facultades de Ciencias (en plural) de la Información (*Communication Science*). Una lectura atenta de la narrativa seminal de López Yepes (1995), publicada originalmente en 1978 bajo el título "Teoría de la documentación", podría interpretarse como una apología (o argumentación, según la perspectiva) de su integración en el campo de la Comunicación (con continuas referencias a la Teoría matemática de Shannon y Weaver o, por ejemplo, citas como "Ciencias de la Información (*Information Sciences*) vienen a unir las áreas teóricas de las ciencias de la comunicación en general con las aplicaciones de recuperación de la información" (p. 163), donde debe notarse el plural de ese "ciencias"/"sciences"). A pesar de la desestructuración de los planes de los estudios en los ochenta y posteriormente mucho más profunda en los años 2000 con el Plan Bolonia, la formación de algunas prácticas de organización del conocimiento/documentación en medios de comunicación en la actualidad todavía debe buscarse en los planes de estudios de Grados Periodismo y

Comunicación Audiovisual (García-Gutiérrez y Martínez-Ávila, 2014). La transformación de los planes de estudios en LIS en España, pasó por diferentes estados a partir de los años ochenta, a resaltar la diplomatura en Biblioteconomía y Documentación y licenciatura en Documentación hasta la ejecución del Plan Bolonia, y los actuales Grados en Información y Documentación. Con este último momento transformativo parece silenciarse una denotación "bibliotecológica" que sin embargo sí que se mantiene en sus contenidos.

En el caso de Brasil, la tabla de áreas de conocimiento del Consejo Nacional de Desarrollo Científico y Tecnológico (CNPq - *Conselho Nacional de Desenvolvimento Científico e Tecnológico, Conselho Nacional de Pesquisa* hasta 1971), delimitan intelectualmente y establecen los representámens del campo. Aquí (CNPq, [2013]) el término inclusivo, Ciencia de la Información (*Ciência da Informação*, 6.07.00.00-9), ampara en un segundo nivel jerárquico a la Teoría de la Información (*Teoria da Informação*, 6.07.01.00-5), la Biblioteconomía (*Biblioteconomia*, 6.07.02.00-1) y la Archivística (*Arquivologia*, 6.07.03.00-8). Como campos "hermanos", pero fuera de la Ciencia de la Información, se presenta la Comunicación (*Comunicação*, 6.09.00.00-8, que incluye la *Teoria da Comunicação*, 6.09.01.00-4, y *Jornalismo e Editoração*, 6.09.02.00-0) y, sorprendentemente, la Museología (*Museologia*, 6.08.00.00-3), la cual es tradicionalmente considerada juntamente y no separada de la Archivística y la Biblioteconomía y Documentación como "las 3 Marías" (Smit, 1993). Numerosos estudios y autores difundidos en Brasil han señalado la Biblioteconomía y la Ciencia de la Documentación como dos campos diferentes aunque con estrechos lazos y puntos en común. En este sentido, por ejemplo, Tefko Saracevic ha señalado que: "El campo común entre la Biblioteconomía y la Ciencia de la Información, que es bastante fuerte, consiste en la compartición de su papel social y su preocupación común con los problemas de efectiva utilización de registros gráficos. Pero también existen diferencias significativas en algunos aspectos críticos [...] Todas estas diferencias demuestran la conclusión de que la Biblioteconomía y la Ciencia de la Información son dos campos diferentes, con una fuerte relación interdisciplinar y no un único campo, en el que uno consiste en la manifestación especial del otro" (Saracevic 1996, 49; Saracevic 1999, 1058, traducción propia). La consideración de de la Biblioteconomía como un aspecto diferente pero con puntos en común con la Ciencia de la Información también ha sido señalada por Barbosa et al. (2000), que además incluyen entre los argumen-

tos y justificaciones para el cambio de nombre de la *Escola de la Universidade Federal de Minas Gerais* (UFMG) las directrices curriculares de la educación nacional de Brasil: "las directrices curriculares de la educación nacional, del Ministerio de Educación (*Ministério da Educação* - MEC), incluyen la biblioteconomía, la archivística y la museología como disciplinas de la Ciencia de la Información" (p. 82).

Birger Hjørland (2000a) también hablaba a propósito del cambio de nombre en 1997 de la '*Royal School of Librarianship*' en Copenhague por '*Royal School of Library and Information Science*', de los orígenes y peculiaridades terminológicas del área. En el caso de la prestigiosa universidad danesa, este cambio fue visto como una evolución (transformación) natural en la que no hubo una sustitución del término Biblioteconomía sino una adición del término Ciencias de la Información. Sin embargo, este respeto terminológico-nominal por la Biblioteconomía tampoco es la tendencia mundial: la mencionada sustitución de la Biblioteconomía y Documentación en España por los Grados en Información y Documentación sería un ejemplo, la cada vez más divergente división en *ischools* (i por *information*, nada que ver con el prefijo de mercadotecnia de Cupertino) y *library schools* en los Estados Unidos sería otro, pese a mantenerse abiertamente los contenidos y programas sobre Biblioteconomía en las *ischools* y siendo éstas acreditadas por la *American Library Association* - ALA.

En el trabajo de Hjørland, se estudiaban los problemas de etiquetar los campos de la "biblioteca, documentación e información" con especial énfasis en los términos "información" y "documento", además de las influencias introducidas por el término "información" en el campo bibliotecario (o biblioteconómico) a mitades del siglo XX. La importancia de análisis como éste puede justificarse desde la importancia del análisis de los nacimientos de prácticas discursivas en los análisis de discurso de Foucault (1972), referencial metodológico adoptado para el presente estudio, hasta la observación de Vakkari en la Biblioteconomía que dice que "concepciones al respecto de la estructura y el alcance de una disciplina son siempre constructos sociales que determinan la inclusión de ciertos objetos en ese dominio y la exclusión de otros" (Vakkari, 1994, p. 1). Como punto de partida (emergencia) del término Biblioteconomía ('*library science*'), y aquí, aunque no es objeto del presente estudio, sí debe notarse las connotaciones terminológicas en lengua inglesa de las dos acepciones más comunes del término Biblioteconomía que son "*library science*", traducida a veces

como biblioteconomía-ciencia, y "*librarianship*"), Hjørland (2000a, p. 27) lo establece en el siglo XIX, cuando Martin Schrettinger introdujo el concepto "*Bibliothekswissenschaft*" en un libro de texto de 1808 (Schrettinger, según Kunze y Rückl 1974, p. 267), y en 1893 se formó el '*Department of Library Science*' en el Armour Institute en Chicago (el cual, según Solberg 2000, imitó la New York State Library School en Albany, reconociendo la ALA a ambas como líderes). Sin embargo, respecto al hito de la "*bibliothekswissenschaft*" (literalmente "*library engineering*", o "ingeniería bibliotecaria"), incluso desde la propia comunidad alemana nuestros días se ha discutido y cuestionado hasta nuestros días (por ejemplo Hauke et al. 2005) su equivalencia con la biblioteconomía como ciencia ("*library science*") y la línea genealógica de otros autores internacionales como Melvil Dewey, Paul Otlet, Suzanne Briet o S. R. Ranganathan que serán analizados a continuación.

### 3. La construcción de la Biblioteconomía como ciencia

Tal como señala Dias, en su análisis terminológico y de las relaciones entre la Biblioteconomía y las Ciencias de la Información (2000), la tradición de la Biblioteconomía ("*librarianship*") en el desarrollo de prácticas aplicables a los problemas de organizar y acceder a las informaciones (conocimiento) contenidas en los documentos se remonta a los tiempos de la biblioteca de Assurbanipal, que ya poseía una especie de catálogo de los libros que contenía (p. 72). Puede añadirse aquí que son claves en la construcción del área, en un primer momento, las instituciones, y en este caso las bibliotecas de la Antigüedad, como por ejemplo desde la biblioteca de Nínive en Mesopotamia hasta la Biblioteca de Alejandría, y en un segundo momento las personas encargadas de esas instituciones, es decir los bibliotecarios, como por ejemplo Calímaco, los monjes copistas de las bibliotecas de los monasterios medievales, etc. Sin embargo, la aparición de la Biblioteconomía como ciencia ("*library science*") surgió en un tercer momento, cuando algunos de esos problemas en organización y acceso de la información en bibliotecas fueron tornándose más complejos e hizo falta algo más que soluciones intuitivas de forma hubo una necesidad de sistematización de conocimientos como teorías y metodologías que llevaron a la formalización de un área de estudios.

En los Estados Unidos, el contexto de esta emergencia es la existencia de una larga tradición de desarrollo de práctica en bibliotecas, principalmente públicas, dentro de una política

de instrucción de inmigrantes, operarios, etc. con vistas a la formación de una ciudadanía (Ortega, 2004). Entre tanto, en 1876 surgieron dos instrumentos volcados a los resultados de problemas prácticos en bibliotecas: las Reglas para elaborar un catálogo-diccionario de Charles Cutter y la Clasificación Decimal de Dewey. Uno de los aspectos que hacen tan distintiva la aportación de Dewey en relación a la ciencia fue su método de trabajo. Aquí Dias señala (p. 72) que la tradición de esfuerzo reflexivo o teórico propia de la investigación universitaria puede remontarse, por lo menos, a Melvil Dewey con el desarrollo de su sistema de Clasificación Decimal en 1876, ya que para desarrollarlo estudió los sistemas de clasificación existentes e hizo visitas in situ a las bibliotecas que los utilizaban (utilizando una sistemática muy similar al proceso de creación científica). Hjørland también establece el nacimiento de la Biblioteconomía (y Ciencia de la Información, textualmente "LIS") como campo de estudio en 1876 cuando Melvil Dewey publicó su sistema de clasificación y las escuelas de Biblioteconomía ("*librarianship*") comenzaron a emerger (Hjørland 2003, p. 88). Birger Hjørland también resalta la importancia de la clasificación, como área, a partir de dicho momento: "Formalmente hablando, también ha sido una materia académica en los programas de Biblioteconomía y Ciencia de la Información [LIS] desde que Melvil Dewey (1851-1931) estableció la primera escuela de 'economía bibliotecaria' [library economy] en los Estados Unidos en 1876" (Hjørland 2012, p. 299).

Otro evento importante en la construcción de la Biblioteconomía como ciencia es la fundación de la *Graduate Library School* de la Universidad de Chicago, en la década de 1930 (Dias 2000, p. 72). La comúnmente llamada Escuela de Chicago, con la participación de otros agentes clave en la construcción del concepto de Biblioteconomía, como Jesse Shera y Peirce Butler aunque esta vez en un discurso académico, tenía como objetivo la formación de bibliotecarios en un nivel de pos-grado. Los estudios universitarios sobre la organización y acceso a la información en bibliotecas eran ofrecidos a estudiantes que ya tenían una formación universitaria previa en cualesquiera otra área estaban especializados (como por ejemplo Física, Química, Ciencias Sociales, etc.). La filosofía de estos estudios de pos-grado era la de formar bibliotecarios especializados a través de la enseñanza de la ciencia bibliotecaria, en un segundo momento, a personas que ya se habían especializado previamente en una área de conocimiento en la que pretendidamente iban a ejercer. Según Dias, el primer director de la Escuela, Louis Round Wilson, orientó la investigación princi-

palmente para el campo de las Ciencias Sociales, incluso privilegiando la aplicación de su metodología de investigación respectiva. Esta escuela de pos-grado en Biblioteconomía, y su enseñanza volcada a la investigación, fue el detonante de una explosión de investigación tanto dentro como fuera de las escuelas de Biblioteconomía en los temas más diversos. Solo a nivel de doctorado, Busha y Harter contabilizan 472 tesis en el periodo 1930-1972 (Busha y Harter 1980, p. 95). El contexto de esta emergencia de la Biblioteconomía como ciencia, o necesidad de estos estudios, era un periodo de entre-guerras (1919-1939) en el que ganan espacio las bibliotecas especializadas, con características específicas en términos de usuarios, colección, formas de organización, etc. lo cual exigía una mayor complejidad aun de teorías y procedimientos (Meadows 1990, p. 59; Hjørland 2000a, p. 28).

#### 4. La relación de la Biblioteconomía con la Documentación

Paralelamente a la aparición y crecimiento de la Biblioteconomía como ciencia también surge la Documentación, la cual, como área de estudios, también tuvo su origen en un momento de crecimiento de la producción científica mundial, en el contexto de la denominada segunda revolución industrial (Guimarães 2008, p. 34).

El término Documentación surge al inicio del Siglo XX, como consecuencia del crecimiento de colecciones especializadas en instituciones de investigación, para denominar una actividad ejercida a partir de documentos de las más diversas formas y en instituciones más amplias e híbridas que las bibliotecas tradicionales. En este sentido, la analogía entre las dos áreas es reseñable, si es que esta nueva etiqueta no se trata simplemente de una nueva denominación para la transformación del concepto. Para Bradford, la Documentación es "(...) el arte de recoger, *clasificar* y volver fácilmente accesibles los registros de toda actividad intelectual. Es el proceso por el documentalista puede proporcionar al especialista con la literatura existente en el campo de su investigación, con la finalidad de que pueda tomar pleno contacto con las obras anteriores en su terreno y, de esa forma, evitar dispersión de esfuerzo en la realización de una tarea ya ejecutada" (Bradford 1961, p. 68, traducción y énfasis propio).

Dias precisamente destaca que la Documentación corresponde, en cierta medida, con una Biblioteconomía ejercida en áreas especializadas de conocimiento (Dias 2000, p. 73), lo cual por otra parte también corresponde con lo que

significaba la Biblioteconomía como ciencia desde la fundación de los estudios de postgrado en la Escuela de Chicago. Es también irónico que uno de los mayores precursores de la Documentación en Europa, Paul Otlet (Bélgica, 1868-1944), con su *Traite de Documentation* (1934) en el cual introdujo formalmente el término Documentación en el universo de las ciencias, también mirara hacia uno de los padre de la Biblioteconomía en los Estados Unidos, y más concretamente también de la clasificación moderna, Melvil Dewey, para traducir y desarrollar la Clasificación Decimal (futura CDU) junto a Henri La Fontaine.

La artificial división entre clasificación para bibliotecas (Clasificación Decimal de Dewey) y clasificación para documentación (Clasificación Decimal Universal) que se promulgó desde determinados círculos en los posteriores años no dejó de ser un reflejo de resistencia económica editorial por parte de los editores de la primera ante otras clasificaciones técnicamente más adelantadas (es decir, facetadas, como la CDU o la Clasificación de Ranganathan) que amenazaban su nicho de mercado. Por otra parte, tal como señala Hjørland (2000a, p. 27), el propio Ranganathan utiliza indistintamente el término Biblioteconomía ("*library science*") en algunos textos (1948, 1957) con el de Documentación en otros. El estudioso de Paul Otlet W. Boyd Rayward dice que "el término 'documentación' es un neologismo inventado por Otlet para designar lo que hoy en día tendemos a llamar Almacenamiento y Recuperación de la Información. De hecho, no es excesivo reivindicar el *Traité* como uno de los primeros libros de texto en ciencia de la información" (Rayward 1994, p. 238, traducción propia). Por otra parte, Meadows también señala la importancia del rol y necesidades de las bibliotecas especializadas durante el periodo de entre-guerras, contraponiéndolas a las bibliotecas públicas, académicas y nacionales (Meadows 1990, p. 59). Sin embargo, y pese a que éste parece un argumento análogo al de la construcción de la Documentación, Meadows también resalta las diferencias entre la Documentación y la Biblioteconomía centrándose en el 'documento': "Las principales diferencias [entre '*library science*' y '*documentation*'] están identificadas en las áreas de la bibliografía y lo que fue llamado 'documentación'. Exactamente las diferencias entre estos nuevos 'documentalistas' y los bibliotecarios tradicionales no estaban bien definidas. Sin embargo, había un consenso general en que los documentalistas no estaban incumbidos solo por el tratamiento físico de documentos, sino, hasta un punto más amplio que los bibliotecarios tradicionales, con la explotación de la información contenida en los

documentos". Debe notarse, por lo tanto, que se explicita que las principales diferencias se encuentran en el documento y su naturaleza, no mencionándose por lo tanto diferencias en otras áreas relacionadas con la organización del conocimiento como pudiera ser la clasificación.

## 5. La relación de Documentación y la Ciencia de la Información en la construcción de la Biblioteconomía

Cuando se habla del documento como objeto de la Documentación y/o la Ciencia de la Información, uno de los principales autores/obras que se reconocen es Michael Buckland con su fundamental artículo "*Information as a thing*" (1991). Sin embargo, muy anteriormente, la francesa Suzanne Briet, reflexionando principalmente sobre la obra de Otlet, ya conceptúa el documento como objeto de la Documentación en unos términos muy similares a los que han sido aceptados posteriormente para la información como objeto de la Ciencia de la Información (carácter probatorio, conocimiento materialmente fijado, etc.). De hecho, las características de ese nuevo momento, o concepto, emergente son tan semejantes que no pocos estudios han sido dedicados a aclarar las diferencias.

La aparición de la Ciencia de la Información está normalmente establecida en el contexto de revolución científica y técnica que siguió el final de la Segunda Guerra Mundial, en un momento de gran auge tecnológico y científico con una explosión de información registrada (documentos) que necesitaban ser organizados. En ese contexto, Vannevar Bush publica su artículo "*As we may think*" (1945), que destacaba, entre otras cosas, la importancia de los registros ('*records*') para la ciencia, la contribución tecnológica al acceso a tales registros, la necesidad de crear "memorias auxiliares" para el tratamiento y almacenamiento de esa información. Tal como indica Barreto (2002, p. 69), la aportación de Bush fue indicar una mudanza de paradigma para el área de información en ciencia y tecnología que envolvía a profesionales, instrumentos de trabajo para almacenamiento y recuperación de la información y representación de la información para procesamiento y almacenaje y recuperación.

Aunque, tal como indican diversos autores (por ejemplo Ingwersen 1992a citado por Oliveira 1998; Foskett 1996; Dias 2000), el término Ciencia de la Información ya fuera utilizado en Inglaterra desde 1958 con la fundación del *Institute of Information Scientists*, el uso y popularización del término en los Estados Unidos data de 1962, durante la celebración del congreso en

el *Georgia Institute of Technology* en el cual participaron científicos, investigadores, especialistas en información y por supuesto también bibliotecarios y documentalistas, incluyendo representantes oficiales de la *American Library Association*, la *American Association of Library Schools* y el *American Documentation Institute*. Diversos autores (por ejemplo Taylor 1966 citado en Garcia 2002) se refieren a las conferencias de Georgia como el marco fundacional de la Ciencia de la Información, debido no solo a la relevancia de su contenido sino a la naturaleza de sus asistentes. Sin embargo, pocos años después de la celebración del Congreso, una de estas asociaciones oficialmente representadas en el congreso, el *American Documentation Institute* (fundado en 1937) cambió su nombre por *American Society for Information Science* (en 1968) y, finalmente, por *American Society for Information Science and Technology* (en el año 2000). Estos cambios de nombre han sido señalados por Hjørland (2000a, p. 28) como una señal inequívoca de un uso más extendido del término "Ciencia de la Información" en la actualidad.

Algunas definiciones icónicas de Ciencia de la Información ("*information science*") son: "disciplina que investiga las propiedades y comportamiento de la información, las fuerzas que gobiernan su flujo y los medios para procesarla buscando el mejor acceso y uso posibles" (Borko, 1968, p.3, traducción propia); la similar definición dada por la *American Society for Information Science* (ASIS) en 1975: "La Ciencia de la Información está incumbida con la generación, acumulación, organización, interpretación, almacenamiento, recuperación, divulgación, transformación y uso de la información, con particular énfasis en las aplicaciones de tecnologías modernas en estas áreas. Como disciplina, busca crear y estructurar un cuerpo de conocimiento científico, tecnológico y de sistemas relacionado con la transferencia de información. Tiene componentes tanto de ciencia pura (teórica) que investigan la materia sin considerar la aplicación, y componentes de ciencia aplicada (práctica), que desarrollan servicios y productos (Griffith 1980, p.5 citado en Hjørland 2000a, p. 29, traducción propia). Sin embargo, tal como también indica Hjørland, "la fundación de ASIS no resolvió los problemas teóricos relacionados con la materia de CI, que siempre ha sido un concepto impreciso" (p. 29). Un estudio de más de 700 definiciones de "ciencia de la información" ("*information science*") y sus antecedentes entre 1900 y 1981, realizado por Alvin M. Schrader (1983), determinó que "la literatura de la ciencia de la información es caracterizada por caos conceptual.

Este caos conceptual emana de la variedad de problemas en la literatura definitoria de la ciencia de la información: citas de definiciones previas sin sentido crítico; mezcla del estudio y la práctica; reclamación obsesiva de su estatus científico; una visión muy reducida de la tecnología; no consideración de la literatura con la etiqueta de científica o tecnológica; inapropiadas analogías; definición circular; y, la multiplicidad de nociones vagas, contradictorias y a veces bizarras de la naturaleza del término "información" (p. 99, traducción propia). Aunque los aspectos de organización del conocimiento y clasificación siempre se encuentran implícitos en las definiciones y fundamentos de Ciencia de la Información (especialmente cuando se habla de acceso, tratamiento y representación de la información en relación a la recuperación), sin embargo, no suelen ser siempre los aspectos históricos que las narrativas oficiales resaltan de esa nueva transformación del área o emergencia de nuevo concepto. Si normalmente se afirma que la organización de la información y la recuperación de la información son dos caras de la misma moneda, la Ciencia de la Información parecía enfocarse en la revelación de la cara que había permanecido más oculta, la de la recuperación, dejando implícitamente el concepto de clasificación como una característica identificativa de un estado anterior del área como es la Biblioteconomía.

## 6. Discusión sobre la clasificación como elemento clave de la Biblioteconomía y las Ciencias de la Información

Puede decirse que el proceso de clasificación es un aspecto natural e inherente de todos los seres humanos que está presente en todos los aspectos de su conocimiento del mundo. Francisco Javier García Marco indicaba al respecto que "la clasificación no es sólo una operación fundamental en las bibliotecas, en los archivos y en otros centros de información. La clasificación es un aspecto clave del comportamiento humano" (García Marco 1996, p. 110).

Sin embargo, cuando se estudia la percepción de la clasificación en el ámbito académico y universitario, ésta se encuentra casi siempre conectada de forma casi exclusiva a la formación práctica en el ámbito de las bibliotecas y la Biblioteconomía, siendo comúnmente excluida y posiblemente infrautilizada en otras áreas de la Ciencia de la Información en las que su enseñanza e investigación también podrían aportar valor. Incluso asumiendo un vínculo predominante entre la clasificación y el área de la Biblioteconomía, la relación de la Biblioteconomía con la Ciencia de la Información, respecto a la clasi-

ficación y la organización del conocimiento, también es un vínculo que debe ser clarificado para el desarrollo del estudio que aquí se presenta, ya que una de las premisas del estudio es la potencial utilidad de la clasificación otras áreas de la Ciencia de la Información. Sin embargo, de tratarse la Biblioteconomía de un subconjunto (o subclase) de la Ciencia de la Información, esto significaría también que existen aspectos de la Biblioteconomía que, ya que pertenecientes también a la Ciencia de la Información, no serán exclusivos de aquella y sí que podrán ser aplicados o contenidos en otros aspectos de ésta, como por ejemplo la Archivística, la Museología, u otras áreas de la Ciencia de la Información donde existan necesidades de organización y acceso a la información.

En este sentido, el presente estudio también parte de la premisa de que la clasificación no va unida a la Biblioteconomía per se, entendiéndose esta como los aspectos específicos e identificativos de su aplicación en bibliotecas, sino a los aspectos y problemas que van ligados a la Biblioteconomía pero también a otras manifestaciones de la Ciencia de la Información, como por ejemplo la organización del conocimiento o el acceso a la información. En este sentido, Dias indica que: "separadamente, la biblioteconomía y la ciencia de la información deben ser vistas como especialidades de ese todo mayor. La separación es justificada por las diferencias que efectivamente existen, pero hay un problema común: el acceso a la información" (Dias 2000, p. 67, traducción propia). Dias también señala que "la ciencia de la información, por su parte, se subdivide en especialidades como teoría de la información, procesos de comunicación, representación de la información, teoría de la clasificación y archivística" (Dias 2000, p. 69). Y ya que en esa clasificación existen aspectos que no son excluyentes como la representación de la información y la archivística, o la teoría de la información y los procesos de comunicación, también parece plausible afirmar que la teoría de la clasificación podría encontrarse en otros aspectos de esa tipología o de su raíz la Ciencia de la Información.

Por otra parte, Dias también identifica la clasificación como una de las funciones altamente profesionales de las bibliotecas, aunque también afirmando que: "una comparación, aunque superficial, entre la biblioteca y los nuevos sistemas que eventualmente podrán sustituirla, justifica formular la hipótesis de que prácticamente todas esas funciones permanecerán necesarias y deberán ser realizadas de una forma o de otra, sea por intermediarios humanos, sea por sistemas automatizados. [...] Está claro que

la localización de información en la red presupone algún tipo de clasificación" (Dias 2000, p. 71, traducción propia). Queda implícito pues que el concepto de clasificación es considerado por Dias como una de las funciones profesionales que son necesarias en otros tipos de sistemas áreas más amplias que la de la Biblioteconomía, como sería la Ciencia de la Información. En este sentido, la clasificación también aparece listada por Birger Hjørland como una de las funciones de trabajo de tanto bibliotecarios como de documentalistas como de especialistas de información, es decir, de varios diferentes subtipos de lo que se entendería por Ciencia de la Información: "Las funciones de trabajo típicamente relacionados con lo esencial de la competencia de los bibliotecarios, documentalistas y especialistas en información se ocupan de la búsqueda de información en bases de datos, en Internet, en bibliotecas, etc. También incluyen búsqueda de información organizada en diferentes tipos de sistemas, tales como los sistemas de clasificación enumerativos, sistemas de clasificación basados en facetas, sistemas de texto libre, etc." (Hjørland 2000b, p. 507, traducción propia). Y En este sentido, García Marco también incluye el concepto de clasificación entre los procesos básicos de la Documentación que podrían atribuirse a sus competencias profesionales: "La clasificación es una fase esencial de proceso documental, porque conecta cada documento con el conjunto documental al que pertenece a partir de una organización conceptual previa" (García Marco 1996, 141).

Por otra parte, se debe ser consciente también de que la inclusión de aspectos como la investigación sobre clasificación en la Ciencia de la Información tampoco es algo que haya sido unánimemente aceptado por la comunidad científica, algo que, como Hjørland indica, también podría afectar a otros aspectos como la "teoría de la información", que otros investigadores adscribirían exclusivamente a la Ciencias Informáticas (Computación): "El enfoque analítico de facetas de Ranganathan contiene una teoría temática, que yo llamaría una teoría. Sin embargo, otros autores, entre ellos Ellis (1996), no cuentan la investigación en clasificación como parte de la CI. [...] Sin embargo, lo que se llama 'teoría de la información' no es considerado por muchas personas -este autor incluido- como una teoría de CI, sino una teoría en ciencias de la computación (comparar con Wersig, 1996)" (Hjørland 2000b, p. 518, traducción propia). Este aspecto podría estar relacionado también con los distintos vertientes y significados que podría tener el concepto de clasificación en el área, siendo su estudio y relaciones con el campo de la Ciencia de la Información una de

los objetivos de este estudio. Por ejemplo, Hjørland, en su "*Lifeboat for Knowledge Organization*", hace una distinción entre los conceptos de clasificación, clasificación de las ciencias y sistema de clasificaciones. A este respecto Hjørland indica solo el caso de clasificación de las ciencias que se trata de uno de los conceptos centrales en LIS ("Biblioteconomía y Ciencia de la Información"), mientras que solo en el caso de sistema de clasificaciones se explicita que se trate de un concepto de *Library Science* (Biblioteconomía como ciencia) y es incluido en su "*Core Concepts in LIS*". Hjørland define la clasificación como "el proceso de asignar elementos o unidades a las clases (o tipos o taxones) según algunos criterios. El sistema de clases en el que las unidades pueden ser clasificadas se llama sistema de clasificación" (Hjørland 2010, traducción propia).

Respecto a la clasificación, Hjørland también indica que: "Clasificamos cosas, personas, etc., en nuestra vida cotidiana, en nuestra vida profesional, en las ciencias. En la enseñanza clasificamos los temas que se imparten en función de su dificultad y el rango en el que son adecuados para ser enseñado. En las bibliotecas clasificamos los documentos con el fin de facilitar la recuperación de información, la navegación y otras formas de obtener documentos relevantes. Los criterios de clasificación (y organización del conocimiento) están por lo tanto profundamente relacionados con la relevancia. No son tanto las cualidades 'inherentes' de los objetos tanto como lo son los criterios pragmáticos que determinan los criterios de clasificación" (Hjørland 2010, traducción propia). Finalmente, Hjørland también señala que una división bastante extendida entre tipos de sistemas de clasificación y formas de clasificar es la distinción entre clasificación Aristotélica y la Teoría de prototipos de Rosch (1978), a la vez que expone las divergencias entre los puntos de vista sobre la clasificación y sus teorías. Respecto a los sistemas de clasificación, Hjørland remarca su vínculo con las clasificaciones bibliotecarias y con la organización del conocimiento, un concepto que, sin embargo, desde el punto de vista amplio de la Ciencia de la Información, no tendría porque ser exclusivo ni de la Biblioteconomía ni únicamente de las clasificaciones bibliotecarias: "El vocablo organización del conocimiento en Biblioteconomía y Ciencia de la Información [LIS] está muy ligada a la clasificación bibliotecaria. [...] El origen de la frase organización del conocimiento frase en Biblioteconomía y Ciencia de la Información [LIS] está claramente relacionada con sus obras, según la cual la clasificación de libros es básicamente organización del conocimiento y el conocimiento necesario para clasificar libros

proviene de la producción del conocimiento, de la cual los libros son la expresión tangible" (Hjørland 2009, traducción propia).

Sobre el concepto de clasificación de las ciencias, Hjørland indica: "La clasificación de las ciencias [*classification of the sciences*] es un campo interdisciplinario relacionado con la construcción de sistemas de clasificación bibliotecarios como la Clasificación Decimal Dewey (que es principalmente una clasificación de disciplinas), así como el mapeo de la ciencia en cienciometría (cf., atlas de la ciencia [*atlas of science*]). No debe confundirse con la clasificación en las ciencias [*classification in the sciences*]: cómo los científicos clasifican sus objetos, por ejemplo, cómo las plantas son clasificadas en la botánica (cf., clasificación científica y taxonomía [*scientific classification and taxonomy*]), aunque estos temas están relacionados a un nivel más profundo, porque las ciencias son en parte organizadas por las estructuras que descubren en los objetos que investigan (cf., organización social del conocimiento [*social organization of knowledge*])" (Hjørland 2008, traducción propia). A continuación, Hjørland cita a Francis Miksa (1998) cuando afirma, por una parte, que el movimiento de clasificar el universo del conocimiento que surgió renovado en el Siglo XVII, y que gozó de enormes proporciones durante el Siglo XIX, terminó su interés después del comienzo del Siglo XX; y, por otra parte, que no existe una fuerte relación entre el viejo movimiento de clasificar el conocimiento y las ciencias con el movimiento de clasificar las bibliotecas que surgió a finales del S. XIX. Hjørland resalta la importancia de los puntos de Miksa, indicando que "estas dos cuestiones: porqué la investigación en la clasificación de las ciencias disminuyó y porqué la investigación en clasificación en Biblioteconomía y Ciencia de la Información (LIS) permanece prácticamente sin relación con este campo, son importantes de considerar porque, para él, unos fundamentos apropiados de la Biblioteconomía y la Ciencia de la Información probablemente deberían estar relacionados con estas cuestiones" (Hjørland 2008, traducción propia).

## 7. La clasificación en el universo ISKO

Respecto a su relación con el área de la Organización del Conocimiento, uno de los principales foros internacionales para la difusión de la investigación y desarrollo en el área son las actividades de la ISKO (*International Society for Knowledge Organization*), siendo la revista *Knowledge Organization* uno de los principales canales de publicación, discusión y conformación de los conceptos. La revista *Knowledge*

*Organization* fue fundada en 1973 con el título *International Classification*, pasándose a llamar *Knowledge Organization* en 1993 (volumen 20) y siguiendo con su publicación hasta nuestros días. Aun después del cambio de nombre, la revista sigue definiéndose en la cabecera como una "revista internacional dedicada a la Teoría de los Conceptos, la Clasificación, la Indexación y la Representación del Conocimiento", lo que deja patente que pese al cambio de nombre no se ha dejado de lado el interés y dedicación de sus contenidos hacia la clasificación. En el primer número del año 2000, Hope Olson revisaba y reafirmaba en una editorial la misión, finalidad y objetivos originales de la revista, tal como fueron declarados en el primer número. Entre los estatutos originales ratificados por Olson, cabría destacar para el interés del presente estudio el punto 1.3 de la misión, donde se explicita que la misión de la revista *Knowledge Organization* es: "encontrar una base común para las metodologías de clasificación que han sido desarrolladas en diferentes campos de materia: biblioteconomía, ciencia de la información, ciencias naturales, administración industrial, lingüística y filosofía" (Olson 2000, p. 1, traducción propia). Es por tanto que, tanto en la misión original de la revista como en la ratificación del año 2000 de Olson, se sugiere que existen metodologías de clasificación no solo en Biblioteconomía sino también en la Ciencia de la Información, siendo además uno de los objetivos de la revista el encontrar una base común para estos.

Por otra parte, la relación entre la clasificación y la organización, dentro del contexto de la revista *Knowledge Organization*, su título y su enfoque, también fue discutida por una editorial de Hope Olson en el 2001, concluyendo que "quizás por lo tanto el cambio de *International Classification* a *Knowledge Organization* señaló la sofisticación de nuestro campo como uno que abarca varios medios, no solo la clasificación, y busca armonía entre ellos" (Olson 2001, p. 3, traducción propia). Esta editorial también fue comentada en subsiguientes cartas al editor en las que se analizaron y discutieron los diferentes matices entre ambos términos y la pertinencia de mantener el término clasificación en la cabecera de la revista (aun fuera del título). Como conclusión, la inclusión del concepto de clasificación dentro de la revista *Knowledge Organization* y el entorno ISKO no solo parece partir desde el punto de vista de la Biblioteconomía, sino, tal como la misión de la revista indica, también desde la Ciencia de la Información y otras áreas afines para las que se busca una participación de artículos. En este sentido, debe resaltarse también la voluntad de la revista por armonizar y

encontrar una base común para una metodología y teoría de la clasificación que se presupone a todas ellas, que de haberse cumplido los puntos de la misión tanto de la fundación de la revista como de su ratificación, deberían existir en la literatura de *Knowledge Organization* y otras publicaciones de ISKO, como las actas de sus congresos.

## 8. Consideraciones finales

En este trabajo se ha estudiado la construcción histórica del discurso de Biblioteconomía como ciencia en relación con otros discursos como el de la Documentación y la Ciencia de la Información en visas de determinar el rol que juega de la organización del conocimiento, y más específicamente la clasificación, en la construcción de esas áreas y la posibilidad de convergencia tanto en enseñanza como de investigación.

En un análisis histórico-conceptual de la Ciencia de la Información, Rafael Capurro resalta que la Ciencia de la Información tiene su origen en dos vertientes: la Biblioteconomía clásica (centrada en los contenidos de los mensajes), y la Computación Digital (representada por el trabajo de Vannevar Bush, más centrada en la Comunicación) (Capurro, 2003, p.8). A estas vertientes americanas sería interesante añadir una tercera quizás de más arraigo en Europa, la Documentación (representada por Otlet y más centrada en los documentos). Pese a sus diferencias y los autores que las defienden, tal como indican Barbosa et al. (2000, p. 84), diversos teóricos contemporáneos han presentado argumentos a favor de la idea de la Ciencia de la Información y la Biblioteconomía como campos convergentes, que en realidad deberían constituir una única disciplina (por ejemplo Wilson 1983, Lancaster 1984, Ford 1990, Miksa 1992, Ingwersen 1992b, Wersig 1992 y Vakkari 1994). Reforzado por la no hegemonía de estas prácticas discursivas, tal como también reconocen Barbosa et al. (2000, p. 84), los procesos estudiados por la Biblioteconomía tradicional pueden ser generalizados para otros contextos más allá de la institución bibliotecaria, sin limitaciones académicas entre áreas sobre lo que es pertinente o no para los campos. Sin embargo, de aceptarse la continuidad de la formación discursiva también se estarían eliminando los límites impuestos entre las áreas, entre sus identidades. En el campo de la clasificación, uno de los argumentos que se han detectado sobre el establecimiento del hito de emergencia de la Biblioteconomía (y la clasificación), en contraposición al hito posterior de la aparición de la Ciencia de la Información, es el método de trabajo de Melvil Dewey en la

elaboración de su sistema, acercado a los métodos científicos.

Algunos argumentos y posibles líneas en la investigación futuras es el argumento de existencia de epistemologías subyacentes en los antecedentes a la Clasificación de Dewey. Por ejemplo, Pauline Rafferty (2001) estudió los valores subyacentes en los sistemas de clasificación deconstruyendo en ese sentido la división entre las clasificaciones que supuestamente tienen una base metafísica, como Dewey, y aquellas que supuestamente son meras ordenaciones prácticas en pos de la conveniencia inmediata, como la clasificación de los libreros de París desarrollada por Brunet. En este sentido, si se difuminan las líneas entre los diferentes momentos de la clasificación como agente determinante de la construcción de la ciencia (la Biblioteconomía), parece plausible también permeabilizar las discontinuidades discursivas entre sus diferentes momentos históricos, etiquetadas también como Biblioteconomía y Ciencia de la Información, lo que desestabilizaría también el argumento de la supresión de la investigación y enseñanza de la clasificación en otros curso y áreas del campo como por ejemplo la Archivística.

## Referencias

- Barbosa, Ricardo Rodrigues; Cendón, Beatriz Valadares; Caldeira, Paulo da Terra; Bax, Marcello Peixoto. (2000). Novo nome e novo paradigma: da biblioteconomia à ciência da informação. *Perspectivas em Ciência da Informação*, 5:n.especial (jan./jun. 2000) 81-91.
- Barreto, A. A. (2002). A condição da informação. *São Paulo em Perspectiva*, 16:3, 67-74.
- Borko, H. (1968). Information science: what is it? *American Documentation*, 19:1, 3-5.
- Bradford, S.C. (1961). *Documentação*. Rio de Janeiro: Fundo de Cultura.
- Briet, S. (1951). *Qu'est-ce que la documentation?* Paris: Industrielles et Techniques.
- Buckland, Michael K. (1991). *Information and information systems*. New York: Greenwood.
- Bush, V. (1945) As we may think. *Atlantic Monthly*, 176:1, 101-108.
- Busha, C.H.; Harter, S.P. (1980). *Research methods in librarianship: techniques and interpretation*. New York: Academic Press.
- Capurro, Rafael (2003). Epistemologia e ciência da informação. En: *Encontro nacional de pesquisa em ciência da informação*, 5., 2003, Belo Horizonte. Anais... Belo Horizonte: UFMG.
- CNPq (2013). *Tabela de Áreas do Conhecimento* Recuperado 4-05-2015 de <http://www.cnpq.br/documents/10157/186158/TabeladeAreasdoConhecimento.pdf>
- Dias, E. W. (2000). Biblioteconomia e ciência da informação: natureza e relações. *Perspectivas em Ciência da Informação*, 5:n. especial (jan./jun. 2000) 67-80.
- Ford, B. (1990). The library as a locus. En: *Information science: the interdisciplinary context*. New York: Neal-Schuman, 115-131.
- Foskett, A.C. (1996). *The subject approach to information*. 5th ed. London: Library Association.
- Foucault, Michel (1972). *The archaeology of knowledge and the discourse on language*. New York: Pantheon Books.
- Garcia, Joana Coeli Ribeiro. (2002). Conferências do Georgia Institute of Technology e a Ciência da Informação: "de volta para o futuro" *Informação & Sociedade: Estudos* 12:1. Recuperado 4-05-2015 de [www.brapci.ufr.br/download.php?dd0=13364](http://www.brapci.ufr.br/download.php?dd0=13364)
- García-Gutiérrez, Antonio; Martínez-Ávila, Daniel (2014). Formación crítica de documentalistas en medios de comunicación. *El profesional de la información*, 23:5, 493-500.
- García Marco, Francisco Javier (1996). Contexto y determinantes funcionales de la clasificación documental, *Scire*, 2:1, 109-145.
- Griffith, B.C. (1980). *Key papers in information science*. New York: Knowledge Industry Publications.
- Guimarães, José Augusto Chaves. (2008). Ciência da informação, arquivologia e biblioteconomia: em busca do necessário diálogo entre o universo teórico e os fazeres profissionais. En: Guimarães, José Augusto Chaves; Fujita, Mariangela Spotti Lopes. (eds.). *Ensino e Pesquisa em Biblioteconomia no Brasil: a emergência de um novo olhar*. Marília: Fundepe.
- Hauke, Petra; Grünwald, Jana; Kaden, Ben; Kaufmann, Andrea; Kindling, Maxi. (2005). Library Science - quo vadis? (Re)Discovering „Bibliothekswissenschaft“ En: *World Library and Information Congress: 71th IFLA General Conference and Council "Libraries - A voyage of discovery" August 14th - 18th 2005*, Oslo, Norway
- Hjørland, Birger. (2000a). Documents, memory institutions and information Science. *Journal of documentation*, 56:1, 27-41.
- Hjørland, Birger. (2000b). Library and information science: practice, theory, and philosophical basis. *Information Processing and Management*, 36, 501-531.
- Hjørland, Birger. (2003). Fundamentals of Knowledge Organization. *Knowledge Organization*, 30:2, 87-111.
- Hjørland, Birger. (2008). Classification of the sciences. En: *Core Concepts in LIS*. 2008. Recuperado 4-05-2015 de [http://www.iva.dk/bh/Core%20Concepts%20in%20LIS/articles%20a-z/classification\\_of\\_the\\_sciences.htm](http://www.iva.dk/bh/Core%20Concepts%20in%20LIS/articles%20a-z/classification_of_the_sciences.htm)
- Hjørland, Birger. (2009). Classification system (Library science). En: *Lifeboat for Knowledge Organization*. Recuperado 4-05-2015 de [http://www.iva.dk/bh/lifeboat\\_ko/CONCEPTS/classification\\_systems.htm](http://www.iva.dk/bh/lifeboat_ko/CONCEPTS/classification_systems.htm)
- Hjørland, Birger. (2010). Classification. En: *Lifeboat for Knowledge Organization*. Recuperado 4-05-2015 de [http://www.iva.dk/bh/lifeboat\\_ko/CONCEPTS/classification.htm](http://www.iva.dk/bh/lifeboat_ko/CONCEPTS/classification.htm)
- Hjørland, Birger. (2012). Is classification necessary after Google? *Journal of Documentation*, 68:3, 299-317.
- Ingwersen, Peter. (1992a). Conceptions of information science. En: Vakkari, P.; Cronin, B. (eds.) *Conceptions of library and information science*. London: Taylor Graham, 299-311.
- Ingwersen, Peter. (1992b) Information and information science in context. *Libri*, 42: 2, 99-135.
- Kunze, H.; Rückl, G.(1974) *Lexikon des Bibliothekswesens. Band 1*. Leipzig: VEB Verlag für Buch und Bibliothekswesen.
- Lancaster, F. (1984). Implications for library and information science education. *Library Trends*, 32, 337-347.

- López Yepes, José. (1995). *La Documentación como Disciplina: Teoría e Historia*. Segunda edición actualizada y ampliada. Pamplona: Ediciones Universidad de Navarra.
- Meadows, A.J. (1990). Theory in information science. *Journal of Information Science*, 16, 59–63.
- Miksa, Francis. (1992). Library and information science: two paradigms. En: *Conceptions of library and information science*. London: Graham Taylor, 229-252.
- Miksa, Francis. (1998). *The DDC, the Universe of Knowledge, and the Post-Modern Library*. Albany, NY: Forest Press.
- Oliveira, M. de A. (1998). investigação científica na ciência da informação: análise da pesquisa financiada pelo CNPq (Tese – Doutorado). Brasília: UnB, Departamento de Ciência da Informação e Documentação.
- Olson, Hope A. (2000). Editorial. *Knowledge Organization*, 27:1/2, 1-3.
- Olson, Hope A. (2001). Editorial Classification or Organization – What's the Difference? *Knowledge Organization*, 28:3, 1-3.
- Ortega, C.D. (2004). Relações históricas entre Biblioteconomia, Documentação e Ciência da Informação. *Datagramazero*, 5:5.
- Otlet, Paul. (1934). *Traite de documentation: le livre sur le livre: theorie et pratique*. Bruxelles, Editions Mundeaneum.
- Rafferty, Pauline. (2001). The Representation of Knowledge in Library Classification Schemes. *Knowledge Organization*, 28:4, 180-191.
- Ranganathan, S.R. (1948). *Preface to library science*. Delhi: University of Delhi.
- Ranganathan, S.R. (1957). *The five laws of library science*. 2nd ed. London: Blunt and Sons.
- Rayward, W.Boyd. (1994). Visions of Xanadu: Paul Otlet (1868–1944) and hypertext. *Journal of the American Society for Information Science*, 45:4, 235-250.
- Rosch, E. (1978). Principles of categorization. En: Rosch, E; Lloyd, B. B. (eds.), *Cognition and categorization*. Hillsdale, NJ: L. Erlbaum Associates, 27-48.
- Saracevic, Tefko. (1996). Ciência da informação: origem, evolução e relações. *Perspectivas em Ciência da Informação*, 1:1, 41-62.
- Saracevic, Tefko. (1999). Information science. *Journal of the American Society for Information Science*, 50:12, 1051-1063.
- Schrader, Alvin M. (1983). *Toward a theory of library and information science*. Vol. 1–2. Ann Arbor: University Microfilms International.
- Schrettinger, M. (1808). *Versuch eines vollständigen Lehrbuchs der Bibliothekswissenschaft*. Band 1–2. München.
- Smit, J. (1993). O documento audiovisual ou a proximidade entre as 3 Marias. *Revista Brasileira de Biblioteconomia e Documentação*, 26:1/2, 81-85.
- Solberg, Winton U. (2000). *The University of Illinois, 1894-1904: The Shaping of the University*. Chicago: University of Illinois Press.
- Taylor, Robert S. (1966). Professional aspects of information science and technology. *Annual Review of Information Science and Technology* 1: 15-40.
- Vakkari, P. (1994). Library and information science: its content and scope. *Advances in Librarianship*, 18, 1-55.
- Wersig, G. (1992). Information Science and theory: a weaver bird's perspective. En: *Conceptions of library and information science*. London: Graham Taylor, 201-217.
- Wersig, G. (1996). Information theory. En: Feather, J.; Sturges, P. (eds.) *International encyclopedia of library and information science*. London & New York: Routledge, p. 220-227.
- Wilson, P. (1983). Bibliographical R&D. En: *The study of information*. New York: Wiley, 389-398.

---

# A dimensão conceitual da organização do conhecimento no universo científico da ISKO: uma análise de domínio a partir dos congressos de ISKO-Brasil e ISKO- Espanha

*The conceptual dimension of knowledge organization in the scientific universe of ISKO: a domain-analysis of ISKO-Brazil and ISKO-Spain*

---

**Daniela Fernanda de Oliveira Matos (1), José Augusto Chaves Guimarães (2), Maria Cláudia Cabrini Grácio (3)**

(1) (2) (3) Departamento de Ciência da Informação – Universidade Estadual Paulista – UNESP - Av. Hygino Muzzi Filho, 737 – Marília – 17525-900 – SP – BRASIL. Correio eletrônico: gpfapoi@gmail.com

## Resumo

Considerando o papel nuclear da organização do conhecimento para a base epistemológica da Ciência da Informação, objetivou-se analisar como vem se construindo a dimensão conceitual, a partir dos pesquisadores da área. Para tanto, levantaram-se os artigos publicados nos anais dos congressos dos capítulos brasileiros e espanhóis da International Society for Knowledge Organization, com a incidência dos termos organização do conhecimento/organização da informação nos títulos, resumos, palavras chave e títulos de seções das comunicações apresentadas nos eventos realizados. A partir da análise de conteúdo (Bardin, 2008), analisaram-se as definições presentes nos artigos recuperados, tendo por base as categorias natureza, objeto, objetivo, instrumentos e diálogos interdisciplinares. A seguir, foram analisados os referentes teóricos dos artigos analisados, por meio da análise de citações e das redes de citação, geradas por meio do software PAJEK. Dentre outros aspectos, os resultados revelaram que, a questão conceitual constitui tema ainda tímido no universo analisado (19% dos trabalhos na ISKO-Brasil e 5% na ISKO-Espanha), o que poderia sinalizar para uma sedimentação e uma superação dessas questões, mas, dada a diversidade de abordagens conceituais quanto ao objeto da área, observa-se que o tema ainda se encontra em construção.

**Palavras-chaves:** Organização do conhecimento.; ISKO; Análise de domínio.

## Abstract

Considering the core role of knowledge organization for the epistemology of information science, the aim of this work is to analyze how the conceptual dimension is treated by the researchers from the area. Therefore, the records of the Congress of the Brazilian and Spanish chapters of the International Society for Knowledge Organization were analyzed from the incidence of the terms knowledge organization / organization of information in the titles, abstracts, keywords and headlines, the sections presented in events performed. Then analyzed the definition published from the content analysis (Bardin, 2008), based on the nature categories, object, objective, interdisciplinary instruments and dialogues as well as the theoretical references were analyzed by means of citation analysis and the construction of citation networks using the software PAJEK. The results have shown, among other things that, in the group studied, the conceptual question is still a shy theme in the universe analyzed (18.82% in the ISKO-Brazil and 4.95% in the ISKO-Spain), which could signal to sedimentation and the overcoming of these issues, but given the diversity of conceptual approaches to the subject area, it is observed that the issue is still under construction.

**Keywords:** Knowledge organization; ISKO; Domain analysis.

## 1. Introdução

A organização do conhecimento desempenha papel nuclear na Ciência da Informação, apresentando avanço teórico notadamente a partir da criação da ISKO, em 1989. No entanto, observa-se que esse campo ainda se encontra em fase de consolidação, em busca da construção e delimitação de sua própria base conceitual e seus limites, em virtude de distintas influências teóricas. Tal aspecto se revela, dentre outros, por uma diversidade no tocante ao seu aspecto conceitual, notadamente com relação a seu objeto, o que leva à necessidade de analisar de que forma o ambiente científico da ISKO vem construindo /delimitando essa dimensão conceitual, a partir de seu discurso científico oficial em âmbito internacional.

Desse modo, a presente pesquisa busca verificar a presença das questões conceituais de organização do conhecimento na literatura da ISKO e identificar a base desses conceitos quanto aos referentes teóricos que os subsidiam e, ainda, a dialogicidade científica que se estabelece nesse domínio.

Em relação ao universo dos congressos internacionais da ISKO, Guimarães et al (2014) observaram que a organização do conhecimento constitui área ou campo de conhecimento e/ou de aplicação que se relaciona à construção de discursos especializados. Sua natureza operacional se encontra ora voltada à organização e representação de conceitos, ora voltada à busca de informação e ao acesso aos conhecimentos, em especial ao diálogo com a recuperação da informação.

Tendo por objeto a estrutura do conhecimento contida nos documentos (conhecimento registrado, socializado e publicado), com especial ênfase aos conceitos e a sua modelagem, assim como às atividades discursivas em domínios científicos e às práticas sociais e culturais específicas, a organização do conhecimento vale-se de instrumentos como *sistemas de classificação*, tesouros, vocabulário e outras linguagens de indexação. Nesse contexto, as abordagens predominantes se relacionam às questões cognitivas, tecnológicas, lógico-conceituais, socioculturais e de gestão, o que sinaliza para os diálogos mais prementes na área.

Nesta oportunidade, e em continuidade à investigação de Guimarães et al (2014), esta pesquisa objetiva analisar a questão conceitual da organização do conhecimento nos congressos realizados pelos capítulos brasileiro e espanhol da ISKO.

A vista disso, parte-se de uma abordagem sucinta da organização do conhecimento enquanto campo de estudos para chegar ao contexto da *International Society for Knowledge Organization* – ISKO, seja em sua organização internacional, seja em seus capítulos nacionais ou regionais.

Consecutivamente, descrevem-se os procedimentos metodológicos, com a coleta dos dados a partir dos anais dos congressos dos capítulos brasileiro e espanhol da ISKO. Os resultados são apresentados, analisados e discutidos sob duas vertentes das onze vertentes de análise de domínio propostas por Hjørland (2002): a) epistemológica, a partir da análise de conteúdo das definições apresentadas, e bibliográfica, a partir das redes de citação que subsidiam essa produção.

## 2. A Organização do Conhecimento e a atuação científica da International Society for Knowledge Organization (ISKO)

No âmbito da Ciência da Informação, a organização e representação do conhecimento apresenta natureza mediadora, configurando-se em um conjunto de processos que estabelecem a intermediação entre um conhecimento que, uma vez produzido, foi materializado e socializado, de tal forma que esse conhecimento possa servir de base para a geração de um novo conhecimento. Uma vez materializado e socializado, este novo conhecimento pode igualmente ser objeto de nova organização e representação, caracterizando aquilo que se pode denominar como fluxo helicoidal da informação (Guimarães, 2008).

Desse modo, como bem recorda Garcia Marco (1997, p.8) a Organização do conhecimento surge como ramo de estudos mais notadamente a partir da década de 90 do século XX,

*(...) na encruzilhada das denominadas ciências cognitivas, no campo de encontro entre as Ciências do Conhecimento Humano (Neurociência, Psicologia e Epistemologia), Ciências da Informação e da Comunicação (incluindo a Semiótica e a Linguística), Matemática (incluindo a Lógica e as Linguagens formais) e a Ciência da Computação.*

O termo *Organização do Conhecimento* foi introduzido por Henry Evelyn Bliss, a partir de suas obras “*The organization of Knowledge and systems of sciences*”, de 1929, e “*The organization of knowledge in libraries*”, de 1933. Referindo-se à precursora obra de Evelyn Bliss, “*A organização do conhecimento e o sistema das*

ciências", de 1931, Ingetraut Dahlberg (1993, p.211) afirma que:

*Lidamos com um volume de conhecimento que foi coletado, aprofundado e amadurecido ao longo de séculos o que, no entanto, apenas em nossos dias tem sido reconhecido como uma área autônoma de conhecimento necessitando encontrar seu próprio lugar na sociedade de modo a ser reconhecida no âmbito do sistema das ciências.*

Posteriormente, o termo foi resgatado e discutido por Dagobert Soergel (1971) e Ingetraut Dahlberg (1993). Para Dahlberg (2008), a Organização do Conhecimento possui natureza científica e destina-se a sistematizar unidades de conhecimento (conceitos) a partir dos elementos ou características que lhes são inerentes. A isso, se alia a aplicação de conceitos e de classes de conceitos que possam permitir a transferência de conteúdos (assuntos). Para a autora (Dahlberg, 1993), duas concepções de conhecimento devem ser consideradas: enquanto processo individual, que não é transferível, e aquele que pressupõe um consenso, a partir de um registro, sendo, este último, objeto da organização do conhecimento. Assim, para a Ciência da Informação, o termo Organização do Conhecimento se dedica às atividades de classificar, indexar e representar o conhecimento por meio de registros com o objetivo específico de suprir a necessidade de informação (BU-FREM, 2008).

No âmbito investigativo, a Organização do Conhecimento tem se dedicado a estudar leis, princípios e procedimentos nos quais se sustenta o conhecimento, em cujo contexto se destaca a atuação da *International Society for Knowledge Organization* – ISKO, instituída em 1989, na Alemanha, por iniciativa de Ingetraut Dahlberg, como uma sociedade de pesquisadores que abordam a temática de Organização do Conhecimento. Sua missão consiste em promover o avanço teórico e aplicado da Organização do Conhecimento em diversos campos e em diferentes maneiras, congregando pesquisadores de distintos países, em áreas como Ciência da Informação, Linguística, Filosofia, e Ciência da Computação entre outras.

Em termos de atividades científicas da ISKO, destacam-se as conferências internacionais, que ocorrem nos anos pares, as conferências nacionais ou regionais em anos ímpares, além de duas publicações de grande impacto científico na área: a revista trimestral *Knowledge Organization* e a série *Advances in Knowledge Organization*.

Ponto alto da atuação científica da ISKO, as conferências internacionais vêm se realizando

bienalmente na seguinte conformidade: Darmstadt, 1990; New Delhi, 1992; Copenhagen, 1994; Washington, 1996; Lille, 1998; Toronto, 2000; Granada, 2002; Londres, 2004; Viena, 2006; Montréal, 2008; Roma, 2010; Mysore, 2012; e Cracóvia, 2014, com o próximo congresso internacional, em 2016, a realizar-se no Rio de Janeiro.

Em suas relações internacionais, a ISKO mantém atividades colaborativas com a UNESCO, a Comissão Europeia, a ISO (*International Standardization Organization*), a IFLA (*International Federation of Library Associations*), a ASIST (*American Society for Information Science and Technology*), entre outras.

Em termos operacionais, a ISKO desenvolve suas atividades a partir de capítulos nacionais ou regionais que, por sua vez, realizam congressos bienais. Hoje, têm-se os seguintes capítulos: Germânico, Norte-americano, Francês, Espanhol, Italiano, Brasileiro, Polonês, Escandinavo, Maghreb e Índia.

Para os fins da presente investigação, foram analisados os capítulos espanhol e brasileiro, por serem capítulos que vem mantendo um importante dialogicidade científica.

O capítulo espanhol, um dos primeiros a ser criado, remonta a 1991 e, a partir de dezembro de 2013, transformou-se em capítulo ibérico, abrigando também pesquisadores portugueses. Até o momento, foram realizados onze congressos, cujos temas foram: *Organización del conocimiento en Sistemas de información y documentación* (Madrid, 1993, e Getafe, 1995 e 1997), *La Representación y la Organización del conocimiento en sus distintas perspectivas: su influencia en la recuperación de información* (Granada, 1999), *La Representación y Organización del conocimiento: metodologías, modelos y aplicaciones* (Alcalá de Henares, 2001), *Tendencias de investigación en Organización del conocimiento* (Salamanca, 2003), *La dimensión humana de la organización del conocimiento* Barcelona (2005), *La interdisciplinariedad y la transdisciplinariedad en la Organización del conocimiento Científico* (León, 2007), *Nuevas Perspectivas para la difusión y Organización del conocimiento* (Valencia, 2009), *20 años del capítulo español de ISKO* (Ferrol, 2011) e *Informação e/ou Conhecimento: as duas faces de Jano* (Porto, 2013).

Criado em 2007, o capítulo brasileiro da ISKO surgiu a partir de um grupo de pesquisadores do Grupo de Trabalho 2- Organização e Representação do Conhecimento-, da Associação Nacional de Pesquisa em Ciência da Informação

(ANCIB). Até o momento, foram realizados dois congressos cujos temas foram: *Desafios e perspectivas científicas para a organização e representação do conhecimento na atualidade* (Brasília, 2011) e *Complexidade e Organização do Conhecimento: desafios de nosso século* (Rio de Janeiro, 2013).

### 3. Metodologia

A pesquisa desenvolveu-se a partir do conjunto de atas dos congressos dos onze capítulos espanhol (1993, 1995, 1997, 1999, 2001, 2003, 2005, 2007, 2009, 2011 e 2013) e dois capítulos brasileiros (2011 e 2013) da ISKO (*International Society for Knowledge Organization*). A escolha dos referidos capítulos deu-se pelo fato de a Espanha ser um dos capítulos internacionais com os quais os pesquisadores brasileiros da área vêm tradicionalmente mantendo maior dialogicidade.

A investigação centrou-se na abordagem da análise de domínio que, no âmbito internacional da Ciência da Informação, vem sendo tradicionalmente trabalhada, em termos teóricos e aplicados, por Hjørland & Albrechtsen (1995); Moya Anegón & Herrero Solana (2001), Hjørland (2002, 2004), Tennis (2003), e Smiraglia (2011), dentre outros. Constitui uma importante abordagem para caracterização e avaliação da ciência, na medida em que permite identificar as condições pelas quais o conhecimento científico se constrói e se socializa. Para tanto, tem-se por domínio a concepção de Smiraglia (2012, p.114) como "... um grupo que apresenta uma base ontológica que revela uma teleologia subjacente, um conjunto de hipóteses comuns, assim como um consenso epistemológico sobre as abordagens metodológicas, sociais e semânticas a serem utilizadas". Desse modo, valeu-se de duas das onze abordagens de análise de domínio previstas por Hjørland (2002): a epistemológica e a bibliométrica.

Tendo por base o percurso metodológico utilizado por Guimarães, Martínez Ávila, Sales e Alencar (2014), partiu-se de um levantamento na coleção completa dos referidos Anais, tomando por base a incidência do termo "*knowledge organization*" (e seus correspondentes em espanhol e português) nos títulos, palavras-chave, resumos e títulos de seções das comunicações publicadas.

Dos 85 trabalhos apresentados na ISKO-Brasil e 586 apresentados na ISKO-Espanha (11 edições: 1993 a 2013), totalizando 671 trabalhos, a partir da incidência dos termos nos títulos, palavras-chave, resumos e títulos de seções, recu-

perou-se um conjunto de 33 trabalhos na ISKO-Brasil e 149 na ISKO-Espanha, totalizando 182 trabalhos. Em seguida, procedeu-se à leitura desse material, observando-se que nem todos traziam, efetivamente, conceituações, definições ou considerações teóricas sobre a organização do conhecimento. Dessa forma, chegou-se a um corpus final de 16 trabalhos na ISKO-Brasil e 29 na ISKO-Espanha, totalizando 42 comunicações.

A seguir, realizou-se a leitura das 42 comunicações que compuseram o corpus final, analisando-as quanto aos conceitos, definições ou considerações teóricas acerca da natureza da organização do conhecimento.

A dimensão epistemológica foi abordada a partir do universo conceitual da literatura analisada. Dessa forma, no tocante às definições, procedeu-se à Análise de Conteúdo (Bardin, 2003), cuja escolha se justifica pelo fato de seus procedimentos possibilitarem uma análise com base em inferências extraídas de conteúdos de documentos – a partir de uma interpretação controlada por meio de variáveis ou indicadores, que proporcionam maior liberdade ao analista, sem que se perca a objetividade da investigação, e por se tratar de uma análise pautada em definições concebidas por autores/investigadores, ou seja, argumentos registrados textualmente. Dividiu-se em três fases: pré-análise, exploração do material e, tratamento dos resultados (inferências e interpretações). Para tanto, a delimitação do corpus respeita as quatro regras básicas previstas por Bardin (2003): exaustividade (na medida em que todas as definições encontradas, quando da coleta de informações, foram consideradas materiais úteis à análise, sem processo seletivo); representatividade (assegurada pela escolha do canal de comunicação científica, pois os trabalhos publicados nos Anais da ISKO são de autoria de pesquisadores e teóricos que representam uma parcela significativa da produção internacional em organização do conhecimento); homogeneidade (as definições analisadas são relativas a um mesmo tema, assim, são homogêneas tanto na estrutura textual quanto no assunto); e pertinência do material (as fontes são as próprias ideias explicitadas pelos autores/pesquisadores em suas definições).

As definições foram analisadas a partir das seguintes categorias: Natureza; Objeto, Objetivo, Instrumentos, Processos, Diálogos interdisciplinares e Perspectivas ou abordagens.

Considerando o fato de que todos os autores das comunicações publicadas nos Anais da ISKO são pesquisadores e ou teóricos da área,

cujas concepções contribuem efetivamente para a construção do universo epistemológico da organização do conhecimento em âmbito internacional, optou-se por nomeá-los em suas definições. Dessa forma, o universo das definições aqui levantadas é composto pelo seguinte conjunto de autores:

- a) ISKO-Brasil: Andrade; Berti Junior; Cervantes; Rodrigues, 2011; Miranda; Paranhos; Oliveira; Paes, 2011; Dodebei, 2011; Brascher, 2011, 2013; Ohly, 2011; Barite, 2011; Smiraglia, 2013; Alvarenga, 2013; Fujita, 2013; Varela; Barbosa, 2013.
- b) ISKO-Espanha: Garcia Marco, 1993; Jaenecke, 1995; Saz, 1995; Jiménez, 2001; López-Huertas, 2003; Bufrem, Breda, Silva, Fecchio, 2003; San Segundo, 2007; Miranda, 2007; Alfaya-Lamas, 2011; Sánchez, Blázquez, Rodríguez, 2011; Guimarães, Ferreira, Freitas, 2011; Tognoli, Milani, Barros, 2011; Café, Agustín-Lacruz, Barros, 2011; Ohly, 2011; Almeida, 2013; Alvares Júnior, Saldanha, 2013; Dal'evedove, Almeida, Fujita, 2013; Martins; Moraes, 2013; Murguia; Sales, 2013; Semidão; Almeida; Moreira, 2013.

A seguir, e de modo a caracterizar o domínio, a partir das referências (referentes teóricos) presentes nos trabalhos selecionados, realizou-se a análise das citações e geraram-se as redes de citação para os capítulos brasileiros e espanhol, de modo a identificar diálogos, consonâncias, correntes teóricas e metodológicas, entre outros.

Por meio da análise de citações, tornou-se possível analisar o impacto e a influência dos autores do domínio científico e visualizar as escolas teóricas. Dessa forma, constituiu-se em especial ferramenta para que, a partir de análise qualitativa, se possa melhor compreender o universo epistemológico do domínio estudado, uma vez que a frequência de citação permitiu identificar a proximidade de conteúdos e de como a estrutura de conhecimento da área é reconhecida por seus pesquisadores. Segundo Smiraglia (2011), as citações definem um domínio.

Para fins de análise das citações relativas aos trabalhos que trouxeram definições de organização do conhecimento, foram excluídas as autocitações e as entidades coletivas, chegando-se: na ISKO-Brasil a um conjunto de 108 autores citados com total de 142 citações, dos quais apenas 15 (13,9%) obtiveram um mínimo de 2 citações; e na ISKO-Espanha, a um conjunto de 345 autores citados com um total de 535 citações, dos quais apenas 72 (20,9%) obtiveram um mínimo de 2 citações.

Para fins de determinação do conjunto de autores mais citados para a geração das redes de citação, aplicou-se a Lei do Elitismo de Price, segundo a qual:

*A distribuição da produtividade dos autores numa coordenada cartesiana é uma distribuição tão inclinada, que inspirou Price (1963) a propor a Lei do Elitismo. Segundo esta lei, se  $k$  representa o número total de contribuintes numa disciplina,  $\sqrt{k}$  representaria a elite da área estudada, assim como o número de contribuintes que gera a metade de todas as contribuições (URBIZAGÁSTEGUI ALVARADO, 2008, p.1).*

Desse modo, foram analisados os 15 autores mais citados na ISKO-Brasil, (o que recai sobre um mínimo de 2 citações) e 19 autores na ISKO-Espanha (com um mínimo de 4 citações cada)

A partir desses conjuntos de autores mais citados, geraram-se duas redes de citação, a fim de permitir uma melhor visualização do comportamento do domínio estudado. A rede de citação gerada para a ISKO-Brasil é composta por 31 autores, sendo 16 pesquisadores citantes relativos às definições apresentadas de O.C. e os 15 autores mais citados. A rede de citação gerada para a ISKO-Espanha é composta por 55 autores, sendo 36 pesquisadores citantes relativos às definições apresentadas de O.C. neste evento e os 19 autores mais citados.

Para fins de visualização das redes de citação, utilizou-se o software PAJEK, descrito por Adamic (2008).

#### 4. Apresentação, análise e discussão dos resultados

A Tabela 1 apresenta a distribuição diacrônica dos trabalhos sobre o tema analisado, apresentados nos capítulos espanhol e brasileiro, em relação ao total de artigos publicados em cada edição do evento.

A partir da análise da Tabela 1, observa-se que parte significativa (19%) dos trabalhos apresentados no congresso da ISKO-Brasil tratam da conceituação da Organização do Conhecimento, sendo distribuída igualmente entre as duas edições (2011 e 2013) do congresso. No capítulo espanhol, observa-se que uma presença mais tímida (5%) do tema no geral das 11 edições realizadas do congresso, em especial nas nove primeiras edições. Por outro lado, destaca-se que a maioria (59%) dos trabalhos apresentados com conceituação da organização do conhecimento foi apresentada nas duas últimas edições da ISKO-Espanha (2011 e 2013). To-

davia, observa-se que a preocupação com o conceito de organização do conhecimento está presente desde a sua primeira edição, em 1993, e com a maioria (59%) foi apresentada nas duas edições (2011 e 2013).

| Ano          | Total de artigos publicados |            |            |            | Total de artigos com conceitos de KO |            |           |            |
|--------------|-----------------------------|------------|------------|------------|--------------------------------------|------------|-----------|------------|
|              | Brasil                      |            | Espanha    |            | Brasil                               |            | Espanha   |            |
|              | Nº                          | %          | Nº         | %          | Nº                                   | %          | Nº        | %          |
| 1993         | -                           | -          | 23         | 4          | -                                    | -          | 3         | 10         |
| 1995         | -                           | -          | 26         | 4          | -                                    | -          | 3         | 10         |
| 1997         | -                           | -          | 34         | 6          | -                                    | -          | 1         | 3          |
| 1999         | -                           | -          | 55         | 9          | -                                    | -          | -         | -          |
| 2001         | -                           | -          | 42         | 7          | -                                    | -          | 1         | 3          |
| 2003         | -                           | -          | 76         | 13         | -                                    | -          | 2         | 7          |
| 2005         | -                           | -          | 49         | 8          | -                                    | -          | -         | -          |
| 2007         | -                           | -          | 62         | 11         | -                                    | -          | 2         | 7          |
| 2009         | -                           | -          | 87         | 15         | -                                    | -          | -         | -          |
| 2011         | 39                          | 46         | 40         | 7          | 8                                    | 50         | 6         | 21         |
| 2013         | 46                          | 50         | 92         | 16         | 8                                    | 50         | 11        | 38         |
| <b>Total</b> | <b>85</b>                   | <b>100</b> | <b>586</b> | <b>100</b> | <b>16</b>                            | <b>100</b> | <b>29</b> | <b>100</b> |

Tabela 1. *Distribuição diacrônica das publicações nos capítulos espanhol e brasileiro.* Fonte: elaborado pelos autores

Para fins de análise e após a aplicação da metodologia apresentada, a partir dos 16 artigos apresentados na ISKO-Brasil, chegou-se a um conjunto de 79 trechos relativos a 28 definições e dos 29 artigos que compuseram o corpus na ISKO-Espanha, 51 trechos os quais foram objeto de análise.

Segundo a análise realizada nos artigos publicados nos Anais da ISKO-Brasil, o objeto da Organização do Conhecimento (O.C.) está relacionado à classificação conceitual, à teoria, à indexação e representação/terminologia de conceito (Andrade; Berti Junior; Cervantes; Rodrigues, 2011; Ohly, 2011; Fujita, 2013).

Para alguns autores (Miranda; Paranhos; Oliveira; Paes, 2011), seu objeto são os fundamentos científicos e o desenvolvimento das técnicas de planejamento, construção, uso, gestão e avaliação das habilidades e ferramentas empregadas nos sistemas de informação para tratamento, armazenamento e recuperação de documentos. Guimarães (2001) e Barité (2009) entendem o

conhecimento socializado ou registrado, o saber acumulado que se documenta por meio dos processos de publicação, edição, registro em patentes ou em formas de socialização dos saberes similares como objeto da O.C. Tal definição está relacionada à informação como conhecimento (*information-as-knowledge*) e à informação como coisa (*information-as-thing*), duas aproximações à noção de “conhecimento materializado” e registrado (Barité, 2011).

Segundo Smiraglia (2013), o objeto da O.C. é uma comunidade de discurso onde uma investigação rigorosa e autoconsciente acontece em relação àquilo que é conhecido e suas várias ordenações e sequências, tanto as naturais ou heurísticas, como aquelas que são impostas. Para Brascher (2013), as representações de domínios do conhecimento são objeto de estudo da O.C.

Com relação aos objetivos, os processos de representação, organização e recuperação ganham um destaque importante a partir das definições levantadas. Para Andrade, Berti Junior, Cervantes e Rodrigues (2011), o objetivo da O.C. é possibilitar a recuperação de objetos e conteúdos informacionais nos ambientes digitais, bem como estruturar e arranjar sistematicamente unidades do conhecimento (os conceitos) de acordo com os elementos de conhecimento (Andrade; Berti Junior; Cervantes; Rodrigues, 2011; Ohly, 2011).

Ainda, segundo Ohly (2011), os objetivos da O.C. são organizar unidades de conceitos de conhecimento e todos os tipos de objetos (minerais, plantas, documentos, pinturas, objetos de museus, etc.), relacionados a termos ou categorias particulares, para capturar o que é conhecido sobre o mundo em alguma forma ordenada, permitindo seu futuro compartilhamento com outros.

Ortega (2013) advoga a favor da promoção do acesso ao conhecimento, visando seu uso e nova produção, bem como para elaborar representações que possam ser significadas e manipuladas a favor de um público determinado. Ainda, segundo a autora, a organização do conhecimento deve ser capaz de fornecer respostas à sociedade para o problema do crescimento dos documentos, por meio de critérios científicos, como uma extensão da capacidade de organização e transferência daquilo que o ser humano sabe.

Para Alvarenga (2013), a O.C. deve prover uma estrutura ou esquema para armazenar e organizar dados, informação ou conhecimento sobre o mundo e sobre pensamentos, visando à com-

preensão, recuperação ou descoberta, ao raciocínio e a outros propósitos.

Segundo Brascher (2013), delimitar o significado de termos no contexto de seus domínios e estabelecer relações conceituais que auxiliam a posicionar um conceito no sistema conceitual é objetivo da Organização do Conhecimento. E, por último, Miranda, Paranhos, Oliveira e Paes (2011) definem a geração de um novo conhecimento como objetivo final da O.C.

Os processos apresentam já um maior consenso entre os autores pesquisados, uma vez que a descrição, indexação e classificação ou representação figuram entre os processos nucleares da O.C. Segundo Barité (2001), esses processos consistem na construção, desenvolvimento e gestão de sistemas de organização do conhecimento e outras estruturas conceituais, e nas atividades de classificação e indexação, para a recuperação temática de documentos, recursos, dados e todo tipo de informação.

Para Fujita (2013), os processos também são a descrição documentária, indexação e classificação, conceitos ligados à representação da informação apontados como processos por Varela e Barbosa (2013). Ainda segundo os autores, a criação de recursos mediadores do acesso ao conhecimento também é entendido como um processo da O.C.

Como instrumentos, Andrade, Berti Junior, Cervantes e Rodrigues (2011) apontam a linguagem do sistema de informação (controlada).

No tocante à natureza da O.C., observaram-se, ainda, mais divergência. Alguns a caracterizam como ciência (Andrade; Berti Junior; Cervantes; Rodrigues, 2011; Ohly, 2011) ou disciplina científica (Miranda; Paranhos; Oliveira; Paes, 2011) e outros a definem como área temática (Ohly, 2011), domínio (Smiraglia, 2013), conjunto de procedimentos e ferramentas que ajudam os seres humanos a compreender e ordenar o caos cognitivo (Ortega, 2013), atividade-meio para o objetivo ou fim da CI (Ortega, 2013) e de caráter social, cognitivo e mediador (Varela; Barbosa, 2013).

Seus diálogos interdisciplinares apresentam as novas tecnologias (Brascher, 2011; Ortega, 2013), sua fundamentação teórica à Indexação (Fujita, 2013) e as inter-relações de pesquisa com domínios além dos tradicionalmente contemplados (Brascher, 2011), ultrapassando barreiras disciplinares e permitindo a fertilização de outras áreas do conhecimento (Dodebei, 2011).

Nos Anais da ISKO-Espanha, é possível perceber um grupo de autores que caracterizam a

Organização do Conhecimento, no tocante à sua natureza, enquanto uma disciplina científica aplicada que ainda carece de uma exploração mais efetiva de sua base teórica (Garcia Marco, 1993; Saz, 1995; Sánchez, Blázquez, Rodriguez, 2011; Guimarães, Ferreira, Freitas, 2011); (Miranda, 2007; Ohly, 2011), e outro grupo que a caracteriza enquanto ciência (Alfaya-Lamas, 2011, Ohly, 2011, Café; Agustín-Lacruz; Barros, 2011, Semidão; Almeida; Moreira, 2013). Sua compreensão enquanto um campo de estudos relacionado à informação também se faz presente em algumas obras (Alfaya-Lamas, 2011, Dal'evadore; Almeida; Fujita, 2013, Almeida, 2013).

Outros autores possuem uma visão única e diferenciada da natureza da O.C. Para Currás (1997), ela possui uma versão documentária e informativa, para Café, Agustín-Lacruz e Barros (2011) é uma área de estudos interdisciplinar, enquanto para Semidão, Almeida e Moreira (2013) é uma área temática.

San Segundo (2007) a caracteriza como um conjunto de subdisciplinas dentro do marco mais amplo das ciências da Documentação ou Informação. Para Tognoli, Milani e Barros (2001), a O.C. é um marco teórico-conceitual (Tognoli, Milani, Barros, 2011). Para Almeida (2013), ela é tanto uma atividade quanto um fenômeno. Martins e Moraes (2013) a consideram um processo de modelagem conceitual, ao passo que Murguia e Sales (2013) acreditam que a O.C. seja um fazer de natureza operacional.

Segundo as definições encontradas nos anais da ISKO-Espanha, o objeto de estudo da Organização do Conhecimento está ligado basicamente ao processo de organização. Assim, para Alfaya-Lamas (2011), Café, Agustín-Lacruz e Barros (2011), Ohly (2011) e Semidão, Almeida e Moreira (2013), o objeto da O.C. é o arranjo estrutural e sistemático das unidades de conhecimento (conceitos), segundo elementos de conhecimento inerentes (características) e a aplicação de conceitos e classes de conceitos ordenadas para a atribuição dos conteúdos de referências de todos os tipos que valem a pena conhecer.

Para Murguia e Sales (2013), é a organização de unidades de conhecimento ou organização de conceitos e, para Guimarães, Ferreira e Freitas (2011), é a organização da informação em registros bibliográficos, incluindo índices de citação, texto completo e internet. Segundo López-Huertas (2003), o objeto são as leis, princípios e procedimentos pelos quais se estrutura o conhecimento especializado. Ainda, para Alfa-

ya-Lamas (2011), seu objeto é a natureza e qualidade dos processos de organização do conhecimento e dos sistemas de organização do conhecimento.

Para Tognoli, Milani e Barros (2011), é o conhecimento registrado e socializado. Para Almeida (2013), são os objetos existentes que impactam diretamente a vida dos cientistas, das pessoas e das comunidades envolvidas e, por último, para Murguia e Sales (2013), os objetos de estudo da O.C. são os conceitos e as estruturas conceituais formalizadas instrumentalmente nos sistemas de organização do conhecimento, tais como os sistemas de classificação, os tesouros e as ontologias.

O objetivo da O.C. está fundamentalmente ligado às questões referentes à recuperação e acesso à informação registrada, como demonstram as definições a seguir. Segundo Garcia Marco (1996), o objetivo da O.C. é facilitar o acesso ao conjunto de conhecimento da sociedade, que adote e mantenha algumas perspectivas suficientemente interdisciplinares, assim como melhorar a circulação da informação dentro e através dos sistemas mediadores e aperfeiçoar a circulação do conhecimento em nossas sociedades (Garcia Marco, 1993). Jaenecke (1995) e Alfaya-Lamas (2011) consideram o acesso como objetivo primordial. Para Tramuillas Saz (1995), os objetivos são estabelecer mecanismos capazes de representar a realidade e oferecer os mecanismos necessários para que qualquer pessoa seja capaz de acessar, representar e compreender, com quase toda segurança em diferente contexto, a realidade 'original'.

Jiménez (2011), Bufrem, Breda, Silva e Fecchio (2003) e Alfaya-Lamas (2011) priorizam o acesso e a recuperação da informação, ao passo que Semidão, Almeida e Moreira (2013) focam o armazenamento eficiente, recuperação e processamento da informação codificada em texto e em outros artefatos humanos. Para López-Huertas (2003), a criação de sistemas que respondam a uma representação e organização linguística-conceitual do conhecimento é o objetivo da O.C. A organização e representação do conhecimento também são objetivos que estão presentes nas definições de Alfaya-Lamas (2011) e Martins e Moraes (2013).

No tocante aos diálogos interdisciplinares, as novas tecnologias e a inteligência artificial se fazem presente nos textos de Garcia Marco (1993) e Sánchez, Blázquez e Rodríguez (2011). Os diálogos com a Biblioteconomia também estão presentes nos textos de Ohly (2011) e Álvares Júnior e Saldanha (2013). A linguísti-

ca e a teoria da Classificação também são apresentadas por Garcia Marco como diálogos com a O.C.

Bufrem, Breda, Silva e Fecchio (2013) destacam as relações com a gestão do conhecimento e a inteligência competitiva. Para San Segundo (2007), há uma confluência de disciplinas com metodologias e categorias heterogêneas, onde a investigação transdisciplinar deve supor o intercâmbio de suas metodologias e postulados epistemológicos. Por último, Almeida (2013) destaca que há uma necessidade do contato com outras disciplinas para a execução de seus produtos e processos.

Os processos estão fundamentalmente ligados à classificação, indexação, catalogação e representação do conhecimento, bem como à elaboração de resumos (Bufrem, Breda, Silva, Fecchio, 2003; Alfaya-Lamas, 2011; Currás, 1997; Guimarães, Ferreira, Freitas, 2011; Alfaya-Lamas, 2011; Tognoli, Milani, Barros, 2011), e ao tratamento temático da Informação (Murguia; Sales, 2013).

A Figura 1 apresenta a rede de citação (citante/citado) relativa aos 15 autores mais citados na ISKO-Brasil, que fundamentam os trabalhos apresentados, definições, conceitos, objetivos e diálogos relativos à O.C. Destaca-se que os autores, considerados mais citados, receberam pelo menos duas citações. Na figura, os círculos pretos correspondem aos autores citantes e os azuis aos autores citados. O tamanho dos círculos dos autores citados é proporcional ao número total de citações recebidas pelo autor e a espessura dos segmentos de reta, proporcional ao número de citações do citante em relação ao citado.

Observa-se na Figura 1 que os autores com maior número de citações recebidas pela comunidade citante das duas edições da ISKO-Brasil foram I. Dahlberg (10 citações) e B. Hjørland (7 citações), seguidos por J.A.C. Guimarães e A.J. Cañas, ambos com 4 citações recebidas.

A rede de citação citante/citado gerada para o capítulo ISKO-Brasil apresenta densidade igual a 27%, correspondente a 65 citações (segmentos de reta) ocorridas das 240 possibilidades (citação dos 16 citantes para os 15 citados), o que indica uma coesão moderada entre citantes e citados. Esta intensidade de coesão se explica, em parte, por um baixo número de citantes em relação aos autores citados: Fugmann, Mai e Naves, citados por somente um dos pesquisadores; e Soergel e Smit, citados por dois dos autores citantes, entre os 16 citantes.

Entre os autores citados, destacam-se o número de pesquisadores que citam I. Dahlberg, J.A.C. Guimarães, B. Hjørland, C. Beghtol e F. Pinho, evidenciados pelo grande número de vetores dos citantes em relação a esses citados.

As relações mais fortes entre os investigadores citantes e autores citados são observadas de B.M.N. Cervantes, M.R.A. Rodrigues, D.W.B. Junior e I.A. de Andrade para os autores citados A.J. Cañas, J.D. Novak e M.A. Moreira, evidenciando assim o forte respaldo teórico desta comunidade citante em relação aos autores citados. Ainda, destaca-se a intensa citação de Dahlberg por M. Barité.

A Figura 2 apresenta a rede de citação citante/citado gerada pelos 36 investigadores citantes em relação aos 19 autores mais citados na ISKO-Espanha, correspondendo a ter recebido pelo menos 4 citações. Os círculos (pontos) pretos correspondem aos investigadores citantes e os vermelhos aos autores citados. O tamanho dos círculos dos autores citados é proporcional ao número de citações recebidas e a espessura dos segmentos de reta, proporcional ao número de citações do citante em relação ao citado.

Assim como na ISKO-Brasil, na rede de citação gerada a partir da ISKO-Espanha, os autores com maior número de citações recebidas pela comunidade citante foram B. Hjørland (26 citações) e I. Dahlberg (23 citações), seguidos por J.E. Mai (10 citações) e M. Barité e J.C. Gardin, ambos com 8 citações recebidas.

A rede de citação citante/citado gerada para o capítulo ISKO-Espanha apresenta densidade igual a 17%, correspondente a 116 relações citante-citado ocorridas, o que indica uma coesão fraca entre citantes e citados, explicada parcialmente por ser um universo grande, dado o maior número de citantes nesta comunidade. Comunidades maiores têm mais baixa probabilidade de apresentar altas densidades de rede. Os autores C.S. Peirce, E. Jimenez Contreras e J.C. Gardin foram citados por um único investigador cada: C.C. Almeida, M.J. López-Huertas e A.G. Jiménez, respectivamente.

Por outro lado, os autores B. Hjørland, I. Dahlberg e M. Barité foram citados por muitos investigadores participantes da ISKO-Espanha: 18 pesquisadores citaram o primeiro e 14 pesquisadores citaram os outros dois autores.

Destaca-se que os autores B. Hjørland, I. Dahlberg, J.E. Mai e A. Garcia-Gutierrez foram citados pelos investigadores participantes dos dois capítulos - Brasil e Espanha-, configurando, desse modo, referencial teórico comum às duas

comunidades científicas no tocante ao conceito de O.C.

As relações mais intensas entre os investigadores citantes e autores citados são observadas de C.C. Almeida para a autora I. Dahlberg (8 citações) e T.L. Thellefsen (5 citações); de E. Alfaya Lamas LFAYA para B. Hjørland (6 citações); de M.J. López-Huertas para E. Jimenez Contreras (5 citações); e de H.P. Ohly para I. Dahlberg (5 citações), evidenciando assim o forte embasamento teórico destes pesquisadores em relação aos autores citados.

Destacam-se, no capítulo ISKO-Espanha, as autoras M.J. López-Huertas e H.P. Ohly, as quais configuram-se tanto como investigadoras citantes como autoras citadas, constituindo ao mesmo tempo pesquisadores participantes como referencial teórico desta comunidade científica. A investigadora M. L. Huertas cita dois autores Hjørland e Jimenez-Contreras e H.P. Ohly cita os autores Capurro, Dahlberg e Mai.

Analisando os investigadores participantes (citantes) dos dois capítulos - Brasil e Espanha - observam-se duas pesquisadoras em comum, a saber: M.S. Fujita e M.L. C. Miranda, o que sugere que estas duas investigadoras brasileiras têm se inserido e trabalhado questões relativas à O.C. em ambas as comunidades.

Por fim, de forma comparativa às duas redes de citação - Figuras 1 e 2 - relativas às questões (objetivos, objetos, procedimentos e instrumentos), destacam-se os autores M. Barité e J.A.C. Guimarães. Barité configurou-se como investigador citante na capítulo brasileiro e autor entre os mais citados no capítulo espanhol e Guimarães aparece entre os autores mais citados no capítulo brasileiro e investigador citante na ISKO-Espanha.

## 5. Conclusão

A ISKO que se constitui num espaço de interlocução com a temática de Organização do Conhecimento (OC), procurou-se analisar a dimensão conceitual de Organização do Conhecimento/Organização da Informação através dos Eventos/Anais da ISKO Espanha no período de 1993-2013 e Brasil 2011 e 2013, e posteriormente realizar um estudo bibliométrico das fontes que sustentam tais conceitos.

A questão conceitual constitui tema ainda tímido no universo analisado (18,82% dos trabalhos na ISKO-Brasil e 4,95% na ISKO-Espanha), o que poderia sinalizar para uma sedimentação e uma superação dessas questões, mas, dada a diversidade de abordagens conceituais quanto ao

objeto da área, observa-se que o tema ainda se encontra em construção.

Como se destaca nas redes de citações, Hjørland e Dalberg, tanto no Brasil como na Espanha, figuram como pontos de centralidade no que tange aos referentes teóricos, o que sinaliza para um movimento da área rumo a uma dimensão mais teórica, com forte abordagem epistemológica e sócio-cognitiva, o que se reitera na análise de conteúdo, em que os processos preponderaram sobre as demais categorias de análise, revelando uma preocupação de natureza teórico-metodológica.

## Referências

- ADAMIC, L. Introductory social network analysis with Pajek. Ann Arbor (MI): School of Informaiton. University of Michigan, 2008. Disponível em: [http://ocw.mit.edu/courses/economics/14-15j-networks-fall-2009/assignments/MIT14\\_15JF09\\_pajek.pdf](http://ocw.mit.edu/courses/economics/14-15j-networks-fall-2009/assignments/MIT14_15JF09_pajek.pdf) . Acesso em 14.07.2014.
- BARDIN, L. L' analyse du contenu. 7. ed. Paris: PUF, 2003. 296 p.
- BUFREM, L.S. Práticas de organização e divulgação da produção intelectual em ciência da informação no Brasil. Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação, Florianópolis, v. 13, n. esp., p. 36-53, 1º sem. 2008.
- DAHLBERG, I. Knowledge organization: its scope and possibilities. Knowledge Organization, Würzburg, v. 20, n. 4, p. 211-222, 1993.
- DAHLBERG, I. Teoria do conceito. Ciência da Informação, Rio de Janeiro, v.7, n. 2, p. 101-107, 1978.
- DAHLBERG, I. The Information Coding Classification (ICC): A modern theory-basead fully-faceted, universal system of knowledge field. Axiomathes. v.18, n.2, p.161-176, 2008.
- GARCIA MARCO, F. J. Avances en Organización del Conocimiento en España. In: GARCIA MARCO, F. J. (ed.) Organización del conocimiento en sistemas de información y documentación. Getafe: Libr. General, 1997.
- GMÜR, M (2003). Co-Citation Analysis and the Search for Invisible Colleges: A Methodological Evaluation. *Scientometrics*, Vol. 57, No. 1, pp. 27-57.
- GUIMARÃES, J.A.C. ; SALES, R.; MARTINEZ-ÁVILA, D. ; ALENCAR, M.F. The conceptual dimension of knowledge organization in the ISKO proceedings domain: a Bardinian analysis. In: Wieslaw Babik. (Org.). Knowledge organization in the 21st century: between historical patterns and future perspectives. 1ed.Wurzburg: Ergon Verlag, 2014, v. , p. 101-106.
- GUIMARÃES, J.A.C.; FERREIRA, G.M.; FREITAS, M.F.M. Correntes teóricas do tratamento temático da informação: uma análise de domínio da presença da catalogação de assunto e da indexação nos congressos de ISKO-España. In: Maria Carmem Péres Pais; María G. Bonome. (Org.). Actas del X Congreso de ISKO-ESPANHA - 20 years of ISKO Spanish Chapter. Coruña: Universidad da Coruña - Serviço de Publicaciones, 2011, v. , p. 181-194.
- GUIMARÃES, José Augusto Chaves. A dimensão teórica do tratamento temático da informação e suas interlocuções com o universo científico da International Society for Knowledge Organization (ISKO). Revista Ibero-Americana de Ciência da Informação, - cidade -, v. 1, n. 1, 2008.
- HJØRLAND, B. Domain analysis: a socio-cognitive orientation for Information Science research. Bulletin of the American Society for Information Science and Technology, v. 30, n.3, feb./mar. 2004. Disponível em: <http://www.asis.org/Bulletin/Feb-04/hjorland.html>.
- HJORLAND, B. Domain analysis in information Science: eleven approaches traditional as well as innovative. Journal of Documentation, v. 58, n. 4, 2002, p. 422-462.
- HJØRLAND, B.; ALBRECHTSEN, H. Toward a new horizon in information science: domain-analysis. Journal of the American Society for Information Science, v. 46, n. 6, p.400-425, 1995.
- ISKO. About ISKO. International society knowledge organization. Disponível em: <http://www.isko.org/chapters.html> . Acesso em 03 jun. 2014.
- MOYA-ANEGÓN, F.; HERRERO-SOLANA, V. Análisis de dominio de la revista mexicana de investigación bibliotecológica. Información, cultura y sociedad, n. 5, 2001, p. 10-28.
- PAJEK. Introductory social network analysis with Pajek. School of Information- University of Michigan. Disponível em: [http://ocw.mit.edu/courses/economics/14-15j-networks-fall-2009/assignments/MIT14\\_15JF09\\_pajek.pdf](http://ocw.mit.edu/courses/economics/14-15j-networks-fall-2009/assignments/MIT14_15JF09_pajek.pdf) > Acesso em 16 jul. 2014.
- SMALL, H. On the shoulders of Robert Merton: towards a normative theory of citation. *Scientometrics*, v.60,n.1, p.71-79, 2004.
- SMIRAGLIA, R.P. The Epistemological Dimension of Knowledge Organization. In: Vera Dobedei, José Augusto Chaves Guimarães. (Orgs.). Complexidade e Organização do Conhecimento: desafios de nosso século – Rio de Janeiro: ISKO-Brasil ; Marília: FUNDEPE, 2013, v.2, p.17-25.
- SMIRAGLIA, R.P., ed. 2011. Proceedings from North American Symposium on Knowledge Organization, Vol. 3. Toronto, Canada.
- SMIRAGLIA, R.P. 2012. Epistemology of domain analysis. In Smiraglia, Richard P. and Hur-Li Lee, eds. Cultural frames of knowledge. Würzburg: Ergon-Verlag, p. 114.
- SOERGEL, D. 1971. Dokumentation und Organisation des Wissens: Versuch einer methodischen und theoretischen Grundlegung und Beispiel der Sozialwissenschaften [ Documento: desafios de nosso século: Towards a Methodological and Theoretical Foundation for the Social Sciences]. Berlin, DEU: Dunker & Humbolt; 1971. 380p. ISBN: 3428024419.
- TENNIS, J.T. Two Axes of Domain Analysis. *Knowledge Organization*, v. 30, n.3/4, p.191-195, 2003.
- URBIZAGÁSTEGUI ALVARADO, R. Elitismo na literatura sobre a produtividade dos autores. *Ciência da Informação*, v.38, n.2, Maio/Ago. 2009. Disponível em: <http://dx.doi.org/10.1590/S0100-19652009000200006> .
- VANZ, S.A.S.; CAREGNATO, S.E. Estudos de citação: uma ferramenta para entender a comunicação científica. *Em Questão*: revista da Faculdade de Biblioteconomia e Comunicação da UFRGS, Porto Alegre, v. 9, n. 2, p. 295-307, jul./dez. 2003.

## Anexo 1

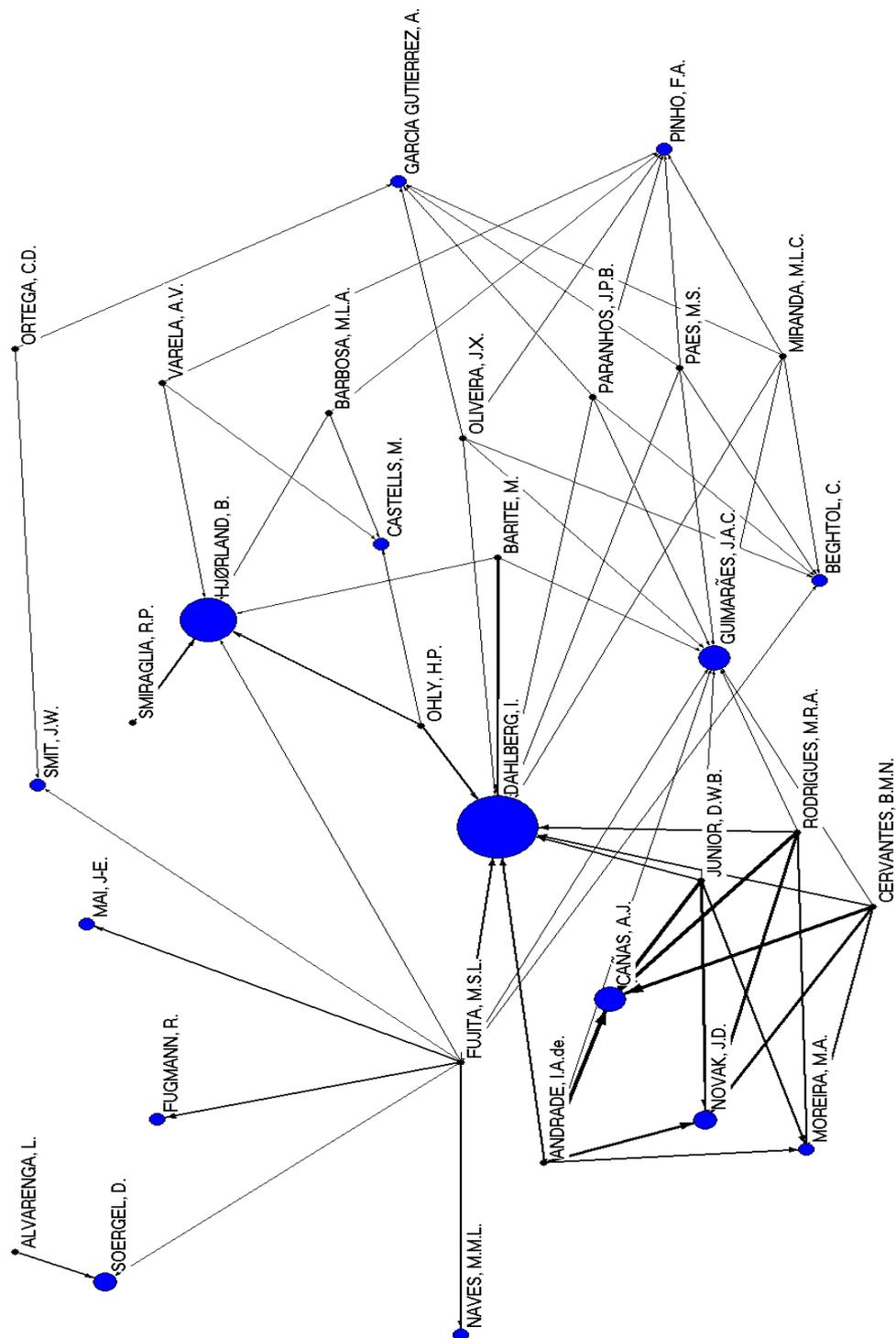


Figura 1: Rede de citação da ISKO-Brasil. Fonte: Elaborado pelos autores

Legenda: círculos (pontos) pretos = pesquisadores citantes; círculos azuis = autores citados.

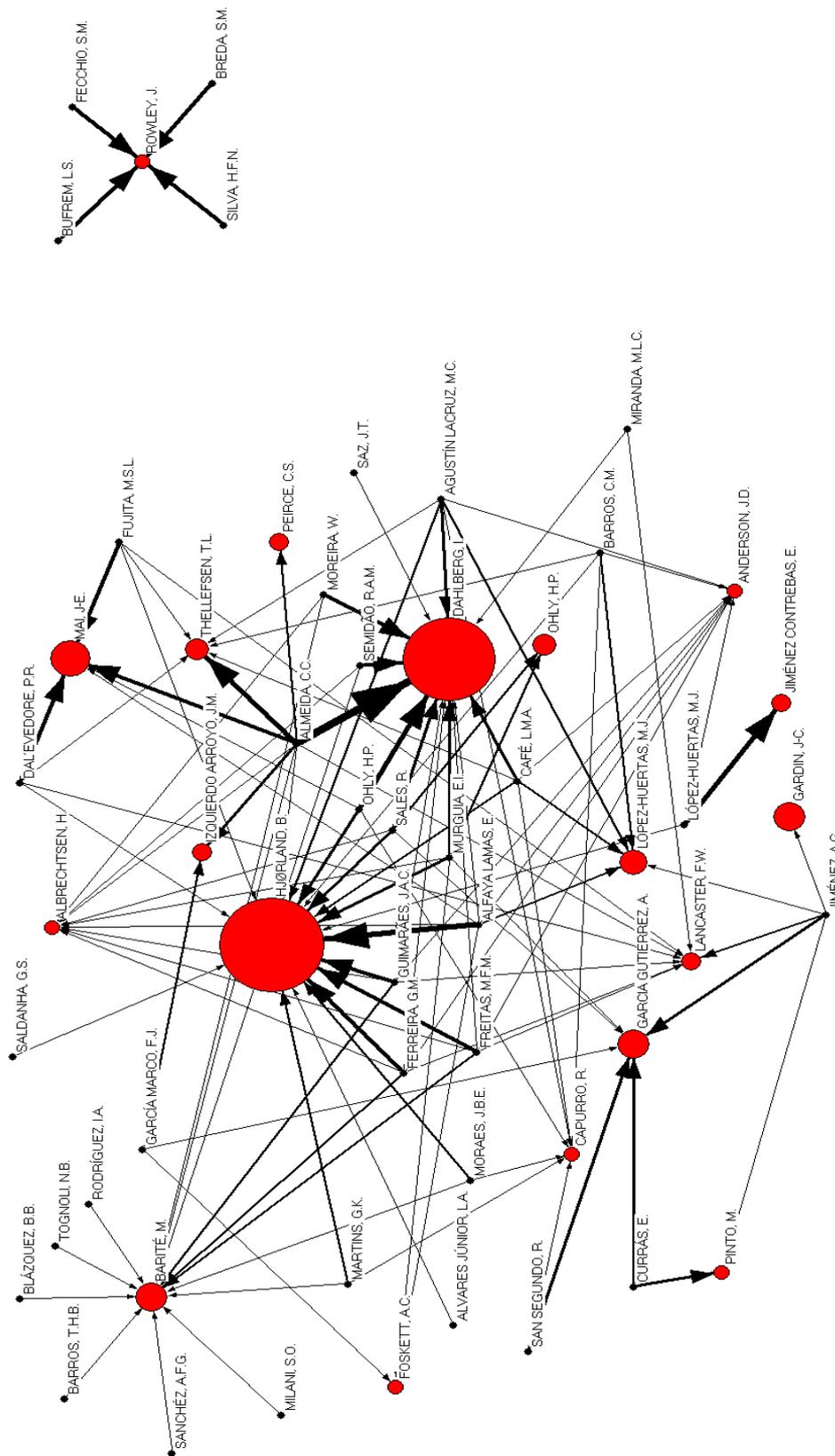


Figura 2: Rede de citação da ISKO – Espanha. Fonte: Elaborado pelos autores  
 Legenda: Círculos pretos - autores citantes; Círculos vermelhos - autores citados.

---

# Organização da informação e do conhecimento no contexto da Ciência da informação: da análise terminológica à reflexão epistemológica

*Information and knowledge organization on Information Science context: from terminological analysis to epistemological reflection.*

---

Daniel Abraão Pando (1), Carlos Cândido de Almeida (2)

(1) Programa de Pós-Graduação em Ciência da Informação - UNESP, Marília - SP, danielabraopando@gmail.com

(2) Departamento de Ciência da Informação, PPGCI - UNESP – Marília-SP, carlosalmeida@marilia.unesp.br

## Resumo

Os termos Organização do conhecimento e Organização da informação ainda geram uma série de indagações sobre suas definições, aplicações e usos nas diversas áreas que compõem a disciplina uma vez que os mesmos têm uma amplitude semântica muito grande e existem diversas perspectivas de análise, domínios e concepções que são aplicadas às diversas áreas que trabalham com eles. Buscamos nesse estudo apresentar algumas considerações epistemológicas sobre o desenvolvimento dessas áreas e como as mesmas são tratadas no contexto da Ciência da informação a partir de uma perspectiva terminológica refletindo sobre as diferenças conceituais, teóricas e históricas entre os campos de Organização da informação e Organização do conhecimento. Pode-se constatar através de uma pesquisa bibliográfica que alguns autores se referem a estes campos como sinônimos, justapostos e sobrepostos, tendo alguns autores optado pelo termo Organização da Informação, enquanto outros por Organização do conhecimento. A definição destes conceitos pode auxiliar os especialistas a reconhecer a especificidade epistemológica da organização da informação.

**Palavras-chaves:** Organização da informação. Organização do conhecimento. Análise terminológica. Epistemologia

## 1. Introdução

A constituição da Ciência da Informação (CI) desde seu surgimento tem provocado muitas discussões no que diz respeito à sua fundamentação como campo de estudo uma vez que suas raízes são objeto de controvérsias e tem sido historicamente marcada por conflitos e falta de consenso sobre suas origens, objetivos, terminologia e abrangência (Abdalla; Kobashi, 2012; Fernandes; Lima-Marques, 2013; Smit, Tálamo; Kobashi, 2004; Barreto, 2002). Se a disciplina como um todo ainda padece de con-

## Abstract

Knowledge organization and Information organization terms still generate a number of questions about their settings, applications and uses in several areas that make up the discipline since they have a very large semantic amplitude and there are several analytical perspectives, domains and conceptions which are applied to various fields which work with them. We aimed in this study present some epistemological considerations about the development of these areas and how they are treated in the context of Information science from a terminology perspective reflecting on the conceptual, theoretical and historical differences between the fields of Information organization and Knowledge organization. It can be verified through a literature search that some authors refer to these fields as synonyms, juxtaposed and overlaid, with some authors who have chosen the term Information organization, while others by Knowledge organization. The definition of these concepts may assist experts to recognize the epistemological specificity of the organization of information.

**Keywords:** Information organization. Knowledge organization. Terminological analysis. Epistemology

sensos(1) embora essa seja uma palavra que cada vez mais vem sendo usada com extremo cuidado, o mesmo se pode dizer de alguns campos que fazem parte de sua constituição, como a Organização do conhecimento (OC) que vem sendo vista por alguns autores como sub-área ou sub-domínio da CI (Abdalla; Kobashi, 2012; Smiraglia, 2011) e por outros como uma área independente, sendo elevada à condição de uma nova ciência (Dahlberg, 2006) que mantém com a CI relações interdisciplinares e com ela dialoga constantemente (Almeida, 2009).

O mesmo se pode afirmar também da Organização da informação (OI), vista por Ortega (2013) como uma subárea da Ciência da informação e considerada por Barreto (2002) e Smit (2009) como o núcleo duro da área.

Por serem áreas de extrema importância que vêm consolidando-se no escopo das ciências que trabalham com a questão informacional, tanto a OC quanto a OI têm-se constituído em importantes agendas de pesquisas que são realizadas nos mais diferentes contextos em nível local e internacional.

Não obstante a esta constatação, podemos notar através de uma breve revisão da literatura que a mesma tem algumas questões que ainda não conseguiram atingir um consenso. Uma dessas questões diz respeito justamente à relação entre os termos Organização da informação e Organização do conhecimento em que ainda persiste uma falta de discernimento entre os mesmos (Café; Barros; Santos, 2014). Apesar de serem dois termos fundamentais que estão na essência da constituição da Ciência da informação, ainda geram uma série de indagações sobre suas definições, aplicações e usos nas diversas áreas que compõem a disciplina (Brascher; Café, 2008), uma vez que têm uma amplitude semântica e existem diversas perspectivas de análise, domínios e concepções que são aplicadas às diversas áreas que trabalham com eles, como apontaram Lima e Álvares (2012), Brookes (1980), Wilson (2002), Le Coadic (2004), Zins (2007), entre outros.

Considerando que eles ainda carecem de fundamentações terminológicas e epistemológicas, assim como uma maior delimitação de suas áreas de abrangência e que, segundo Almeida (2009) parece haver um impasse entre essas duas áreas, pelo menos em âmbito conceitual, justificamos a realização deste estudo como forma de refletir sobre essa temática que vem destacando-se nas agendas de pesquisadores de várias partes do mundo.

Portanto, buscamos neste estudo apresentar algumas considerações epistemológicas e terminológicas sobre o desenvolvimento dessas áreas e como são tratadas no contexto da Ciência da informação, bem como refletir sobre as diferenças conceituais, teóricas e históricas entre os campos de Organização da informação e Organização do conhecimento tendo como argumento central que estes termos ainda não estão claramente definidos e geram conflitos em âmbito teórico e conceitual.

Podemos antever através de uma pesquisa bibliográfica que alguns autores referem-se a estes campos como: 1) sinônimos, para Tristão,

Fachin e Alarcon (2004), Rizzi (2008) e Abreu e Monteiro (2010); 2) complementares, como em Almeida (2011); 3) distintos, segundo Brascher e Café (2008), Cervantes (2009) e de forma preferencial. Assim, alguns autores optam pelo termo Organização da Informação - Salton (1968), Taylor (2004) e Svenonius, (2000) e outros pelo termo Organização do conhecimento, como em Barité (2001), Hjørland (2008, 2012), Rabello e Guimarães (2006). A continuação da pesquisa pretendemos discutir os aspectos históricos e epistemológicos da Organização da Informação, frente aos desafios de uma possível cientificidade.

## 2. Organização da informação e do conhecimento: aspectos históricos e conceituais

O termo Organização do conhecimento (*Knowledge Organization*) como é empregado atualmente é relativamente recente, surgindo por volta da década de 70. Sua fundamentação histórica no entanto surgiu nas primeiras décadas do século 20 ao ser utilizado entre outros por Henry E. Bliss, Charles A. Cutter, W.C. Berwick Sayers e Ernest C. Richardson (Hjørland, 2008). Dentre esses autores, Bliss é o autor de três obras relacionadas aos estudos da Organização do Conhecimento, sendo elas: *Organization of Knowledge (1927)*, *Organization of knowledge in libraries and subject approach to books (1933)* e *A system of bibliographic classification (1935)* que foram consideradas precursoras para a constituição da área.

Dalhberg (2006) no seu artigo sobre *Knowledge Organization: a new science*, e em uma entrevista concedida ao periódico *Knowledge Organization (2007)*, relata como se deu historicamente a consolidação do termo como ele é atualmente conhecido. De acordo com a referida autora, o grupo de pesquisadores da qual fazia parte, procurava um nome mais abrangente do que classificação (primeiro nome pelo qual a área era conhecida). A primeira ideia para o nome foi pensada a partir da tradução de *Wissensordnung*(2), título do seu livro de 1974. Como a tradução para o inglês não era adequada ela propôs o uso da expressão que foi utilizada por Bliss, em 1929, *Organization of Knowledge*, mas na ordem inversa *Knowledge Organization* (Lima; Alvares, 2012).

Vale destacar que foi a partir de 1989, com a fundação da *International Society for Knowledge Organization* (ISKO), considerada como a primeira sociedade acadêmica internacional dedicada à teoria e prática da organização do conhecimento (McIlwaine; Mitchell, 2008), sob

a presidência de Ingetraut Dahlberg, que se começou a pensar mais detidamente em relação a essa temática. Desde sua fundação, a referida instituição tem voltado aos estudos que abordam aspectos interdisciplinares na questão de organização do conhecimento e tratamento da informação destacando, nesse contexto, os trabalhos da própria Ingetraut Dahlberg, de Claire Begthol, Francisco Javier Garcia Marco, Miguel Angel Esteban Navarro, Mário Barité e Birger Hjørland, dentre outros.

In this sense, the International Society for Knowledge Organization – ISKO has a special contribution since it is considered a space for scientific discussions about knowledge organization as well as by the study of different methods and approaches in order to facilitate access to knowledge by the society (Guimarães; Oliveira; Grácio, 2012, p. 31).

De acordo com Barité (1997, p. 106), a Organização do Conhecimento é “uma disciplina de formação recente que estuda as leis, os princípios e os procedimentos pelos quais se estrutura o conhecimento especializado em qualquer disciplina”. De certo modo é uma disciplina de convergência teórico-metodológica, pois se nutre de elementos da Linguística, da Documentação, da Informática e da Comunicação (Barité, 2000). Além disso, mantém vínculos com os âmbitos que se ocupam da produção, do pensamento científico (Ciência da Ciência, Filosofia da Ciência, Sociologia da Ciência, e de sua classificação - a Taxonomia e a nomenclatura entre outros). Todo empenho realizado em função da Organização do Conhecimento justificamos porque a humanidade só avança na medida em que se sistematiza o saber acumulado para cumprir determinados propósitos e só através do estabelecimento de consensos sobre este saber, se torna possível o intercâmbio, a comunicação, o debate, a difusão e os modos de circulação do conhecimento especializado (Barité, 2001).

A narrativa da evolução da área confunde-se em sua origem e em seus primeiros passos, com as histórias particulares da Ciência da Informação (e mais especificamente da Biblioteconomia), contribuindo para formação e desenvolvimento destas disciplinas, desde a parcela que se ocupa dos princípios e das técnicas de tratamento documental de conteúdo (Barité, 2001). O estudo a respeito da Organização do Conhecimento ao longo do tempo, passou por distintos interessados. No período pré-científico da antiguidade foram os filósofos, como Aristóteles, que tiveram (e ainda têm) grande influência nas bases lógicas e epistemológicas do pensamento ocidental (Barité, 2001; Miranda, 1999). A partir da separação entre ciência e

filosofia, a tentativa de sistematização do conhecimento coube primeiro aos monges e em seguida aos primeiros cientistas laicos responsáveis pelas grandes revoluções intelectuais da história da humanidade. Na era Moderna passou a ser alvo dos enciclopedistas como Avicenna, Hugo de São Vitor, Vincente de Beuvais, Bartholomaeus Angelicus e também de educadores como Comenius, J.H. Alsted, W. Ratke (Amorim Neto, 2013).

A partir do século XIX, com a expressiva generalização das bibliotecas públicas como suporte educativo e cultural da expansão econômica, impulsionada pela Revolução Industrial, começa-se a pensar com mais profundidade na necessidade de criação de instrumentos específicos para o armazenamento e a recuperação de documentos de um modo estruturado sendo também a partir do “terceiro quartel do século XX [...] preocupação também dos Cientistas da Informação” (Miranda, 1999, p. 68). Até o presente muitos filósofos, epistemólogos, cientistas e historiadores da ciência moderna continuam preocupados por organizar o cada vez mais complexo mapa do saber humano (Barité, 2001).

Hjørland (2003) lembra que o conceito da Organização do Conhecimento é muito amplo, pois este se encontra envolvido com a divisão social do trabalho (disciplinas); as instituições sociais (universidades); as linguagens e sistemas simbólicos; os sistemas conceituais e teóricos; e ainda, com as literaturas e gêneros. O autor acredita que em muitas ocasiões a Ciência da Informação tem ignorado o amplo significado das teorias da Organização do Conhecimento, essenciais para a construção de eficientes sistemas de organização de informação.

[...] a organização do conhecimento procura, então, oferecer um modelo conceitual adequado com as diversas práticas e atividades sociais vinculadas com o acesso ao conhecimento e pretende operar como instrumento de tratamento da informação e de gestão de uso da informação, abarcador e integrador dos fenômenos e aplicações vinculados com a estruturação, a disposição, o acesso e a difusão do conhecimento socializado(3) (Barité, 2001, p. 39-40, grifos do autor).

No mesmo sentido, Rabello e Guimarães (2006, p.9) argumentam que a OC “apresenta como principal núcleo epistemológico o conhecimento registrado, ou seja, aquele materializado em documentos”. Como disciplina, dá conta do desenvolvimento de técnicas para a construção, a gestão, o uso e a avaliação de classificações científicas, taxonomias, nomenclaturas e linguagens documentárias. Aporta, por outra parte,

metodologias de uso e recuperação por linguagem natural (Barité, 2001).

Esta visão integral do conhecimento é de suma importância já que a ela se associam as classificações filosóficas ou científicas do saber com as classificações destinadas à organização de documentos em bibliotecas, arquivos e outras unidades de informação abrindo, assim, maiores perspectivas para um importante desenvolvimento disciplinar e interdisciplinar no âmbito da Biblioteconomia, da Documentação e da própria Ciência da Informação.

Quanto a área de Organização da informação, podemos afirmar de acordo com Medeiros (2010, p.40) “que [a mesma] tem suas origens principalmente na Documentação e também na Biblioteconomia” e no contexto da Ciência da informação, segundo Almeida (2009), é um núcleo de pesquisa fundamental visto que para a disseminação da informação é preciso primeiro organizá-la. Ela é composta por duas dimensões relativas ao tratamento da informação – a dimensão descritiva, voltada aos elementos relativos à forma dos documentos (como na catalogação descritiva) e a dimensão temática, voltada aos conteúdos informacionais. Sua história conceitual nos moldes atuais, manifestada notadamente na catalogação, na classificação, na indexação e na análise documental, pode começar a ser contada a partir do surgimento da Classificação Decimal de Dewey – CDD em 1876, e da obra *Rules for a Dictionary Catalog* de Cutter (Café; Sales, 2010).

Tanto Barreto (2002) quanto Smit (2009), consideram a organização da informação como o núcleo da Ciência da informação sendo um dos seus papéis fundamentais o estabelecimento de arranjos a partir de pontos de vista, não somente no contexto das bibliotecas, mas no de instituições de diversas naturezas (Andrade, 2010, p. 16). Nesse sentido, Smit assim se manifesta:

[...] a organização da informação não constitui somente uma imperiosa necessidade para que o acesso a mesma possa ser ativado, mas é condição *sine qua non* para o sistema de informação “faça sentido”, ou seja, que o mesmo cumpra seu papel social. Informação acumulada, sem organização, não é nada mais do que um conjunto de informações que “nada dizem”. Em função da discussão acima venho considerando que a organização da informação constitui o “núcleo duro” da área, aquilo que a diferencia em relação às outras áreas que trabalham com a informação, concorrendo substancialmente para a constituição da identidade da Ciência da informação (Smit, 2009, p. 62).

A função da OI é possibilitar a recuperação das informações, para que elas possam servir como instrumento para tomada de decisões, produção e disseminação do conhecimento em vários contextos como o acadêmico e o empresarial, por exemplo, através de procedimentos adequados. Lima e Álvares (2012, p. 35) afirmam que “o principal objetivo da organização da informação é recuperar objetos informacionais, que são informações registradas nos mais variados suportes – textos, imagens, registros sonoros, representações cartográficas e páginas web, entre outros”.

Café e Sales (2010) consideram que a Organização da informação é abordada na Ciência da informação sob dois aspectos: enquanto espaço investigativo que fornece os pressupostos teóricos e metodológicos ao tratamento da informação e enquanto atividade operacional inerente ao fazer profissional relativo ao tratamento da informação. Os autores afirmam ainda que os saberes da OI no espaço investigativo alicerçam os fazeres dessa organização como atividade aplicada. No entanto, essa atividade não deve ser vista apenas como um trabalho meramente técnico porque como afirma Andrade (2010), definitivamente não o é uma vez que coloca em jogo operações cognitivas e intelectuais, não se caracterizando como um mero fazer operacional.

Muitas vezes, as atividades de organização da informação são vistas apenas como tarefas técnicas, descontextualizadas dos objetivos institucionais. Nem sempre se considera o papel social da OI na circulação do conhecimento considerando as necessidades dos usuários da informação (Andrade, 2010, p. 127)

A OI tem como constituição epistemológica abordagens do tratamento temático da informação a partir dos fundamentos da catalogação de assuntos, da classificação, da indexação e da análise documental. Nesse sentido, devemos considerar a relevância indiscutível de suas práticas no interior da CI tanto no que diz respeito à análise temática da informação quanto na análise dos atributos físicos dos documentos (Almeida, 2009).

### 3. Organização da informação e Organização do conhecimento: análise terminológica

A definição dos termos fundamentais de qualquer ciência para o seu adequado desenvolvimento é uma questão defendida por vários autores das mais variadas disciplinas. Nesse sentido, a importância da Terminologia ganhou força no século XIX com a internacionalização pro-

gressiva da ciência, por exemplo, botânicos (1867), zoólogos (1889), e químicos (1892) expressam essa preocupação em encontros internacionais (Cabré, 1993, p. 21 *apud* Tálamo; Lenzi, 2006, p. 5). De acordo com Garrido e Café (2011, p. 4) “a terminologia de um domínio científico, seu uso corrente e suas mudanças que ocorrem com o tempo, determinam o nível de maturidade e também a consolidação do domínio de conhecimento em questão”.

Nessa direção, Benveniste (1989, p. 252) afirma que sem a terminologia não existiria ciência, uma vez que na sua visão “uma ciência só começa a existir ou consegue se impor na medida em que faz existir e que impõe os seus conceitos, através de sua denominação”. Ainda de acordo com o autor a ciência não tem outro meio de estabelecer a sua legitimidade senão por especificar seu objeto denominando-o. Na especificação mesma de seu objeto, o autor sustenta que é preciso denominá-lo e nesse sentido “Denominar, isto é, criar um conceito é, ao mesmo tempo, a primeira e a última operação de uma ciência” (BENVENISTE, 1989, p. 252).

Nessa perspectiva, a Ciência da Informação para a realização de suas atividades de análise, representação, ordenação, armazenamento e recuperação da informação tem a informação terminológica como uma ferramenta imprescindível. Ela surge do estudo dos conceitos sob seus aspectos teóricos e metodológicos e trata ainda de sua representação por meio da língua de especialidade e dos termos inseridos em uma determinada área.

Dessa forma, a terminologia é de fundamental importância uma vez que para se controlar a dispersão e facilitar a comunicação entre os especialistas, é vital que se controle a linguagem. Cabré (1993, p.43 *apud* Tálamo; Lenzi, 2006, p. 5) defende que “[...] a terminologia é uma ferramenta básica para a comunicação especializada (uma boa terminologia dá garantias a seus usuários sobre a precisão e eficácia da comunicação)”. Como poderia ser essa contribuição? No entendimento de Tálamo e Lenzi (2006), a terminologia pode auxiliar nos seguintes processos: permutar informação entre diferentes grupos; fornecer uma descrição adequada sobre um domínio; melhorar o entendimento de um domínio; evitar distorções cognitivas; permitir reuso de conceitos em domínios onde o reuso é importante.

Para Lara (2001, p.10), “[...] a formação de um conceito é fundamental na organização do conhecimento porque provê os meios necessários ao reconhecimento dos objetos e seu agrupa-

mento em unidades significativas num domínio particular.” A fragilidade conceitual compromete a estruturação epistemológica, por isso quando se fala da epistemologia da CI os aspectos conceituais e terminológicos não podem ser deixados de lado (Smit; Tálamo; Ko-bashi, 2004).

Buscando analisar alguns conceitos afeitos aos termos OI e OC, apresentamos nos anexos 1 e 2, dois quadros com definições disponíveis na literatura. Consideramos que a perspectiva terminológica tem instrumental para identificar conceitos, termos e organizá-los de forma sistêmica. Para a realização de qualquer trabalho de natureza terminológica “[...] é necessário a adoção de alguns critérios que delimitem o corpus de trabalho” (Siqueira, 2011, p. 93). Nesse sentido, delimitamos como procedimento metodológico para a análise terminológica dos termos OI e OC as seguintes fases: levantamento bibliográfico, leitura e coleta de termos, análise dos termos, síntese das noções (Siqueira, 2011). Para tanto, adotamos a seguinte ficha terminográfica:

| <i>Termo</i>   | <i>Contexto</i>   | <i>Referência</i>   |
|--|---|---|
| “Palavra, grupo de palavras, que designam uma noção de uma área específica”. | “Parte de um texto ou enunciado, em que está inclusa uma unidade lexical, com a função de contribuir para a determinação do seu significado”. | “Pessoa, organização ou obra de referência que fornece informação ou documentos que servem para atestar o uso de um termo”. |

Quadro 1 – *Ficha de coleta (Siqueira, 2011)*

Apresentamos nos anexos 1 e 2 os dados a partir do levantamento bibliográfico e da coleta e análise dos termos Organização da informação e Organização do conhecimento como especificado anteriormente.

A partir da análise das definições apresentadas pudemos chegar às seguintes constatações: a) a importância dos estudos de classificação que apontam como o cerne do domínio de Organização do conhecimento (esse foi o primeiro nome pelo qual a área foi conhecida); b) alguns termos sinônimos da OC como os apontados por Langridge; c) um ponto de vista tecnopragmático apontado por Vickery que, de acordo com Amorim Neto (2013), restringe os profissionais da informação como meros reagentes que respondem a estímulos a partir de um agente; d) a indicação de que OI e OC supostamente são a mesma coisa e que o uso de

uma ou outra expressão está restrita ao uso pessoal do autor como mostram as definições de Smiraglia e Anderson; e) a relação de complementariedade que existe entre esses dois termos; f) a visão de que a organização da informação tende a ser uma atividade pragmática de arranjo de acervos ou objetos informacionais sem considerar a sua contribuição social para essa questão, como apontado anteriormente;

Podemos constatar, de acordo com Andersen (2002), que a Organização do conhecimento tem uma longa história no contexto da Ciência da informação, no entanto, essa tradição da área tem sido marcada pela busca de técnicas, normas e regras para a OC ao invés de haver se esforçado para uma compreensão mais profunda de seu papel na produção e comunicação do conhecimento na sociedade. O resultado disso, segundo o autor, é uma concepção bastante restrita da OC no contexto da BCI ignorando a ampla organização social do conhecimento. Devemos ressaltar, no entanto, algumas exceções como os estudos de Bliss (1929), Egan e Shera (1952), Patrick Wilson (1968) e Birger Hjørland (1994, 1997), os quais têm defendido o papel da OC na produção e comunicação do conhecimento na sociedade (Andersen, 2002).

Como a análise terminológica não será nosso principal foco no trabalho, passamos à abordagem epistemológica sobre essas duas áreas que se tem tornado centrais no contexto da Ciência da informação.

#### 4. Organização da informação e do conhecimento: reflexão epistemológica

Organização da informação e Organização do conhecimento são utilizados em vários contextos para definir o escopo de atuação de instituições, grupos de pesquisas, cursos e disciplinas em Ciência da Informação. No entanto, o que podemos perceber através das leituras de algumas obras que tratam dessa temática é que muitas vezes esses termos são utilizados sem a devida delimitação do contexto de aplicação, o que revela falta de clareza quanto à delimitação dos mesmos, como destacado por alguns autores (Brascher; Café, 2008; Lima; Alvares, 2012).

Essas temáticas constituem-se em importantes objetos de estudos para a Ciência da Informação e, por esse motivo, a busca de uma visão mais integralizadora da área de Organização do conhecimento é um assunto que vem sendo debatido no contexto da *International Society for Knowledge Organization* (ISKO). López-Huertas e Jimenez Contreras (2004, p.137) já pontuavam que: “*Knowledge organization, on the oth-*

*er hand, is a field whose boundaries are not always clearly defined. It stands as a broad concept, and not all authors make their understanding explicit*”.

Smiraglia (2005, p. 139), editor da revista da referida instituição e seu principal meio de comunicação, já questionava-se: “*What exactly is knowledge organization?*” E ele mesmo constata que: “*It turns out there are many different definitions and not all scholars within the domain agree*”. No mesmo sentido, López-Huertas (2008, p.114) posiciona-se: “*The concept of KO itself is under revision nowadays. LIS scholars argue that the conceptual limits of what has been understood by KO do not correspond with how KO is viewed today. It calls for reconsideration*”.

Conforme relatam McIlwaine e Mitchell (2008), na nona conferência internacional da ISKO, realizada em Viena (2006), durante conversas informais e discussões no seio do Conselho Consultivo Científico, ficou claro que havia uma necessidade de apresentar uma definição compartilhada do campo Organização do conhecimento, uma vez que, embora a maioria dos membros da ISKO seja proveniente dos campos de Biblioteconomia e Ciência da informação, o interesse nessa temática não está restrito a eles. O interesse de outras áreas nessa questão, segundo os autores, acabou levando também a uma confusão sobre o significado de organização do conhecimento e a sua relação com outros campos como a gestão do conhecimento.

Essa questão da relação da OC com outras áreas passa também pela própria denominação do campo. Antes de chegar ao uso corrente de como é conhecido atualmente, *Knowledge Organization*, à área foram sugeridos os seguintes termos de acordo com Anderson (2003 *apud* LIMA; ALVARES, 2012, p. 45): a) indexação e classificação; b) ciência cognitiva; c) descrição e organização; d) recursos de informação; e) ciência da biblioteca e da informação; f) organização de recursos do conhecimento; g) organização da informação.

Como observamos na citação, uma outra questão que ainda não está totalmente resolvida na área é a sua relação com a Organização da informação, que como podemos notar, foi uma das sugestões para a denominação da área. Nesse sentido, no encontro da ISKO realizado em 2012, na Índia, Hjørland aborda essa questão com um artigo intitulado *Knowledge Organization = Information Organization?* em que procura fazer uma análise de quatro termos muito próximos que estão presentes na literatura de

Biblioteconomia e Ciência da informação, quais sejam: *Information organization (IO)*, *Organization of information (OI)*, *Information architecture (IA)* e *Knowledge organization (KO)* e questiona-se se esses termos devem ser considerados sinônimos. O autor, em sua argumentação, afirma que: “*Apparently the data shows that these terms should not be considered synonyms because each of the terms IO, OI, IA and KO produce a different set of high ranked authors, journals and papers*” (Hjorland, 2012, p.8).

No mesmo texto, o autor ainda questiona se: “*Are there differences in meaning or are the different expressions attributable, in part, to what Konrad (2007) termed —poor terminological hygiene?*” Antes desse evento, outros autores já tinham se dedicado a essa questão como Brascher e Café (2008), através de um artigo intitulado *Organização da Informação ou Organização do conhecimento?*

Percebe-se, portanto, que as investigações a respeito dessa temática tem um interesse para o campo de Organização do conhecimento visto que alguns autores como Hjorland (2008, 2012), Brascher e Café (2008), Café, Barros e Santos (2014) e Café, Agustín Lacruz e Barros (2012) estão aprofundando suas pesquisas a respeito da definição e maior sistematização de OI e OC uma vez que, o emprego desses termos tem gerado muitas discussões no que diz respeito ao seu uso por diversos autores, proporcionando significativos debates. Desse modo, inferimos que estamos distantes de um consenso. A esse respeito Lima e Alvares (2012, p. 38) assim manifestam-se:

[...] organização e representação da informação e organização e representação do conhecimento, apesar das distinções de caráter teórico e prático [...] estão em constante interação e possuem fortes interseções. Tal relação ocasiona, as vezes, o uso inadvertido de um termo pelo outro como se fossem sinônimos, já que trazem em seu bojo a dinâmica complexa da interação e interdependência existentes entre conhecimento e informação, que a ciência da informação e as demais disciplinas (ciências cognitivas, linguística, ciência da computação, etc), cada qual com seu foco, estão em constante busca de compreensão.

Embora Vignoli, Almeida e Catarino (2014, p. 124) argumentem que “[...] em um primeiro momento, pode-se considerar que existe apenas uma linha tênue que separa o significado desses termos e que essa diferença poderia não ser tão importante, principalmente se ponderado que alguns autores tratam a informação e conhecimento com sentidos semelhantes”, procuramos aqui argumentar no sentido contrário a essa visão. Não podemos conceber que a dife-

rença entre os termos seja algo menos importante ou irrelevante, visto que pode levar-nos a entendimentos superficiais que não contribuem para o amadurecimento da área. A falta de clareza quanto ao uso dos termos de uma disciplina é uma forma de barrar o crescimento e fundamentação da mesma e também uma forma que prejudica a comunicação científica, sendo esta de fundamental importância para o estabelecimento de qualquer ciência. “Para evitar ruídos no processo de transmissão de informações, a comunicação científica depende da precisão e clareza na delimitação dos termos e respectivos conceitos empregados no sistema referencial de uma área do conhecimento” (Café; Agustín Lacruz; Barros, 2012, p. 286).

O amadurecimento de uma determinada área do conhecimento passa necessariamente pelo uso de termos que sejam bem delimitados. “A definição de um conceito impacta no direcionamento que é dado ao desenvolvimento de pesquisas que partem de determinada definição” (Café; Agustín Lacruz; Barros, 2012, p. 286). Dessa forma, o emprego de determinado termo quando associado a um conceito, a um significado terá consequências na sua relação com os demais termos que constituem um campo científico (Capurro; Hjorland, 2007).

O uso preferencial dos termos OC e OI passa por algumas metodologias da área e subcampos da Ciência da informação. Como exemplo, podemos citar o fato de que em Linguística documental usam-se os termos *Atividade documentária* e *Organização da informação* (4). Em *Análise de Domínio*, por sua vez, prefere-se o termo *Organização do conhecimento* (Andrade, 2010, p. 132-133).

Barité (2001), ao discorrer sobre essa temática, pergunta-se por que não usamos a expressão *Organização da Informação* em lugar de *Organização do Conhecimento*, já que a palavra *informação*, além de possuir significados muito relevantes, também está vinculada ao nome genérico de várias disciplinas afins como *Ciência da informação*. O autor explicita, a partir de uma argumentação sobre o uso de dados, *informação* e *conhecimento*, três ideias que o levam a preferir *Organização do conhecimento*, tal como apresentamos a seguir:

La primera es que no puede haber conocimiento sin información y un proceso subjetivo sobre la misma. La segunda es que la información es una expresión social objetiva, mientras que el conocimiento es un acto individual, y por tanto, subjetivo. La tercera es que toda información es materia prima, y todo conocimiento es un producto final con determinado grado de elaboración.

Portanto, na visão do autor, o conhecimento engloba a informação e uma vez socializado transforma-se em nova informação, o que leva à constatação de que organizamos, na realidade, o conhecimento, visto que a informação seria a matéria-prima da construção do conhecimento e, este se apresenta como um produto finalizado, daí surge a sua preferência pelo uso dessa expressão para designar o campo.

Hjorland, por seu turno, ao tratar do uso de OI ou OC, assim manifesta-se:

Knowledge Organization is one among many contemporary fields which try to play a role in the future environments of communicating and exchanging knowledge. Among the competitors are Knowledge Management and Computer Science. Much knowledge may be shared among such fields, but is important for each field to develop a clear identity and a history of its own. KO has in particular been connected with LIS and has aimed at supporting learning and research activities, which may be one of the important pillars on which to base the field. Another related pillar is the concept of knowledge and theories of knowledge. Knowledge Organization may have a valuable theoretical base in theory of knowledge, which may be the reason why we should stick to this label as the name of our field (Hjorland, 2008, p. 99).

Em outra oportunidade, o referido autor também registra a sua preferência pelo termo OC: *“This study has argued that it might be a good idea to continue to use the term knowledge organization and to connect KO better with other disciplines devoted to the study of knowledge”* (Hjorland, 2012, p. 13).

Para sustentar essa preferência, o autor argumenta que:

My suggestion is, in other words, that the term — knowledge moves us relatively away from fields like information theory and computer science towards fields such as social semiotics, science studies and the study of documents and their role in human activities (— activity theory). I believe that such a —social turn is very important for developing LIS as a scholarly discipline (HJORLAND, 2012, p. 12).

Presumimos que o autor procura distanciar-se do termo informação, sob a alegação de que este estaria muito ligado a campos como a Teoria da informação e a Ciência da computação, os quais também têm no uso do referido termo suas principais preocupações. Dessa forma, a preferência por “conhecimento” evitaria certas confusões semânticas e terminológicas com essas áreas e também o uso do termo “informação” que, na sua opinião, tornou-se popular em Biblioteconomia e Ciência da informação muito mais por causa de seu apelo do que por méritos científicos (Hjorland, 2012).

Rabello e Guimarães (2006), ao tratarem dessa questão também posicionaram-se favoravelmente ao uso da expressão OC e assim justificam a escolha:

Observa-se, ainda, o uso indiscriminado das expressões Organização do conhecimento e Organização da Informação como dispositivos do mesmo processo. No entanto, preferimos utilizar a denominação Organização do Conhecimento, seja por ter sido a oficialmente adotada pela ISKO, seja por considerar, como conhecimento, uma informação contextualizada por sua produção e pelo seu uso, em um processo cíclico.

Taylor (2004, p.99) também soma-se à linha da discordância entre os partidários que acreditam que se organiza informação e aqueles que afirmam que se organiza conhecimento. Para demonstrar essa questão a autora exemplifica da seguinte forma:

There is a running argument between those who believe we are organizing information and those who believe we are organizing knowledge. It seems to me that I can use my knowledge to write a book, but until you read that book, understand it, and integrate it into your own knowledge, it is just information. On the other hand, the book would be a representation of my knowledge, although it would be an imperfect representation, in the sense that some concepts might not be explained as clearly or as fully as I truly understand them. I believe we organize information so that others can find it, read or otherwise absorb it, and use it to add to their own store of knowledge. But it is also fair to say that we organize representations of knowledge.

A autora, nesse sentido, tem uma preferência pela expressão organização da informação e não do conhecimento, embora como ela mesmo afirma, seria justo falar de organização de representações do conhecimento.

O que se pode verificar através desse levantamento e reflexão, é que, às vezes parece que a utilização dos termos OI e OC se refere ao mesmo objeto dado que a escolha do nome responde mais a preferências pessoais e políticas de determinado pesquisador que de uma reflexão sobre a natureza do processo.

Isso pode ser notado na literatura da área quando alguns autores se referem aos mesmos como sendo sinônimos, dentre os quais se podem destacar os trabalhos de Tristão, Fachin e Alarcon (2004), Rizzi (2008) e Abreu e Monteiro (2010).

Por outro lado, Almeida (2011), os considera distintos, porém complementares, o que também defendemos no presente trabalho.

Na mesma direção, Brascher e Café (2008), argumentam que eles não são sinônimos nem têm a mesma finalidade o desenvolvimento dessas áreas. Para fundamentar essa posição, desenvolveram o que elas chamaram de conceito norteador para fazer uma melhor distinção entre esses dois termos. Apoiando-se no pressuposto de que informação e conhecimento são conceitos distintos, afirmam as autoras que OI e OC também o são: organização da informação é um processo que envolve a descrição física e de conteúdo de objetos informacionais tendo como objeto os registros da informação, mas em seu caráter individual. O produto da OI é a representação da informação, na forma da descrição física e descrição temática resultantes dos processos correlacionados – catalogação, classificação, redação de resumos e indexação (ALMEIDA, 2009). As autoras afirmam que: “[...] no contexto da OI e da RI, temos como objeto os registros de informação. Estamos, portanto, no mundo dos objetos físicos, distinto do mundo da cognição, ou das ideias, cuja unidade elementar é o conceito” (Brascher; Café, 2008).

A OC, para as autoras, é um processo de construção de modelos de mundo que se constituem em abstrações da realidade. A intenção é elaborar modelos de mundo via representação do conhecimento de uma coletividade. O produto consequente é a representação do conhecimento em um dado campo, e não somente a representação da informação (Almeida, 2009).

Hjorland (2012, p. 11) pergunta-se: “Does KO = IO? What differences does it make whether we prefer the term knowledge or the term information in LIS and in KO?” Em nossa opinião, faz toda a diferença uma vez que o uso indiscriminado de termos revela falta de amadurecimento científico e conceitual e pode levar o campo a sofrer inconsistência terminológica e epistemológica, cujas consequências mais diretas são o retardamento do avanço de pesquisas e desenvolvimentos científicos comprometendo o processo de institucionalização científica da área. Nesse contexto, Arboit (2014) baseada no pensamento de Whitley (1974), levanta a questão do baixo grau de institucionalização cognitiva que se refere a um baixo grau de ordem intelectual e, correspondentemente, falta de consenso ou compromisso comum. Nesta situação, a autora afirma que cientistas provavelmente vão aderir a valores básicos comuns e crenças a respeito da natureza de um empreendimento científico e possivelmente algumas uniformizações serão percebidas “mas seus trabalhos serão desconexos e desarticulados” (ARBOIT, 2014, p. 102).

Hjorland (2012) defende que o uso dos termos que faz parte do escopo de uma área do conhecimento deve ser feito com critérios essencialmente científicos. A partir do trabalho desenvolvido por Sheila Weber, o autor consta-tou que o uso de termos com apelo popular na mudança do nome de cursos ou áreas do conhecimento para atrair determinados públicos pode ser uma prática perigosa para a consolidação de um campo científico.

Essa é uma questão de interesse para a área uma vez que algumas pesquisas que vem sendo desenvolvidas começam a constatar que caso semelhante pode estar ocorrendo com o uso do termo Organização do conhecimento, senão vejamos. Em pesquisa realizada sobre o conceito de Organização do conhecimento nas revistas brasileiras de Ciência da informação Café, Barros e Santos (2014) constataram que alguns artigos adotam o termo Organização do conhecimento, mas sem abordar o tema propriamente dito e muitas vezes utilizam o termo nas palavras-chave ou no título, mas não o retomam ao longo do texto.

De forma semelhante, o estudo conduzido por Café, Agustín Lacruz e Barros (2012, p. 299), tendo como corpus de pesquisa os artigos publicados no periódico da ISKO, no período entre 2000 e 2010, constatou que “alguns artigos não tratavam do tema OC com especificidade suficiente para que fossem indexados com esse termo [...]”. No mesmo estudo as referidas autoras também constataram que “[...] alguns autores, por exemplo, iniciavam a abordagem de OC apontando a relação entre os conceitos de um domínio de conhecimento e, no andamento da discussão, se referiam da mesma forma à organização da informação” (Café; Agustín Lacruz; Barros, 2012, p.299).

Estudos como esses podem revelar que devido ao peso político do termo OC, ele pode estar sendo usado apenas como forma de propiciar uma maior recuperação de artigos que efetivamente não contemplem tal temática. E isso torna-se preocupante porque o campo tem se consolidado como importante fórum de pesquisas e tem avançado em questões que são centrais para o mesmo. Contudo, a questão do uso dos termos pode gerar fragmentações que seriam mais bem enfrentadas se o campo demonstrar clareza conceitual.

## 5. Considerações finais

Constatamos que a temática envolvendo o uso dos termos OC e OI ainda carece de estudos que possam sistematizar melhor o contexto de uso dos referidos termos. A análise terminológica

ca mostrou-nos que os termos ainda não seguem um padrão e, conseqüentemente, não se tem um consenso sobre o uso dos mesmos. Através da reflexão epistemológica podemos constatar que o uso dos termos não se trata de apenas uma questão de escolha por um ou outro rótulo, pelo contrário, revela que se a área não conseguir estabelecer uma diretriz para o uso apropriado dos termos básicos correrá o risco de se banalizar o uso dos termos OC e OI, deixando-os bem próximos do senso comum.

## 6. Notas

- (1) Ao analisar a obra de Bliss, que procurava por um consenso na comunidade científica e que, ao fazê-lo, acreditou que era possível identificar e mapear uma estrutura básica permanente de uma área de assunto, Hjørland e Albrechtsen (1995) ressaltam que a existência de uma estrutura básica provou estar errada e que é ingênuo considerar que o consenso garante a verdade. Porém, os autores alertam que “isso não significa automaticamente rejeitar a construção de consenso como um método e que uma importante característica da área temática pode ser o grau de estabilidade, o grau de consenso entre os pesquisadores num dado tempo (Hjørland; Albrechtsen, 1995, p. 403).
- (2) Algo como “ordem do conhecimento” a partir de uma tradução livre.
- (3) Essa expressão é utilizada para exprimir o mesmo conceito utilizado por Guimarães (1994) *conhecimento registrado* e Jaenecke (1994) *conhecimento documental* (Rabello; Guimarães, 2006, p.14).
- (4) Na Linguística documentária, o termo Atividade documentária em alguns contextos é mais amplo do que Organização da informação, quando engloba todos os aspectos da Documentação, desde a criação do documento para o sistema documentário (emissão) até a utilização da informação pelo usuário (recepção). Em outros, o termo Atividade Documentária é praticamente sinônimo de Organização da Informação, referindo-se aos processos e às operações utilizadas para registrar a informação (que incluem em Análise Documentária) de forma que ela possa ser recuperada, comunicada e interpretada. (Andrade, 2010, p. 133).

## Referências

- Abdalla, R.B.; Kobashi, N.Y. (2012). Ciência da informação: reflexões sobre a constituição de um campo científico. In: Guimarães, J.A.C.; Dobedei, V. (Orgs) (2012). *Desafios e perspectivas científicas para a organização e representação do conhecimento na atualidade*. Marília: ISKO-Brasil: FUNDEPE. 54-58.
- Abreu, J. G.; Monteiro, S. D. (2010). Matrizes da linguagem e a organização virtual do conhecimento. *Ciência da Informação*, 2(39), 9-26. Recuperado 02-02-2015 de <http://revista.ibict.br/ciinf/index.php/ciinf/article/view/1788/1363>
- Almeida, C.C. (2009). *Peirce e a Organização da Informação*: contribuições teóricas da Semiótica e do Pragmatismo. Tese (Doutorado em Ciência da Informação), Faculdade de Filosofia e Ciências, UNESP, Marília-SP.
- Almeida, C.C. (2011). Sobre o pensamento de Peirce e a organização da informação e do conhecimento. *Liinc em Revista*, 7:1 (Março 2011), 104-120.
- Amorim Neto, M.R. (2013). *Organização do conhecimento na Ciência da informação: uma análise métrica nos periódicos brasileiros (1972-2012)*. 2013. Dissertação (Mestrado em Ciência da Informação), Universidade Federal Fluminense, Rio de Janeiro.
- Andersen, J. (2002). Communication technologies and the concept of Knowledge Organization: a mediumtheory perspective. *Knowledge Organization*, 29:1, 29-39.
- Andrade, J. (2010). *A Linguística Documentária e a Análise de Domínio na Organização da Informação*. Dissertação (Mestrado em Ciência da Informação) Escola de Comunicações e Artes, Universidade de São Paulo (USP), São Paulo.
- Arboit, A.E. (2014) *O processo de institucionalização socio-cognitiva do domínio de organização do conhecimento a partir dos trabalhos científicos dos congressos da ISKO*. Tese (Doutorado em Ciência da Informação) Faculdade de Filosofia e Ciências, UNESP, Marília-SP.
- Barité, M. (1997). *Referenciales teóricos vigentes en el área de tratamiento temático de la información y su expresión metodológica*. Relatório técnico do II encontro de dirigentes dos cursos Superiores de Biblioteconomia dos países do MERCOSUL e I Encontro de Docentes de Biblioteconomia e Ciência da Informação do MERCOSUL, Buenos Aires, 27 – 29 nov. 1997. Porto Alegre : ABEED.
- Barité, Mário (2000). Los conceptos y su representación: una perspectiva terminológica para el tratamiento temático de la información. *Scire*, 6:1, (Enero-Junio 2000) 31-53.
- Barité, M. (2001). Organización del conocimiento: un nuevo marco teórico-conceptual en Bibliotecología y Documentación, *Organização do Conhecimento*. En: Carrara, K. (Org.) (2001). *Educação, universidade e pesquisa*. Marília: Unesp Marília Publicações; São Paulo: FAESP, 35-60.
- Barreto, A.A. (2002). O tempo e o espaço da Ciência da Informação. *Transinformação*, 14:1 (Janeiro/Junho 2002), 17-24.
- Benveniste, E. (1989). *Problemas de Linguística geral II*. Campinas: Pontes.
- Brascher, M.; Café, L. (2008). Organização da Informação ou Organização do Conhecimento? In: Encontro Nacional de Pesquisa em Ciência da Informação, 9, 2008, São Paulo, *Anais...* São Paulo: ANCIB, 2008.
- Brascher, M; Victorino, M. (2009) Organização da Informação e do Conhecimento, Engenharia de Software e Arquitetura Orientada a Serviços: uma Abordagem Holística para o Desenvolvimento de Sistemas de Informação Computadorizados. *DataGramaZero - Revista de Ciência da Informação*, Rio de Janeiro, 10:3 (Junho 2009). Recuperado 25-01-2015 de [http://www.dgz.org.br/jun09/Art\\_03.htm](http://www.dgz.org.br/jun09/Art_03.htm)
- Brookes, B.C. (1980). The foundations of information Science. Part I. Philosophical aspects. *Journal of Information Science*, v.2, 125-133.
- Bufrem, L. S. (2004). Levantando significações para significantes: da gestão do conhecimento a organização do saber. *Revista Eletrônica Biblioteconomia Ciência da Informação*, Florianópolis, n. esp., 1º sem.
- Café, L.M.A.; Agustín Lacruz, M.D.C.; Barros, C.M. (2012) Organização do conhecimento: análise conceitual. *Actas del X Congreso ISKO Capítulo Español*. Universidade da Coruña (Espanha) 283-302.

- Café, L.M.A.; Barros, C.M.; Santos, V.C. (2014). O conceito de Organização do Conhecimento nas revistas brasileiras de Ciência da Informação. *Revista Interamericana de Bibliotecologia*, 37:3 (Setembro-Dezembro 2014) 201-214.
- Café, L.M.A.; Sales, R. (2010). Organização da informação: conceitos básicos e breve fundamentação teórica. En: Robreo, J.; Brascher, M. (Orgs). *Passeios no Bosque da Informação: estudos sobre Representação e Organização da Informação e do Conhecimento*. Brasília, DF: Ibiict.
- Capurro, R.; Hjørland, B. The concept of information. *Perspectivas em Ciência da Informação*, Belo Horizonte, 2007, 12:1 (Janeiro/Abril 2007) 148-207.
- Cervantes, B.M.N. (2009). *A construção de tesouros com a integração de procedimentos terminográficos*. Tese (Doutorado em Ciência da Informação) Faculdade de Filosofia e Ciências, UNESP, Marília-SP.
- Dahlberg, I. (2006). *Knowledge Organization*. Recuperado 05-02-2015 de [http://www.iva.dk/bh/lifeboat\\_ko/CONCEPTS/knowledge\\_organization\\_Dahlberg.htm](http://www.iva.dk/bh/lifeboat_ko/CONCEPTS/knowledge_organization_Dahlberg.htm).
- Dahlberg, I. (2006). Knowledge Organization: a new science? *Knowledge Organization*, 33:1, 11-19.
- Dahlberg, I. (2008) Interview with Ingetraut Dahlberg - December 2007. *Knowledge Organization*, 35:2-3, 82-85.
- Esteban Navarro, M. A. (1996). El marco disciplinar de los lenguajes documentales: la Organización del Conocimiento y las ciencias sociales. *Scire*. 2:1, 93-107.
- Fernandes, G.L.; Lima-Marques, M. (2013). Considerações sobre a abrangência disciplinar da Ciência da informação: fundamentos, modelos e métodos para a organização e representação da informação. In: Ribeiro, Fernanda; Maria Elisa Cerveira (Orgs.). (2013). *Informação e/ou Conhecimento: as duas faces de Jano*. I Congresso ISKO Espanha e Portugal / XI Congreso ISKO España, 213-228.
- Fujita, M. S. L. (2001). Organização do Conhecimento: algumas considerações para o tratamento temático da informação, Organização do Conhecimento. In: Carrara, K. (Org.) (2002). *Educação, universidade e pesquisa*. Marília: Unesp Marília Publicações; São Paulo: FAESP, 29-34.
- García Marco, F.J. (1995). Los contenidos y la secuencia docente de la organización y representación del conocimiento: una propuesta interdisciplinar. *Organización del Conocimiento en Sistemas de Información y Documentación*. v. 1, p. 219-228. Recuperado 03-02-2015 de <http://dialnet.unirioja.es/descarga/articulo/2341341.pdf>
- Garrido, I.S.; Café, L.M.A. (2011). *Organização da informação: uma análise conceitual*. Florianópolis, SC: Universidade Federal de Santa Catarina.
- Guimarães, J.A.C.; Oliveira, E.T; Grácio, M.C.C. (2012). Theoretical Referents in Knowledge Organization: A Domain Analysis of Knowledge Organization Journal. En: Neelameghan, A.; Raghavan, K. S. (Eds.). *Categories, contexts and relations in knowledge organization: Proceedings of the Twelfth International ISKO Conference* (Mysore, India, August 6-9, 2012). Advances in knowledge organization, no. 13. Würzburg: Ergon, 31-38.
- Hjørland, B. (2003). Fundamentals of knowledge organization. *Knowledge Organization*, 30:2, 87-111.
- Hjørland, B. (2012). Knowledge Organization = Information Organization? NEELAMEGHAN, A.; RAGHAVAN, K. S. (Eds.). *Categories, contexts and relations in knowledge organization: Proceedings of the Twelfth International ISKO Conference* (Mysore, India, August 6-9, 2012). Advances in knowledge organization, no. 13. Würzburg: Ergon, 8-14.
- Hjørland, B. (2008). What is Knowledge Organization (KO)? *Knowledge Organization*, 35:3/2, 86-101.
- Hjørland, B.; Albrechtsen, H. (1995). Toward a new horizon in information science: domain-analysis. *Journal of the American Society for Information Science*, 46:2, 400-425.
- Kobashi, Nair Yumiko. (2007). Fundamentos semânticos e pragmáticos da construção de instrumentos de representação de informação. *Datagramazero*, Rio de Janeiro, 8:6 (Dezembro 2007). Recuperado 02-02-2015 de [http://www.datagramazero.org.br/dez07/Art\\_01.htm](http://www.datagramazero.org.br/dez07/Art_01.htm).
- Lara, M.L.G. (2001). Linguagem documentária e terminologia. *Transinformação*, Campinas, 16:3 (Setembro/Dezembro, 2004) 231-240.
- Langridge, Derek W. (2006) *Classificação: abordagem para estudantes de biblioteconomia*. Rio de Janeiro: Interciência.
- Le Coadic, Yves-François. (2004). *A Ciência da Informação*. Brasília: Briquet de Lemos Livros.
- Lima, J.L.O.; Álvares, L. (2012). Organização e representação da informação e do conhecimento. En: Álvares, L. (Org.) *Organização da informação e do conhecimento: conceitos, subsídios interdisciplinares e aplicações*. São Paulo: B4 Ed.
- López-Huertas, M.J. (2008). Some Current Research Questions in the Field of Knowledge Organization. *Knowledge Organization*, 35:2-3, 113-136.
- López-huertas, M. J.; Contreras, Evaristo J. (2004). Spanish research in Knowledge Organization (1992-2001). *Knowledge Organization*, 31:3, 136-150.
- Martinez Ávila, D. (2014). *Organização de conhecimento: conceitos fundamentais*. Notas de aula ministrada no Programa de Pós-Graduação em Ciência da Informação – UNESP, Marília-SP. (68 slides)
- McIlwaine, I.C.; Mitchel, J.S. (2008). Preface to Special Issue - "What is Knowledge Organization". *Knowledge Organization*, 35:2/3, 79-81.
- Medeiros, G.M. (2010). *Organização da informação em repositórios digitais: implicações do auto-arquivamento na representação da informação*. Dissertação (Mestrado em Ciência da Informação) Centro de Ciências da Educação, Universidade Federal de Santa Catarina, Florianópolis-SC.
- Miranda, M.L.C. (1999). A organização do conhecimento e seus paradigmas científicos: algumas questões epistemológicas. *INFORMARE*, Cad. Prog. Pós-Grad, Rio de Janeiro, 5:2 (Julho-Dezembro 1999) 64-77.
- Ortega, C.D. (2013). Aspectos teóricos, procedimentais, normativos e pragmáticos como categorias para uma epistemologia da organização da informação. En: Guimarães, J.A.C.; Dobedei, V. (Orgs) (2013). *Complexidade e organização do conhecimento: desafios de nosso século*. Rio de Janeiro: ISKO-Brasil; Marília : FUNDEPE, 2013. 34-38.
- Rabello, Rodrigo; Guimarães, José Augusto Chaves. (2006). *A relação conceitual entre conhecimento e documento no contexto da Organização do Conhecimento: elementos para uma reflexão*. En: VII Encontro Nacional de Pesquisa em Ciência da Informação (ENANCIB), Marília-SP, 19 - 22 nov. 2006.
- Rizzi, I.R.F. (2008). *A paz nos instrumentos de Organização da Informação: uma análise dos conceitos de paz e guerra, da Cultura de Paz e dos Estudos para Paz na Classificação Decimal de Dewey*. Dissertação (Mestrado em Ciência da Informação), Faculdade de Filosofia e Ciências, Marília-SP, UNESP.
- Salton, G. (1968). *Automatic information organization and retrieval*. New York: McGraw-Hill.

- Siqueira, J.C. (2011). *As noções de documento e de informação: uma abordagem terminológica*. Dissertação (Mestrado em Ciência da Informação) Escola de Comunicações e Artes, Universidade de São Paulo, USP, São Paulo.
- Smiraglia, R.P. (2002). The progress of theory in knowledge organization. *Library Trends*, 50:3, 330-349.
- Smiraglia, R.P. (2005). About Knowledge Organization: An Editorial. *Knowledge Organization*, 32:4, 139-140.
- Smiraglia, R.P. (2006). Whither Knowledge Organization An Editorial. *Knowledge Organization*, 33:1, 8-10.
- Smiraglia, R.P. (2011) *Domain Coherence Within Knowledge Organization: People, Interacting Theoretically, Across Geopolitical and Cultural Boundaries*. Recuperado 02-02-2015 de [http://www.cais-acsi.ca/proceedings/2011/73\\_Smiraglia.pdf](http://www.cais-acsi.ca/proceedings/2011/73_Smiraglia.pdf)
- Smiraglia, R.P. (2012). Organización del conocimiento: algunas tendencias en un dominio emergente. *El profesional de la información*. 21:3 (Mayo-Junio 2012) 225-227.
- Smit, J.W. (2009). Novas abordagens na organização no acesso e na transferência da informação. In: Silva, H.C.; Barros, M.H.T.C. (Orgs.) *Ciência da Informação: múltiplos diálogos*. Marília: Oficina Universitária Unesp.
- Smit, J.W.; Tálamo, M.F.G.M.; Kobashi, N.Y. (2004). A determinação do campo científico da Ciência da Informação: uma abordagem terminológica. *DataGramaZero*, 5:1 (Fevereiro 2004) Recuperado 02-03-2015 de [http://www.dgz.org.br/fev04/F\\_I\\_art.htm](http://www.dgz.org.br/fev04/F_I_art.htm)
- Svenonius, E. (2000). *The intellectual foundations of information organization*. Cambridge: The MIT Press.
- Tálamo, M.F.G.M.; Lenzi, L.A.F. (2006). Terminologia e documentação: a relação solidária das organizações do conhecimento e da informação no domínio da inovação tecnológica. *DataGramaZero - Revista de Ciência da Informação*, Rio de Janeiro. 7:4 (Agosto 2006) Recuperado 02-02-2015 de [http://www.dgz.org.br/ago06/F\\_I\\_art.htm](http://www.dgz.org.br/ago06/F_I_art.htm).
- Taylor, A.G. (2004). *The organization of information*. Englewood: Libraries Unlimited.
- Tristão, A.M.D.; Fachin, G.R.B.; Alarcon, O.E. (2004). Sistema de classificação facetada e tesouros: instrumentos para organização do conhecimento. *Ciência da Informação*, 33:2. Recuperado 02-12-2014 de <http://revista.ibict.br/index.php/ciinf/article/view/88/82>
- Universidade Federal do Rio Grande do Sul – UFRGS - Glossário da Área de Organização e Tratamento da Informação. Recuperado 03-02-2015 de <http://www.ufrgs.br/fabico/documentos-comgrads/glossario-de-ciencia-da-informacao>
- Vickery, B. (2008). *On 'knowledge organization'*. Recuperado 08-02-2015 de <http://www.lucius.me.uk/knowlorg.htm#start>.
- Vignoli, R.G.; Almeida, P.O.P.; Catarino, M.E. Folksonomias como ferramenta da organização e representação da informação. *Rev. digit. bibliotecon. cienc. inf.* São Paulo, 12:2 (Maio-Agosto 2014) 120-135.
- Wilson, T. (2002) The nonsense of "Knowledge management". *Information Research*, 1:1 (Oct. 2002) Recuperado 01-02-2015 de <http://informationr.net/ir/8-1/paper144.html>
- Zins, Chaim. (2007) Conceptions of Information Science. *Journal of the American Society for Information Science and Technology*, 58:3,335-350.

## Anexo 1 – Ficha terminográfica sobre a organização do conhecimento

| Termo   | Contexto  | Referência  |
|---|---|---|
| Organização do conhecimento   | “baseada em práticas e atividades significativas de armazenagem de recuperação de documentos no decorrer da sistematização do conhecimento humano, a Organização do Conhecimento tem suas origens na criação, por bibliotecários, de sistemas de classificação como instrumentos de organização temática de documentos para armazenagem”.   | Fujita (2001, p.29)                                       |
|   | “Knowledge organization, at least as it is practiced inside the domain of library and information science, has been largely (up to now) the province of the construction of tools for the storage and retrieval of documentary entities.”   | Smiraglia (2002, p. 331)                                  |
|   | “In the Library and Information Science community (LIS) Knowledge Organization (KO) means especially the organization of information in bibliographical records, including citation indexes, full text records and the Internet”.   | Hjørland (2003, p. 87)                                    |
|   | “The description of documents, their content, features and purpose, and the organization of these descriptions so as to make these documents and their parts accessible to persons seeking them or the messages that they contain. Knowledge organization encompasses every type and method of indexing, abstracting, cataloguing, classification, records management, bibliography and the creation of textual or bibliographic databases for information retrieval ... In the context of library and information science, the organization of knowledge (often called the organization of information) is the organization of documented messages in which knowledge or information is represented.”  | Anderson (1996, p. 336 <i>apud</i> Andersen, 2002, p. 32) |
|   | “Knowledge Organization is the science of structuring and systematically arranging of knowledge units (concepts) according to their inherent knowledge elements (characteristics) and the application of concepts and classes of concepts ordered by this way for the assignment of the worth knowing contents of referents (objects/subjects) of all kinds.”   | Dahlberg (2006, online)                                   |
|   | “La Organización del Conocimiento, em particular, procura brindar subsídios teóricos (y retroalimentarse) con todo lo relativo al <i>tratamiento de la información</i> , particularmente el <i>tratamiento temático de la información</i> , y de un modo menos específico – pero no menos importante – a la <i>gestión del uso social de la información</i> (grifos do autor)”.   | Barité (2001, p. 38)                                      |
|   | “a categoria temática Organização do Conhecimento (Knowledge Organization) relaciona-se originalmente às tarefas de classificar, indexar e representar o conhecimento por meio de registros informatizados para atender as necessidades mais urgentes de informação”.   | Bufrem (2004, p. 2)                                       |
|   | “[...] la Organización del Conocimiento es la disciplina específica que se dedica dentro de la Ciencia de la Información Documental al estudio de los fundamentos teóricos del tratamiento y la recuperación de la información y a la construcción, mantenimiento, uso y evaluación de los instrumentos lógico-lingüísticos más adecuados para controlar los procesos de representación, clasificación, ordenación y almacenamiento del contenido informativo de los documentos con el fin de permitir su recuperación y comunicación. Se ocupa, por tanto, de los principios y herramientas puestos en acción para la gestión del conocimiento humano desde una triple perspectiva: su representación, su organización y su comunicación documental”.              | Esteban Navarro (1996, p. 97-98),                         |
|   | “Knowledge organization (sometimes also called information organization) is a key sub-domain of information science, which is devoted to the conceptual order of knowledge. In the broadest sense KO is the arena in which the heuristics of ordering knowledge are studied. More narrowly, within information science, KO is the arena in which classification and ontology, thesauri and controlled vocabulary, epistemology and warrant are studied and in which applications are developed and tested (often, resource description is included as well)”.   | (Smiraglia, 2011, p. 01)                                  |
|   | “La organización del conocimiento es una disciplina científica – sujeta al método científico – de carácter aplicado – pues tiene um objetivo práctico específico: optimizar la circulación del conocimiento em nuestras sociedades -, y cuyo apoyo teórico se encuentra distribuído por um amplio número de ciências que van desde la Lingüística, passando por la Teoria de la Clasificación hasta llegar a la Inteligencia Artificial o Informática Avanzada”   | García Marco (1995, p.220)                                |
|   | A ORGANIZAÇÃO DO CONHECIMENTO é a expressão mais abrangente para designar a função da biblioteca desempenhada pela classificação. Indica a habilidade não apenas para identificar itens de informação específicos e definidos de forma precisa, mas também para demonstrar a completa gama de assuntos disponíveis na biblioteca e suas relações entre si. Essa é uma função educacional muito importante das bibliotecas e tem sido um pouco negligenciada nos modernos estudos de classificação. É significativo o fato que enquanto escritores do pré-guerra usavam a expressão organização do conhecimento, os escritores modernos a têm substituído, frequentemente, por expressões como RECUPERAÇÃO DA INFORMAÇÃO ou ARMAZENAGEM E RECUPERAÇÃO DA INFORMAÇÃO. | Langridge (2006, p.19)                                    |
| To organise knowledge is to gather together what we know into a comprehensive organised structure, to show its parts and their relationships. This is the work of scholars and encyclopaedists. It is not the role of the information profession. Our tasks are to make knowledge (whether organised or unorganised) available to those who seek it, to store it in an accessible way, and to provide tools and procedures that make it easier for people to find what they seek in those stores. | Vickery (2008, online)  |   |

| Termo | Contexto   | Referência             |
|-------|--|------------------------|
|       | In the narrow meaning Knowledge Organization (KO) is about activities such as document description, indexing and classification performed in libraries, bibliographical databases, archives and other kinds of “memory institutions” by librarians, archivists, information specialists, subject specialists, as well as by computer algorithms and laymen. KO as a field of study is concerned with the nature and quality of such knowledge organizing processes (KOP) as well as the knowledge organizing systems (KOS) used to organize documents, document representations, works and concepts. Library and Information Science (LIS) is the central discipline of KO in this narrow sense (although seriously challenged by, among other fields, computer science). In the broader meaning KO is about the social division of mental labor, i.e. the organization of universities and other institutions for research and higher education, the structure of disciplines and professions, the social organization of media, the production and dissemination of “knowledge” etc. | Hjorland (2008, p. 86) |

## Anexo 2 – Ficha terminográfica sobre organização da informação

| Termo                            | Contexto  | Referência   |
|----------------------------------|---|--|
| <b>Organização da informação</b> | “a organização da informação é um processo de arranjo de acervos tradicionais ou eletrônicos realizado por meio da descrição física e de conteúdo (assunto) de seus objetos informacionais”.  | Café e Sales (2010, p. 118)  |
|                                  | “É um processo que envolve a descrição física e de conteúdo dos objetos informacionais. O produto deste processo descritivo é a representação da informação, entendida como um conjunto de elementos descritivos que representam os atributos de um objeto informacional específico”.   | (Brascher; Café, 2008, p. 06)  |
|                                  | “Organização da Informação é um termo genérico utilizado em estudos de informação para incorporar a ciência da organização do conhecimento com as aplicações de recuperação de informação bibliográfica”.   | (Martinez Avila, 2014)   |
|                                  | “Organização da Informação é o conjunto de processos (representação descritiva e temática) que objetivam a organização de documentos, visando a posterior recuperação da informação”.   | (Universidade Federal do Rio Grande do Sul – UFRGS – Glossário da Área de Organização e Tratamento da Informação). |
|                                  | “Resumidamente, pode-se afirmar que o objetivo da organização da informação é dar suporte ao fluxo de tratamento e recuperação dos objetos informacionais”.   | (Brascher; Victorino, 2009, p. 04)   |
|                                  | Organização que é realizada com fins de promoção do acesso [ao conhecimento] visando seu uso e nova produção, ou seja, a atividade de elaborar representações que possam ser significadas e manipuladas a favor de um certo público.  | Ortega (2013, p.34)  |
|                                  | A Organização da informação, enquanto campo disciplinar tem como uma de suas preocupações mais importantes propor princípios e métodos para representar “[...]conhecimento institucionalizado e funcionalizado como informação” (Abril, 2004, p. 9). Em outras palavras, procura-se criar métodos e instrumentos para fabricar informação documentária. Indexar, resumir e construir linguagens de representação são os termos técnicos que denominam essas operações. Com efeito, mais do que nunca, a informação é indexada por palavras (justapostas, relacionadas graficamente em mapas estáticos ou dinâmicos) que são também utilizadas para busca, ou seja, para indexar a pergunta do usuário | Kobashi (2007, online)   |

---

# Referentes teóricos basilares na ISKO-Ibérico e suas contribuições para a Organização do Conhecimento (OC)

Fundamental theoretical references *in* ISKO-Ibérico and their contribution to Knowledge Organization (KO)

---

**Doutoranda Maria Aparecida Pavanelli (1) Prof. Dra Ely Francina Tannuri de Oliveira (2), Prof. Dra. Leilah Santiago Bufrem (3), Doutorando Marcos Aparecido Rodrigues do Prado (4) Doutorando Bruno Henrique Alves (5)**

- (1) Doutoranda do Programa de Pós-Graduação em Ciência da Informação da UNESP-Marília-Brasil, cidinhapavanelli@gmail.com  
(2) Pesquisadora 2 do CNPq e Livre-docente do Programa de Pós-Graduação em Ciência da Informação da UNESP-Marília-Brasil, etannuri@gmail.com  
(3) Pesquisadora 1 do CNPq e Docente do curso de Pós Graduação em Ciência da Informação da Universidade Estadual Paulista (UNESP)-Marília e da Universidade Federal de Pernambuco, santiagobufrem@gmail.com  
(4) Doutorando do Programa de Pós-Graduação em Ciência da Informação da UNESP-Marília-Brasil, marcospraddo@yahoo.com.br  
(5) brhenriquealves@gmail.com

## Resumo

O objetivo geral desta pesquisa consiste em identificar os referentes hegemônicos a partir do "Sistema de Classificação da Literatura em Organização do Conhecimento" de Dahlberg (1993), utilizando-se como fonte de dados os trabalhos da ISKO-Ibérico, nas últimas cinco edições do evento. Especificamente se propõe: 1) agrupar as diferentes pesquisas apresentadas neste recorte temporal do ISKO-ibérico, segundo o Sistema de Classificação da autora; 2) identificar os principais pesquisadores que contribuíram para o desenvolvimento da Organização do Conhecimento (OC) em relação aos referentes teóricos basilares, segundo as classes da literatura em estudo; 3) apresentar os principais referentes teóricos utilizados segundo o "Sistema de Classificação da Literatura em Organização do Conhecimento" de Dahlberg. Como procedimento de pesquisa levantaram-se 320 trabalhos completos dos anais do ISKO-Ibérico, apresentados no período em estudo, sendo: 43 trabalhos de Barcelona, em 2005; 60 de León, em 2007; 87 de Valencia, em 2009; 38 de Ferrol, em 2011 e 92 do Porto, em 2013, por considerá-los evidências concretas que registram as relações entre os principais conceitos, autores e temáticas. Os 320 trabalhos foram agrupados com base no "Sistema de Classificação da Literatura em Organização do Conhecimento" de Dahlberg (1993) distribuídos em dez categorias, utilizando-se o título, resumo e palavras-chave. Procedeu-se a análise, destacando-se as categorias mais frequentes e os autores mais produtivos. Do total de 509 autores foram identificados 40 mais produtivos, com pelo menos 3 trabalhos cada

um. Estes foram associados às categorias do CSKOL, para posteriormente gerar a rede de relações entre as duas variáveis, a partir da matriz 40x9 (autores e categorias). Utilizou-se o *software Ucinet*, para gerar a rede de relações *two-mode*, entre as variáveis. Analisaram-se, para esses autores mais produtivos, as categorias mais contempladas, ou seja, aquelas que constituem os referentes teóricos basilares segundo Dahlberg. Concluiu-se, destacando-se como grande foco de pesquisa deste *corpus* aqueles referentes aos "Fundamentos teóricos e problemas gerais de Organização do Conhecimento" (Grupo 1 de Dahlberg).

**Palavras-chave:** Epistemologia em Organização do Conhecimento; Referentes teóricos na ISKO-Ibérico; Classificação de Dahlberg; Teoria dos conceitos.

## Abstract

The general objective of this research is to identify the hegemonic references from Dahlberg's "Classification System for Knowledge Organization Literature" (1993), using the papers of ISKO-Ibérico as source of data, in the last five editions of the event. Specifically, the study aims to: 1) group the different research presented in this period of time at ISKO-Ibérico, according to the classification system of the author; 2) identify the key researchers who contributed to the development of Knowledge Organization (KO) in relation to the fundamental theoretical reference, according to the literature in study; 3) present the main theoretical references used according to Dahlberg's "System of Literature Classification in Knowledge Organization". As research procedure, we

retrieved 320 complete papers from ISKO-Ibérico's proceedings, presented during the period in study, as follows: 43 papers from Barcelona in 2005; 60 from León in 2007; 87 from Valencia in 2009; 38 from Ferrol in 2011 and 92 from Porto, in 2013; considering them concrete evidence that record the relationships between the main concepts, authors and themes. The 320 papers were grouped based on Dahlberg's "Classification System for Knowledge Organization Literature" (1993), divided into ten categories, using the title, abstract and keywords. We proceeded to analysis, highlighting the most frequent categories and the most productive authors. From the total of 509 authors, we identified the 40 most productive, with at least three papers each. These authors were associ-

ated with CSKOL categories, to, subsequently, generate the network of relationships between two variables from 40x9 matrix (authors and categories). Ucinet software was used to generate the two-mode network relations between the variables. We analyzed the most contemplated categories by the most productive authors, i.e., the categories that constitute the fundamental theoretical references according to Dahlberg. It is concluded, highlighting as the main focus of this research corpus, the categories related to "Theoretical Foundations and general problems of Knowledge Organization" (Dahlberg's Group 1).

**Keywords:** Epistemology in Knowledge Organization; Theoretical references in ISKO-Ibérico; Dahlberg's classification; Concept Theory.

## 1. Contextualização da pesquisa

Na comunidade da Ciência da Informação (CI), o termo Organização do Conhecimento (OC) significa a melhor maneira de construir a organização da informação em registros bibliográficos, incluindo índices de citação, texto completo, entre outros. Considera-se que a maneira mais otimizada de utilizar estes recursos, se faz pela recuperação da informação (RI). A OC é, portanto, um conceito mais amplo que informação, pois é "conhecimento organizado" (HJØRLAND, 2003). Segundo o autor, este conceito amplo de OC, baseia-se em teorias amplas e abrangentes, que não podem ser ignoradas, pois os principais atores em CI são os produtores de conhecimento, entre eles os autores, os usuários e os intermediários. Cada um destes atores traz consigo certos pré-entendimentos, visões, conceitos e linguagens, adquiridas principalmente durante sua socialização. O sucesso da interação, com registros bibliográficos é o foco da CI e depende fundamentalmente destes pré-entendimentos, conceitos e linguagens.

Destaca-se também, a conceituação de Barité (2001), para quem o objeto de estudo da OC é o conhecimento socializado, como disciplina que cuida do desenvolvimento de técnicas para a construção, gestão, uso e avaliação de classificações científicas, taxonomias, nomenclaturas e linguagens documentais. Por outro lado, a OC volta-se às metodologias de uso e recuperação por linguagem natural. É a essa visão integral do conhecimento que se associam as classificações filosóficas ou científicas do saber como as classificações destinadas à organização de documentos em bibliotecas, arquivos e outras unidades de informação, ampliando possibilidades para um importante desenvolvimento disciplinar e interdisciplinar no âmbito da Biblioteconomia e da Documentação.

Considerando que a OC é um domínio em construção dinâmica, com foco na consolidação de

um estatuto epistemológico consistente, busca-se nesta pesquisa a definição das suas principais linhas de pensamento, de forma a estabelecer relações entre seus pesquisadores e as temáticas dos trabalhos apresentados nas edições internacionais dos eventos promovidos pela *International Society for Knowledge Organization* (ISKO). O domínio de OC tem sido construído buscando o aporte na concepção filosófica do conceito de Dahlberg (1993), fundante da Teoria dos Conceitos, que compartilha com a teoria dos sistemas de OC e tem como missão construir uma ponte entre a lexicografia especializada e o sistema de ordenação de conceitos.

Observe-se, ainda, que a *International Coding Classification* (DAHLBERG, 1995) e outros sistemas de classificação, tais como, o *Classification Research Group*, o *Broad System of Ordering* (BSO) e a *BLISS Bibliographic Classification*<sup>2</sup> (1972) foram influenciados pela Teoria dos níveis integrativos, pois sua filosofia mostra-se adequada a sistemas interdisciplinares, fundamentando pesquisas para o avanço de sistemas de classificação. Segundo a teoria dos níveis integrativos, o mundo se organiza em níveis de complexidade crescente: cada nível só pode existir a partir do nível anterior, o que não significa que as propriedades sejam herdadas nos níveis subsequentes. Embora discutida desde o século XIX, seu emprego como base para OC surge nos anos 60 do século passado.

Complete-se ainda que a ideia de OC foi originalmente apresentada pelo bibliotecário americano Henry Evelyn Bliss, responsável por dois livros marcantes para o surgimento do termo referente à OC, a saber: *The Organization of Knowledge and the System of the Sciences* e *The Organization of Knowledge in Libraries*, títulos publicados respectivamente nos anos de 1929 e 1933 (DAHLBERG, 2006, p. 11). A influência destes dois trabalhos foi determinante para que posteriormente a OC fosse promulga-

da como um domínio de estudo autônomo, para atuar nas práticas científicas. Assim, a OC ganhou força e interlocução na comunidade acadêmica internacional. Sua proeminência consolidou-se em 1974 com a criação da *Society for Classification*, uma instituição “[...] de natureza interdisciplinar, voltada para questões relacionadas com a Classificação (estudo e aplicação de suas abordagens)” (GUIMARÃES, 2012, p. 85). Destaca-se que esta entidade foi o embrião da ISKO.

O caráter interdisciplinar é a base que sustenta e orienta a OC desde seu surgimento. Entretanto, o desenvolvimento dos preceitos aplicados na organização e representação do conhecimento, conforme argumenta Pinho (2009), foi influenciado pela filosofia e seus sistemas ontológicos de divisão do conhecimento humano. É neste dinâmico processo de sistematizações, que a classificação representa a codificação da linguagem humana. Afinal, “um sistema de classificação do conteúdo temático de um documento é uma linguagem codificada” (CURRÁS, 2010, p. 28). Apesar da sua importância na OC, a classificação não é o objeto que norteia seu interesse primordial.

Se por um lado, Gomes (2009, p. 61) enfatiza a falta de clareza relativa às “[...] fronteiras da OC, o que é uma característica da atividade científica atual, ou seja, não há limites claros nas atuais atividades científicas da área”. Por outro lado, Dahlberg (2006) atribui entre os objetos da OC os elementos relacionados à Teoria do Conceito e aos princípios que envolvem as práticas de classificação e de indexação como recursos indispensáveis à representação do conhecimento. Logo, a classificação se beneficia direta e indiretamente do conjunto de atividades desenvolvidas pela OC, para ampliar seus aspectos teóricos e práticos, estando “presente na construção dos tesouros, que passam a ser um novo meio de manipulação da informação a partir dos anos 60” (CAMPOS; GOMES, 2003 p. 151). Como sistemas hierárquicos, os tesouros apresentam termos recorrentes em contextos específicos, sendo o termo um “[...] signo verbal que designa um referente, ou, num nível maior de abstração, que denota o conceito” (CAMPOS; GOMES, 2006, p. 349). O uso dos tesouros em sistemas especializados de informação garante a aplicação de ferramentas imprescindíveis para instrumentalizar os múltiplos potenciais da OC. Assim, os tesouros constituem aportes fundamentais ao processo de classificação e indexação.

Vale destacar que os procedimentos elaborados para efetuar a classificação consistem em identificar o conceito, para representá-lo em uma

estrutura sistematizada. Desse modo, Dahlberg (1978b) associa à sistematização do conceito o emprego descritivo de terminologias. Estes recursos são designados por termos ou palavras que representam o conjunto de símbolos, objetos contextualizados pelos elementos da linguística. A autora menciona que tal processo resulta dos mecanismos de representação do conhecimento humano e arquitetado na dimensão ontológica. E esta dinâmica adota recursos do vocabulário, da sintaxe e de outras lógicas comunicativas para cooperar na aplicação dos elementos utilizados na linguagem como um fator preponderante para o estabelecimento de convenções terminológicas. Dahlberg (1978a, p. 12) compreende que a “formação dos conceitos é a síntese dos predicados necessários e verdadeiros a respeito de determinado objeto”. O conceito constitui um elemento preponderante nas estruturas componentes da OC, pois conjuga toda uma forma de pensamento para que se atinja a elaboração do conhecimento (CURRÁS, 1995).

Concebidos como unidades de pensamento e conhecimento, os conceitos integram-se para compor relações semânticas, contribuindo, portanto, à consolidação de fundamentos para a interação com o mundo. As relações entre conceitos permitem a construção de identidades, de forma que nossas teorias, categorizações e paradigmas resultam na classificação de coisas e conhecimento sobre as coisas, de acordo com os mesmos princípios básicos (classes).

Um descritor ou uma classe representa um conceito, e a RI é essencialmente preocupada com as relações semânticas entre consultas, representações de documentos e textos. A teoria de conceitos, com seus significados e semânticas é, provavelmente, um dos domínios de pesquisa de mais difícil complexidade. Dai a importância de se entender, por exemplo, as diferentes visões entre as teorias tradicionais baseadas no positivismo lógico e nas visões alternativas, baseadas em teorias pragmáticas. A importância dos conceitos é corroborada na comunidade da OC, tanto do ponto de vista de seu valor como elemento cognitivo, quanto pela sua institucionalização social, bem ilustrada pelo periódico *Knowledge Organization*, cujo subtítulo expressa o fato de ser internacional, além de ser devotado “à Teoria do Conceito, classificação, indexação, e representação do conhecimento” (HJØRLAND, 2003).

A construção de sistemas classificatórios e de tesouros foi estruturada com base em Ranganaathan, um dos teóricos basilares da OC, confirmando assim que a teoria de conceitos é aceita de modo amplo e que desempenha um impor-

tante papel no processo de institucionalização cognitiva dos fundamentos da área em questão (DAHLBERG, 1993). Além de sua contribuição para a fundação da ISKO, Dahlberg (1993) cria na ISKO, com base na teoria dos conceitos, o instrumento *Classification Scheme for Knowledge Organization Literature* (CSKOL), que agrupa em dez categorias, os principais conceitos do domínio de OC.

É importante apontar que as contribuições de Dahlberg são notáveis e relevantes para o entendimento dos recursos relacionados com a OC. Esta pesquisadora foi quem “[...] sistematizou e formalizou a teoria pertinente à Organização do Conhecimento com base na Teoria Geral da Terminologia, de Eugene Wüster, e na Teoria da Classificação Facetada, de Ranganathan: a Teoria do Conceito” (MEDEIROS, 2010, p. 45-46).

Compreender o impacto e a dimensão das contribuições de Dahlberg à OC consiste no esforço de identificar os princípios que regem e fundamentam este domínio. Assim, os trabalhos de Dahlberg são imprescindíveis para vislumbrar os recursos instrumentais utilizados nos processos de organização, interpretação e estruturação do conhecimento humano. Para Dahlberg (1993), o Sistema de Classificação da literatura em OC assim se apresenta:

- GRUPO 0 - Este se refere à divisão de forma dos documentos, na literatura em OC. No entanto, este não é o objeto desta pesquisa, sendo assim, nenhuma pesquisa foi neste grupo classificada.
- GRUPO 1- Fundamentos teóricos e problemas gerais da OC, com os seguintes conteúdos: Ordenação e Organização do Conhecimento, conceptologia em OC, Matemática e OC, Teoria dos Sistemas e OC, Psicologia e OC, Ciência e OC, problemas em OC, Pesquisa da Classificação, História da OC;
- GRUPO 2 - Sistemas de Classificação e Tesouros, estrutura e construção. Apresentam-se os seguintes conteúdos: Estruturas e elementos de Sistemas de Classificação e Tesaurus (SC&T), Construção de SC, Relações entre conceitos, taxonomias, notações e códigos, manutenção, atualização e armazenagem em C&T, concordância entre linguagens de indexação; avaliação de sistemas de classificação e tesaurus;
- GRUPO 3 - Metodologia de Classificação e Indexação, subentendendo os seguintes conteúdos: Teoria da Classificação e indexação, Análise de assunto, técnicas de Classificação e indexação, Classificação e indexação automática, Ordenação Manual e automática, Geração de índices e programas, reclassificação, Avaliação de classificação e Indexação, Codificação e Decodificação;
- GRUPO 4 - Sistemas de Classificação e Tesaurus, com os seguintes conteúdos: Sistema de Classificação de Dewey, Classificação da Biblioteca do Congresso, assuntos de

leituras, classificação colon, outros sistemas de classificação e tesouros;

- GRUPO 5-Classificação orientada à objetos (Taxonomias), nas diferentes áreas do conhecimento;
- GRUPO 6- Classificações e Tesouros de assuntos específicos;
- GRUPO 7- Representação do conhecimento por linguagens e terminología, com os seguintes conteúdos: problemas gerais de linguagem natural em relação à OC, semântica, processamento automático da linguagem, sistemas de recuperação *on-line* e tecnologias, Problemas de terminología, de gramática, de léxicos/dicionários. Sistemas multilíngues e traduções, trabalho de terminología orientada ao assunto;
- GRUPO 8 - Indexação e Classificação Aplicadas, compreendendo os seguintes conteúdos: problemas gerais, catálogos, diretrizes, regras, Índices, indexação e classificação de dados, títulos, literatura primária, de livros, literatura secundária, materia especiais ou não livros e de assuntos específicos;
- GRUPO 9 - Ambiente da OC, com os seguintes conteúdos: organização profissional e espacial (nível nacional e internacional) bem como itens de educação e treinamento, legislação, Econômicos, uso de serviços, Normalização do trabalho de OC, estudo de usuarios etc.

Acrescente-se ainda que os trabalhos de Dahlberg são frequentemente objetos de diversos estudos que analisam as características e as qualidades das teorias por ela produzidas. A estrutura lógica de representação conceitual já citada (CSKOL), relativa às dimensões temáticas do conhecimento, tornou claras as relações entre o domínio de OC e os fundamentos filosóficos da Teoria dos Conceitos. O CSKOL é utilizado por esta pesquisa para identificar os referentes hegemônicos apresentados em trabalhos nas cinco edições da ISKO-Ibérico. Assim, o exercício de identificação das relações entre as categorias definidas no CSKOL e a produção científica constituinte do corpus desta pesquisa, justifica-se pela ampliação do alcance e da profundidade relativa aos aportes referenciais dos estudos, assim como aos procedimentos metodológicos neles presentes. Abrangendo estudos relativos a classificação, representação, recuperação da informação, construção de tesouros e, ainda, fundamentos para o uso e a aplicação de terminologias, a análise estabeleceu conexões teóricas entre as categorias e entre essas e o corpus.

Assim, além da importância de Dahlberg à OC, o emprego da classificação CSKOL como aporte categorial constitui um recurso que enseja a distribuição temática dos trabalhos apresentados nos eventos da ISKO, e a visualização das novas ideias construídas pela elite dos pesquisadores voltados aos interesses da OC.

## 2. Objetivos e Metodologia

O objetivo geral desta pesquisa consiste em identificar os referentes hegemônicos a partir do “Sistema de Classificação da Literatura em Organização do Conhecimento” de Dahlberg (1993), compreendendo-se que esses se sobressaíram e foram legitimados pelo grupo dos intelectuais, produtores do conhecimento nos trabalhos da ISKO-Ibérico, nas últimas cinco edições do evento. De forma mais específica propõe-se a: 1) agrupar as diferentes pesquisas apresentadas neste recorte temporal da ISKO-Ibérico, segundo as categorias do sistema de classificação da autora; 2) identificar os principais pesquisadores que contribuíram para o desenvolvimento da OC em relação aos referentes teóricos basilares, segundo as classes da literatura em estudo; 3) apresentar os principais referentes teóricos utilizados segundo o Sistema de Classificação da Literatura em Organização do Conhecimento” de Dahlberg.

Como procedimento de pesquisa, levantaram-se 320 trabalhos completos dos anais das edições do ISKO-Ibérico, sendo: 43 trabalhos de Barcelona, em 2005; 60 de León, em 2007; 87 de Valencia, em 2009; 38 de Ferrol, em 2011 e 92 do Porto, em 2013, por considerá-los evidências concretas que registram as relações entre os principais conceitos, autores e temáticas. Os 320 trabalhos foram agrupados com base no CSKOL, distribuídos em nove categorias, utilizando-se o título, resumo e palavras-chave. Procedeu-se a análise, destacando-se as categorias mais frequentes e os autores mais produtivos. Do total de 509 autores, foram identificados os 40 mais produtivos, com pelo menos 3 trabalhos cada um, desconsiderando-se as co-autorias, sendo que estas foram desdobradas. Estes foram associados às categorias do CSKOL, para posteriormente gerar a rede de relações entre as duas variáveis, a partir da matriz 40x9 (autores e categorias). Utilizou-se o *software Ucinet*, para gerar a rede de relações *two-mode*, entre as variáveis. Por fim, relacionaram-se esses autores mais produtivos, às categorias mais contempladas, ou seja, aquelas que constituem os referentes teóricos basilares, segundo Dahlberg.

## 3. Descrição e análise dos resultados

A partir dos autores mais produtivos, apresenta-se a Tabela I, constituída por aqueles que apresentaram pelo menos 3 trabalhos. Destaca-se o primeiro autor com um total de 10 trabalhos, observando-se que a distribuição tem uma amplitude de variação de 3 a 10 trabalhos. A presença de autores advindos de instituições

espanholas e de instituições brasileiras é bastante significativa .

| <i>Autores</i>                  | <i>Nº de trab. publicados</i> |
|---------------------------------|-------------------------------|
| José Augusto C. Guimarães       | 10                            |
| Agnes Hajdu Barat               | 7                             |
| María del Carmen Agustín Lacruz | 7                             |
| Mariângela Spotti Lopes Fujita  | 7                             |
| Elmira Luzia Melo Soares Simeão | 6                             |
| Leilah Santiago Bufrem          | 6                             |
| Marilda Lopes Ginez de Lara     | 6                             |
| Carlos Cândido de Almeida       | 5                             |
| Evelyn Goyannes Dill Orrico     | 5                             |
| José Antonio Moreiro González   | 5                             |
| Aida Varela Varela              | 4                             |
| Ana Lúcia Terra                 | 4                             |
| Juan Carlos Fernández-Molina    | 4                             |
| Luis Rodríguez Yunta            | 4                             |
| Rosa San Segundo                | 4                             |
| Thiago Henrique Bragato Barros  | 4                             |
| Armando Malheiro da Silva       | 3                             |
| Aurora Cuevas Cerveró           | 3                             |
| Blanca Rodríguez Bravo          | 3                             |
| Daniel Martínez Ávila           | 3                             |
| Dulce Amélia de Brito Neves     | 3                             |
| Fabio A. Pinho                  | 3                             |
| Fernanda Ribeiro                | 3                             |
| H. Peter Ohly                   | 3                             |
| Jesús Tramullas                 | 3                             |
| João Batista Ernesto de Moraes  | 3                             |
| José Antonio Salvador Oliván    | 3                             |
| Julia Osca-Lluch                | 3                             |
| Leticia Barrionuevo Almuzara    | 3                             |
| María Peñaranda-Ortega          | 3                             |
| Marilene Lobo Abreu Barbosa     | 3                             |
| Nair Yumiko Kobashi             | 3                             |
| Natália Bolfarini Tognoli       | 3                             |
| Piedad Garrido                  | 3                             |
| Ricardo Eito Brun               | 3                             |
| Rodrigo de Sales                | 3                             |
| Silvana Aparecida B. G. Vidotti | 3                             |
| Sônia Maria Breda               | 3                             |
| Suellen O. Milani               | 3                             |
| Tidra Viana Sorribas            | 3                             |

Tabela I. Autores que mais produtivos

Para melhor visualização do modo de distribuição dos trabalhos de cada um desses autores pelas categorias, apresenta-se a rede *two-mode* da Figura 1, na qual se relacionam os 40 pesquisadores mais produtivos, representados por círculos em vermelho, e as categorias de Dahl-

beg, representadas por quadriláteros azuis. A espessura dos segmentos que unem o pesquisador e a respectiva categoria corresponde à frequência com que ele produziu na temática.

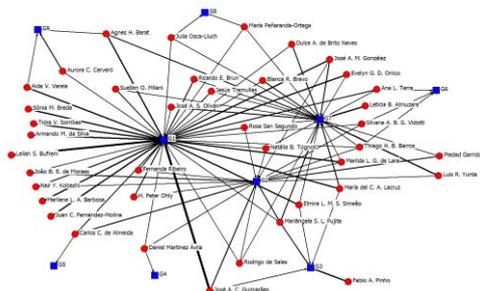


Figura 1. Rede two-mode entre autores e categorias

As categorias mais contempladas são: G1 que se refere aos “Fundamentos teóricos e problemas gerais de Organização do Conhecimento”, e que compreende os seguintes subtemas: Ordenação, Conceitos em Organização do Conhecimento; Teoria dos sistemas, Ciência em Organização do Conhecimento, História de Organização do Conhecimento. Nesta categoria foram encontradas 86 pesquisas, construídas por 36 do total de 40 pesquisadores, sendo esta a categoria de maior destaque. Observe-se que os únicos pesquisadores que não trabalharam em G1, do total dos 40, são: Luis R. Yunta, Fábio A. Pinho, Letícia B. Almizara e Piedad Garrido. Isso posto, compreende-se que a grande preocupação dos pesquisadores da ISKO tem sido as questões relativas aos fundamentos teóricos da Organização do Conhecimento. Os pesquisadores com foco maior no G1 são: Leilah Santiago Bufrem com todos os seus 6 trabalhos nesta categoria e José Augusto Chaves Guimarães, o autor com 10 trabalhos, o mais produtivo de todos presentes na tabela, distribuindo suas pesquisas entre G1 com 7 pesquisas, G2 com 2 pesquisas e G3 com 1 pesquisa.

Observe-se que a pesquisadora Leilah Santiago Bufrem, centra todas suas pesquisas em G1, enquanto o segundo pesquisador citado, apresenta maior diversidade temática, que percorre as categorias G1, G2 e G3, porém com preponderância em G1.

Destaca-se a categoria G7 denominada “Representação do Conhecimento por Linguagens e Terminologias,” compreendendo o seguinte subtema: problemas gerais em linguagem natural em relação à Organização do Conhecimento, semântica, processamento automático da linguagem, problemas de terminología, gramática, léxicos/dicionários. Foram encontrados 37 trabalhos com foco na categoria G7, considerando o recorte dos 40 pesquisadores mais produtivos,

sendo 24 os pesquisadores responsáveis por trabalhos nesta categoria. Destaca-se a pesquisadora Letícia B. Almizara com suas 3 únicas pesquisas centradas nesta categoria. Ainda, o pesquisador Luis R. Yunta, com 3 pesquisas nesta categoria e 1 pesquisa em G2, em um total de 4 pesquisas. Nesta categoria, com 2 trabalhos cada um, apresentam-se os seguintes pesquisadores: Agnes H. Barat, María Del C. A. Lacruz, Mariângela S. L. Fujita, José A. M. González, Blanca R. Bravo, Dulce A. de Brito Neves e José A. S. Oliván. Ainda, os demais 14 pesquisadores associados a esta categoria apresentam cada um deles, apenas 1 trabalho. Observa-se que, em geral, os pesquisadores com foco na categoria G2 têm também foco de pesquisa na categoria G7, indicando uma proximidade temática entre as duas categorias.

Ainda, identifica-se a categoria G2 denominada “Sistema de Classificação e Tesouros, Estrutura e Construção” com o seguinte subtema: Estruturas e Elementos de Sistemas De Classificação De Tesouros SC&T Construção de SC&T Relação entre Conceitos, Taxonomias, Compatibilidade e Concordância entre Linguagem de Indexação, Avaliação de Sistemas de Classificação e Tesouros. Nesta categoria foram encontradas 22 pesquisas, sendo 17 pesquisadores responsáveis por elas, com destaque para José Augusto Chaves Guimarães, Mariângela S. L. Fujita, Carlos C. de Almeida, João B. E. de Moraes e Piedad Garrido todos eles com 2 pesquisas cada um, sendo que os demais 12 pesquisadores apresentaram apenas 1 pesquisa nesta categoria. Ressalte-se assim que, todos os pesquisadores que centram suas pesquisas em G2 têm apenas 1 ou 2 trabalhos nesta categoria.

Na categoria G3 “Metodologia de Classificação e Indexação”, foram encontrados 8 trabalhos com os seguintes subtemas: Teoria da classificação e indexação; Análise de assunto, Técnicas de classificação e indexação, Classificação e indexação, geração de índices e programas, Avaliação de classificação e indexação, entre outras. São 5 os pesquisadores que trabalham no tema em questão, com destaque para Fábio A. Pinho, com 3 trabalhos, além de Mariângela S.L. Fujita, com 2 pesquisas, além de 3 pesquisadores com uma única pesquisa cada um deles.

A categoria G6, cujo tema central é “Classificações e Tesouros de assuntos específicos”, é representada por 3 pesquisadores, cada um deles com uma pesquisa, sendo que os 3 também pesquisam na temática G1, o que sugere que além da temática “Classificação e Tesouros de assuntos específicos” estão também preocupados em pesquisar seus fundamentos.

O G4, denominado “Sistemas de Classificação e Tesouros”, tem as seguintes subclasses: Classificação Decimal Dewey, Classificação da Biblioteca do Congresso e outros Sistemas Universais de Classificação e Tesouros. O pesquisador Daniel Martínez-Avila, é o único centrado nesta temática e ainda nas temáticas de G1 e G2, distribuindo seus 3 trabalhos, nas mencionadas categorias.

Em relação ao G5, cujo tema principal é a “Classificação orientada a Objetos e Taxonomias”, nas diferentes áreas do conhecimento, apresenta-se um único pesquisador, Carlos C. Almeida com 1 trabalho em G5, como também com 2 em G1 e outros 2 em G2, em um total de 5 pesquisas.

A categoria G8 trata de “Indexação e Classificação Aplicada”, com os seguintes subtemas: catálogos, índices e demais assuntos específicos. Apresentam-se dois pesquisadores diretamente ligados ao G8, cada qual com 1 trabalho. Estes apresentam a característica de pesquisarem, também, nas categorias G1 e G7. Observa-se que as categorias G7 e G8, apresentam proximidade temática na medida que trabalham com as questões de linguagem e representação do conhecimento.

A categoria G9, “Ambiente da Organização do Conhecimento” tem como principais subtemas: Estudos de usuário, Normalização no trabalho de Organização do Conhecimento entre outros. São 3 pesquisadores que têm suas pesquisas contempladas nestas temáticas, com destaque para a pesquisadora Aurora C. Cerveró, com 2 pesquisas e que também apresenta 1 pesquisa em G1. Os demais pesquisadores têm uma pesquisa cada um.

Assim, as principais categorias utilizadas segundo o CSKOL, são: G1, G2 e G7. Os demais grupos, apesar de estarem contemplados nos trabalhos da ISKO, têm presença menos significativa

O G1, que apresenta os Fundamentos teóricos e problemas da OC, compreende o foco e o core das pesquisas apresentadas na Isko Ibérico, neste período estudado, segundo o Sistema CSKOL, consignando aproximadamente 27% das pesquisas (em relação ao total de 320) e 90% dos pesquisadores (referente aos autores mais produtivos), atuando neste grupo. Em seguida aparece o G7, que pesquisa as questões relativas à representação do conhecimento por linguagens e terminologias, com 11% das pesquisas e 60 % pesquisadores, no grupo. Por fim, o G2 que trabalha questões relativas ao sistema de Classificação e Tesaurus, estrutura

e construção, com 7% das pesquisas nesta temática e 42% dos pesquisadores neste grupo.

Por outro lado, o G5, que estuda Classificação orientada a objetos, nas diferentes áreas do conhecimento, tendo como produto as Taxonomias, apresentou apenas uma pesquisa e um pesquisador.

Esses resultados sugerem que os pesquisadores da ISKO, estão com foco em questões mais amplas e de fundamentos, relativas às principais questões da OC, com menor interesse de pesquisa em questões mais pontuais, específicas ou restritas.

#### 4. Considerações Finais

Retomando-se a proposta deste estudo, de relacionar as diferentes pesquisas apresentadas no período de 2005 à 2011 da ISKO-Ibérico, ao Sistema de Classificação de Dahlberg, esta pesquisa procurou captar as características do corpus analisado, destacando seus modos de apresentação e procurando revelar as relações entre categorias temáticas e autores, valorizando criticamente suas realizações e relações. É possível afirmar que a categoria 1 (G1), cujo núcleo são os fundamentos teóricos e problemas gerais de OC, é a que congrega maior número de estudos e o núcleo dos pesquisadores mais produtivos do corpus analisado. Destaque-se que alguns dos pesquisadores mais produtivos concentram suas pesquisas em apenas uma das categorias CSKOL, indicando especificidade e aprofundamento de pesquisa. Entretanto, há pesquisadores que transitam entre quatro ou três grupos do CSKOL, sugerindo aparente dispersão temática, embora esses grupos distintos sejam interconectados e interrelacionados entre si.

Como conclusão, destaca-se que grande parte dos pesquisadores focam suas pesquisas para reflexões da área que se voltam sobre ela mesma, em consonância com o G1: Fundamentos teóricos e problemas gerais de OC, sugerindo que a ISKO Ibérico, é um evento que trata de uma área do conhecimento em constante dinâmica de revisão e construção dos seus fundamentos teóricos.

#### Referências

- Barité, M. (2001). Organización del conocimiento: un nuevo marco teóricoconceptual en bibliotecología y documentación. In: Carrara, K. (Org.). *Educação, universidade e pesquisa: III Simpósio em Filosofia e Ciências Marília*. São Paulo: Unesp-Marília-Publicações, FAPESP. 35-60.
- Bliss Bibliographic classification2. (1972). Recuperado 25-06-2015 de <http://www.sid.camp.ac.uk/bca/bcahome.htm>

- Campos, M.; Gomes, H. E. (2003). Organização de domínios de conhecimento e os princípios ranganathianos. *Perspectivas em Ciência da Informação*, 8 (2) 150-163.
- Campos, M.; Gomes, H. E. (2006). Metodologia de elaboração de tesouro conceitual: a categorização como princípio norteador. *Perspectivas em Ciência da Informação*, 11 (3), 348-359.
- Currás, E. (1995). *Tesauros linguagens terminológicas*. Brasília: IBICT. 286 p. ISBN 84-283-1825-5
- Currás, E. (2010). *Ontologias, taxonomia e tesauros em teoria de sistemas e sistemática*. Brasília: Thesaurus. 182 p. ISBN 9788570629685
- Dahlberg, I. (1978a). Fundamentos teórico-conceituais da classificação. *Ciência da Informação*, 6 (2) 9-21.
- Dahlberg, I. (1978b). Teoria do conceito. *Ciência da Informação*, 7 (2) 101-107.
- Dahlberg, I. (1993). Knowledge Organization: Its Scope and Possibilities. *Knowledge Organization*, 20 (4) 211-222.
- Dahlberg, I. (1995). Conceptual structures and systematization. *International Forum of Information and Documentation*, 20 (3) 9-24.
- Dahlberg, I. (2006). Knowledge Organization: a new science? *Knowledge Organization*, 33 (1), 11-19.
- Gomes, H. E. (2009). Tendências da pesquisa em organização do conhecimento. *Pesquisa Brasileira em Ciência da Informação*, 2 (1) 60-88.
- Guimarães, J. A. C. (2012). A dimensão teórica do tratamento temático da informação e suas interlocuções com o universo científico da International Society for Knowledge Organization (ISKO). In: Guimarães, J. A. C.; Dodebei, V. L. D. L de M. (Orgs). (2012). *Desafios e perspectivas científicas para a organização e representação do conhecimento na atualidade*. Marília: ISKO-Brasil: FUNDEPE, 77-99.
- Hjørland, B. (2003). Fundamentals of knowledge organization. *Knowledge Organization*, 30:2, 87-111.
- Medeiros, J. da S. (2010). A construção do conceito: aproximações complementares entre a análise de Michel Foucault e Ingetraut Dahlberg. *Revista ACB: Biblioteconomia em Santa Catarina*, 15 (2) 40-53.
- Pinho, F. A. (2009). Fundamentos da organização e representação do conhecimento. Recife: Ed. da UFPE. 156 p. ISBN 978-85-7315-695-9

Anexo 1

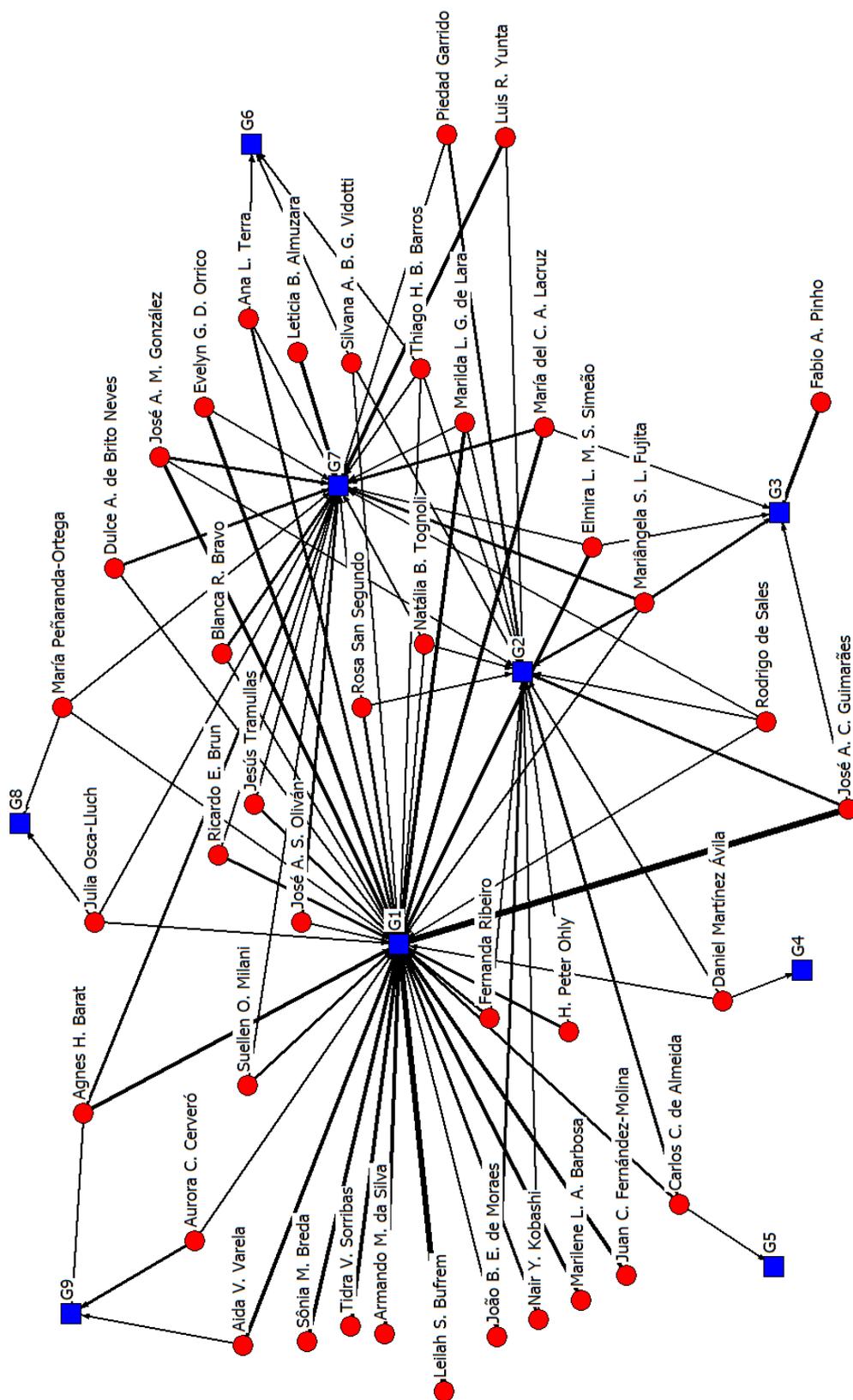


Figura 1. Rede two-mode entre autores e categorias

---

# Organização e representação da informação em arquivos: uma análise a partir da função classificação.

*Organization and representation of archival information: an analysis from a classification of records*

---

**Clarissa Schmidt (1), Johanna Smit (2)**

(1) Universidade Federal Fluminense, Rua Prof. Lara Viléla, 126 - Niterói, RJ, Brasil, CEP: 24210-590, clarissaschmidt@id.uff.br.

(2) Universidade de São Paulo, Avenida Prof. Lúcio Martins Rodrigues, 443 - São Paulo - SP, Brasil, CEP: 05508-020, johannawsmit@gmail.com

## Resumo

A função classificação estabelece-se como uma das abordagens teórico-metodológicas que subsidia a organização do conhecimento arquivístico. Nesse sentido, destaca-se o Princípio da Proveniência como elemento teórico e norteador do desenvolvimento de tal função, podendo ser considerado o primeiro enunciado a pensar a classificação dos documentos de arquivo menos por seus assuntos e temas, e mais por seu contexto de produção orgânico-funcional. A partir do momento que a função classificação institui a primazia do contexto em relação ao conteúdo do documento de arquivo, emerge outro marco importante e condicionante para a construção do campo dos arquivos enquanto área de conhecimento, isto é, uma autonomia em relação às teorias utilizadas em outras áreas, possibilitando diferenciar e organizar o documento de arquivo a partir de parâmetros fundamentais para conferir autonomia e identidade à Arquivística. Embora a função classificação seja consideravelmente analisada na literatura da área, constatam-se divergências quanto à definição da informação a ser organizada – de contexto x de conteúdo. Já abordagens quanto à organização e representação da informação, estas não estão diretamente inseridas em bibliografias desta natureza. Neste contexto, o objetivo deste trabalho é compreender a função classificação, sendo que o foco central de análise consiste em problematizar a informação a ser organizada nesse processo. Propõe-se, ainda, questionar se esta informação é a mesma a ser utilizada no tocante à representação da informação em arquivos. Em termos metodológicos, valemo-nos de estudo bibliográfico e análise epistemológica. Desta maneira, espera-se contribuir para as análises sobre a organização do conhecimento arquivístico.

**Palavras-chave:** Função classificação em arquivos. Organização da informação em arquivos. Representação da informação em arquivos. Documento de arquivo.

## Abstract

Record classification is established as one of the theoretical-methodological approaches which subsidies the organization of archival knowledge. In this way, Principle of Provenance as the theoretical element and guide of the development of such function that could be considered the first enunciation to think the classification of archival records, least for its subjects and themes but mainly for its context of functional-organic production. Since record classification establishes the primacy of context related to contents of the archival documents another important milestone emerges to construct the archival field as an area of knowledge, i.e. an autonomy compared to other area's theories, making it possible to differentiate and to organize the classification of records from reasonable parameters to promote autonomy and identity to Archival science. Although the classification of records has been considerably evaluated in its literature there are divergences related to definition of the information that should be organized – from context versus content. On the other hand, the organization and representation approaches are not directly included in this nature of bibliography. In this way, the objective of the present work is to understand the classification of records, mainly focusing in questioning the information which should be organized during this process. This work also aims to question if this information should be the same used to represent the archival information. In methodological means, we preformed bibliographical studies and epistemological analysis. Thus, contributing to the evaluations of the classification of records and the organization of archival knowledge.

**Keywords:** Classification of records. Organization of archival information. Representation of archival information. Archival documents.

## 1. Introdução

A Arquivologia começa a delinear sua configuração disciplinar a partir de desenvolvimentos técnicos de seu *Fazer* na segunda metade do século XVIII. O arquivista alemão Adolf Brenneke (1953) pensou a trajetória da área a partir dos meios e das maneiras pelas quais os arquivistas foram construindo, ao longo da história, para organizar os documentos nos arquivos. Para o autor, a teoria arquivística é resultado destas diferentes maneiras estabelecidas ao longo dos tempos acerca da organização dos documentos. Com uma perspectiva que tem como referência o desenvolvimento da teoria arquivística, elabora periodização, constituída por três fases, para analisar a trajetória da Arquivologia enquanto área de conhecimento: - Séculos XVI até XVII – Predomínio do método “prático indutivo” para a classificação arquivística; caráter técnico e experimental da área, representada pelos arquivos antigos e medievais (dualismo entre arquivo de expedição e recepção); - Século XVIII – Predomínio do sistema de classificação “teórico dedutivo” devido ao contexto racional e iluminista; tem-se o sistema de classificação já em uma perspectiva de teoria; - Século XIX – Surgimento do Princípio da Proveniência como novo sistema de Classificação arquivística, significando a revolução da teoria pelo estabelecimento de princípios teóricos para a área.

Essa periodização proposta por Brenneke está claramente baseada na classificação como elemento teórico e norteador do desenvolvimento da área, nos remetendo a considerar sua concepção de que a reflexão no campo dos arquivos avança a partir do momento em que encontra um método próprio para organizar e classificar seu objeto de trabalho, que é o documento que está no arquivo, construindo assim, como resultado, elaborações teóricas. Nesta perspectiva, podemos considerar, para classificação de documentos de arquivo, que atribui-se ao historiador Natalis de Wailly, chefe da seção administrativa dos arquivos do Ministério do Interior da França, em 1841, a formulação francesa do Princípio da Proveniência, isto é, o "Princípio do Respeito aos Fundos", o que segundo Michel Duchein (1986) “é a certidão de nascimento da noção de Fundos de arquivos”. Ainda que, de acordo com Martín-Pozuelo Campillos (1996), a construção do Princípio da Proveniência não seja francesa e sim possivelmente alemã, trata-se do primeiro enunciado teórico, marco teórico da Arquivologia, que passa a pensar a classificação dos documentos de arquivo menos por assuntos e temas, e mais por contexto de produção.

Essa formulação francesa do Princípio da Proveniência, “O Respeito aos Fundos”, surge como resposta a problemas práticos, do *Fazer*, se tornando um dos primeiros e talvez um dos principais princípios teóricos da área, um *Saber*, que quase cinquenta anos depois foi cancelado pelo Manual dos Holandeses. De acordo com Bellotto (2005, p. 16), os dois elementos mais importantes para sustentar a teoria da Arquivologia são o Princípio da Proveniência e o Princípio da Organicidade, pois enquanto base da sua teoria configuram a diferença entre esta e outras áreas como a Biblioteconomia e a Documentação, por exemplo. E essa diferença acontece, principalmente, a partir do momento em que a classificação dos documentos de arquivo passa a ser entendida *menos pelos seus assuntos e temáticas para realizá-las a ordem na qual os documentos haviam sido produzidos* (DUCHEIN, 1986, p. 15). *Afinal, a classificação, até as primeiras décadas do século XIX, era elaborada sem levar em conta a origem administrativa dos documentos* (SOUZA, 2007, p. 3). Neste sentido, é fundamental considerarmos que durante esse período a Arquivologia se delinea enquanto área de conhecimento a partir do estabelecimento do Princípio da Proveniência, que surge como resposta a problemas práticos junto aos arquivos, do *Fazer*, tendo como necessidade classificar documentos. Ou seja, este princípio, ao ser erigido como critério norteador da classificação de documentos de arquivo, altera a situação anterior de classificação por assuntos e temas.

Desde então, o campo dos arquivos se delinea enquanto área de conhecimento fundamentalmente a partir da organização de documentos pelo contexto de produção frente ao seu conteúdo. Nessa perspectiva, fica a pergunta: qual(is) informação(ões) organizamos, classificamos e representamos em arquivos?

O objetivo deste trabalho é compreender a função classificação, sendo que o foco central de análise consiste em problematizar a informação a ser organizada nesse processo. Propõe-se, ainda, questionar se esta informação é a mesma a ser utilizada no tocante à representação da informação em arquivos. Em termos metodológicos, valemo-nos de estudo bibliográfico e análise epistemológica. Desta maneira, espera-se contribuir para as análises sobre a organização do conhecimento arquivístico.

## 2. A Função Classificação

Por classificação de documentos de arquivo entende-se, segundo Camargo e Bellotto (1996, p. 30), *a sequencia de operações que, de acor-*

do com as diferentes estruturas, funções e atividades da entidade produtora, visam a distribuir os documentos de um arquivo. É através desta prática que se torna possível dar visibilidade às funções e às atividades do órgão produtor, deixando claras as ligações entre estas e os documentos. Trata-se de uma operação intelectual, lógica. Para Sousa (2014), a classificação possibilita o estabelecimento das relações entre os documentos individuais, que reunidos proporcionam uma representação contínua da atividade; a garantia de que os documentos mantenham sua denominação ao longo do tempo; o estabelecimento de prazos e medidas de conservação e a disponibilidade dos documentos. Vale ressaltar que, neste trabalho, estamos assumindo os termos classificação e arranjo como equivalentes.

Assim, a classificação deve deixar claros os vínculos arquivísticos, os contextos documentais, os procedimentos e a proveniência dos documentos que os caracterizam e identificam de maneira única. Através do ato da classificação, a rede de relações inerentes à natureza de qualquer documento não só dá à luz, como também fica estabelecida e perpetuada. Deste modo, o significado de cada documento em relação com os demais, assim como a estrutura do todo (fundo), podem ser compreendidos e perpetuados ao longo dos tempos (FOSCARINI, 2010).

Desta maneira, podemos afirmar que a função classificação em arquivos é fundamental no tocante ao acesso à informação contida no documento de arquivo, cabendo ao arquivista representar o contexto de produção/acúmulo deste documento, das razões de sua criação, bem como seu trâmite. Isso é importante para manter a qualidade primordial do documento de arquivo, sua função primeira, que é ser prova da atividade que lhe deu origem, sendo que valemo-nos das palavras de Bellotto (2010, p. 161) para quem este documento consegue ser prova *justamente por causa da simbiose indivisível entre o produtor – contexto – gênese – função*. Cabe aos arquivistas representar essa simbiose de maneira confiável, isto é, essa função primeira só pode ser manifestada se os elementos que fundamentam a classificação forem assegurados.

Reconhecidas a importância da classificação e seu papel primordial na construção do campo dos arquivos enquanto área de conhecimento, é inquietante observar que na Arquivologia pouco se dedica a estudá-la. Observa-se ainda, que dentre as funções arquivísticas, pouco se pratica a classificação por contexto de produção funcional dos documentos. E esse cenário, in-

discutivelmente, afeta características e qualidades importantes dos arquivos. Nessa linha de pensamento, Foscarini (2010) afirma que há pouca elaboração teórica sobre a teoria da classificação funcional e uma fraca compreensão sobre os princípios arquivísticos envolvidos nesta tarefa. Coloca ainda que isso se torna ainda mais complicado, pois a Arquivologia não possui clareza terminológica em relação aos termos “função”, “atividades”, “processos de negócios”, dentre outros. Ademais, diz que a classificação é fundamental para a formação do arquivista no contemporâneo, uma vez que a fisicalidade do documento não existe mais. Ela funciona como uma ferramenta bem mais importante que no passado. E para tanto, é fundamental conhecer funções, atividades e transações do criador dos documentos para operar a classificação.

Ainda que estas afirmações sobre o ensino e a prática da classificação não sejam objetos de estudo deste trabalho, é importante colocar tais observações visto que estas relacionam-se com questionamentos sobre qual (is) informação (ões) organizar e representar em arquivos.

Com base no exposto, pode-se dizer que a classificação em arquivos é uma atividade matricial (Souza, 2014), capaz de apresentar o caráter orgânico, os vínculos entre os documentos e quem os produziu, conferindo a estes seu valor primário, isto é, os motivos pelos quais foram criados, possibilitando assim preservarmos seu valor probatório.

### 3. Tipo de informação a ser classificada

Ao considerarmos que os documentos de arquivo devem ser classificados a partir da perspectiva orgânica-funcional, e que o estatuto probatório deste documento não se dá pela natureza do assunto, mas sim pela natureza contextual, afirmamos que na teoria arquivística as informações de contexto são as mais significativas. Nesse sentido, é fundamental respondermos as seguintes indagações: Quais informações devem fundamentar a classificação? O que são informações de contexto e o que são informações de conteúdo?

Ao problematizar sobre recuperação, acesso e uso de documentos de arquivo, Smit (2015, p. 20) discute sobre os diferentes tipos de informação que coexistem nos arquivos, a saber: “1. a informação contida nos documentos; 2. a informação produzida pelo arquivo através da descrição e da classificação e que cumpre dupla função: contextualizar os documentos em termos de proveniência e atividade na qual se enquadram e, ao mesmo tempo, distinguir os

documentos ou conjuntos de documentos dos demais através da descrição e classificação. O trabalho de organização da informação é executado através de duas operações básicas – a agregação em conjuntos e a consequente diferenciação entre conjuntos; 3. a informação produzida pelo arquivo para documentar as transações efetuadas sobre os documentos de modo a preservar sua autenticidade e valor probatório; 4. e, potencialmente ainda, uma informação, igualmente produzida pelo arquivo, sobre a história do órgão acumulador”.

Estes diferentes tipos de informação ocorrem tanto a partir do documento de arquivo, como da gestão, dos processos e das atividades realizadas pelos arquivistas no exercício da profissão. Como objetiva-se refletir sobre o(s) tipo(s) de informação (ções) que fundamentam a função classificação, é necessário compreender que nossa análise se dará a partir das informações inerentes aos documentos de arquivo – o objeto a ser classificado, isto é, informações de contexto (de fora) e de conteúdo (de dentro), o que Smit apresenta nos itens 1 e 2.

As informações de conteúdo, que estão “dentro” do documento de arquivo, são aquelas referentes ao assunto e tema. Já as informações de contexto, ou seja, aquelas que estão “fora” do documento de arquivo, são aquelas derivadas da aplicação do princípio da proveniência e da ordem original, originadas pelas funções e atividades que dão origem ao documento. Para conhecê-las, precisamos saber quem produziu o documento, como, por que, para que e quando.

Nesta linha de pensamento, vislumbramos duas possibilidades informacionais ao documento de arquivo; as de “fora” e as de “dentro”. Todavia, é fundamental ressaltar que além de não existir oposição entre ambas, mas sim complementaridade, devemos sempre priorizar a identificação contextual. Na perspectiva arquivística, a classificação objetiva atribuir sentido aos documentos, ou seja, inseri-los em seu contexto de produção, pois somente contextualizados cumprem sua função.

Sendo assim, entendemos a classificação como uma operação intelectual, que organiza fundamentalmente as informações que estão “fora” do documento de arquivo, e que somente após a realização desta atividade é que se torna possível pensarmos na organização das informações de “dentro”. (SCHMIDT; SMIT, 2013, p. 585). Isso não significa desconsiderar as informações de conteúdo, mas sim diferenciar o papel que ambas as possibilidades informacionais possuem na construção do conhecimento arquivístico.

#### 4. Considerações finais

Se o documento de arquivo nasce para registrar a ação, a função classificação é realizada para representar essa relação, revelando assim o vínculo arquivístico. E como vimos, o documento de arquivo possui duas possibilidades informacionais, isto é, as informações de contexto, de “fora”, e as de conteúdo, de “dentro”. Nesse sentido, no campo dos arquivos a função classificação preocupa-se em representar as informações de “fora” de modo a garantir os vínculos, o caráter orgânico e consequentemente o valor primário e probatório do documento de arquivo. Trata-se de elemento caro à área, pois é o que atribui identidade e autonomia ao conhecimento arquivístico.

Sendo assim, é importante delimitarmos que, na perspectiva da Arquivologia, ao pensarmos na classificação enquanto função arquivística, a informação que deve ser organizada e representada, que efetivamente fundamenta esta classificação, é aquela que nomeamos como informações de “fora”, as de contexto. Essa informação contextual é essencial – pedra de toque – para classificar/organizar os documentos e para recuperá-los. Contudo, não podemos deixar de considerar que é possível complementar esta representação pela informação de “dentro”, a de conteúdo.

Assim, o acesso à informação nos arquivos demanda a representação dos dois tipos de informação (contexto e conteúdo) e o entrelaçamento das duas possibilidades informacionais favorece que arquivos preencham seu papel social, ou seja, a organização do acesso a um determinado tipo de informação.

#### Referências

- BELLOTTO, Heloisa Liberalli (2005). *Arquivística, arquivos e documentos*. São Paulo: Associação de Arquivistas de São Paulo.
- BELLOTTO, Heloisa Liberalli (2010). Da gênese à função: o documento de arquivo como informação e testemunho. In: *Seminário de Estudos da Informação, Documento: gênese e contextos de uso*. Niterói: EdUFF.
- BELLOTTO, Heloisa Liberalli; CAMARGO, Ana Maria de Almeida (1996). *Dicionário de terminologia arquivística*. São Paulo: Secretaria de Estado da Cultura do Estado de São Paulo.
- BRENNEKE, Adolf (1953). *Archivkunde. Ein Beitrag zur Theorie und Geschichte des europäischen Archivwesens*. Leipzig.
- DUCHEIN, Michel. (1986). O respeito aos fundos em arquivística: princípios teóricos e problemas práticos. *Arquivo & Administração*, Rio de Janeiro, v. 10-14, n. 1, pp. 14-36.
- FOSCARINI, Fiorella (2010). La clasificación de documentos basada em funciones: comparación de La teoría y La práctica. *Tabula 13*. Salamanca: Asociación de Archiveros de Castilla y León.

- MARTIN-POZUELO CAMPILLOS, M. P (1996). *La construcción teórica en archivística: el principio de procedencia*. Madrid: Universidad Carlos III y BOE.
- SCHMIDT, Clarissa. M. S. (2012). Arquivologia e a construção do seu objeto científico: concepções, trajetórias, contextualizações. Tese de Doutorado, Escola de Comunicações e Artes, Universidade de São Paulo, São Paulo. Recuperado em 2015-06-29, de <http://www.teses.usp.br/teses/disponiveis/27/27151/tde-02072013-170328/>
- SCHMIDT, Clarissa. M. S.; SMIT, JOHANNA. Organização da Informação e Arquivos: diferentes perspectivas informacionais em torno do documento de arquivo. I Congresso ISKO Espanha e Portugal / XI Congreso ISKO España. *Informação e/ou Conhecimento: as duas faces de Jano*. Recuperado de 2015-06-28, de <http://www.youblisher.com/p/749221-I-Congresso-ISKO-Espanha-e-Portugal-XI-Congreso-ISKO-Espana/> p. 579-588.
- SMIT, Johanna Wilhelmina. Recuperação, acesso e uso dos documentos arquivísticos. *Ciência da Informação*, [S.l.], v. 42, n. 1, jan. 2015. ISSN 1518-8353. Recuperado 7-07-2015 <http://revista.ibict.br/cienciainformacao/index.php/ciinf/article/view/2256>.
- SOUSA, Renato Tarciso Barbosa de (2007). *O arquivo e a gestão da informação*. In: Universidade de Brasília. Faculdade de Ciência da Informação UNB-FCI: Brasília. Recuperado 1-06-2012 de <http://www.cid.unb.br/publico/setores/000/84/materiais/2007/1/529/M%C3%B3dulo%201%20O%20arquivo%20e%20a%20gest%C3%A3o%20da%20informa%C3%A7%C3%A3o.doc> .
- SOUSA, Renato Tarciso Barbosa de (2014). Alguns apontamentos sobre a Classificação de documentos de arquivo. *The Brazilian Journal of Information Science: research trends(BRAJIS)*. Marília:UNESP (Vol.8, N.1-2). Recuperado 2-05-2015 de <http://www2.marilia.unesp.br/revistas/index.php/bjis/article/view/4246>

---

# O “nome próprio” no contexto da Epistemologia da Organização do Conhecimento: um debate filosófico-teórico

*“Own name” in Knowledge Organization Epistemology: a philosophical-theoretical debate*

---

**Naira Christofolletti Silveira (1), Gustavo Silva Saldanha (2)**

(1) Professora Adjunta na Universidade Federal do Estado do Rio de Janeiro (UNIRIO), Av. Pasteur, 458, CCH sala 404, Urca, Rio de Janeiro, Brasil, naira.silveira@unirio.br

(2) Pesquisador Adjunto do Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT), Rua Lauro Muller, 455, sala 401, Botafogo, Rio de Janeiro, Brasil; Professor Adjunto na Universidade Federal do Estado do Rio de Janeiro (UNIRIO), saldanhaquim@gmail.com

## Resumo

A ‘autoria’ permeia diversas áreas do conhecimento e pode ser abordada a partir de diferentes olhares. Neste trabalho ela está relacionada diretamente à Filosofia e à Organização do Conhecimento. Como a ‘autoria’ é um elemento corrente no registro bibliográfico, estudos integrando questões filosóficas, conceituais e culturais à Representação Documentária permitiriam uma reflexão mais crítica sobre o desdobramento do uso da autoria nas práticas informacionais. O objetivo geral do trabalho é compreender a perspectiva da autoria a partir da noção de ‘nome próprio’, tendo como inflexão a discussão filosófica que se debruça sobre essa análise e questões sociais inerentes ao conceito. Trata-se de uma investigação exploratória e teórica, pautada nas pesquisas bibliográfica e documental, composta de dois movimentos: o primeiro, visa compreender a crítica filosófica à noção de ‘nome próprio’; o segundo, procura discutir a noção no contexto de sua formulação na filosofia da Organização e Representação do Conhecimento. A partir de uma crítica panorâmica sobre a discussão filosófica em torno do ‘nome próprio’, pontuamos a emergência do debate no contexto contemporâneo, como o caso dos filósofos do século XX, como Wittgenstein, Barthes e Foucault. Percebe-se que, apesar das mudanças advindas da Modernidade, o ‘autor’ no registro bibliográfico pressupõe ainda hoje o sentido de ‘nome próprio’, pois este personaliza e organiza o conhecimento através da construção de pontos de acessos padronizados.

**Palavras-chave:** Nome próprio. Representação do conhecimento. Organização do conhecimento. Autoria. Representação documentária.

## 1. Considerações iniciais

Para quem medita sobre o inefável, é útil observar que a linguagem pode perfeitamente nomear aquilo de que não pode falar. (Agamben, 2013, p. 103)

## Abstract

The ‘authorship’ is investigated in many knowledge areas and can be approached from different perspectives. In this work it is directly related to philosophy and knowledge organization. As the ‘authorship’ is a recurring element in the bibliographic records, studies integrating philosophical, conceptual and cultural questions in documentary representation allow a more critical reflection on the deployment of the use of authoring in informational practices. The overall goal is to understand the perspective of authorship from the notion of ‘own name’, with the inflection philosophical discussion that focuses on this analysis. This is an exploratory and theoretical research, based on bibliographical and documental research, consisting of two steps: the first, aims to understand the philosophical critique of the notion of ‘first name’; the second, discusses the notion in the context of its formulation in the philosophy of the knowledge organization. From a panoramic critique of the philosophical debate around the ‘own name’, we pointed out the emergence of debate in the contemporary context, as the case of twentieth-century philosophers such as Wittgenstein, Barthes and Foucault. We recognize that, despite the changes arising from Modernity, the ‘author’ in the bibliographic records today still assumes the meaning of ‘own name’, because it personalizes and organizes the knowledge by building standardized access points.

**Keywords:** Own name. Knowledge representation. Knowledge organization. Authorship. Documentary representation.

Conta-se que Beda, pai da historiografia britânica e grande erudito do renascimento da geração de Carlos Magno, recebeu o epíteto de “Venerável” a partir de dois concílios realizados em Aquisgrano (ou Aachen ou Aix-la-Chapelle, capital do Império Carolíngio), nos anos de 816 e

836. Os eventos ocorreram, dentre outras pautas, exatamente para a inclusão dos epítetos em algumas autoridades da Igreja Católica. Posteriormente, o costume se tornaria uma regra. Os epítetos passaram a fazer parte da “política” de nomes próprios da Igreja. (LE GOFF, 2013).

No entanto, conta-se também que, na verdade, foram os anjos que teriam construído o epíteto para completar o epitáfio de Beda. Para a história, existe a dúvida sobre o processo alegórico e o processo normatizador de inclusão do “título” de “venerável” na construção dos nomes das autoridades católicas.

O fato que nos interessa é que “Beda”, ou “Beda, O Venerável”, ou “Venerável Beda” são três formas de invocar a “identidade” de um dos grandes sábios do Medievo, responsável por inúmeras obras no campo da Retórica, da Filologia, da Literatura e da História. Especificamente, estamos diante do argumento central da presente reflexão: um debate sobre a fronteira entre filosofia do “nome próprio”, epistemologia da organização do conhecimento (centro teórico de nossa proposta) e as abordagens aplicadas deste último domínio.

A ‘autoria’ permeia diversas áreas do conhecimento e pode ser abordada a partir de diferentes olhares. Neste trabalho ela está relacionada diretamente à Filosofia e à Organização do Conhecimento. Como a ‘autoria’ é um elemento recorrente no registro bibliográfico, estudos integrando questões filosóficas, conceituais e culturais à Representação Documentária, permitiriam uma reflexão mais crítica sobre o desdobramento do uso da autoria em processos e práticas informacionais.

Para Otlet (1934), o “autor” (e o nome próprio dele resultante) representa o inventor ou criador de uma obra de imaginação ou documental. O pensador belga relaciona em § 231.2 as diferentes modalidades de inserção do nome próprio nos documentos tradicionais, como impressos, manuscritos. Dentre os problemas vinculados às práticas de representação e de acesso ao conhecimento registrado lançadas pela visão otletiana, estão a questão da latinização dos nomes próprios no Medievo e a presença de obras sem autoria.

O cenário que discute a ‘autoria’ no acesso informacional constitui-se o mais amplo desta pesquisa, que busca evidenciar o caráter reflexivo e social da Organização do Conhecimento, especialmente da Representação Descritiva. Reflexões teóricas acerca de questões socioculturais e a representação documental identificam uma tendência e buscam ir além da visão tecnicista e atribuir uma função filosófica e social à

Representação Documentária, capaz de mobilizar uma sociedade que se quer predicada com a noção de ‘conhecimento’, provendo acesso ao documento, respeitando as particularidades de seus usuários e, conseqüentemente, contribuindo para a representação, circulação, uso e produção de saberes.

O objetivo geral é discutir a perspectiva da autoria a partir da noção de ‘nome próprio’, tendo como inflexão a discussão filosófica que se debruça sobre essa análise. Trata-se de uma investigação exploratória e teórica, pautada nas pesquisas bibliográfica e documental, composta de dois movimentos metodológicos de construção da análise: o primeiro, visa compreender a crítica filosófica à noção de ‘nome próprio’; a segunda, procura discutir a noção no contexto de sua formulação na filosofia da organização dos saberes. Para esta reflexão, o primeiro ponto de análise é a discussão sobre ‘conhecimento’ e ‘linguagem’ no diálogo Crátilo de Platão.

A partir de uma crítica panorâmica sobre o debate filosófico em torno do “nome próprio”, pontuamos a emergência do debate no contexto contemporâneo, como o caso dos filósofos do século XX, especialmente, Barthes, Wittgenstein e Foucault. Para o último, a função autor está diretamente relacionada ao contexto social, histórico e econômico de uma sociedade.

No século XVII ou no século XVIII, a função autor começa a ser apagada dos discursos científicos e outro conjunto sistemático de técnicas, métodos, objetos de experimentação, laboratório e data é que atribui validade ao discurso e não mais o indivíduo que o produziu. Entretanto, esse conjunto de elementos não anulou o autor do discurso. O ‘autor’ no registro bibliográfico pressupõe o sentido de ‘nome próprio’, pois personaliza e organiza o conhecimento através da construção de pontos de acesso padronizados.

## 2. O dilema do nome próprio na

### Filosofia: da representação da realidade à representação do conhecimento

Ou o nome será, por assim dizer,  
um conceito lógico?  
(Wittgenstein, 1998, p. 79)

A relação entre nome próprio e representação da informação guarda muito mais elementos que a suposta mera relação de incorporação de um nome em um sistema bibliográfico pode, à primeira vista, vislumbrar. Das preocupações platônicas com a linguagem às inquietações wittgensteinianas e barthesianas do século XX sobre certeza, jogos de linguagem, “morte” do

autor, o nome próprio se estabelece como uma fonte de argumentação filosófica e termina por transversalizar as questões inerentes à Organização do Conhecimento.

Como nos demonstra Marcondes (1985), está no *Crátilo* de Platão a pioneira preocupação com a eliminação das imperfeições das linguagens naturais – na visão de Shannon e Weaver (1975), dois milênios depois, a construção do canal ideal para evitar os ruídos que atuam sobre a mensagem, ou simplesmente, o privilégio da sintaxe sobre a semântica e, principalmente, sobre a pragmática.

Para a filosofia da linguagem, central para o desenvolvimento da Organização do Conhecimento para diferentes autores, como John Budd e Rafael Capurro, a questão do nome próprio iniciada em Platão, toma contornos aporéticos: trata-se de identificar, nesta discussão particular, elementos gerais da própria filosofia em sua totalidade, como questões da tipologia da verdade, do real e do bem.

No *Crátilo* de Platão, identificamos um dos pilares da filosofia da linguagem na Antiguidade. Etimologia e forma têm aqui papel importante. Trata-se de um diálogo sobre as origens da linguagem. Hermógenes e Crátilo dialogam e dialetizam com Sócrates. Para Hermógenes, seria a relação entre nome e coisa uma convenção, enquanto para Crátilo, isto se dá na conformidade com a natureza das coisas. Para Sócrates, é preciso, primeiro, aceitar que existem coisas falsas e verdadeiras. Isto leva à discussão sobre o “verdadeiro” que há na relação entre nome e coisa – ou designação da natureza da coisa e natureza da coisa propriamente dita (Platão, 1963).

Está aqui uma das primeiras classificações da linguagem. Sócrates estabelece, junto de Hermógenes, que o “nome é a mais pequena” parte do discurso (Platão, 1963, 5, p. 13). Uma das conclusões a que chega Sócrates na primeira parte do diálogo, junto a Hermógenes, é que o nome não pode se dar por convenção, pois permitiria a falsidade. “A justeza de um nome, afirmamos nós [Sócrates para Crátilo], é aquilo que mostra uma coisa tal qual é”. (Platão, § 10, 1963, p. 120).

A afirmação do princípio mimético na estrutura da linguagem e sua negação moral é também aqui realizada pelo filósofo da Academia.

Sócrates: Poderás tu afirmar que o nome é uma coisa e aquilo que ele designa é outra?

Crátilo: Sim.

Sócrates: Convirás também que o nome é uma imitação de uma coisa?

Crátilo: Em absoluto. (Platão, § 5-10, 1963, p. 133)

No bojo da tentativa de vinculação de uma relação moral no uso e na determinação dos usos da linguagem, aparece em Platão a diferenciação das partes do discurso, a saber nomes, verbos, frases e, em outro momento, sílabas e letras (Platão, § 20, 1963, p. 136):

Sócrates: portanto, se existe aqui uma distribuição desta natureza, nós queremos chamar a uma *dizer a verdade* e à outra *falar falso*. Nestas circunstâncias e na possibilidade de se poder distribuir os nomes inexactamente e de não atribuir a cada coisa os que lhe pertencem, mas atribuir-lhes, algumas vezes, nomes que não lhes convém, seria também possível, neste caso, fazer o mesmo quanto aos verbos. Ora, se é possível proceder deste modo quanto aos nomes e aos verbos, também o é, necessariamente, quanto às frases; porquanto estas, como eu julgo, não são mais do que uma reunião daqueles elementos”. (Platão, 1963, 15-25, p. 135, grifo nosso)

Para Sócrates, porém, não seria possível concordar com a relação de identidade absoluta entre nomes e coisas, pois isto seria tomar o nome como a própria coisa. Está na passagem abaixo um dos aspectos claros da crítica platônica ao mundo que se estabelece pela linguagem – este, no olhar do filósofo da Academia, um mundo confuso, que não poderia discenir entre coisa e representação da coisa.

Seria, com certeza, coisa para rir, ó Crátilo, o efeito dos nomes sobre os objectos, de que são nomes, no caso de concordarem em absoluto com eles. Tudo seria duplo e não se poderia dizer qual é o objecto e qual o nome. (Platão, § 5-10, 1963, p. 139, grifo nosso)

Surge ainda outro aspecto importante do diálogo para a aproximação entre as artes filológica e bibliológica: a proposta, em *Crátilo*, da possibilidade de construção de um método para “conhecer o mundo” a partir da linguagem. Para Crátilo, que toma os nomes primitivos como idênticos às coisas, nós poderíamos, de maneira inversa à visão de Sócrates, conhecer as coisas pelos nomes.

Sócrates refuta a visão cratílina. Há, aqui, uma crítica da cognoscibilidade das coisas pelos nomes – ou seja, só é possível conhecer a coisa pela própria coisa.

Sócrates: Visto estarem os nomes desavindos e afirmarem uns a sua semelhança com a verdade e reclamarem outros para si o mesmo direito, sobre que base poderemos, de hoje em diante, apoiar-nos, para resolvermos a questão, ou a que recorrer? (Platão, § 5, 1963, p. 154, grifo nosso)

(...)

Baste-nos, por isso, concordar que não é dos nomes que se deve partir, mas que as coisas [...] devem conhecer-se e investigar-se, partindo delas mesmas. (Platão, § 10, 1963, p. 156, grifo nosso)

Dada a argumentação socrática, a visão de Crátilo não é admitida na Antiguidade, mas seus questionamentos podem levar à reconstrução dos motivos que nos conduzem a tomar a linguagem como processo do conhecimento – e, principalmente, como o próprio conhecimento. No entanto, o processo que se estabelece na direção desta mudança de rota é longo, e se inicia com a revisão aristotélica sobre a linguagem, junto do próprio nascimento da Filologia.

Mais do que isto, a postura que a linguagem, como possibilidade do conhecimento, como aponta Crátilo, terá impacto conceitual no surgimento de uma arte bibliológica – e, principalmente, em sua emancipação, já no século XIX. Tal arte vincula-se diretamente ao desenvolvimento de uma teoria bibliográfica ou de uma teoria do conhecimento estruturada em uma racionalidade bibliográfica que, por sua vez, virá constituir o que atualmente tratamos a partir da expressão “organização do conhecimento”.

Os questionamentos sobre o nome próprio (sempre juntos da questão do “nome”, “substantivo”) não deixam de ser objeto do debate filosófico da Antiguidade à Contemporaneidade. A filosofia do século XX também abordará o objeto como um de seus elementos centrais. Seja através da virada lingüística (e, dentro desta, da relação entre filosofia analítica e filosofia da linguagem ordinária), seja a partir das dicotomias entre estruturalistas e pós-estruturalistas, o nome próprio surgirá, aqui e acolá, como elemento de investigação entre filósofos e teóricos das ciências sociais.

Estritamente no âmbito dos filósofos da virada lingüística, esta aderência pode ser visualizada nos debates que atravessam Russel, Quine, Frege e Wittgenstein. Para Russel, podemos reduzir os “nomes próprios” aos nomes comuns, assim como a inversão se faria igualmente válida, ou seja, podemos reduzir os nomes comuns aos nomes próprios. A perspectiva russeliana pressupõe que os nomes próprios resultam de elementos retirados da experiência direta, e tem relação objetiva no contexto da denotação. (Abbagnano, 2000)

Frege, por sua vez, atenta-se para uma distinção entre denotação e sentido, sendo o último algo se dá na apreensão resultante do indivíduo que domina suficientemente uma determinada língua. A denotação responde pontualmente por uma relação de referência ao objeto. Isto resulta na possibilidade de nomes (incluindo nomes

próprios) responderem por sentidos diferentes, mesmo apontando (referindo-se) ao mesmo objeto, ou pessoa. Tal solução fregeiana apresenta a inferência da não existência de pressupostos particulares sobre a natureza da linguagem (Abbagnano, 2000)

O conhecido desdobramento destas questões na filosofia chegará ao complexo pensamento de Ludwig Wittgenstein. Entre os filósofos anteriores da virada lingüística, como o caso russeliano, temos centralmente uma perspectiva lançada sob o ponto de vista logicista. Em outras palavras, a aporia do nome próprio poderia ser resolvida tão somente através da procura por uma demonstração lógica da relação entre nome e indivíduo correspondente. Com a divisão comum das fases do pensamento wittgensteiniano (a saber, um primeiro Wittgenstein da lógica formal, um segundo da linguagem ordinária), tal base logicista começa a receber outras abordagens.

No clássico fragmento § 79 das Investigações Filosóficas (obra nuclear da segunda fase de seu pensamento) resultante de passagens de discussão sobre o nome, tendo como interlocutor, por exemplo, Frege e Russel, Wittgenstein (1979, p. 43-44) manifesta a seguinte proposição: “Moisés não existiu”. A partir desta constatação hipotética, sua discussão procura elencar as consequências semânticas, como, “os israelitas não tiveram nenhum chefe quando deixaram o Egito”, ou “não existiu nenhum homem que tivesse realizado tudo que a Bíblia narra de Moisés”.

Na visão wittgensteiniana, a partir de Russel diríamos que Moisés pode ser definido segundo diferentes descrições, como “o homem que guiou os israelitas”, “o homem que viveu naquele tempo”, “o homem que em criança foi retirado do Nilo pela filha do faraó”. Do mesmo modo, conforme definição selecionada, a frase “Moisés existiu” recebe um sentido diferente. (Wittgenstein, 1979)

Uma das rotas da reflexão wittgensteiniana é afirmar que, diante da assertiva sobre a condição ôntica de Moisés (a existência de alguém considerado “Moisés”), responderá sempre por uma camada aberta de acréscimo de descrições para tal nome próprio. “Terá para mim o nome ‘Moisés’ um determinado uso, sólido e sem equívoco em todos os casos possíveis?”, indaga Wittgenstein (1979, p. 44).

As respostas da reflexão wittgensteiniana (de segunda fase) para as aporias do nome próprio começam a ganhar aqui a conhecida direção do debate da linguagem ordinária. O filósofo admite que o uso de um nome próprio não possui uma certa “significação rígida”, como gostariam os

lógicos, incluindo ele, em sua primeira fase, demarcada pela obra *Tratado Lógico-Filosófico*. A ausência de tal rigidez não impacta, porém, o seu “uso”, ou seja, o nome próprio, por mais impreciso (logicamente) que possa parecer, não torna seu uso um absurdo no cotidiano. (Wittgenstein, 1979)

As observações sobre o uso da linguagem (incluindo a apropriação de nomes próprios), levam Wittgenstein (1979, p. 49) a abordar criticamente a lógica. Em suas palavras, trata-se de observar, de fato, “em que medida a lógica é algo sublime”. Este caráter sublime estaria na direção de uma certa “profundidade especial”, uma “significação universal”, ou seja, da “consideração da lógica” como aquela ciência que “investiga a essência de todas as coisas”. No entanto, a “nova” abordagem wittgensteiniana declara não ser esta a intenção central da filosofia, pelo menos, sua filosofia ordinária. Não é a profundidade o interesse em questão. “Queremos compreender algo que já está diante de nossos olhos. Pois parecemos, em algum sentido, não compreender isto”. (Wittgenstein, 1979, p. 49)

Wittgenstein (1979) encontra aqui dos usos cotidianos do nome próprio à variedade e à dinâmica de sua vivência entre cada comunidade de uso. A perspectiva de segunda fase wittgensteiniana sobre a filosofia dos “nomes próprios” aponta para as diferentes experiências de reconhecimento de como nome próprio foi tecido como objeto da filosofia, em uma primeira impressão, e, principalmente, a multiplicidade de condições em que foi elaborado em cada contexto.

O quadro vasto da multiplicidade nos remete ao conjunto de espaços discursivos de apropriação do “nome próprio”, presentes em domínios como Filologia (incluindo aqui Gramática e Etimologia), Retórica, a própria Lógica, História, Mitologia, Sociologia, Antropologia, Estudos Literários, Direito, dentre outros. Este “mapeamento” nunca concluso oferece uma gama inumerável de indícios para a compreensão do percurso que leva o nome próprio, de seus usos plurais às tentativas de reconhecimento, de reflexão e de padronização no âmbito da Organização do Conhecimento.

Em outras palavras, da relevância histórica presente na “condição filosófica” dos “nomes próprios”, passando por seus questionamentos nos terrenos dos estudos sociais (político, sociológico e antropológico), chegamos à sua relevância pontual no exercício de construção dos experimentos, das abordagens teóricas e das abordagens aplicadas da Organização do Conhecimento na Contemporaneidade. A marca espaço-temporal do século XX, no entanto, nos con-

vida a compreender como “abordagens pós-estruturalistas” não só não abandonaram a discussão sobre o nome próprio, como colocaram-no em evidência em diferentes casos.

Voltar-se para a Organização do Conhecimento nos últimos cem anos e encarar o “nome próprio” como um de seus elementos é também compreender as questões complexas lançadas por filósofos como Michel Foucault, Roland Barthes, François Lyotard, Jacques Derrida, Jean Baudrillard, dentre outros. Não sendo intuito demarcar, neste trabalho, apenas um modo “pós-estruturalista” de apreensão do “nome próprio”, detemo-nos em alguns destes autores para aprofundar o panorama crítico que nos leva ao foco das práticas informacionais.

Podemos elencar alguns dos elementos gerais que posicionam-se na fronteira entre os novos questionamentos sobre o “nome próprio” e as abordagens tensionadas na Organização do Conhecimento, dentre eles:

- A condição da “invenção” do autor (indistigível da inferência do “nome próprio”) na análise foucaultiana a partir do Medieval e os questionamentos das distinções de sua função a partir de então;
- A destituição da centralidade do autor a partir das perspectivas de iluminação do leitor como edificador do sentido;
- A fragmentação da “condição ôntica” do autor e, por conseguinte, da “evidência” do “nome próprio”, diante das práticas de produção colaborativa no mundo pós-*web*.

Estas fundações da Contemporaneidade, pós-wittgensteinianas, definem a dificuldade de tratamento da aparente condição objetiva do “nome próprio” e entrelaçam-se nos dilemas corretos da Organização do Conhecimento. É preciso reconhecer, como procuramos desenvolver, a seguir, que os laços, os nós e os fios dispersos que subsistem na filosofia dos “nomes próprios” têm, na Organização do Conhecimento, não apenas um reflexo (a manifestação aplicada de suas problemáticas), mas também suas possibilidades de aparecimento e desenvolvimento. Em outros termos, percebemos que os horizontes da organização, da representação e do acesso a conteúdos e continentes sempre estiveram imbricados no debate filosófico e informacional quando tratamos dos “nomes próprios”.

### 3. O nome próprio na Organização do Conhecimento

Allonymes. On donne ce nom à ceux qui, en publiant un ouvrage, se cachent sous le nom véritable de quelque auteur de réputation, et

cherchent à leur attribuer des ouvrages qu'ils n'on pas faits. (Peignot, 1802, p. 12)

Para Esteban Navarro e García Marco (1995), a Organização do Conhecimento é uma disciplina dedicada ao estudo e desenvolvimento de fundamentos e técnicas de planejamento, construção, gestão, uso e avaliação de sistemas de descrição, catalogação, ordenação, classificação, armazenamento, comunicação e recuperação de documentos. Portanto, ela atua em três perspectivas: representação, organização e comunicação do conhecimento humano.

A autoria pode estar contida tanto na organização quanto na representação do conhecimento. Embora pareçam semelhantes, a organização e a representação possuem conotações próprias.

A organização da informação é, portanto, um processo que envolve a descrição física e de conteúdo dos objetos informacionais. O produto desse processo descritivo é a representação da informação, entendida como um conjunto de elementos descritivos que representam os atributos de um objeto informacional específico (Brascher e Café, 2008, p. 5).

Enquanto que,

A organização do conhecimento, por sua vez, visa à construção de modelos de mundo que se constituem em abstrações da realidade. Esses dois processos produzem, conseqüentemente, dois tipos distintos de representação: a representação da informação, compreendida como o conjunto de atributos que representa determinado objeto informacional e que é obtido pelos processos de descrição física e de conteúdo [...] (Brascher e Café, 2008, p. 6).

Considerando tais diferenças, considera-se que a ação de identificar a autoria, estabelecer e padronizar os pontos de acesso a um registro bibliográfico está contido na Organização do Conhecimento, e a transcrição da autoria na área 1 da ISBD, conforme consta na página de rosto de livro, seria organização da informação.

O documento a ser representado pode fornecer os dados para a organização da informação referente à autoria de modo mais claro que à organização do conhecimento autoral, este último que se relaciona diretamente a questão filosófica de "nome próprio" e organização do conhecimento.

Em 1968, Roland Barthes publicou o artigo "A morte do autor". Nesse texto, Barthes (1988) afirma que o escritor é um sujeito social, historicamente constituído, que produz um texto com base em outros textos.

A função autor é também abordada em "O que é um autor?", comunicação de Foucault à Société

Française de Philosophie, em 1969. Para ele, alguns "[...] discursos são providos da função 'autor', ao passo que outros são dela desprovidos" (Foucault, 2006, p. 46).

É importante destacar que o surgimento do autor na Idade Média coincide com o surgimento de catálogos mais estruturados. Antes do Medievo, existiam tabletes de argilas e listas com dados bibliográficos, mas se desconhece se tais listas eram mecanismos de busca e recuperação da informação ou se eram apenas listas de inventários (Garrido Arilla, 1996). O processo de elaboração de catálogos mais estruturados é decorrência de ações realizadas séculos antes. No século II, por exemplo, a organização bibliográfica tornou-se mais específica e individualizada, "[...] quando Galeno, médico grego, fez a relação de seus trabalhos para que não fossem confundidos com os de outros autores" (Caldeira, 1984, p. 261).

Nesse caso, a introdução da autoria foi um mecanismo de organização que permitiu a personificação do conhecimento, além de um importante elemento de individualização de documentos. Desde então, as listas com dados bibliográficos tornaram-se mais elaboradas e iniciou-se um processo de identificação dos documentos, no qual são incluídos elementos descritivos com a finalidade de fornecer uma representação minuciosa e fidedigna, distinguindo um documento dos demais presentes no acervo.

No entanto, somente a partir da Idade Média é que a identificação dos autores tornou-se recorrente. Porém, essa constatação não significa a inexistência de autores antes desse período. Ela revela a alteração da relação entre o autor e o texto, ocorrida durante o Medievo, e o desenvolvimento dos catálogos que começaram a incorporar, em seus elementos descritivos, o nome do autor para identificar uma obra.

No século VIII, nas bibliotecas medievais, surge uma das primeiras listas com dados sobre o título e autor; tal lista está contida em um livro. Não é possível observar nenhuma ordenação, nem de classificação nem alfabética; essa lista se aproximava mais de um inventário que de um instrumento de busca e recuperação. Somente no século IX é que surgem catálogos mais bem estruturados. A biblioteca de Reichenau, na Alemanha, compilou vários catálogos entre os anos de 822 e 842. O primeiro deles mantinha as obras de um mesmo autor reunidas, outro continha informações sobre o número dos volumes ou rolos e as obras nele contidas. Em 831, o mosteiro St. Requier, na França, compilou um catálogo que também apresentava o conteúdo dos volumes ou rolos e fazia entrada

para o autor, porém sem ordem perceptível (Strout, 1956).

O catálogo continuou evoluindo e se aprimorando. Em 1247, uma lista da biblioteca de Glastonbury Abbey adicionava uma designação variada à descrição dos livros, como por exemplo: inútil, legível, bom – provavelmente revelando as condições físicas dos mesmos. Em 1372, na Inglaterra, os frades agostinianos de York organizaram uma lista na qual as obras de um autor eram separadas quando os assuntos eram distintos. Em 1558, o catálogo do convento de Bretton, em Yorkshire, incluiu a entrada para os nomes dos editores e tradutores (Strout, 1956).

As primeiras ocorrências sistemáticas e ordenadas alfabeticamente de nomes de autores encontram-se nos índices dos livros e autores proibidos, estabelecidos no século XVI [...]. É isso que Foucault chama de “apropriação penal dos discursos” – o fato de poder ser perseguido e condenado por um texto considerado transgressor (Chartier, 1999, p. 34).

No século XVII várias publicações abordavam o desenvolvimento do catálogo. Na França, Gabriel Naudé destacou que a função mais importante do catálogo era encontrar livros e identificá-los bibliograficamente. Ele recomendava um catálogo dividido em duas seções, uma por assunto e outra por autores (Strout, 1956).

Naudé apresenta a biblioteca como uma instituição necessariamente pública e universal. Pública no sentido de abertura a todos e universal por conter todos os autores (ele não escreveu “livros” mas “autores”) que tivessem escrito sobre a diversidade dos assuntos interessantes ao ser humano, em particular as artes e ciências (Coelho, 1997, p. 77).

Na época em que Naudé utilizou o termo “autores”, a função do autor estava consolidada e a obra era intimamente ligada ao autor. Sob o desdobramento desta visão moderna nascente, na qual se enquadra o ponto de vista de Naudé, podemos encontrar a crítica barthesiana. Barthes (1988, p. 66) chama a atenção que a imagem da literatura na cultura corrente está tiranicamente centralizada no autor, sua pessoa, e “[...] a explicação da obra é sempre buscada do lado de quem a produziu, como se através da alegoria mais ou menos transparente da ficção, fosse sempre afinal a voz de uma só e mesma pessoa, o autor, a entregar a sua ‘confidência’”.

Ao acompanhar as mutações do catálogo, percebe-se muitas vezes a ausência da identificação do autor. Os catálogos mais elaborados com as questões autorais surgem durante a

Idade Média. Coincidência ou não, foi nessa época que a censura às obras e aos autores se tornou rigorosa. Neste período, a obra era condenada junto com seu autor, portanto, a identificação era necessária para a concretização da censura.

Provavelmente, a ausência ou presença do autor não se relacionam somente às questões de censura; o autor não surgiu apenas na medida em que era necessário puni-lo durante a Inquisição. O “surgimento” do autor também está relacionado com a necessidade do reconhecimento de um autor pela sociedade.

O autor é uma personagem moderna, produzida sem dúvida por nossa sociedade na medida em que, ao sair da Idade Média, com o empirismo inglês, o racionalismo francês e a fé pessoal da Reforma, ela descobriu o prestígio do indivíduo ou, como se diz mais nobremente, da “pessoa humana” (Barthes, 1988, p. 66).

Ainda nas relações da passagem do Medievo para o Mundo Moderno, Chartier (2008) nos revela uma alteração concreta no comportamento do autor, ao comentar sobre o livro oferecido ao príncipe. “No livro, a dedicatória ao príncipe é, inicialmente, uma imagem [...] são os frontispícios que representam o ‘autor’, ajoelhado, oferecendo ao príncipe [...] um livro ricamente encadernado, que contém a obra da qual ele é criador [...]” (Chartier, 2008, p. 186).

Uma vez que o autor passou a ser reconhecido como responsável pelo seu discurso, tornou-se um elemento descritivo muito importante na representação da informação. O surgimento da função “autor” e o olhar da sociedade sobre ele alteraram a identificação da autoria nos catálogos. Após a Idade Média, a presença da autoria fica cada vez mais evidente. Existem duas justificativas para o acréscimo da autoria como pontos de acesso nos catálogos:

- a) para encontrar um documento específico quando o autor é conhecido;
- b) para encontrar quais documentos de um determinado autor há na biblioteca (Needham, 1964).

Existem outras características que justificam a inclusão da autoria nos catálogos: a identificação e a seleção de documentos semelhantes, com autorias diferentes. Com a multiplicação “geométrica” da produção documental, século após século, após o Quatrocentos, tornou-se cada vez mais necessário elencar elementos que possam ajudar na organização, identificação e seleção de documentos. Os autores e outros responsáveis, como tradutores, editores,

entre outros, são elementos fundamentais para a busca, recuperação e seleção de documentos.

A função autor, portanto, está diretamente relacionada ao contexto social, histórico e econômico de uma sociedade. Durante um período e em determinadas sociedades, os textos tidos como literários (narrativas, contos, epopeia, entre outros), eram recebidos anonimamente sem que o reconhecimento de sua autoria fosse requerido. Atualmente, a função autor em uma obra literária desempenha um papel fundamental. No século XVII ou no século XVIII, a função autor começa a ser apagada dos discursos científicos e outro conjunto sistemático de técnicas, métodos, objetos de experimentação, laboratório e data é que atribui validade ao discurso e não mais o indivíduo que o produziu (Foucault, 2006).

Needham (1964) aponta que a busca por assunto é enfatizada no meio científico e tecnológico, no qual o autor tem sido menos importante. Principalmente após o século XIX, o crescimento da produção científica e o aumento das buscas por assunto culminaram no desenvolvimento da Documentação.

Como a Documentação prioriza o tratamento da informação por assunto, Wersig (1993) a considera como um indicador da despersonalização do conhecimento. Ao priorizar a busca por assunto, o autor deixou de ser o critério mais relevante para recuperação. Embora ele permaneça como elemento importante, faz parte de um conjunto de elementos e, nem sempre, será o elemento principal para a busca, recuperação e seleção de documentos pelos usuários.

Foucault (2006) restringiu as características da “função autor” considerando o autor de um texto, de um livro, àquele que se atribui uma criação textual. Porém, reconhece que a função autor é apenas uma das especificações das diferentes funções que um sujeito pode assumir em relação ao discurso.

As quatro características da função autor são assim resumidas:

[...] a função autor está ligada ao sistema jurídico e institucional que encerra, determina, articula o universo dos discursos; não se exerce uniformemente e da mesma maneira sobre todos os discursos, em todas as épocas e em todas as formas de civilização; não se define pela atribuição espontânea de um discurso ao seu produtor, mas através de uma série de operações específicas e complexas; não reenvia pura e simplesmente para um indivíduo real, podendo dar lugar a vários “eus” em, simultâneo, a várias posições-sujeitos que classes diferentes de indivíduos podem ocupar (Foucault, 2006, p. 56-57).

Das quatro características elencadas por Foucault (2006), a primeira e a segunda vinculam-se intrinsecamente às questões da Organização do Conhecimento e podem ser visualizadas na representação de um documento.

Os códigos de catalogação, um dos instrumentos que norteiam a organização (pontos de acesso para responsabilidade) e representação (transcrição das áreas) da autoria nos permitem verificar que o autor do discurso pode “ocupar” vários espaços. Inclusive, uma mesma pessoa pode possuir várias identidades (pseudônimos, por exemplo, ou então quando uma pessoa ora exerce um cargo público e seu discurso passa a ser identificado como um discurso institucional) a fim de “organizar” seu discurso, ou seja, utiliza-se do conceito de “nome próprio” para seus múltiplos discursos.

#### 4. Reflexões críticas sobre o nome próprio hoje

Mencionar o nome sem uma prece ou sem a fórmula dissimulada de modestia parece muito raro na literatura do alto alemão médio. [...] O esclarecimento da questão não é tão necessário quanto poderia parecer. Contribuiu para nosso conhecimento do conceito que de si fazia o homem medieval. (Curtius, 2013, p. 661)

Se tomarmos um ponto de partida histórico-arbitrário, para identificarmos um “quando e onde” a questão do nome próprio torna-se elementar para as práticas de organização, representação e acesso aos conteúdos e continentes, Alexandria nos lega uma série de indícios que podem se tornar critérios para tal decisão. Não procuramos aqui discuti-los. Nosso ponto de inflexão (que simultaneamente apresenta-se como um critério) trata-se da chamada “questão homérica”. Entre as práticas de construção de uma remota ciência da linguagem, estão os experimentos de organização do conhecimento, reconhecidos principalmente a partir do erudito, filólogo e bibliotecário Calímaco.

Nesse período, também chamado Alexandrino, que se estende aproximadamente do ano 322 a 146 a.C., eruditos de toda procedência se sucederam como bibliotecários da famosa Biblioteca da Alexandria, que, com os seus 490 mil volumes e os 43 mil colocados, por falta de espaço, no Museu Serapeum contíguo à biblioteca, se tornou o maior centro de cultura helênica da Antiguidade. A ordenação e catalogação dessas obras levantaram problemas pertinentes à sua autenticidade, à vida dos autores e posteriormente à preparação dos textos para o público e para as escolas. (Spina, 1994, p. 66)

Passados dois milênios, entre o século XIX e o XX, o desenvolvimento de teorias, abordagens e

instrumentos de organização e de uso da informação continuaram tendo o nome próprio como uma centralidade. É exemplo a preocupação de Peignot (1802a,b) (Figura 1) e de Paul Otlet (1934) com a descrição de documentos e o imperativo do nome próprio (Figuras 2 e 3).

| P S E                                 |   | 141 |
|---------------------------------------|---|-----|
| NOMS SUPPOSÉS.                        | NOMS VRAIS.   |     |
|                                       | de Constantinople, qui se fit moine.                              |     |
| <i>Chrysippus</i> .....               | Libert Fromond.   |     |
| <i>Cleophilus, Octavius</i> .....     | François de Fano;   |     |
| <i>Clevier (Thomas du)</i> .....      | Bouav. Desperiers,  |     |
| <i>Clouset ou Duclouset</i> .....     | Jean Coustel et Isaac le Maître.                                  |     |
| <i>Clopiere</i> .....                 | Voltaire.   |     |
| <i>Cocaius Merlinus</i> .....         | Théophile Folengi.  |     |
| <i>Colet Champenois, Claude</i> ..    | Gilles Boileau.   |     |
| <i>Colvinus</i> .....                 | Louis Dumoulin.   |     |
| <i>Columba</i> .....                  | Jean Coster.  |     |
| <i>Comes ou de Comitibus</i> ....     | Majoragius.   |     |
| <i>Conchilax</i> .....                | { Pamphile d'Alexandrie, selon Suidas; et Nicandre, selon Galien. |     |
| <i>Constantius, Marcus-Antonius</i> . | Etiennè Gardiner.   |     |
| <i>Copus, Alanus</i> .....            | Nicolas Harpsfeld (1).  |     |
| <i>Corallus et Cordatus</i> .....     | Ulric Hutten.   |     |
| <i>Corona ou Coronæus, Jea</i> ..     | Jean-Etienne Menochius;   |     |
| <i>Covelle, Robert</i> .....          | Voltaire.   |     |
| <i>Crapin</i> .....                   | Jean Tristan de S.-Amant.   |     |
| <i>Crassinius</i> .....               | Charles Sigonius.   |     |
| <i>Critobulus</i> .....               | S. Jérôme.  |     |
| <i>Crox (Florent de)</i> .....        | Jean Lepelletier.   |     |
| <i>Cubstorf</i> .....                 | Voltaire.   |     |
| <i>Cyprianus, Carthag</i> .....       | Erasmè.   |     |

Figura 1. Pseudônimos identificados por Peignot.  
Fonte: Peignot (1802b)

Os verbetes peignotianos estruturalmente se preocupam com a necessidade de reconhecimento de uma espécie de “vida particular dos nomes próprios”, de um lado, e, de outro, sua “vivência pública”, um “jogo de linguagem” que resulta na variação de significantes e de significados conforme cada “forma de vida”, cada classe social, cada ambiente, cada contexto. A Figura 1 apresenta parte da longa lista de nomes verdadeiros e nomes “artificiais” identificados por Peignot (1802b).

A “subjetividade” atentada por Paul Otlet (1934) evoca a complexidade do nome próprio no contexto social (não apenas hoje) e, como consequência, seus desafios no desenvolvimento da Organização do Conhecimento. Influência clara do advogado belga, Gabriel Peignot, em 1802, já atentava-se igualmente para a questão do nome próprio em verbetes como *allonymes*, *cryptonymes*, *pseudonymes*, *hétéronymes*, *anonymes*. As Figuras 2 e 3 demonstram a teoria aplicada do “nome próprio” na construção das fichas catalográficas em formato manual.

Figura 2. Modelo teórico da ficha catalográfica de Otlet  
Fonte: Otlet (1934)

Figura 3. Modelo aplicado da ficha catalográfica de Otlet  
Fonte: Otlet (1934)

Atualmente, a IFLA tem desenvolvido vários modelos conceituais a partir de ferramentas computadorizadas para o universo bibliográfico, iniciando com os *Functional Requirements for Bibliographic Records* (FRBR), relacionado ao universo bibliográfico de modo geral, depois com o *Functional Requirements for Authority Data* (FRAD), modelo conceitual especificamente relacionado à responsabilidade (ou autoria), e por último com os *Functional Requirements for Subject Authority Data* (FRSAD), relacionado aos temas abordados em documentos. Os dois primeiros modelos abordam a autoria na organização e representação da informação e do conhecimento, conforme apresentado a seguir.

Os FRBR definem o registro bibliográfico

como um agregado de dados associados a entidades descritas em catálogos de bibliotecas e bibliografias nacionais. Incluso naquele agregado de dados estão os elementos de dados descritivos, como aqueles definidos nas Descrições Bibliográficas Internacionais Normalizadas (ISBDs); os elementos de dados usados nos cabeçalhos para pessoas, entidades coletivas, títulos e assuntos, que funcionam como instrumentos de armazenagem ou entradas de índices; os outros elementos de dados usados para organizar um arquivo de registros, como os números de classificação; as notas como resumos ou sumários; e os dados específicos [relativos] a coleções de bibliotecas, tais como números de acesso e chamada (IFLA, 1998, p. 7, tradução nossa).

Nesta definição podemos afirmar que a autoria está presente no registro bibliográfico (representação documental) em dois momentos distintos: na organização, quando localizada nos pontos de acesso; e, na representação, quando compõem a descrição bibliográfica, seguindo as áreas presentes nas *Internacional Standard Bibliographic Description* (ISBD).

O foco deste trabalho não estaria na transcrição literal da autoria nas áreas de descrição bibliográfica e sim nos pontos de acesso, pois são eles que organizam o conhecimento consolidado por indivíduos. Neste sentido, o conceito de “nome próprio” permeia a representação da pessoa (autor) contida no registro de autoridade, que é definido como

um agregado de informações sobre uma instância de uma entidade da qual seu nome é usado como um ponto de acesso controlado para citações ou registros bibliográficos em um catálogo de biblioteca ou arquivo bibliográfico (IFLA FRANAR, 2007, p. 1, tradução nossa).

O estabelecimento de registro de autoridade possui um papel fundamental na organização do conhecimento, pois permite identificar o sujeito do discurso, em suas várias identidades bibliográficas. Com o intuito de melhor organizar e representar os autores, surge o *Virtual International Authority File* (VIAF), que integra vários catálogos e registros de autoridade. Independentemente da grafia ou da forma que o autor é conhecido, todas as variações de seu nome podem levar ao autor do discurso.

Neste contexto, na forma adotada para compor os pontos de acesso de um registro bibliográfico, o “autor” seria identificado unicamente por um código numérico - no VIAF seria o VIAF ID. Há também outro número que busca tornar único o autor, o *International Standard Name Identifier* (ISNI).

Essas iniciativas, como pode-se observar, são tentativas que organizam os discursos e atrelam historicamente aos seus autores, relacionando pessoas e documentos. Assim como o livro possui o ISBN, o periódico o ISSN, a pessoa teria o ISNI.

Acredita-se que o desenvolvimento de modelos conceituais e códigos identificadores para os autores seja uma busca ou um retorno aos fundamentos teóricos da disciplina a fim de encontrar soluções para os problemas contemporâneos de Representação Documental. O autor passa a ser compreendido como uma das possíveis funções de relacionamento entre a entidade pessoa e as entidades do Grupo 1 dos FRBRs que representa um documento, a saber, “Obra, Expressão, Manifestação e Item”, não

mais como as várias funções sob o termo “autor”.

Durante muito tempo essa concepção agregadora atendeu às demandas, porém não é mais satisfatória. Inicialmente, a figura do autor era suficiente para identificar um documento. Com o passar do tempo, outras funções, como o compilador e a entidade coletiva, assumiram o papel do “autor” para tornar a descrição mais precisa, mais detalhada, até o momento em que o “autor” e a “autoria” (quaisquer indicações de responsabilidade por determinado recurso) passaram a ser tão amplos, abrangendo tantas possibilidades, que, ao invés de facilitar o tratamento documental, o banalizou. Tudo poderia ser considerado como “autor” na elaboração do registro bibliográfico, porém esta concepção não é compartilhada nem pelo usuário nem pelos criadores do documento.

A trajetória da autoria e do autor revela que as primeiras concepções permeavam soluções práticas para a elaboração de registros bibliográficos, organização e recuperação da informação. Ou seja, as soluções centravam-se na facilidade de processamento documental. Hoje a concepção adotada com os modelos conceituais busca integrar a visão da produção do conhecimento com o usuário.

Paralelamente à preocupação teórica e aplicada, o reflexo do imperativo do “nome próprio” é facilmente reconhecido em um contexto social que transversaliza as teorias e as práticas de Organização do Conhecimento. Mesmo em uma sociedade estruturada na ideia e no uso da imagem pessoal como “identidade” (fotografias de faces, por exemplo), o “nome próprio” continua sendo elemento-chave, “ponto de acesso”, fonte de uma alteridade ainda inerente ao sujeito.

Ferramentas privadas como Facebook, e públicas, como o diretório de currículos da Plataforma Lattes demonstram como a necessidade da imagem não substitui, necessariamente, o imperativo do “nome próprio”. As Figuras 4 e 5 ilustram essa discussão.



Figura 4. Exemplo de perfil da empresa Facebook  
Fonte: Facebook. Disponível em: <<https://www.facebook.com>>. Acesso em 30 jun.2015.



Figura 5. Exemplo de perfil na Plataforma Lattes do CNPq  
Fonte: CNPq. Disponível em: < <http://lattes.cnpq.br/>>

O avanço de uma “sociedade das imagens” (e, principalmente, uma “sociedade das imagens faciais”, ou, ainda, uma *selfie society*), está intimamente relacionado aos jogos de linguagem que manipulam nomes e imagens de rostos, perfis, etc. Já em 1934, Otlet destacava a necessidade de alguns autores relacionarem seus nomes junto de fotografias (retratos) nos livros. Ao mesmo tempo em que identifica-se movimentos com o intuito de ressaltar a diferença entre representar o autor de um documento (representação da informação) e do autor como ponto de acesso (organização do conhecimento), verifica-se uma dificuldade em adequar a representação documental com o grande volume de produção de documentos e a necessidade de alimentar base de dados.

Tal *selfie society* já está “manifestada” na Organização do Conhecimento, porém com as incorreções e desafios das transformações culturais em curso. Um exemplo típico é o citado caso do VIAF. Trata-se de um catálogo de registros de autoridade de diferentes agências bibliográficas nacionais e/ou locais. Sua função central é identificar e descrever autores, permitindo a inclusão de uma imagem relativa ao registro.

Uma das críticas atuais nas relações entre “nome próprio” e “imagem identitária” está na aplicação conflitante encontrada no contexto do VIAF. Ao contrário da correspondência esperada, o que encontramos, em geral, como o modelo da Figura 6, dentro do catálogo de autoridade é, no lugar da “imagem referente ao nome” (uma fotografia da face do autor), a imagem do recurso bibliográfico (um livro do autor). Percebe-se aqui o jogo entre o imperativo do autor diante do imperativo do recurso, fonte de uma das mais clássicas discussões do século XX, com o mencionado trabalho de Roland Barthes sobre a pretensa “morte do autor”.



Figura 6. Registro de autoridade - Barthes  
Fonte: VIAF. Disponível em: < [http://www.viaf.org/viaf/44295060/#Barthes,\\_Roland.>](http://www.viaf.org/viaf/44295060/#Barthes,_Roland.>)

Na procura atual (permitida sistemicamente pelo VIAF), o “autor” tem sua condição não apenas “reconhecida”, como “existenciada” a partir da “imagem” que o identifica. Em certa medida, trata-se de uma espécie de “retorno” ou “afirmação” da presença da imagem da autoria, para além do nome, dentro da obra, como destacado por Otlet (1934). Reconhecemos a diferença de uma imagem adotada em uma publicação, com objetivos diversos, como o comercial, e sua aplicação em um sistema internacional de autoridades. O fato nos chama a atenção, uma vez que, no segundo caso, estamos diante de um processo e de um produto intelectual, o VIAF, de fundamentação do “nome próprio”, atrelando-o a uma imagem identitária no mundo atual.

Outra crítica relacionada ao âmbito dos “nomes próprios” na Organização do Conhecimento está nos metadados. Existem vários metadados que indicam o autor em apenas um de seus campos. Neste caso, ou a representação da informação contida no documento ou a organização do ponto de acesso terá lacuna. Isto ocorre nos repositórios digitais que utilizam-se do Dublin Core, em seu registro possui apenas um campo para indicar o autor. Esse campo pode utilizar-se do controle e padronização do nome do autor, estabelecido pelo “registro de autoridade”. Neste caso, a forma do autor como consta na página de rosto de um documento não será transcrita. Caso seja adotada a forma como consta no documento, não haverá padronização de pontos de acesso, resultando em dispersão do discurso.

A identificação das variadas formas (grafia, nome completo ou abreviado, pseudônimos, identidades diferentes para uma pessoa que ocupa um cargo político etc., já anunciadas no Oitocentos pela visão peignotiana) deve ser contemplada. Neste sentido, observa-se claramente a questão do “nome próprio” da filosofia na Organização do Conhecimento.

## 5. Considerações finais

L'auteur joint souvent à son nom ses propres titres, qualités, notamment ceux de sa profession ou ceux des ses titres scientifiques qui forment son autorité quant à l'ouvrage. (Otlet, 1934, p. 113)

Um dos modos mais claros de se atentar para uma “epistemologia da organização do conhecimento” vislumbrada em uma abordagem do “nome próprio” partícipe de nosso pensamento informacional está na expressão de Paul Otlet (1934, p. 251) no fragmento § 251.22 (parte integrante de 251.2. L'auteurs et l'oeuvre): “L'auteur est l'élément subjectif de la connaissance”.

Ao observar a história dos catálogos e dos instrumentos utilizados para organizar e representar documentos, o autor sempre foi um elemento presente. Algumas vezes, sua identificação simplificada foi o bastante. Atualmente, verifica-se ainda a importância de padronização para busca e recuperação da informação. Tendo em vista os embates da filosofia e de diferentes ramos do saber, podemos reunir alguns elementos que confluem para uma reflexão crítica sobre o “nome próprio” na teoria e na prática da organização, da representação e do acesso ao conhecimento registrado.

Reconhecendo a Lógica como espaço discursivo fundamental entre Recuperação da Informação e Filosofia, percebemos, como já tratado no debate sobre a filosofia analítica (passando por Russel, Frege e Wittgenstein), o imperativo essencialista, de representação acurada de um nome próprio (o nome de um indivíduo representando objetivamente sua obra).

De outro lado, considerando a tradição filológica dos estudos biblioteconômico-informacionais, chegamos também até outro ponto da fronteira mencionada. Para a Filologia a preocupação com a autenticidade do documento passa por questões como autenticidade da obra, genuidade do documento e relações paratextuais. Este, no entanto, é apenas o princípio de uma longa atividade discursiva nas Humanidades e nas Ciências Sociais sobre o nome próprio, passando pela Mitologia, pela História, pela Ciência Política, pela Retórica, pelo Direito, pela Sociologia, pela Antropologia, dentre outros saberes.

No domínio da Sociologia, por exemplo, os potenciais problemas vinculados ao nome próprio que provêm ou repercutem na Organização do Conhecimento são inumeráveis. É o exemplo, nos estudos de gênero e comunicação científica, dos desafios de representação de nomes próprios femininos, o que dificulta a identificação de gênero e a possibilidade de análise dos

dilemas da relação entre os sexos. Otlet (1934) destacava o delicado costume de supressão do nome de autores de obras de mulheres, nomes que desapareceram para sempre, em seu ponto de vista.

Outros elementos sociais afetam diretamente as possibilidades de compreensão do nome próprio dentro da Organização do Conhecimento, como adoção de determinados sobrenomes de poder, variações culturais em razão de casamentos ou conversões religiosas ou ainda títulos civis, profissionais, anônimos, pseudônimos, plágios.

Se tomássemos, ainda, a Retórica como ponto de partida, retornaríamos à “questão Beda” e à construção dos epítetos no contexto da Igreja Católica nos fins do primeiro milênio. Este e aqueles exemplos demonstram a complexidade e a relevância dos estudos sobre o “nome próprio” e o seu lugar na epistemologia e na história da Organização do Conhecimento.

Diante de tais desafios, este estudo, de caráter exploratório e teórico, buscou debater o pensamento sobre o “nome próprio” desenvolvido dentro e fora da Organização do Conhecimento, para iluminar a discussão neste último caso. A travessia nos mostrou a extensão da questão e as necessidades de diálogos disciplinares resultados da vastidão epistemológica do objeto de estudo. Reconhecemos, em certa medida, uma pretensa “invisibilidade do aparente”, ou uma “neutralidade do múltiplo”.

Em outros termos, percebemos como o “nome próprio” vivencia, a princípio, uma condição de elemento preciso e objetivo, de “fácil” delimitação e formalização. Tratar-se-ia de um elemento lógico em sua natureza ou de um caso de padronização simplificável. Esta primeira impressão presente na discussão do Crátilo, clássico diálogo platônico que debate a origem dos nomes, choca-se rapidamente com a turbulenta, social e culturalmente fundada história dos “nomes próprios” e dos modos de seu uso. Esta dinâmica carece, pois, de um trato crítico da Organização do Conhecimento, responsável por pensar e por implementar formas de identificação, ordenação e acesso aos conteúdos e conteúdos, atividades estas que ora reconhecem, ora fundam, ora apagam o “nome próprio” como elemento sócio-cultural, antes de um “ponto de acesso”.

## Referências

- Abbagnano, Nicola. (1998). *Dicionário de filosofia*. 4. ed. São Paulo: Martins Fontes.
- Agamben, Giorgio. (2013). *A ideia de prosa*. Belo Horizonte: Autêntica.

- Barthes, R. (1988). *O rumor da língua*. São Paulo: Brasiliense.
- Brascher, M.; Café, L. (2008). Organização da informação ou organização do conhecimento? In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 9., São Paulo. *Anais...* São Paulo: ANCIB; USP.
- Caldeira, P. da T. (1984). A situação do Brasil em relação ao controle bibliográfico universal. *R. Esc. Bibliotecon.* UFMG, Belo Horizonte, v. 13, n. 2, p. 260-283, set.
- Chartier, R. (1999). *A aventura do livro: do leitor ao navegador*. São Paulo: Ed. da Unesp.
- Coelho, T. (1997). *Dicionário crítico de política cultural*. São Paulo: Iluminuras.
- Curtius, Ernst. (2013) *Literatura europeia e idade média latina*. São Paulo: EDUSP.
- Esteban Navarro, M.A.; García Marco, F.J. (1995). Las primeras jornadas sobre organización del conocimiento: organización del conocimiento e información científica. *Scire*, Zaragoza, v. 1, n. 1, p. 149-157. Disponível em: <<http://ibersid.eu/ojs/index.php/scire/article/view/1038/1020>>. Acesso em: 01 jun. 2011.
- Foucault, M. (2006). *O que é um autor?* 6. ed. Lisboa: Vega Passagem.
- IFLA FRANAR (2007). *Functional Requirements for authority Data*. Versão de 1 abril de 2007. Disponível em: <<http://www.ifla.org/VII/d4/Franar-ConceptualModel-2ndReview.pdf>>. Acesso em: 16 jun. 2007.
- IFLA. (1998). *Functional requirements for bibliographic records: final report*. UBCIM Publications - New Series, vol. 19. München: K. G. Saur. Disponível em: <[www.ifla.org](http://www.ifla.org)>. Acesso em: 16 ago. 2006.
- Le Goff, Jacques. (2013). *Homens e mulheres da Idade Média*. São Paulo: Estação Liberdade.
- Marcondes, Danilo. (1985). A questão das linguagens artificiais nos séculos XVII e XVIII. In: Simpósio Fluminense de Lógica, Filosofia e Teoria da Ciência, 2, 1985, Niterói. *Anais...* Niterói.
- Needham, C. D. (1964). *Organizing knowledge in libraries: an introduction to classification and cataloguing*. London: Andre Deutsch.
- Otlet, Paul. (1934). *Traité de documentation: le livre sur le livre; théorie et pratique*. Bruxelas, Editions Mundaneum.
- Platão. (1963). *Crátilo: diálogo sobre a justiça dos nomes*. Lisboa: Sá da Costa.
- \_\_\_\_\_. (2008). *A República*. 11. ed. Lisboa: Fundação Calouste Gulbenkian, 2008.
- Peignot, G. *Dictionnaire raisonné de bibliologie*, tomo I. Paris: Chez Villier, 1802a.
- \_\_\_\_\_. \_\_\_\_\_. tomo II. Paris: Chez Villier, 1802b.
- Shannon, C.; Weaver, W. (1975). *A Teoria matemática da comunicação*. São Paulo: Difel.
- Spina, S. (1994). *Introdução à edótica: crítica textual*. São Paulo: Ars Poética; Edusp.
- Strout, R. F. (1956). The development of the catalog and cataloging codes. *The Library Quarterly*, v. 26, n. 4, p. 254-275, Oct. Disponível em: <http://www.jstor.org/stable/4304573>>. Acesso em: 13 ago. 2012.
- Wittgenstein, L. (1980). *Culture and value*. Oxford: Blackwell.
- \_\_\_\_\_. (1990). *Da certeza*. Lisboa: ed. 70.
- \_\_\_\_\_. (1979). *Investigações Filosóficas*. 2. ed. São Paulo: Abril Cultural.
- \_\_\_\_\_. (1998). *Cadernos 1914-1916*. Lisboa: Edições 70..
- \_\_\_\_\_. (1992a). *O Livro azul*. Lisboa: Ed.70.
- \_\_\_\_\_. (1992b). *O Livro castanho*. Rio de Janeiro: Edições 70.
- \_\_\_\_\_. (2005). *Observações filosóficas*. São Paulo: Loyola.
- \_\_\_\_\_. (1968). *Tractatus logico-philosophicus*. São Paulo: Ed. Nacional; Edusp.
- \_\_\_\_\_. (2002). *Tratado Lógico Filosófico; Investigações filosóficas*. 3. ed. Lisboa: Fundação Calouste Gulbenkian.

**Hito 1: Modelos de organización del conocimiento. Modelaje de sistemas y ciclo de vida de la información**

**Hito 2: Entornos y sistemas para la organización del conocimiento. Diseño, normalización e interoperabilidad de sistemas de información abiertos.**

**Hito 3: Instrumentos para el acceso y transparencia de la información: gestión de documentos, plataformas digitales y archivos**

**Hito 4: Usuarios y usos de la información. Patrones vinculados a la transparencia y el gobierno abierto**

**Hito 5: Gobernanza y Auditoría de información en la organización del conocimiento.**

**Hito 6: Representación del conocimiento: procesos y herramientas en sistemas abiertos y multilingües**

**Hito 7: Epistemología en la organización del conocimiento**

**Hito 8: Igualdad, género y diversidad en la organización del conocimiento**

---

# Sesgo y universalidad: un enfoque histórico-conceptual

*Bias and Universality: a historical-conceptual approach*

---

**Stephanie Colombo (1)**

(1) Facultad de Información y Comunicación. Universidad de la República, Emilio Frugoni 1427, Montevideo, Uruguay, stephanie.colombo@fic.edu.uy

## Resumen

Dentro de la Organización del Conocimiento y más precisamente en los Sistemas de Organización y Representación del Conocimiento se puede identificar una cuestión denominada sesgo (bias). Se entiende a este como el conjunto de connotaciones de un tópico que marca una inclinación hacia una concepción ideológica. El sentido de esta connotación ha ido variando a lo largo de la historia, en algunas oportunidades con una carga negativa que es necesario eliminar y en otras con una carga positiva, útil para poder representar las necesidades de información de una comunidad. Se desarrollarán aquí los tres enfoques del sesgo de acuerdo con el tipo de carga que representan; sesgo negativo o sesgo con carga negativa, sesgo neutro o sesgo con carga neutra y sesgo positivo o sesgo con carga positiva. Por otra parte, la necesidad y pertinencia de la universalidad frente a la necesidad de enfoques más locales, está siendo discutida también en el entorno de la Organización del Conocimiento. En esta ponencia se estudiará la interacción de los términos sesgo y universalidad. Se dejarán planteadas también, dos discusiones que pueden servir para desarrollar más la literatura referente al sesgo. Una de ellas refiere a los elementos que determinan la carga del sesgo. Es decir, qué es lo que incide para que un sesgo sea positivo, negativo o neutral. Por otro lado, ¿los conceptos de sesgo y universalidad son necesariamente opuestos?, ¿no puede existir un sistema que sea universal y sesgado a la vez?

**Palabras clave:** Sesgo. Universalidad. Sistemas de organización y representación del conocimiento. Sesgo positivo. Sesgo negativo. Sesgo neutral.

## 1. Introducción

Dentro de la Organización del Conocimiento y más precisamente en los Sistemas de Organización y Representación del Conocimiento se puede identificar los concepto de sesgo y universalidad.

El primero de ellos, el sesgo, más reconocido por su término en inglés *bias*, representa una inclinación o una postura frente a determinada ideología o enfoque cultural y en algunos casos

## Abstract

Within the Knowledge Organization and more precisely in the Knowledge and Representation Organization Systems it is possible to identify an issue called bias. It is understood as the set of topic's connotations that marks an inclination towards an ideological conception. The meaning of this connotation has varied throughout history, at times with a negative charge to be removed and others with a positive charge, useful to represent the information needs of a community. The three approaches of bias based on the type of charge will be developed here; negative bias or bias with a negative charge, neutral bias or bias with a neutral charge and positive bias or bias with a positive charge. Moreover, the need and relevance of universality versus the need for more local approaches is also being discussed in Knowledge Organization. In this paper the interaction of the terms bias and universality will be considered. It also be presented two discussion that can be used to develop the bias's literature. One concerns which elements determine the positive, negative or neutral charge of bias. On the other hand, are the concepts of bias and universality necessarily opposites? Is it possible to create a system that is universal and biased at the same time?

**Keywords:** Bias. Universality. Knowledge and Representation Organization Systems. Positive bias. Negative bias. Neutral bias.

el tratamiento de un tema desde otro enfoque disciplinario.

El sesgo se manifiesta en la estructura de los Sistemas de Organización del Conocimiento ya sea por la disposición de sus términos, las relaciones entre ellos, la elección (o incluso la omisión) de los mismos. De esta forma, se puede decir que un sistema es sesgado cuando en su estructura prevalece cierto o ciertos enfoque por sobre otro u otros.

Por otro lado, el concepto de universalidad tiene como base semántica abarcar todo el conocimiento producido. De este modo, un sistema de organización del conocimiento universal, es aquel cuya estructura permite representar todo el conocimiento existente al momento de ser publicado.

A pesar de ser términos fuertemente implantados en la Organización del Conocimiento, es complejo encontrar definiciones precisas de ambos términos. A continuación se estudiarán ambos conceptos y cómo se vinculan en la Organización del Conocimiento.

## 2. Sesgo

El concepto de sesgo no es nuevo en la literatura del área, pero sí son relativamente nuevos, en comparación con la historia de la Organización del Conocimiento, algunos de sus enfoques.

Se pueden identificar tres enfoques del concepto de sesgo de acuerdo con el tipo de carga que representan; ellos son el sesgo negativo o sesgo con carga negativa, el sesgo neutro o sesgo con carga neutra y el sesgo positivo o sesgo con carga positiva.

Entre los trabajos que se enfocan en el sesgo con carga positiva y/o negativa se pueden citar a Olson y Schlegl (1999), Feinberg (2007), Hjørland (2008), Mai (2010), Milani y Guimarães (2011), Barité y Colombo (2013), Milani (2014), entre otros.

Mientras que los trabajos que suelen recoger el estudio del sesgo con carga neutra son principalmente producidos en torno a la obra de Ranganathan. Por citar algunos, Ranganathan (1958, 1960), *Indian Standard Institution* (1963), Satija (2011), entre otros.

Dentro de las definiciones que se centran en la carga positiva y negativa se encuentran:

Connotación de un tópico (término, descriptor, palabra clave, etc.) que marca una tendencia, una inclinación o la expresión de una concepción ideológica. Puede manifestarse en la terminología, en la estructura o en las relaciones conceptuales establecidas en un sistema de organización del conocimiento. (Barité, 2015).

El glosario de la *American Library Association* (ALA) define sesgo como

[t]endencia de una estimación a desviarse en una dirección por la media de un conjunto de valores a partir de un valor verdadero. (Young, 1988, p. 31, traducción propia)

Mientras que las centradas en la carga neutra lo definen como:

“Una relación de fase entre dos términos indicando que un tema ha sido tratado como fondo para aquellos usuarios que son especialistas en otro campo” (Terminology of documentation, 1976, p. 123, traducción propia)

“En clasificación, donde un tema es descrito (usualmente en una manera relativamente elemental) en beneficio de los que trabajan en, o se relaciona con otro campo, un documento es clasificado bajo el tema presentado, no bajo la persona para la que es escrito, por ej. anatomía para fonaudiólogos iría bajo anatomía y no bajo fonaudiólogos” (Harrod, 1990, p. 57, traducción propia)

Por su parte, el diccionario de la lengua general define sesgo como:

“1. Torcido, cortado o situado oblicuamente. [...] 4. Oblicuidad o torcimiento de una cosa hacia un lado, o en el corte, o en la situación, o en el movimiento.” (Real Academia Española, 2014, p. 2003).

### 2.1. Sesgo neutro

El concepto de sesgo neutro o sesgo con carga neutra fue incorporado a la disciplina por S.R. Ranganathan. Dicho enfoque de sesgo representa una herramienta en un sistema de clasificación, aunque puede ser extrapolado a cualquier sistema de organización y representación del conocimiento.

El mismo se utiliza para representar una relación dentro de las relaciones de fase (*phase relation*) donde un tema es tratado desde la perspectiva de un especialista de otro tema.

En palabras de Ranganathan una relación sesgada (*bias relation*)

“indica que la exposición de la primera fase es sesgada hacia la segunda fase, esto es, la exposición está especialmente en sintonía ya sea por la selección o el arreglo de los tópicos o el énfasis o el modelo o los ejemplos u otro medio hacia las necesidades de un especialista en el campo formado por la segunda fase” (Ranganathan, 1960, p. 1.56, traducción propia).

Por ejemplo, cuando en una relación sesgada, la primera fase es estadística y la segunda fase es ingenieros, implica que el tema estadística se encuentra tratado desde la perspectiva o para usos de un ingeniero.

Este enfoque es denominado neutro, ya que no posee ningún tipo de connotación, ni positiva, ni negativa. Se utiliza para representar perspectivas disciplinarias. Este sesgo se puede encontrar en las obras a procesar y se representa a través de los sistemas de organización del conocimiento. A diferencia de los enfoques desarrollados a continuación, este tipo de carga de sesgo no se percibe en las estructuras de los sistemas de organización del conocimiento, sino

que debe ser creado por el profesional de la información a partir de ellos, para representar de forma más pertinente el tratamiento de los temas.

## 2.2. Sesgo negativo y sesgo positivo

Por otro lado, se encuentra el sesgo negativo o con carga negativa. Esta carga de sesgo es la predominante en la literatura de Organización del Conocimiento. Este tipo de sesgo ocurre cuando, según la disposición de los términos, las relaciones entre ellos o incluso la elección de los términos a ser incluidos en un sistema de organización del conocimiento realza, predomina o deja invisible determinada perspectiva por sobre otra. A su vez, el desarrollo de los temas no es útil para representar el conocimiento de la colección a la que se aplica el sistema de organización del conocimiento y tampoco refleja las necesidades de información de la comunidad de usuarios de dicha colección. Este tipo de enfoque se lo suele también vincular con prejuicios.

En la producción científica del área, es más común identificar trabajos sobre sesgo negativo con la última característica mencionada. Es decir, relacionar la negatividad del sesgo con tendencias más allá de evaluar la utilidad del sistema.

Por último, pero no menos importante, se encuentra el sesgo positivo o sesgo con carga positiva. Este tipo de sesgo es generalmente creado de forma activa para reflejar las características culturales y/o ideológicas de una comunidad.

En oposición, el sesgo negativo no tiende a ser creado de forma consciente. Sino que por el contrario, refleja la forma de pensar tanto de los autores (cuando el sesgo se encuentra en los documentos) como de los profesionales de la información (al momento de crear los sistemas de organización del conocimiento e incluso al aplicarlos). Estos, como seres culturales insertos en un espacio y tiempo, conllevan ciertos sesgos que pueden terminar siendo plasmados al momento de crear un documento o un sistema o de ubicar un documento dentro de una colección.

Para Olson desde el momento que se elige un nombre por sobre otro se está generando sesgo.

“La elección de los términos para [los] conceptos a menudo introducen un sesgo evidente o, más comúnmente, [un sesgo] sutil, marginaciones insidiosas.” (Olson, 2002, p.6)

Por consiguiente, tanto la carga negativa como la carga positiva, tienen íntima relación con la

utilidad para representar los temas de una colección y las necesidades de información de su comunidad de usuarios. De esta forma y en resumen, la característica de negativo o positivo del término también se vincula con la utilidad del sistema de organización del conocimiento para representar los temas de determinada colección y las necesidades de información de sus usuarios.

Sin embargo, Broughton mencionaba con respecto al sesgo cultural y político de la Lista de Encabezamientos de la Biblioteca del Congreso

“esos encabezamientos solo vienen a existir sobre la base de necesidad. Estos encabezamientos en particular existen para adaptarse a una literatura producida antes de que tales actitudes se convirtieran en inaceptables, y, debido a que la literatura seguirá existiendo, también lo harán los encabezamientos de materia ofensivos. Es un problema del material más que del sistema utilizado para organizarlo.” (Broughton, 2012, p.13, traducción propia)

Con esta contribución se deja aquí planteada la discusión de si el sesgo es negativo porque es ofensivo o negativo porque no sirve para reflejar determinada realidad. En el caso planteado por Broughton, esos encabezamientos tendrían un sesgo negativo desde el momento que son ofensivos, pero también tendrían un sesgo positivo ya que fueron creados activamente para relejar la realidad de los documentos y son útiles para representar los temas dentro de los documentos.

## 3. Universalidad

Se puede rastrear el concepto de universalidad desde la época del Iluminismo y la creación de la Enciclopedia. Esta obra tenía como fin, reunir todos los saberes del mundo. Desde entonces, la idea de universalidad fue la idea predominante en la Organización del Conocimiento. Ya decía Olson que, “[e]n Organización del Conocimiento, la universalidad es el concepto dominante” (Olson, 1994, p. 73, traducción propia).

Actualmente, se puede percibir una discusión entre la importancia de sistemas universales que puedan ser aplicados en todo el mundo, frente a sistemas locales, los cuales se aplican en determinadas comunidades.

A pesar de ello, no suele encontrarse definición alguna en diccionarios especializados del área. Sino que por el contrario, el término universalidad suele aparecer de forma adjetivada. De este modo, no se define *universalidad* pero sí *alfabetización universal* (Martínez de Souza, 2004), *alfabeto universal* (Martínez de Souza, 2004), *bibliografía universal* (Massa de Gil, 1971; Ter-

minology of documentation, 1976b; López Yepes, 2004), *biblioteca universal* (Terminology of documentation, 1976b), *clasificación universal* (Terminology of documentation, 1976b), *diccionario universal* (Martínez de Souza, 2004), *enciclopedia universal* (Martínez de Souza, 2004), *sistema universal* (Barité, 2015), entre otros. En todos ellos, subyace la idea de reunir toda la producción existente, es decir, alcanzar la mayor exhaustividad posible.

Este concepto es consistente con el utilizado en el lenguaje general. El Diccionario de la Real Academia Española define universalidad como "cualidad de universal" (Real Academia Española, 2014, p. 2197). Mientras que bajo universal se encuentran las siguientes acepciones

"2. Que comprende o es común a todos en su especie, sin excepción de ninguno. 3. Que lo comprende todo en la especie de que se habla. 4. Que pertenece o se extiende a todo el mundo, a todos los países, a todos los tiempos." (Real Academia Española, 2014, p. 2197)

#### 4. Discusión

En la literatura reciente se tiene en buena estima, la idea de generar un sesgo positivo, en el sentido de que pueda representar mejor las necesidades de información de un sector de la sociedad.

Para Feinberg (2007, traducción propia) la "elaboración de un sistema universal para todos los documentos es un error; en cambio, se necesita analizar e interpretar como los diferentes grupos definen y describen su información."

Sin embargo, Olson (1994) sostiene que a pesar de que el nuevo enfoque cognitivo reconoce las diferencias entre individuos y aún más entre comunidades, no se abandona la idea de universalidad.

La autora menciona que "una posible razón es que la presunción de la necesidad de un lenguaje universal que nos dejó Cutter crea expectativas de un sistema capaz de incluir todo y seguimos operando sobre la base de esa presunción. Como resultado de ello, buscamos un medio universal de expresión, rechazando la diversidad del lenguaje." (Olson, 2002, p. 143, traducción propia).

De este modo, y siguiendo esta línea de pensamiento, no habría lugar para un sistema con sesgo positivo que pudiera reflejar diversidad.

A raíz de esto, se podría decir que ambos conceptos no pueden coexistir en un mismo sistema de organización del conocimiento. Feinberg (2007) sostiene que la presencia del sesgo pone en tela de juicio la creencia en sistemas universales. Por tanto, un sistema de organización del

conocimiento no podría ser al mismo tiempo universal y sesgado.

Sin embargo, ambos conceptos no son necesariamente opuestos. Nada impide que exista un sistema que aspire a la universalidad, y que a su vez sea sesgado. Es decir, que exponga dicha universalidad desde la perspectiva de una ideología o cultura determinada. No se debe confundir el concepto de universalidad con el de objetividad.

Mientras que el sesgo no aparezca en su forma de omisión, no hay elementos que limiten desarrollar un sistema universal sesgado.

#### 5. Conclusiones

A través de lo aquí expuesto se puede observar que el concepto de sesgo todavía se encuentra en debate dentro de la organización del conocimiento. Por otro lado, el concepto de universalidad se encuentra más establecido, aunque es complejo encontrar una definición concreta del mismo.

Se puede observar que en las últimas décadas, y quizás en consecuencia también al auge de los estudios culturales, se ha comenzado a cuestionar la pertinencia de los sistemas universales. De esta forma, se prefieren los sistemas del conocimiento más focales, que representen determinada comunidad frente a sistemas más generales.

Se dejan planteadas aquí dos discusiones que pueden servir para desarrollar la literatura en esta área. Por un lado, ¿qué determina el enfoque del sesgo? Es decir, el carácter negativo o positivo del sesgo, está dado por lo ofensivo que pudiera resultar el sistema para sus usuarios o el punto se encuentra en el grado de utilidad para presentar determinadas posturas. Pareciera ser, que el sesgo es negativo cuando es ofensivo, pero es positivo cuando ayuda a reflejar cierta realidad.

Por otro lado, ¿los conceptos de sesgo y universalidad son opuestos? No es posible generar un sistema que sea a la vez universal y sesgado. El concepto de universalidad, tiene necesariamente intrínseco el concepto de objetividad.

#### Referencias

- Barité, M. (2015). *Diccionario de Organización del Conocimiento: Clasificación, Indización, Terminología*. Montevideo: CSIC; Prodic. [No publicado].
- Barité, M; Colombo, S. (2013). Garantía Cultural y Sesgo Positivo: Aproximación Teórico-Metodológica. En: *Informação e/ou Conhecimento: as duas faces de Jano. I Congresso ISKO Espanha e Portugal / XI Congresso ISKO España*. Porto: Faculdade de Letras da Universidade do Porto, 1381-1394. Recuperado 20-05-2015, de

- <http://www.youblisher.com/p/749221-I-Congresso-ISKO-Espanha-e-Portugal-XI-Congreso-ISKO-Espana/>.
- Broughton, V. (2012). *Essential Library of Congress Subject Headings*. Londres: Facet Publishing.
- Feinberg, M. (2007). Hidden bias to responsible bias: an approach to information systems based on Haraway's situated knowledges. *Information Research*. 12(4). Recuperado 20-05-2015, de <http://www.informationr.net/ir/12-4/colis07.html>.
- Harrod, LM. (1990). Bias phase. En: Prytherch, R. (Comp.). *Harrod's librarians glossary of terms used in librarianship, documentation and the book crafts and reference book*. 7th ed. Aldershot: Gower, 56.
- Hjørland, B. (2008). Deliberate bias in Knowledge Organization? En: Arsenault, C.; Tennis, JT. (eds.). *Cultural and Identity in Knowledge Organization*. Würzburg: Ergon. (Advances in Knowledge Organization, 11), 256-261.
- Indian Standards Institution. (1963). *Indian Standard: IS:2550-1963: Glossary of classification terms*. New Delhi: Indian Standards Institution.
- López Yepes, J. (ed.). (2004). *Diccionario enciclopédico de ciencias de la documentación*. Madrid: Síntesis.
- Mai, J.-E. (2010). Classification in a social world: bias and trust. *Journal of Documentation*. 66(5) 627-642.
- Martínez de Sousa, J. (2004). *Diccionario de bibliología y ciencias afines*. 3a. ed. Gijón: TREA.
- Massa de Gil, B. (1971). *Dizionario tecnico di biblioteconomia: italiano, spagnolo, inglese*. México: Trillas.
- Milani, SO. (2014). *Bias na representação de assunto: Uma Discussão de Oposições Binárias nos Funcional Requirements for Subject Authority Data (FRSAD)* [Tesis de doctorado]. Marília: Faculdade de Filosofia e Ciências, Universidade Estadual Paulista.
- Milani, SO; Guimarães, JAC. (2011). Biases in knowledge representation: an analysis of the feminine domain in Brazilian indexing languages. En: Smiraglia, RP. (ed.). *Proceedings from North American Symposium on Knowledge Organization*, 94-104. Recuperado 20-05-2015, de <http://journals.lib.washington.edu/index.php/nasko/article/download/12794/11275>.
- Olson, HA. (2002). *The power to name: locating the limits of subject representation in libraries*. Dordrecht: Kluwer
- Olson, HA.; SCHLEGL, R. (1999). Bias in subject access standards: a content analysis of the critical literature. En: *Information Science: Where Has It Been, Where Is It Going?*; June 9 - 11, 1999; Sherbrooke, Quebec. Montreal: CASI, 236 -247. Recuperado 20-05-2015, de [http://www.caais-acs.ca/proceedings/1999/Olson\\_1999.pdf](http://www.caais-acs.ca/proceedings/1999/Olson_1999.pdf).
- Ranganathan, SR. (1958). Library Classification Glossary. *Annals of Library Science*. 5(3) 65-112.
- Ranganathan, SR. (1960). *Colon Classification: Basic Classification*. 6a ed. Bangalore: Sarada Ranganathan Endowment for Library Science.
- Real Academia Española. (2014). *Diccionario de la lengua española*. 23a ed. Barcelona: Espasa.
- Satija, MP. (2011). *A guide to the theory and practice of Colon Classification*. New Delhi: Ess Publications.
- Terminology of documentation (1976) =Terminologie de la documentation = Terminología de la documentación. Gernot, W.; Ulrich, N. (comp.). Paris: Unesco.
- Young, H. (ed.). (1988). *Glosario ALA de bibliotecología y ciencias de la información*. Madrid: Díaz de Santos.

---

# Interações entre Semiótica da Cultura e Organização do Conhecimento: conceitos integradores

*Interactions between semiotics culture and knowledge organization: integrating concepts*

---

**Mona Cleide Quirino da Silva FARIAS (1), Carlos Cândido de ALMEIDA (2)**

(1) Doutoranda no Programa de Pós-Graduação em Ciência da Informação – PPGCI, Grupo de Pesquisa Fundamentos Teóricos da Informação – GPFTI, Faculdade de Filosofia e Ciências - UNESP, Av. Hygundo Muzzi Filho, 737 – Marília (17.525-900) - SP – BRASIL. monaquirino@gmail.com / monaquirino@marilia.unesp.br

(2) Docente do Programa de Pós-Graduação em Ciência da Informação – PPGCI, Grupo de Pesquisa Fundamentos Teóricos da Informação – GPFTI, Faculdade de Filosofia e Ciências - UNESP, Av. Hygundo Muzzi Filho, 737 – Marília (17.525-900) - SP - BRASIL. carlosalmeida@marilia.unesp.br

## Resumo

A linguagem é uma expressão do conhecimento pela qual podemos estabelecer diálogos nos mais variados contextos. Entendemos o conceito de linguagem a partir da Semiótica da Cultura, teoria russa que estuda a linguagem entre outros sistemas de signos. A linguagem é um fenômeno articulador que condiciona relações transculturais e no âmbito das abordagens socioculturais da Organização do Conhecimento, como núcleo de investigações que ressaltam discussões com a necessidade de compreender a cultura e o contexto das comunidades discussivas, a linguagem pode favorecer a projeção de novos conhecimentos. Desse núcleo destacamos Antônio García Gutiérrez, Clare Beghtol e Michèle Hudon. Na Semiótica da Cultura destacamos Yuri M. Lotman, Irene Machado, Schnaiderman. Assim, investigamos os conceitos das abordagens socioculturais da Organização do Conhecimento por uma perspectiva teórico-conceitual da Semiótica da Cultura. Como resultado encontramos as relações teóricas entre os conceitos cultura, linguagem, texto, contexto, tradução e sistemas modelizantes e os conceitos fundamentais à organização do conhecimento: transculturalismo, ética transcultural, garantia cultural, hospitalidade e multilinguismo. Concluímos que os conceitos investigados com base na Semiótica da Cultura revelam à Organização do Conhecimento relações interdisciplinares no âmbito das abordagens socioculturais.

**Palavras chave:** Semiótica da Cultura. Organização do Conhecimento. Linguagem. Abordagens Socioculturais.

## 1. Introdução

A linguagem é uma expressão do conhecimento e nesse contexto entendemos o conhecimento como uma construção coletiva que reflete uma determinada cultura, por isso o consideramos produto da sociedade. Contudo, é necessário

## Abstract

Language is an expression of knowledge in which we can establish dialogues in several contexts. The concept of language is this study understood from the Semiotics of Culture, a Russian theory that studies the language and other sign systems. Language is an articulating phenomenon that determines cross cultural relations and in the context of socio-cultural approaches to Knowledge Organization, as nucleus of research that highlight discussions with the necessity of understanding the culture and the context of discussive communities, the language can promote the projection of new knowledge. From that nucleus we highlight Antonio García Gutiérrez, Clare Beghtol and Michèle Hudon. In Semiotics of Culture we highlight Yuri M. Lotman, Irene Machado, Schnaiderman. Thus, we investigated the concepts of sociocultural approaches in Knowledge Organization by a theoretical and conceptual perspective of Semiotics of Culture. As a result we find the theoretical relationships between the concepts culture, language, text, context, translation and modeling systems and the key concepts of knowledge organization: transculturalism, transcultural ethics, cultural guarantee, hospitality and multilingualism. We conclude that the concepts investigated based on the Cultural Semiotics reveal the Organization of Knowledge interdisciplinary relationship in socio-cultural approaches.

**Keywords:** Semiotics of Culture. Knowledge Organization. Language. Socio-cultural approaches.

reconhecer que, a dinâmica da cultura envolve diversas linguagens e elementos culturais, e o conhecimento retrata uma hibridéz cultural e linguística, tendo em vista os conteúdos construídos e compartilhados em diversas comunidades.

Neste estudo, consideramos a Organização do Conhecimento como um campo de mediações que nos revela uma diversidade de temas no âmbito da área da Ciência da Informação. Para Dahlberg (2006), compete à Organização do Conhecimento tratar de aspectos voltados à teoria do conceito. Contudo, também podemos citar outros temas que figuram nos estudos da Organização do Conhecimento, a saber: análise e discussões sobre linguagens documentárias e terminologias (Lara, 2011), teorias semânticas, análise conceitual (Brascher e Café, 2011), teoria da classificação, entre outros.

Por abordagens socioculturais entendemos um dos núcleos de estudo da Organização do Conhecimento que contemplam investigações, as quais oferecem uma leitura crítica sobre a cultura e seu contexto social. Nesta linha de estudos, podemos citar teóricos como García Gutiérrez (1998, 2002), Beghtol (2002), Guimarães e Fernández-Molina (2003), Hudon (1997), Hjørland (2003, 2007, 2013) e Campbell (2000, 2010).

Consideramos que as abordagens socioculturais no contexto da Organização do Conhecimento fornecem um contributo para a análise das questões éticas, linguísticas, políticas e sociais que se direcionam à configuração de sistemas de organização e representação do conhecimento. Também entendemos que as reflexões trazidas pelos estudos da Semiótica como “[...] ciência que estuda o signo em geral; todos os signos que formem linguagens ou sistemas” (Beuchot, 2004, p. 7) são fundamentais à Organização do Conhecimento.

No entanto, a Semiótica possui diversas correntes. As mais conhecidas, por vezes entendidas como principais são a Semiótica Peirceana e a Semiologia Estruturalista. A concepção de signo de Peirce (1990, p.4), afirma que o signo é qualquer coisa que conduz a alguma outra coisa (seu interpretante). Nesta linha, Coelho Netto (2010, p.56) destaca que signo é aquilo que sob determinado aspecto representa algo para alguém. Outra concepção de signo foi atribuída por Saussure (1970, p. 81), quando explicou que o signo linguístico constitui-se de uma unidade psíquica de duas faces: a imagem acústica e o conceito, mais propriamente, o significante e o significado. Não obstante, as diversas nuances destas correntes podem não compreender todo o espectro de temas semióticos.

Nesse estudo, destacamos a Semiótica da Cultura, conhecida também como Semiótica Russa, a qual desenvolveu discussões sobre os sistemas de signos (cultura, texto cultural, códigos culturais e linguagens), em meados da década de 1960. Um dos teóricos que destacamos é

Yuri M. Lotman (1990, 1999, 2002, 2003a, 2003b, 2011 e 2013), com a análise da semiótica, isto é, um espaço de desenvolvimento dos sistemas de signos da cultura, um ambiente semiótico sujeito a processos de modelizações e transcódificações.

De antemão, inferimos que os conceitos da Semiótica da Cultura podem inserir questões importantes nas discussões sobre linguagem e cultura, as quais contemplam problemas de ordem dos aspectos multiculturais da linguagem. Para tanto, a análise do conhecimento de uma comunidade discursiva deve rever as condições éticas da representação do conhecimento, a matriz multicultural da sociedade, o respeito às diferenças e a garantia linguística de uma comunidade. Logo, a análise semântica dos conceitos, bem como as categorias utilizadas para organizar o conhecimento, devem revelar estas particularidades culturais. Nesse sentido, entendemos que linguagem é um modo de expressão, que pode ser representada por diversos signos que não seja somente o signo linguístico (Machado, 2003, p. 163). Por isso, as reflexões produzidas na Semiótica da Cultura devem ser tomadas como objeto pela Organização do Conhecimento.

Entendemos que tais discussões podem se somar às discussões sobre linguagem e cultura que potencializem e contribuam com os estudos contemplados pelas abordagens socioculturais, a saber, as questões que envolvem a linguagem, os aspectos éticos e socioculturais da Organização do Conhecimento. Para tanto, objetivamos delinear os diálogos conceituais possíveis entre as abordagens socioculturais e a perspectiva teórica da Semiótica da Cultura.

## **2. Abordagens socioculturais da Organização do Conhecimento**

As abordagens socioculturais representam investigações que ressaltam a necessidade em compreender a cultura e o contexto das comunidades discursivas. Neste estudo, destacamos as abordagens de Beghtol (2002), Hudon (1997, 1999) e García Gutiérrez (1998, 2002a, 2002b, 2002c, 2004 e 2008), cujo propósito foi apresentar os temas garantia, hospitalidade cultural, multilinguismo e ética transcultural da mediação. Vale destacar que há diversos autores como Barité (2001, 2011), López-Huertas (2006, 2008 e 2010), Smiraglia (2002), que colaboram com análises e discussões socioculturais. Porém, decidimos enfatizar as ideias dos autores citados anteriormente pelo fato de identificarmos uma aproximação entre as suas investigações e a Semiótica da Cultura. Assim, objetivamos

nesta seção destacar que as abordagens socio-culturais desencadeiam questões de ordem antropológica, ética e multilíngue, portanto, são consideradas relevantes no contexto desta análise. Desse modo, entendemos que tais investigações podem auxiliar a reflexão sobre os processos de organização e representação do conhecimento sob uma vertente sociocultural. A seguir, introduzimos as abordagens conhecidas como “Ética transcultural da mediação”, “Garantia Cultural” e “Multilinguismo”.

### 2.1. García Gutiérrez e a ética transcultural da mediação

A ética transcultural da mediação é um tratamento transcultural à representação do conhecimento. Ela busca refletir a complexidade existente nas culturas e permite considerar a diversidade de elementos culturais que de forma direta ou indireta estão imbuídos em processos de representação do conhecimento. Guimarães et al. (2005) compreendem a ética transcultural como um fundamento, não somente implica uma desconstrução do campo teórico da Organização do Conhecimento, mas enuncia o surgimento de um novo paradigma.

A abordagem de uma ética transcultural encontra-se relacionada com a temática da “epistemografia interativa”, também discutida por Antônio García Gutiérrez. A epistemografia interativa consiste em uma proposta de abordagem crítica que se opõe à epistemologia em seu caráter tradicional, a qual destaca o conhecimento em um plano ordenado e elitista como entendido pelo autor. Nesse sentido, a epistemografia interativa implica em um novo paradigma cujo propósito é introduzir reflexões que abarquem uma dimensão ética, política e sociocultural no conhecimento à sua organização (García Gutiérrez, 1998, 2002, 2004, 2006).

A ética transcultural da mediação busca também introduzir na prática do profissional mediador uma visão alargada, isto é, transcultural sobre questões de ordem sociocultural, o que implica em uma consciência frente ao que será representado, sem deixar escapar uma ética sobre os aspectos que caracterizam a pluralidade cultural em diversas comunidades.

A discussão da epistemografia interativa retoma também às Linguagens Epistemográficas de García Gutiérrez (1998). As linguagens epistemográficas são consideradas linguagens associativas cujo objetivo é a representação plural de elementos como culturas e línguas diversas. Nesse sentido, García Gutiérrez (2004) considera que a linguagem codifica o pensamento e isso implica considerarmos que o organizador

do conhecimento não escapa aos mecanismos estéticos, emotivos, passionais e cognitivos que venham a se impor sobre sua formação assim como de suas práticas.

García Gutiérrez (1998) destaca que as linguagens epistemográficas possuem duas funções: a normalização do vocabulário e a sugestão de alternativas na representação de informações. Vale ressaltar que o nome linguagem epistemográfica foi cuidadosamente escolhido para representar todo esse mundo de convergência. Por um lado, inspirada pelo adjetivo expressivo de 'Epistemologia prática' de Gardin, que reflete tanto seus métodos de construção assim como suas áreas de aplicação.

Nesse sentido, entendemos que as linguagens epistemográficas consistem em instrumentos de representações cognitivas processadas a partir de construções lógico-semânticas e discursivas. Também compartilham de uma base léxica se aproximando da linguagem natural, o que propõe superar a redução da língua e ainda permitir o uso de substantivos, adjetivos etc., em busca de introduzir no instrumento de representação um caráter transcultural que garanta o respeito frente às questões socioculturais como as de caráter étnico, de gênero entre outras. No entanto, em um processo de organização e representação do conhecimento o profissional, mediador, não é um sujeito omissor, e por isso no domínio de suas decisões e escolhas para a representação este carrega consigo valores, crenças, os quais estão imbuídos na própria linguagem natural da qual os profissionais possuem.

Consideramos que a ética transcultural da mediação relaciona-se a um tratamento transcultural à representação de um determinado conhecimento, pelo qual devemos considerar a diversidade de elementos complexos (éticos, políticos, linguísticos, históricos e socioculturais) que estão imbuídos nas culturas/contextos das comunidades/usuários e, dessa maneira, necessitam ser contemplados na representação do conhecimento. Desse modo, não podemos deixar de reconhecer que na configuração de um determinado sistema de organização e representação é necessária atenção às questões culturais, ou melhor, as diferenças culturais. Conforme Smith (2001, p. 534), “[...] as diferenças culturais formam os modos com os quais as pessoas se relacionam com a informação e o seu papel na sociedade.” As diferenças culturais evidenciam cada vez mais a dimensão cultural que os indivíduos em suas comunidades e contextos possuem e, dessa maneira, estabelecem interações.

Vale destacar que a ética transcultural surge como uma proposta teórica no contexto da Organização do Conhecimento, tendo em vista a compreensão de cultura. Partindo do olhar sobre a cultura é que chegamos a noção de transculturalismo conceito mencionado nas abordagens da ética transcultural da mediação de García Gutiérrez, com a qual é possível analisar a transculturalidade em diversas culturas, o que permite entender a cultura como um sistema aberto, dialógico e dinâmico, propondo uma análise crítica sobre a ideia de multiculturalismo. Na esteira dessa discussão e refletindo dessa forma sobre os sistemas de organização e representação do conhecimento, García Gutiérrez propõe um conceito que supere a noção de multiculturalismo que defenda a necessidade de uma mudança de paradigma proposta por meio do conceito de transculturalismo.

A ética transcultural da mediação, no âmbito da Organização do Conhecimento e das abordagens socioculturais, busca compreender a cultura em um plano mais abrangente, sem priorizar particularidades culturais, pois a cultura é entendida como um campo híbrido e destaca-se por suas peculiaridades simbólicas.

## 2.2. Clare Beghtol e a garantia cultural

A concepção de garantia cultural proposta por Clare Beghtol manifesta uma preocupação de caráter ético sobre os sistemas de classificação e representação do conhecimento. Implica uma reflexão crítica sobre os conteúdos informacionais representados, ou aqueles sujeitos à representação. Destacamos que o termo garantia cultural deriva de garantia literária, tal como destacado por Hulme (1911). Nesse sentido, a garantia cultural requer que cada sistema de classificação como, por exemplo, CDU, represente certa cultura, independente de que esta seja a cultura de um país, ou de uma etnia, um grupo, um domínio de artes, um partido político, uma religião e/ou língua (Beghtol, 1986, 2001).

A garantia cultural prevê fomentar o acesso de informações, e a hospitalidade cultural está para o cuidado com a diversidade cultural e lingüística, tendo em vista a recepção de informações sobre diversas comunidades. O conceito de hospitalidade estabeleceu-se como uma das capacidades que possui uma notação de admitir novos conceitos de forma adequada e para acomodá-los nas relações corretas com outros conceitos. Nesse contexto, a hospitalidade cultural busca garantir às diversas culturas, em respeito de suas particularidades, éticas, lingüísticas e simbólicas. Para tanto, é fundamental uma atuação ética por parte do profissional

da informação, tendo em vista que a noção de garantia e hospitalidade cultural deve prezar pelo respeito a todos os grupos étnicos e lingüísticos.

Ademais, o cuidado para com a integração de diversidades de culturas demonstra o quão a proposta de garantia cultural pode tornar complexas as discussões no âmbito dos processos e sistemas de organização e representação do conhecimento. Desse modo, Beghtol (2002) afirma que representar e organizar não consistem em processos simples. Nesse sentido, a garantia cultural está voltada necessariamente às questões do acesso do conteúdo. E tal acesso parte do reconhecimento das diversas culturas mediante procedimentos de representação que primem por um tratamento ético dos conteúdos culturais.

Outros conceitos são importantes à discussão da garantia cultural, por essa razão, Beghtol estende à análise de determinados conceitos como, por exemplo, globalização e cultura. Sobre a globalização, a autora sustenta que este conceito retoma a questão das tecnologias de informação, assim como seu desenvolvimento e velocidade em regimes capitalistas. A cultura supõe uma análise transdisciplinar que parte de uma diversidade de elementos que retratam a história, os valores, as línguas, as práticas sociais, a memória coletiva, etc. Nesse sentido, Beghtol apresenta a sua preocupação não apenas com o conceito de cultura, mas com a significado que este sugere quando do tratamento deste e de demais elementos que devem ser considerados no momento de uma representação.

A noção de garantia cultural se torna relevante por provocar reflexões em torno das questões socioculturais. Desse modo, não podemos pensar em representação sem antes reconhecermos o pressuposto da diversidade de culturas, das linguagens e dos contextos das comunidades. A cultura possui peculiaridades simbólicas, lingüísticas, valores que compartilham de crenças, visto que é uma dinâmica transformadora, ao mesmo tempo em que cria conhecimentos, transforma-os. Para tanto, devemos atentar para estas características ao buscar estabelecer sistemas de organização e representação do conhecimento.

Nesse escopo, entendemos que a garantia cultural permite que os indivíduos pertencentes a culturas distintas tenham acesso a informações diversas. Por meio dessa perspectiva de integração entendemos que toda cultura é constituída por diversos elementos e estes devem ser reconhecidos principalmente quando buscamos

organizar e representar as culturas das comunidades.

### 2.3. Michèle Hudon e o multilinguismo

O conceito de multilinguismo destaca as reflexões dos tesouros multilíngues propostas nos estudos de Michèle Hudon. Para Ducrot e Todorov (1988, p.67) “Diz-se que um indivíduo é multilíngüe (bi, trilingüe...) se domina várias línguas [...]”. Nesse sentido, o multilinguismo destaca a diversidade de línguas e de culturas que devem ser reconhecidos nos processos de representação.

Hudon (1999) considera a língua como um dos fatores que influenciam as propostas científicas, culturais e de negócios. Nesse contexto, há a predominância da língua correspondente de uma nação dominante. Nesse sentido, Michèle Hudon discorre sobre a necessidade de instrumentos de representação como os tesouros multilíngues que possam auxiliar o processo de indexação, permitindo que os documentos sejam indexados em várias línguas, além do documento original. Com isso, Hudon defende que o tesouro multilíngue passa a atuar como um instrumento que propiciará a relação entre culturas, resultando na comunicação interlinguística (Hudon, 1997, p. 85).

Hudon (1997, p.86) enfatiza que para a elaboração de um tesouro multilíngue, é necessário um tratamento igualitário das línguas, pois este instrumento reflete o universo de conceitos como também de termos utilizados em cada cultura e língua representadas. Desse modo, entendemos que, na elaboração de um tesouro multilíngue, deve-se considerar as múltiplas línguas para que se democratize o acesso aos conteúdos informacionais. De fato não é uma tarefa simples, pois a diversidade de línguas existentes, embora se aproximem em alguns aspectos, carregam em si uma particularidade de signos que são necessários à comunicação, e que consideram também uma diversidade cultural. Consideramos que, conhecer os instrumentos e elaborar os sistemas que organizem e representem o conhecimento é uma atribuição do profissional que agrega valor ao campo da Organização do Conhecimento no contexto da Ciência da Informação.

Hudon (1997, p. 84) ressalta que há uma dimensão cultural definida para o processo de elaboração dos tesouros multilíngues. Nesse sentido, os tesouros podem posteriormente tornar-se mais apropriados para se referir a tesouros multiculturais; ao invés de tesouros multilíngues. Há também uma dimensão política para a construção de tesouros multilíngues,

especialmente ao lidar com línguas que não se apresentam contextualmente, com o mesmo prestígio, isto é, que apresentam uma condição superior a demais línguas.

Entende-se que os tesouros multilíngues visam um tratamento igualitário entre as línguas. Nesse aspecto, os tesouros buscam contemplar a questão da diversidade linguística e destacam-se por possibilitar um tratamento ético comparado aos demais instrumentos de representação. Tal proposta reflete também na possibilidade de comunicação entre várias línguas. Além disso, sugere instrumentos que conectem e facilitem a comunicação entre culturas e línguas distintas. Assim, Hudon (1999) menciona que, no contexto das tecnologias de informação, tem-se conseqüentemente o aumento do fluxo informacional. Nesse sentido, evidencia-se também as barreiras geográficas as quais já encontram-se de algum modo superadas pela dinâmica das tecnologias de informação. Porém outra também ainda necessita ser enfrentada: as barreiras linguísticas.

Vale registrar que os tesouros multilíngues possuem a intenção de servir principalmente à indexação, tal como instrumentos auxiliares na recuperação em sistemas de informação multilíngues, ou seja, sistemas de informação com o domínio em diversas línguas. A disposição de um tesouro multilíngue permite que os documentos sejam indexados em uma ou em várias línguas dispostas em um centro de informação. Nesse sentido, as buscas podem ser realizadas em diversos idiomas, na maioria das vezes utilizando uma linguagem do usuário. Tal instrumento busca aproximar diversas culturas e facilitar a “comunicação interlinguística” (Hudon, 1997. p.85).

Com base nas discussões socioculturais de García Gutiérrez, Beghtol e Hudon, elaboramos uma síntese dos conceitos identificados nessas abordagens que se encontra no (Quadro I do Anexo 1).

### 3. A Semiótica da Cultura e seus conceitos

A Semiótica da Cultura é um campo que busca a análise das linguagens existentes entre a natureza e cultura (Machado, 2003, p. 25). É um campo transdisciplinar que dialoga com diversas disciplinas, tais como: Teoria Literária, Linguística estrutural, Semiótica, Crítica da Arte, Cibernética, Teoria da Informação e da Comunicação, Antropologia, Neurolinguística e Etnologia, entre outras.

A Semiótica da Cultura desenvolveu suas discussões sobre os sistemas de signos (cultura,

texto cultural, códigos culturais e linguagens), em meados da década de 1960 na Escola de Tártu-Moscú, da Universidade de Tártu, Estônia. Nesse campo, a linguagem é considerada como um dispositivo que se expressa por signos que não possuem, necessariamente, relação direta com os signos linguísticos.

O propósito norteador da Semiótica da Cultura é compreender a comunicação como um processo semiótico, sendo assim, a cultura é entendida como um conjunto unificado de sistemas sógnicos, constituídos por uma dinâmica de códigos e linguagens operados em um contínuo semiótico.

A cultura é um grande texto, aspecto elementar da semiótica moderna, principalmente pelo seu caráter dinâmico e dialógico. A cultura é considerada um texto aberto, pois carrega em si sistemas sógnicos diversos que em sua complexidade dialogam e transformam-se constantemente.

Objetivamos nessa seção apresentar a Semiótica da Cultura. Compreende-se que a análise da linguagem, assim como da cultura no âmbito dessa abordagem semiótica depende necessariamente do reconhecimento de diversos outros sistemas de signos. A seguir, apresentamos alguns dos conceitos relativos à Semiótica da Cultura.

### 3.1. Linguagem

A linguagem na Semiótica da Cultura consiste em qualquer sistema de signos que sirva de veículo para a comunicação, assim como para a produção de cultura em seu sentido amplo (Machado, 2007, p. 28). A linguagem é elemento primordial para representação de um determinado contexto porque passa a atuar como potencializadora da cultura. Dessa maneira, também considera a linguagem um elemento definidor do conhecimento, pois partimos da questão de que sem linguagem não é possível criar e comunicar conhecimento.

A linguagem é também um dispositivo dialógico utilizada pelos sujeitos em virtude do estabelecimento de interações. Lotman (2011), em seu artigo "The place of art among other modelling systems" afirma que a linguagem ocorre no espaço da semiosfera não apenas em substâncias materiais (objetos tangíveis) mas também pode apresentar-se por meios cinéticos. Em "On the semiosphere" (2005), o autor acrescenta que sem semiosfera uma linguagem não só não funciona, como nem sequer existe. Segundo Torop (2005, p.166), a abordagem de linguagem que se evidencia na Semiótica da Cultura, é

resultante de uma ampliação da concepção de linguagem verbal discutida por Jakobson (1985) em "Metalanguage as a linguistic problem". Assim, para a escola fundadora da Semiótica da Cultura, como para os integrantes da Escola de Tártu-Moscú, a linguagem verbal passa a configurar-se como um sistema de nível primário (seria a própria língua).

As linguagens, no sentido da Semiótica da Cultura, são sistemas modelizantes. Os sistemas modelizantes são "sistemas relacionais constituídos por elementos e por regras combinatórias no sentido de criar uma estruturalidade que se define, assim como uma fonte ou um modelo" (Machado, 2003, p. 167). Para a Semiótica da Cultura, a linguagem é um dispositivo que se expressa mediante signos que não possuem necessariamente uma relação direta com os signos linguísticos, mas estes são relevantes à expressão dos signos de caráter não linguístico. Por sua vez, a língua é uma instituição social (Saussure, 1970).

A linguagem é uma forma de comunicação e organização do conhecimento humano. A linguagem, independente de sua unidade expressiva (seja por meio de palavras, de sons ou até imagens), é um sistema de símbolos. Nesse sentido, a linguagem permite que as relações humanas sejam objetivadas. Tais relações são estabelecidas por uma ordem social, isto é, baseiam-se nas necessidades humanas da relação uns com os outros e, desse modo, a linguagem acaba por anteceder a ação comunicativa (Calefo, 2009, p. 71-72).

A linguagem é expressão híbrida, principalmente se observada por uma perspectiva dialógica e de domínio sociocultural. Sobre tal perspectiva recordamos Bakhtin (2004), em sua obra "Marxismo e filosofia da linguagem" em que considera a linguagem uma entidade, um signo ideológico-social, o qual compartilha um processo de interação social que refrata determinadas realidades. Linguagem é compreendida como sistema simbólico que promove a comunicação.

Todo e qualquer elemento simbólico utilizado pelo homem é decorrente de uma necessidade de interação, de comunicação, isto é, requer o relacionamento com os outros. Por isso ressaltamos a função da linguagem no sentido da Semiótica da Cultura que, como elemento de comunicação, a linguagem acontece a partir do diálogo entre os sujeitos e das interações socioculturais estabelecidas entre grupos. Por meio da linguagem construímos o cotidiano, na medida em que dispomos de uma interpretação da realidade. Assim, como a linguagem significa e ganha sentido na construção do conhecimento,

o contexto e a cultura, e todos os seus processos de tradução e criação de textos, atuam como transformadores das linguagens.

A linguagem é a expressão da cultura. Desse modo, para a Semiótica da Cultura, a linguagem é uma forma de expressão que não se restringe apenas à expressão de signos linguísticos, mas integra uma diversidade de signos através da arte, imagem, som, entre outros códigos.

### 3.2. Cultura

A cultura para a Semiótica da Cultura é um conjunto de informações não-hereditárias armazenadas e transmitidas por grupos em domínios diferenciados (Machado, 2003, p.157; Schnaiderman, 2010, p. 31). A cultura em seu aspecto funcional é acumulativa, pois organiza as informações e as conserva, preserva tradições e possibilita a construção do conhecimento.

A cultura é um texto de ampla complexidade. É definida também como o espaço semiótico em constante transição dos signos culturais que interagem por meio de linguagens. De modo geral, a cultura é algo peculiar ao *homo sapiens* (Barát, 2008, p. 92). Sobre isso a autora entende que a cultura é própria do *homo sapiens*, justamente pelo uso de objetos materiais como parte integrante do seu comportamento. A cultura compreende a linguagem, ideias, crenças, costumes, códigos, instituições, ferramentas, técnicas, obras de arte, rituais, cerimônias, entre outros elementos. Desse modo, a cultura é tanto homogeneidade como heterogeneidade.

De acordo com Torop (2002), a cultura tem seus próprios sistemas de signos com os quais comunica-se com seus membros. Assim, uma possibilidade para entender uma cultura é aprender suas linguagens, pois as línguas das culturas tendem a mudar e os seus sinais são ambíguos. A cultura também pode ser percebida como um processo que encontrar-se-á sempre em constante transformação, resultante das relações entre os indivíduos e as suas necessidades humanas de interação.

Para Lotman (1999, p.109, tradução nossa) “A cultura em seu conjunto pode ser considerada como um texto”. Os textos são produções dinâmicas e dependem de uma cultura, uma língua e das linguagens, assim como do contexto para que sejam compreendidos.

A cultura é um elemento dinâmico e dialógico. Também consideramos ser a cultura um sistema aberto e desse modo nos faz refletir sobre os elementos signícos como parte constitutiva de toda e qualquer cultura.

### 3.3. Texto

texto é produto da cultura e pode ser algo materializado ou não. O texto, do ponto de vista da Semiótica da Cultura, pode ser uma obra de arte, assim como será também uma música, um romance desde que transmita mensagens. Desse modo, o texto consiste em um elemento semiótico criado através de linguagens diversas, isto é, são construídos e expressados a partir da conceptualização de linguagens que possui caráter dinâmico e busca constituir-se na representação de uma cultura.

Nesse sentido, o texto implica uma construção individual e coletiva, no instante em que se põe à análise e às interpretações. O texto estende-se para a criação de diversos novos textos que ocorre na medida em que um texto, por exemplo, uma obra de arte, será interpretado por sujeitos distintos. Lotman (1999) considera que um texto pode ser construído dentro de outro texto, e tal construção evidencia um mecanismo de interpretação que envolve o autor do texto “original” e o leitor do mesmo. Assim, na cultura temos um constante fluxo de reservas específicas de texto com perda de códigos, ou seja, com a ausência de características que identificam uma cultura e leva a processos de criação de novos códigos, muitas vezes entendidos como reconstruções, como por exemplo, as lembranças (Lotman, 2005, p. 215). Para Lotman (1999, p. 109) texto é um espaço organizado de maneira homogênea, mas que considera a introdução de múltiplos elementos que são agregados de outros textos.

A característica sistêmica da cultura é favorável à construção de textos expressados por diversas linguagens. Nesse sentido, o texto pode apresentar-se tanto em um nível verbal, como em um nível de abstração, o qual procede da interpretação dos sujeitos envolvidos com o texto, sendo deste modo mais complexa a sua interpretação em relação ao texto escrito. Lotman (1990, p. 9-19) *apud* Machado (2011, p. 92) entende texto como um mecanismo de geração de sentidos caracterizado por três funções: a função comunicativa, a função memorial e a função criativa. Vale destacar que a função comunicativa corresponde à capacidade que as linguagens possuem para a comunicação dos textos. A função memorial do texto implica a restauração da memória da cultura, das lembranças de um passado cultural. A função criativa retrata justamente a capacidade que o texto possui de produzir significados, isto é, de transformar as mensagens, dadas no texto, em outras mensagens atribuindo novos sentidos ao

texto. Nesta última função é considerada a manifestação simultânea de várias linguagens.

### 3.4. Contexto

Na Semiótica da Cultura o contexto é entendido a partir da concepção de texto. De modo que o contexto representa um determinado texto. Entendemos que o contexto também representa o momento histórico, político, social e cultural no qual os textos da cultura são produzidos e comunicados. Desse modo, o contexto possui uma relação com a ideia de espaço semiótico, isto é, um espaço em que os sistemas de signos da cultura se comunicam. Segundo Lotman (1990) a semiosfera é então esse espaço no qual a semiose acontece.

Para Torop (2005, p. 169) a semiosfera é um conceito que permite que a Semiótica da Cultura possa chegar a uma nova compreensão do holismo, isto é, de uma análise holística dos processos dinâmicos. Na Semiótica da Cultura, o termo semiosfera congrega as ciências que estudam cultura, a cultura converte-se em semiótica – o desejo de encontrar uma linguagem de descrição que poderia ser traduzida para unificar diferentes linguagens disciplinares e interdisciplinares.

A semiosfera é o espaço semiótico, fora dela a semiose não pode existir. Em resumo, sem atos semióticos distintos não obteremos um universo semiótico. Muito pelo contrário, só a existência de um tal universo - a semiosfera - faz o verdadeiro ato signatário específico (Lotman, 2005, p. 208).

Em suma, o contexto é compreendido como o momento histórico, social e cultural de representação dos textos da cultura. Por essa razão, entendemos o contexto como um espaço de identificação dos textos, de produção e de reconstrução de novos textos que é possível em consequência de um processo de recepção, interpretação e de tradução.

### 3.5. Tradução

A tradução na Semiótica da Cultura nos remete à noção de “tradução da tradição”, um mecanismo de conversão da própria cultura. Consideramos a tradução uma atividade interpretativa e a própria cultura é resultado da tradução. Nesse sentido, a tradução não ocorre apenas em um nível linguístico, como a tradução entre línguas. Torop (2002, p. 603) diz que a Semiótica da Cultura afirma que a cultura é a tradução, e também que a tradução é cultura. A tradução consiste em uma atividade que explica os mecanismos da cultura.

No universo teórico da Semiótica da Cultura, ressaltamos a concepção de tradução da tradição que, segundo Machado (2003, p. 30), “[...] é o encontro entre diferentes culturas a partir do qual nascem códigos culturais que funcionam como programa para ulteriores desenvolvimentos”. A tradução da tradição implica um processo interpretativo dos aspectos culturais. Em outras palavras, é uma ação semiótica que possibilita a interpretação e a compreensão de uma determinada cultura.

Sobre a noção de tradução Torop (2002 p. 593) entende que traduzir é uma atividade e tradução é um resultado desta atividade, porém são conceitos inseparáveis da cultura. Desse modo, a capacidade que a cultura tem de traduzir é uma especificidade da própria cultura. A cultura opera através da atividade de tradução. No entanto, a tradução para a Semiótica da Cultura implica um processo e/ou atividade interpretativa de alta complexidade.

Lotman (1990, p. 143) entende que um “elementar ato de pensar é a tradução”, e diante de uma perspectiva lógica, considera que “o mecanismo elementar de traduzir é o diálogo”. O diálogo evidencia-se na própria relação entre os indivíduos em uma determinada cultura. É no diálogo que a linguagem assume a função de potencializar a representação dos textos da cultura. A representação é uma atividade de tradução.

Compreendemos que para a atividade de traduzir implica conhecer a cultura, compreender como os seus sistemas de signos funcionam, pois a cultura tem seus próprios sistemas de signos ou linguagens com base nos quais os membros da cultura possam se comunicar. Uma possibilidade para entender uma determinada cultura é aprender as linguagens da mesma, assim como os sistemas de sinais que operam dentro da cultura. Outra possibilidade seria uma aproximação da cultura através de eventos e textos que ligam diferentes sistemas de signos, mas que possuam um significado geral que este pode ser descrito (Torop, 2002, p.600).

### 3.6. Sistemas modelizantes

Os sistemas modelizantes são sistemas de representação nos quais se destacam as linguagens e os códigos culturais. Segundo Lotman (1978) *apud* Machado (2007, p.29) a linguagem, além de um sistema de comunicação, é também um sistema modelizante. Com base nisso e no aspecto da representação do texto em análise de sua função comunicativa, por exemplo, a linguagem, enquanto um sistema modelizante e de comunicação, é responsável pela transmissão da mensagem do texto da cultura.

A linguagem, por exemplo, é um sistema modelizante que estabelece relações entre os indivíduos/comunidades e a comunidade receptora, as quais se encontram em uma mesma esfera cultural, fazem uso da mesma língua, porém os valores, costumes, crenças e mitos imbuídos nessa mesma esfera cultural possuem uma complexa variedade e o desenvolvimento de linguagens. Contudo, o uso desse tipo de sistema modelizante que é a linguagem necessita contemplar uma diversidade de expressões que não somente a expressão verbal, mas também a expressão visual, sonora, gestual.

A linguagem é um instrumento mediador e deve facilitar os processos de interpretação dos textos da cultura. Vale destacar que mesmo havendo o uso de apenas um código, por exemplo, a língua, é preciso que tal código seja passível de tradução que contemple o conteúdo do texto, pois a cultura possui em seu sistema de signos uma variedade de sistemas como códigos, linguagens, símbolos etc. Assim, para compreender as mensagens produzidas é necessário o compartilhamento de todos os sistemas de códigos, de linguagens, de símbolos, que garantam a diversidade cultural presente nesses textos.

Os sistemas modelizantes na Semiótica da Cultura são classificados em dois grupos: sistemas modelizantes de nível primário e sistemas modelizantes de nível secundário. No primeiro está a língua, a qual se identifica por sua estrutura verbal; a língua atua como um sistema modelizante primário porque a partir deste é possível o desenvolvimento de diversos outros sistemas como literatura, mitos os quais são construídos em analogia dos sistemas primários (Machado, 2003, p. 167-168). Assim, o código verbal é entendido como um sistema modelizante primário por constituir uma base estrutural. Outros tipos de linguagens como a arte, o mito, entre outras são sistemas secundários, porém possuem uma estrutura distinta da estrutura verbal e possuem a capacidade de desenvolver posteriormente uma estruturalidade (Machado, 2007, p. 29).

A ideia de estruturalidade corresponde à qualidade textual da cultura sem a qual as mensagens não podem ser reconhecidas, armazenadas e divulgadas (Machado, 2003, p. 158). A língua, para Jakobson (1975), é o código que permite a comunicação, e sem um código comum não há comunicação efetiva. Nesse contexto, o autor propõe a função metalinguística dos discursos, e afirma que os enunciados só podem ser alcançados se o “emissor e destinatário” compartilham o mesmo código.

A linguagem, em seu aspecto dinâmico e expressivo é denominada um sistema modelizante secundário, como já mencionado. Nesse aspecto, a linguagem caracteriza-se como um sistema modelizante de segundo grau por não possuir uma estrutura como a língua, mas há estruturalidades, isto é, uma lógica que implica na compreensão de diversas expressões que contemplam os diversos fenômenos da cultura.

Com base em Torop (1999, p. 10) a linguagem natural, por exemplo, consiste em um sistema de modelagem primária em relação à realidade, isto é, seu contexto, enquanto o sistema de modelagem secundário, como uma linguagem de descrição, relaciona-se com todas as outras linguagens das artes e da cultura como, por exemplo, mitologia, religião, normas de comportamento etc. Consideramos que as linguagens traduzem-se em sistemas abertos da e na cultura. Assim, entendemos que os sistemas modelizantes pressupõem processos de tradução dos textos culturais que ocorrem na construção de novos traços da cultura.

Sobre os conceitos apresentados elaboramos uma síntese que pode ser visualizada no (Quadro II do Anexo 2).

#### **4. Interações entre a Semiótica da Cultura e a Organização do Conhecimento**

Entendemos que tanto as abordagens socioculturais da Organização do Conhecimento como a perspectiva conceitual trazida pela Semiótica da Cultura apresentam alternativas teóricas para organizar o conhecimento, mesmo entendendo que possuem interesses distintos. Vale ressaltar que as abordagens socioculturais da Organização do Conhecimento desenvolvem investigações em sistemas de organização e representação do conhecimento na criação de instrumentos que garantam o acesso à informação e ao conhecimento na perspectiva de um tratamento ético do conteúdo realizado pelo profissional.

Nesse sentido, a Organização do Conhecimento deve respeitar as condições simbólicas destas esferas no momento de representar a informação e organizar o conhecimento. Um tesouro, por exemplo, somente pode ser elaborado se houver o respeito da estrutura linguística dos conceitos e temas de uma comunidade discursiva, pois tal comunidade dispõe de uma cultura materializada em texto, compreendidos como sistemas de signos sociais oriundos de uma diversidade de linguagens.

A Semiótica da Cultura tem como base as discussões sobre linguagem e todo o desenvolvi-

mento de signos da cultura. No entanto, inclina-se às abordagens socioculturais por subsidiar os conceitos de linguagem e cultura. Desse modo, a Semiótica da Cultura produz reflexões que abrangem as perspectivas histórica, antropológica, sociológica e linguística.

Na construção dessa temática e das discussões traçadas no desenvolvimento deste estudo, verificamos um potencial diálogo cuja linha de convergência apresenta-se por uma perspectiva de temas de modo que o direcionamento se dá por um viés sociocultural. Nesse sentido, entendemos que as abordagens destacadas nos dois campos de estudos não convergem necessariamente para uma relação conceitual. Porém, envolvem temáticas que se aproximam e poderiam abrigar uma futura relação interdisciplinar.

Hjørland (2008, p. 98 tradução nossa) evidencia o caráter interdisciplinar da Organização do Conhecimento quando relata que

A Organização do Conhecimento não é apenas algo que a Biblioteconomia pode fazer sem considerar a pesquisa em outros domínios, por exemplo, ciência da computação, lingüística e processo natural da língua, teoria do conhecimento, teoria da organização social, etc. Em particular, um entendimento da natureza do conhecimento, cognição, língua e organização social são decisivos para a compreensão da Organização do Conhecimento e, assim, para a habilidade de planejar, avaliar e usar os processos de organização do conhecimento e sistemas de organização do conhecimento. Muitos campos podem ter um interesse na definição de questões do conhecimento ou podem ser consideradas disciplinas relacionadas.

O campo da Organização do Conhecimento é um espaço com capacidade de produzir investigações, de caráter teórico e conceitual. Assim, consideramos que nesse contexto transitam variados domínios que, em geral, contribuem com o seu desenvolvimento e foi sob uma perspectiva interdisciplinar que trouxemos os conceitos de linguagem, cultura, texto, contexto, tradução e sistemas modelizantes por entender que estes compartilham com temas socioculturais e por isso dialogam com os conceitos de ética transcultural, transculturalismo, garantia cultural, hospitalidade cultural e multilinguismo.

No conceito de linguagem, por exemplo, consideramos não somente a língua como um elemento necessário para a comunicação e a representação de um conhecimento, mas também entendemos que as diversas formas de linguagem: visual, verbal, sonora, artística etc. devem ser consideradas pelo fato de estarem presentes em várias línguas. Desse modo, consideramos que realizar análises críticas subsidia uma visão ética e promove a garantia cultural, de-

monstra não apenas o interesse no tratamento e organização de um conhecimento, mas implica também reconhecer a origem desse conhecimento.

Com base nas análises das abordagens socioculturais compreendemos que é necessário conhecer as culturas, como também suas linguagens e contextos. Para a Organização do Conhecimento entender as culturas implica em um exercício de compreensão dos contextos das comunidades usuárias, pois quando examinamos os aspectos culturais do conhecimento, aceitamos que o conhecimento é consequência das relações sociais e encontra-se imerso a uma diversidade de valores, crenças, costumes, linguagens e códigos que devem ser contemplados para sua representação em um sistema de informação.

O conhecimento pode ser entendido próximo à noção de texto da Semiótica da Cultura. O conhecimento como produto social e cultural é reconhecido em sua dinâmica e contexto, e permite sua recepção via a comunicação/representação. A aproximação que liga a ideia de conhecimento (como um produto fruto de uma construção social) a de texto (produto cultural) supõe que todo texto é passível de interpretações, isto é, de tradução, não sendo diferente com o conhecimento, pois quando um conhecimento é representado este já passou por um processo de tradução em termos semióticos. Esse processo interpretativo requer uma consciência ética sobre os valores e esquemas contidos no conhecimento.

No domínio da Semiótica da Cultura, o fenômeno da tradução nos permite também uma conexão com as abordagens socioculturais da Organização do Conhecimento, sobretudo, ao que corresponde a noção da garantia cultural (Begtthol, 2002). Na garantia cultural, cada indivíduo/comunidade deve ter traduzido adequadamente sua cultura, de maneira que priorize os elementos que identificam a mesma, e sejam refletidos com base em seu contexto a partir do sentido que recebe dos sujeitos nas comunidades. Posto isso, é necessário ressaltar que o processo de tradução na Ciência da Informação ainda implica a passagem de uma língua para outra. Para Lancaster (1993, p. 16) “[...] o processo de tradução envolve a representação da análise conceitual mediante um termo ou termos extraídos de um vocabulário”.

Contudo, o conceito de tradução utilizado deriva da perspectiva linguística e entendemos que na área de análise documental, por exemplo, a tradução é utilizada para explicar o processo de representação documentária (Lara, 1999). Essa

representação busca o acesso aos conteúdos documentais. Nesse sentido, a tradução é a passagem da língua do documento para a língua do sistema. O mecanismo de passagem de uma língua para outra também envolve uma atividade interpretativa, mas uma interpretação que se limita a um nível semântico e que considera mudanças em tempo e espaço sobre os sentidos atribuídos aos termos.

Nesse contexto, a ideia de tradução acaba por não ser considerada em sua amplitude, como um processo de tradução cultural, do texto e cultura dos autores e de sua comunidade discursiva para a cultura do sistema da informação e de sua comunidade usuária. No entanto, entendemos que a noção de tradução, na Ciência da Informação, adquire um caráter reducionista e limitador, e pressupomos ser complicado, por exemplo, entender a configuração de novos textos na cultura, de novos conhecimentos. A tradução, na Semiótica da Cultura, enquanto uma atividade interpretativa, é compreendida como resultado da cultura, isto é, das trocas culturais, simbólicas entre os grupos e comunidades discursivas.

Nesse escopo, sobre tradução destacamos a discussão das barreiras linguísticas tratadas por Hudon (1999), assim como a defesa do multilinguismo em que a autora considera respeitar a diversidade de línguas, por isso há necessidade de propor condições para que os indivíduos não fiquem presos a determinismos linguísticos. A autora também expõe a necessidade de um instrumento de representação que disponha de uma diversidade linguística cujo propósito seja o acesso à informação.

Em outra parte dessa discussão encontra-se a garantia de acesso por parte dos sujeitos/usuários com respeito a sua cultura e seu contexto. A garantia cultural busca não somente o acesso a informação, mais a representação democrática de uma determinada cultura. No desdobramento desse diálogo, também está o conceito de ética transcultural de García Gutiérrez (1998, 2002, 2004, 2005 e 2006) e a concepção de transculturalismo que propõem análises transculturais dos fenômenos culturais.

Na Semiótica da Cultura a linguagem é pensada com relação à denominação de linguagem epistemográfica, linguagem documental, isto é, a linguagem como elemento expressivo do conhecimento e de uma cultura. A linguagem modeliza-se no intuito de representar e socializar um conhecimento, assim como um texto com ênfase nos contextos culturais. Portanto, consideramos que as noções destacadas, tanto no que se refere às abordagens socioculturais co-

mo da Semiótica da Cultura, nos direcionam a uma comparação interdisciplinar.

Nesse contexto, a linguagem é elemento comunicador entre um sistema de signos culturais, traduzidos por meio de um sistema que venha representar o conhecimento dos indivíduos, grupos e comunidades. A linguagem deve ser analisada não apenas por sua diversidade linguística, mas também de sua diversidade cultural imbuída na configuração do conhecimento, pois as comunidades em uso das linguagens requerem garantias, tendo em vista a apropriação tanto da cultura como do conhecimento.

A linguagem é um mecanismo potencializador do conhecimento, e em termos semióticos é um instrumento decodificador da cultura. Vale ressaltar que existem várias outras concepções na Semiótica da Cultura que poderiam ser mencionadas e exploradas nessa discussão, porém buscamos destacar as que mais representavam a temática desta pesquisa.

Para tanto, buscamos evidenciar as interações entre as perspectivas investigativas mencionadas neste estudo em que o intuito foi verificar a partir das discussões das abordagens socioculturais da Organização do Conhecimento e da Semiótica da Cultura diálogos interdisciplinares.

## 5. Considerações finais

Organizar e representar um conhecimento são atividades de tradução da cultura. Desse modo, entender a origem de qualquer conhecimento implica uma atividade semiótica. Assim, conhecer os contextos é conhecer as culturas e as linguagens. É ainda ter o conhecimento dos códigos, símbolos e linguagens que caracterizam e identificam os indivíduos e as comunidades.

Nesse sentido, consideramos que a linguagem é um instrumento cultural potencializador do conhecimento humano (Semprini, 1999) é o “próprio objeto do conhecimento” (Mendonça, 2000, p. 65). Sendo esta objeto do conhecimento, é prudente afirmar que existem diversas formas de linguagens e por isso entendemos que em cada linguagem haverá uma manifestação distinta do conhecimento, pelo fato de existir uma diversidade de culturas.

A concepção de cultura no contexto da Organização do Conhecimento não pode ser fechada e determinada pelas relações interdisciplinares já estabelecidas (como, aliás, não deve ser em nenhum contexto). Assim, devemos valer-nos de uma perspectiva alargada quanto ao problema da diversidade cultural na sociedade. A cultura é em si diversificada e tal diversificação

pode ser entendida a partir do contexto em que tratamos a cultura e a sua multiplicidade de elementos os quais caracterizam a própria cultura. Na Semiótica da Cultura, verificamos que os conceitos de linguagem, cultura, texto, contexto, tradução e sistemas modelizantes tratam de considerar os contextos culturais para uma análise semiótica da linguagem na cultura.

Sustentamos o argumento de que a concepção de linguagem constitui-se um elemento de aproximação com as perspectivas socioculturais da Organização do Conhecimento, por ser necessária para a produção e comunicação do conhecimento, tendo como objeto de análise os contextos culturais que condicionam a interpretação dos sujeitos. Verificamos que o conceito de linguagem é destacado nas discussões sobre “diversidade linguística” em seu aspecto multilíngue evidenciado por Hudon (1999), assim como a noção de “garantia e hospitalidade cultural” de Begthol (2002) e ainda na abordagem sobre transculturalidade tratada por García Gutiérrez (2002a; 2002b; 2002c). Essas análises possuem estreita relação com as teorias culturais e da linguagem e podem dar margem para uma perspectiva semiótica da produção, representação e organização da linguagem em nosso campo.

Ressaltamos que a Semiótica da Cultura fornece aportes teóricos à Organização do Conhecimento, visto que viabiliza a valorização dos aspectos culturais, e, particularmente, da diversidade cultural nas discussões sobre a linguagem em sentido amplo e nas questões particulares das linguagens documentárias, no domínio da Linguística Documentária e da Análise Documentária.

O objetivo deste estudo foi investigar as possíveis interações de ordem teórico-conceituais fornecidas pela Semiótica da Cultura para as abordagens socioculturais da Organização do Conhecimento. Desse modo, a perspectiva da linguagem, além dos demais conceitos mencionados no desenvolvimento deste estudo com foco na teoria da Semiótica da Cultura atuam como elementos interdisciplinares e podem sugerir diversos temas de estudos para o campo da Organização do Conhecimento, mais precisamente, àqueles que estejam inclinados às abordagens socioculturais.

A concepção da linguagem pode ser entendida como elemento de aproximação na temática abordada, com as perspectivas socioculturais da Organização do Conhecimento. Desse modo, é necessário entender que em outros contextos, isto é, em outros campos de estudo, a linguagem será necessária para a produção e

comunicação do conhecimento, tendo em vista a análise dos contextos culturais que possam permitir a interpretação dos sujeitos/comunidades/usuários.

Compreender a cultura e os contextos para a organização e representação do conhecimento é necessário não apenas para o registro, mas para que a cultura encontre-se em tal registro, isto é, que seja contemplada tanto a sua diversidade cultural como linguística e, desse modo, seja garantida a comunicação dos conteúdos da cultura, das comunidades e contextos.

Entendemos que a linguagem é o centro desta discussão nas abordagens socioculturais da Organização do Conhecimento, desde que seu contributo seja favorecer o entendimento dos contextos culturais que estabelecem relação entre a cultura/texto e a comunidade/coletivo. Pois a linguagem é um mecanismo expressivo que as comunidades e culturas utilizam na socialização do conhecimento.

Consideramos que o campo da Organização do Conhecimento contempla estudos que se aproximam em diversos aspectos e podem desse modo estabelecer inúmeros diálogos, sejam em um nível teórico ou conceitual, mas que implicam em análises interdisciplinares. Porém, com este estudo enfatizamos que o diálogo se destaca em um nível temático e não necessariamente conceitual. No entanto, entendemos que os conceitos apresentados não convergem para um mesmo sentido, mas destacam em suas análises que pode haver relações teóricas de caráter interdisciplinar e podem desencadear *a posteriori* diversas análises com base nos aspectos socioculturais.

Em suma, a linguagem é elemento de identificação das culturas e entendemos que a noção de cultura da Semiótica da Cultura será melhor incorporada na Organização do Conhecimento se considerarmos o contexto como central, na medida em que há uma diversidade de contextos que garante que um determinado conhecimento produzido coletivamente seja organizado e posteriormente representado com o objetivo de recuperar a informação.

## Referencias

- Bakhtin, M. (2004). *Marxismo e filosofia da linguagem*. 10.ed. São Paulo: Hucitec.
- Barát, Ágnes Hajdu. (2008) *Knowledge Organization in the Cross-cultural and Multicultural Society*. En: ARSENAUT, C. ; TENNIS, J. T. (Ed.). *Cultural and identity in knowledge organization*. *Advances in Knowledge Organization*. Canada: ERGON-Verlag, v. 11. p.84-90
- Barité, M. (2011). *La garantía cultural como justificación en sistemas de organización del conocimiento: aproximación*

- ción crítica. *Palabra Clave* (la Plata. En línea), Universidad de la Sabana.
- Barité, M. (2001). Organización del conocimiento: un nuevo marco teórico-conceptual em Bibliotecología y Documentación. In: CARRARA, K. (Org.). *Educação, universidade e pesquisa*. Marília: Unesp-Marília-Publicações, p. 35-60.
- Begthol, C. (1986). Semantic Validity: concepts of warrant in bibliographic classification systems. *Library Resources & Technical Services*. v. 30, n. 2, p. 109-125.
- Begthol, C. (1995). Domain analysis, literary warrant, and consensus: the case of fiction studies, *Journal of the American Society for Information Science and Technology*, 46, n. 1, p. 30-44.
- Begthol, C. (2001). Relationships in Classificatory structure and meaning. In C.A. Bean & R. Green, (eds.). *Relationships in the organization knowledge*. Dordrecht, Netherlands: Kluwer, p. 99-113.
- Begthol, C. (2002). Universal concepts, cultural warrant, and cultural hospitality. In: López-Huertas, M. J. (Ed.). **Challenges in knowledge representation and organization for the 21 century**: integration of knowledge across boundaries. Würzburg: ERGON-Verlag. p.45-49. (Advances in Knowledge Organization, 8).
- Begthol, C. (2005). Ethical decision-making for knowledge representation and organization systems for global use. *Journal of the American Society for Information Science and Technology*, New York, v. 56, n. 9, p. 903-912.
- Beuchot, M. (2004). *La semiótica: teorías del signo y el lenguaje em la história*. México: Fondo de Cultura Económica, 2004, p. 7.
- Brascher, M. Café, L.M.A. (2011.). Organização do conhecimento: teorias semânticas como base para estudos e representação de conceitos. *Inf. Inf.*, Londrina, v.16. n.3. p. 25-51, jan./jun.
- Calefato, P. (2009). Language in social reproduction: sociolinguistics and sociosemiotics. *Sign Systems Studies*. v. 37, n. 1/2.
- Campbell, G. (2000). Queer theory and the creation of contextual subject access tools for gay and lesbian communities. *Knowledge Organization*, v.27, n. 3, p. 122-131.
- Campbell, G. (2010). Tensions between language and discourse in north american knowledge organization, *Knowledge Organization*, v. 37, n. 1, 2010. p. 51-57.
- Coelho Neto, J.T. (2010). *Semiótica, Informação e Comunicação*. São Paulo: Perspectiva.
- Dahlberg, I. (2006). Knowledge organization: a new science? *Knowledge Organization*, Frankfurt, 33 (1). p.11-19.
- Ducrot, O.; Todorov, T. (1988). *Dicionário enciclopédico das ciências da linguagem*. 2.ed. rev. amp. São Paulo: Perspectiva.
- García Gutiérrez, A. (1998). *Principios de lenguaje epistemográfico: la representación del conocimiento sobre Patrimonio Histórico Andaluz*. Sevilla: Instituto Andaluz del Patrimonio Histórico, (Cuadernos técnicos, 3).
- García Gutiérrez, A. (2002a). Knowledge organization from a "culture of the border": towards a transcultural ethics of mediation. In: López-Huertas, M. J.(eds.). *Challenges in knowledge representation and organization for the 21st century: integration of knowledge across boundaries*. Würzburg: ERGON-Verlag, pp. 5L6-522.
- García Gutiérrez, A. (2002b). *La memoria subrogada: mediación, cultura y conciencia em la red digital*. Granada: Editorial de la Universidad de Granada, 2002b. Archivo em meio magnético.
- García Gutiérrez, A. (2002c). Exomemoria y cultura de frontera: hacia una ética transcultural de la mediación. *VII Congreso Internacional sobre Organización del Conocimiento celebrado em Granada (España)*, Granada España.
- García Gutiérrez, A. (2004). *Otra memoria es posible: estrategias descolonizadoras del archivo mundial*. Sevilla: Universidad de Sevilla.
- García Gutiérrez, A. (2006). Científicamente favelados: uma visão crítica do conhecimento a partir da epistemografia. *Transinformação*, Campinas, v.18, n.2, p.103-112, maio/ago.
- García Gutiérrez, A. (2008). *Outra memória é possível: estratégias descolonizadoras do arquivo mundial*. Petrópolis: Vozes.
- Guimarães, J. A. C.; Fernández-Molina, J.C. (2003). Los aspectos éticos de la organización y representación del conocimiento em la revista Knowledge Organization. In: Frías Montoya, J.A. and Travieso, C. de., *Tendencias de investigación em organización del conocimiento*. Salamanca: Universidad de Salamanca. pp. 809-16.
- Guimarães, J. A. C. et al. (2005) Aspectos éticos em organización y representación del conocimiento: un análisis de la bibliografía científica em busca de una categorización preliminar de valores. In: Gascón, J.; Burguillos, F.; Pons, A. (Ed.). *La dimensión humana de la organización del conocimiento*. Barcelona: Universitat de Barcelona. p. 278-285.
- Hjørland, B. (2003). Fundaments of knowledge organization. *Knowledge Organization*, v.30, n. 2, p.87-111.
- Hjørland, B. (2007). Semantics and Knowledge organization. *Annual Review of Information Science and Technology*, 41(1). 367-405. Recuperado em agosto, 2013 de em:[http://www.academia.edu/1912560/Semantics\\_and\\_knowledge\\_organization](http://www.academia.edu/1912560/Semantics_and_knowledge_organization)"[http://www.academia.edu/1912560/Semantics\\_and\\_knowledge\\_organization](http://www.academia.edu/1912560/Semantics_and_knowledge_organization).
- Hjørland, B. (2013). Theories of knowledge organization-Theories of knowledge. *Knowledge Organization*, v. 40, n. 3, p 169-181.
- Hudon, M. (1997). Multilingual thesaurus construction: integrating the views of different cultures in one gateway to knowledge and concepts. *Knowledge Organization*. v. 24, n.2, p84-91..
- Hudon, M. (1999). Accessing documents and information in a world without frontiers. *The Indexer*, London, v. 21, n. 4, p. 156-159.
- Hudon, M. (2003). True and tested products: thesauri on the web. *The Indexer*, London, v. 23, n. 3, p. 115-119, ISKO. About ISKO. Recuperado em agosto, em: <<http://www.isko.org/about.html>"<http://www.isko.org/about.html>>
- Hulme, E. W. (1911.). *Principles of book classification*. Library Association Record, London, v. 13, p. 354-358.
- Jakobson, R.(1985). Metalanguage as a linguistic problem. In: Jakobson, R. *Selected Writings*. v. 7. p. 113-12. The Hague: Mouton.
- Jakobson, R. (1995). *Lingüística e Comunicação*. São Paulo: Cultrix.
- Lancaster, F. W. ( 1993). *Indexação e resumos: teoria e prática*. Brasília: Brinquet de Lemos/Livros.
- Lara, M.L.G. de.(1999). *Representação e linguagens documentárias: bases teórico-metodológicas*. São Paulo, 1999. 208f. Tese (Doutorado em Ciência da Comunicação)-Escola de Comunicação e Artes, Universidade de São Paulo, São Paulo.
- Lara, M.L.G. de. (2011). Conceitos de Organização e representação do conhecimento na ótica das reflexões do grupo tema. *Inf.Inf.*, Londrina, v.16, n. 3. p.92-121, jan/jun.
- López-Huertas, M. J.( 2006). Análisis del dominio interdisciplinar para la representación y organización del conocimiento. In: González de Gómez, M. N.; Orrico, E. G. D.

- Políticas de memória e informação: reflexões na organização do conhecimento.* Natal: EDUFRRN, p. 209-235.
- López-Huertas, M. J. (2008). Cultural impact on Knowledge Representation and Organization in a Subject Domain. In: Arsenaut, C.; Tennis, J. T. (Ed.). *Cultural and identity in knowledge organization.* Würzburg: Ergon Verlag. v.11. p. 304-346. (Advances in Knowledge Organization, 11).
- López-Huertas, M. J. (2010). Epistemological dynamics in scientific domains and their influence in knowledge Organization. In: Gnoli, C.; Mazzocchi, F. (Ed.). *Paradigms and conceptual systems in knowledge organization.* Italy: ERGON VERLAG. p. 91-97. (Advances in Knowledge Organization, 12).
- Lotman, Y. M. (1990). *Universe of the mind: a semiotic theory of culture.* London: I.B. Tauris & CO. LTD.
- Lotman, Y. M. (1999). *Cultura y explosión: lo previsible y lo imprevisible en los procesos de cambio social.* España: Gedisa editorial.
- Lotman, Y. M. (2002). Umwelt and semiosphere. *Sign Systems Studies.* v.30, n.1, p.33-40.
- Lotman, Y. M. (2003a). Sobre el concepto contemporáneo de texto. *Entretextos. Revista Electrónica Semestral de Estudios Semióticos de la Cultura*,(2). Recuperado em julho, 2013 de em:<<http://www.ugr.es/~mcaceres/entretextos/pdf/entre2/lotman.pdf>><http://www.ugr.es/~mcaceres/entretextos/pdf/entre2/lotman.pdf>
- Lotman, I. M. (2003b). El símbolo en el sistema de la cultura. *Entretextos. Revista Eletrónica Semestral de Estudios Semióticos de la Cultura*.n.2,Nov. Recuperado em julho de 2013 em:<<http://www.ugr.es/~mcaceres/entretextos/pdf/entre2/escritos/escritos4.pdf>>
- Lotman, I. M. (2005). On the semiosphere. *Sign Systems Studies.* v. 33, n. 1, p. 205-229.
- Lotman, I. M. (2011). The place of art among other modelling systems. *Sign Systems Studies.* v. 39, n. 2/4 p. 249-270.
- Lotman, I. M. (2013). On the dynamics of culture. *Sign Systems Studies.* v. 41, n. 2/3 p. 355-370.
- Machado, I. (2003). *Escola de Semiótica: a experiência de Tártu - Moscou para o estudo da cultura.* São Paulo: Ateliê Editorial. 189p.
- Machado, I. (2007). *Semiótica da Cultura e Semiosfera.* São Paulo: Annablume/Fapesp.
- Mendonça, E. S. (2000). Linguística e ciência da informação: estudos de uma intersecção. *Ci. Inf.*, Brasília, v. 29, n. 3, p. 50-70, set./dez.
- Peirce, C. S.(1990). *Semiótica.* 2. ed. São Paulo: Perspectiva.
- Saussure, F. de. (1970). *Curso de lingüística geral.* 2. ed. São Paulo: Cultrix.
- Semprini, A. (1999). *Multiculturalismo.* Bauru, SP: EDUSC, 178 p.
- Schnaiderman, B. (1979). *Semiótica Russa.* São Paulo: Perspectiva.
- Schnaiderman, B. (2010). *Semiótica Russa.* 2.ed. São Paulo: Perspectiva.
- Smiraglia, R. P.(2002). The progress of theory in knowledge organization. *Library Trends,* v. 50, n. 3, p. 330-349.
- Smith, M. M.(2001). Global information justice: rights, responsibilities, and caring connections. *Library Trends,* Champaign, v. 49, n. 3, p. 519-537. Disponível em:<<http://web.b.ebscohost.com/ehost/pdfviewer/pdfviewer?sid=440dacb2-5035-4c36-b6c2-23920712b55f%40sessionmgr113&vid=2&hid=126>>
- Torop, P. (1999). Cultural semiotics and culture. *Sign Systems Studies.* v. 27. p. 9-23.
- Torop, P. (2002). Translation as translating as culture. *Sign Systems Studies.* v.30, n.2, p.593-605.
- Torop, P. (2005). Semiosphere and/as the research object of semiotics of culture. *Sign Systems Studies.* v. 33, n. 1, p. 159-173.

## Anexo 1

### *Abordagens socioculturais da Organização do Conhecimento*

| <i>Conceitos</i>                | <i>Síntese</i>  |
|---------------------------------|---|
| Ética transcultural da mediação | É uma proposta conceitual que busca agregar elementos éticos, políticos e socioculturais nos processos de organização e representação do conhecimento.        |
| Transculturalismo               | Reforça sobre a diversidade cultural, e a fusão entre culturas. Nesse aspecto, se opõe a perspectiva do multiculturalismo.                                    |
| Garantia Cultural               | Consiste na proposta de uma garantia de acesso ao conhecimento e desse modo considera os valores, as línguas e as crenças de diversas culturas e comunidades. |
| Hospitalidade Cultural          | Prevê, necessariamente, além da garantia de acesso à informação, a recepção dos conteúdos culturais representados.  |
| Multilinguismo                  | Refere-se ao domínio de várias línguas.   |

Quadro I. *Síntese dos conceitos identificados das abordagens socioculturais da Organização do Conhecimento*

## Anexo 2

### *Semiótica da Cultura*

| <i>Conceitos</i>      | <i>Síntese</i>  |
|-----------------------|---|
| Linguagem             | Sistema modelizante de nível secundário com a função de expressar os textos da cultura por diversas formas: visuais, verbais, sonoras, etc.   |
| Cultura               | É um sistema complexo de signos onde estão presentes os códigos, as linguagens e todos os elementos simbólicos da cultura. Para a Semiótica da Cultura, a cultura é informação, independente que seja materializada ou não, mas desde que possa ser transmitida por meio da memória coletiva.   |
| Texto                 | Consiste na produção da cultura e se expressa por meio de diversas linguagens aprimoradas na dinâmica das culturas e suas comunidades.  |
| Contexto              | Representa o momento histórico, social, político, cultural em que o texto da cultura é produzido. O contexto favorece o conhecimento das culturas e nos permite a partir de diversas perspectivas compreender os textos culturais.  |
| Tradução              | É um processo interpretativo da cultura. Implica tanto em uma atividade, tradução, como no resultado da atividade interpretativa que é traduzir.  |
| Sistemas modelizantes | São sistemas de representação da cultura e atuam na representação da cultura assim como de todos os sistemas de signos desenvolvidos na cultura. Desse modo, são caracterizados sistemas modelizantes de nível primário, a língua e os sistemas modelizantes secundários que são as linguagens. |

Quadro II. *Síntese dos conceitos destacados na Semiótica da Cultura*

---

# Información y crianza con apego en España

*Information and attachment parenting in Spain*

---

**Michela Montesi**

Facultad de Ciencias de la Documentación, Madrid, mmontesi@pdi.ucm.es

## Resumen

La búsqueda de información para la crianza de los hijos es un tema en auge en la literatura tanto médica como social. En este trabajo presentamos los resultados preliminares de una serie de entrevistas con 15 madres procedentes de varias comunidades autónomas de España que se adhieren a la filosofía de la crianza con apego, un estilo de crianza que, entre otras cosas, pretende fundamentarse en la información. El objetivo es documentar y comprender su búsqueda de información. Los resultados ponen de manifiesto aspectos de interés para la investigación sobre comportamiento informacional, como el valor del conocimiento experiencial y afectivo en la toma de decisiones y la evaluación de la información.

**Palabras clave:** Búsqueda de información, crianza con apego, parentalidad, fuentes de información, maternidad, conocimiento experiencial, afecto.

## Abstract

Information seeking within childrearing is a popular topic in the medical and social sciences literature. In this study, we present the preliminary results of a series of interviews with 15 mothers from different autonomous communities in Spain. Participants are practicing a style of childrearing which, among the other things, heavily relies on information, known as attachment parenting. The purpose is to document and understand the search for information within childrearing. Results highlight aspects which might be of interest for research into information behaviour, such as the value of experiential knowledge and affect in decision making and information evaluation.

**Keywords:** information seeking, attachment parenting, parenting, information sources, motherhood, experiential knowledge, affect.

## 1. Introducción

La crianza de los hijos hoy en día representa una etapa existencial de alto consumo de información. Las estadísticas publicadas por el *Pew Research Center* sobre la búsqueda de información médica en la web vienen poniendo de manifiesto que uno de los temas más investigados por los estadounidenses cuando buscan información médica por internet son precisamente el embarazo y el parto (Fox y Duggan, 2013). No es ninguna sorpresa que, en la literatura médica, la búsqueda de información para la crianza sea un tema de interés desde por lo menos principios del siglo XXI. Plantin y Daneback (2009) revisan la literatura médica para comprender cómo los padres utilizan internet para localizar información útil para la crianza de los hijos. Entre los motivos más destacados para ir online, mencionan un mayor grado de conciencia junto con una reducción de apoyo por parte de familiares. Por un lado, las nuevas madres se sienten más aisladas, por otro consideran la información que reciben de sus propias madres como desfasada. Tras el análisis de 94 estudios, los autores concluyen que en la web los padres de hoy en día encuentran a la vez información y apoyo, se interesan por los consejos dados con base en la experiencia, y compensan

la disminución de apoyo por parte de la familia y amigos a través de la interacción con otros padres.

Dada la tremenda presencia de la web social en la búsqueda de información sobre crianza, una revisión más reciente en el ámbito de la literatura médica se centra precisamente en las redes sociales y en el apoyo que prestan a los padres (Niela Vilén et al., 2014). Un mejor conocimiento de este fenómeno puede repercutir en el suministro de determinados servicios sanitarios, así como contribuir a un mejor estado de salud tanto de los niños como de padres y madres. Algunos estudios han incluso llegado a confirmar la existencia de una relación entre presencia en las redes sociales y bienestar de las mujeres, pues a través de la web 2.0 las mujeres conseguirían conectar con su familia extendida y amigos percibiendo un apoyo social más fuerte (McDaniel et al., 2012). La revisión de Niela Vilén et al. (2014) confirma la doble función, informativa y emocional, de la web social. A través de plataformas sociales los padres obtienen tanto información procedente de fuentes profesionales como de experiencias de otros padres, en el momento oportuno. Finalmente los autores destacan que la búsqueda de información sobre crianza comparte con la búsqueda de

información médica una característica destacada que es la presencia más importante de mujeres con respecto a los hombres.

Debido a esta característica de género, otras áreas de investigación como la sociología y los estudios feministas también analizan el tema de la maternidad en relación con la información y la web. Chae (2014), por ejemplo, entiende que la práctica de la “maternidad intensiva”, es decir un estilo de crianza basada en conocimientos a nivel de experto, emocionalmente absorbente, y de alta dedicación, requiere mucha información. Esta “ideología”, cuya conceptualización teórica aparece en un trabajo de Hays (1996), ha sido criticada en la literatura feminista por concebir la maternidad como una etapa en la cual la mujer deja de tener necesidades propias para volcarse alma y cuerpo en los hijos (Medina, S., & Magnuson, 2009). La idea de que la crianza deba tener una base importante de información es presente en otra teoría cada vez más popular en los países occidentales, que es la crianza con apego o respetuosa, y que, a pesar de sus conexiones con la maternidad intensiva, no sería, según Medina y Magnuson (2009), igual de exigente con la mujer. Si bien es verdad que otros autores conciben que la crianza con apego no sería sino uno de los estilos de maternidad intensiva (Faircloth, 2013), queda clara la relación entre estos dos planteamientos y la necesidad de ejercer la parentalidad con base en la información. Los partidarios de la crianza con apego entienden que la relación de apego que se establece durante la infancia entre el niño y sus cuidadores, sean éstos su madre, su padre o cualquier otra figura, tiene repercusiones de por vida en la forma de relacionarse con los demás. El primer principio de crianza de la fundación Attachment Parenting International dice claramente “[...] Tomar decisiones basadas en la información sobre nacimiento, primeros cuidados y prácticas de crianza, es una inversión decisiva en las relaciones entre padres e hijos” (del sitio web de Attachment Parenting International).

En el ámbito de la literatura feminista, es especialmente importante el tema de las mamás blogueras o “mommy bloggers” como se las suele etiquetar. Reproduciendo el género del diario, el blog permitiría a sus autoras dar a conocer la vida diaria de una madre mostrando asimismo la cara fea de la maternidad en contraposición al discurso idealizado de muchos medios de comunicación, siendo en este sentido una actividad liberadora para el género femenino (López, 2009). Por otro lado, el blog da pie a la creación de una comunidad en la que los participantes dan y reciben apoyo y acaban

formando amistad. La interacción alrededor del blog permite a la bloguera descubrirse a sí misma e ir definiendo su identidad en el diálogo con otros miembros de su comunidad (Webb and Lee, 2011). Las temáticas de la literatura feminista aparecen también en un conjunto de estudios realizados en el ámbito de la inteligencia artificial y la interacción hombre-ordenador. Por ejemplo, Gibson y Hanson (2013) indican que la mejora de la auto-estima y la reafirmación de la identidad de las madres representan la aportación principal de las tecnologías de la información y comunicación. Schoenebeck (2013), por otro lado, tras analizar una muestra de posts enviados al sitio web YouBeMom.com, destaca la aportación de este espacio único de interacción en el sentido de ofrecer anonimidad y desinhibición en la vida de sus usuarias.

La búsqueda de información para la parentalidad se discute también en el ámbito de otras ciencias sociales, incluida la pedagogía y el trabajo social, normalmente enfatizando el papel de la tecnología. Radey y Randolph (2009) realizan una encuesta con 1240 padres reclutados en un estado americano para aislar factores demográficos influyentes en el tipo de fuentes de información utilizadas y sobre todo en el uso de internet. Encuentran que el uso más intenso de fuentes de información para la parentalidad (prensa, televisión, libros, familia y amigos, profesores, cursos, e internet) está relacionado con mayores niveles educativos, la edad, y el estado civil, siendo más común entre los mejores educados, los más jóvenes y las personas no casadas. En cuanto al uso de internet, los resultados de Radey y Randolph (2009), por un lado, confirman la existencia de una brecha digital, pues internet es más común entre los más educados, pero por otro aíslan nuevos factores correlacionados con un uso mayor de internet como la condición de no casado. Por otro lado, la pedagogía manifiesta un interés creciente sobre el tema de la búsqueda de información para la parentalidad pues permitiría informar la actuación de educadores y trabajadores sociales, así como el desarrollo de herramientas educativas dirigidas a los padres como el portal Educar en Positivo que describen detalladamente Torres et al. (2014).

Finalmente cabe mencionar unos trabajos realizados sobre el tema en el ámbito de las Ciencias de la Documentación, que quedan de alguna manera desvinculados de los anteriores, a pesar de que la búsqueda de información representa una especialidad bien afianzada en la disciplina. Entre los estudios reseñables sobre el tema, podemos mencionar el de Fisher y Landry (2007) que estudian el comportamiento

informativa de madres amas de casa desde la perspectiva del afecto y de las emociones, y el de Walker (2012), quien constata el creciente interés político-social por la crianza de los hijos en el Reino Unido y describe el universo informativo de 33 padres y madres de Leeds. En la literatura española el tema es escasamente estudiado, aunque podemos mencionar dos trabajos reciente sobre fuentes de información 2.0 para la pediatría (Coronado Ferrer, et al., 2011; Peset Mancebo et al., 2011), y otros realizados en el ámbito de la enfermería (Lima Pereira et al.) y la medicina (Castelló Egea et al., 2011). Por otro lado, existe un cuerpo de literatura importante en el ámbito de las Ciencias de la Documentación que puede proporcionar un marco teórico para el estudio de la búsqueda de información en la crianza de los hijos. Aunque, como mencionábamos anteriormente, toda la literatura sobre comportamiento informativo es relevante en este trabajo, especialmente pertinentes son algunos trabajos realizados recientemente sobre búsqueda de información médica. Genuis (2012) se centra en el universo informativo de un conjunto de pacientes en menopausia, estudiando cómo las pacientes consiguen arreglárselas para comprender información médica contradictoria. Asimismo, el estudio de Veinot (2009) sobre la información entre enfermos de hiv/sida y sus familiares es destacable en cuanto a la teoría que propone sobre comportamiento informativo colaborativo. Si entendemos que la crianza y educación de los niños son una responsabilidad social en la que intervienen múltiples actores aunque con implicaciones diferentes, como madres, padres, familiares, profesionales médicos, profesionales de la educación, entre otros, la información sobre crianza debería contextualizarse en un marco colaborativo. Normalmente se estudia el comportamiento informativo colaborativo dentro de entornos organizacionales (Reddy et al., 2008), aunque recientemente se ha subrayado la necesidad de ampliar el foco de esta especialidad y Shah (2014) propone explícitamente aplicar el concepto de comportamiento informativo colaborativo al ámbito de la familia y la salud (2014: p. 233).

Tras la revisión de la literatura anterior, el objetivo del presente estudio es documentar el comportamiento informativo de un conjunto de madres residentes en varias Comunidades Autónomas y adheridas a la crianza con apego o crianza respetuosa e intentar comprender las circunstancias y características de su búsqueda de información. En una revisión citada al principio (Plantín y Daneback, 2009), los autores concluían comentando sobre la necesidad de un marco teórico que permita comprender este

nuevo fenómeno. El presente trabajo pretende contribuir a través de un análisis cualitativo a la definición de este marco teórico.

## 2. Metodología

A fecha de hoy, se han realizado 15 entrevistas semi-estructuradas, aunque la investigación sigue abierta. Los participantes se localizaron a través de varios canales, en lo específico dos proyectos educativos con sede en Madrid adheridos a la filosofía de la crianza con apego y unos grupos de Facebook que también manifiestan claramente su adhesión a ese estilo de crianza. La razón de centrarnos exclusivamente en madres adheridas a esta filosofía es que la crianza con apego concibe que la información es una condición de base para llevar a cabo una crianza adecuada. Los participantes se reclutaron a través de mensajes publicados en los grupos, contactando con los responsables de los proyectos mencionados arriba o por el boca a boca. Las entrevistas tuvieron una duración de entre 22 y 52 minutos aproximadamente. Se realizaron a través de Skype, se grabaron, se transcribieron, y posteriormente se analizaron con el software NVivo. Los datos se analizaron con el objetivo de desarrollar una teoría que explique el porqué de la búsqueda de información en este conjunto de madres. La propia estructura de las entrevistas ha proporcionado un primer marco de análisis, que se ha ido modificando teniendo en cuenta los temas emergentes. En todo el proceso de recolección y análisis de los datos un elemento importante para tener en cuenta es que las entrevistas se realizaron en cierto sentido desde dentro del colectivo estudiado, pues la investigadora es madre a su vez y participa, más o menos directamente, en los proyectos mencionados. Se trata de un aspecto recurrente en la investigación sobre varios aspectos de la maternidad. Papen (2013), en un estudio sobre prácticas de información de mujeres embarazadas, se incluye a sí misma en la muestra de participantes aportando datos de una auto-etnografía. Cooper y Rogers (2015), en el ámbito de la sociología feminista, analizan algunas de las problemáticas que les han surgido a la hora de investigar a otras madres. El hecho de que el investigador sea un *insider* – otra madre - de alguna manera facilita el trabajo porque aporta mayor conexión con las participantes y mejor capacidad de comprensión y reflexión. Por otro lado, da confianza y seguridad a unas participantes a menudo no familiarizadas con el trabajo de investigación, y contribuye además a crear un clima de reciprocidad especialmente importante en el trabajo de entrevista. Sin embargo, más allá del posible subjetivismo, aceptado en la investigación feminis-

ta, queda por determinar cuán honesto, ético y afectuoso el investigador debería ser, especialmente a la hora de seleccionar los aspectos de la investigación que quiera dejar fuera. En documentación, González-Teruel y Abad-García (2012) consideran que la teoría fundamentada, la estrategia de análisis que utilizamos en el presente estudio, conlleva necesariamente la inmersión del investigador en un universo informacional de las personas estudiadas, para comprender el cómo y el porqué de la interacción con la información. Realizar las entrevistas como un *insider* ha mejorado sin duda alguna la capacidad de comunicación de la investigadora con las participantes. En muchas de las grabaciones, es posible apreciar cómo, al principio del coloquio, las participantes empiezan tentativamente, algunas dudando sobre qué decir o si lo que están diciendo es correcto y bien encaminado. Reafirmar y reformular sus historias u opiniones les ha permitido ganar confianza en que efectivamente estaban contando algo interesante desde un punto de vista informacional. En cierto sentido, el terreno en común con las personas entrevistadas ha guiado el análisis de los datos y de alguna forma condicionado el resultado del trabajo. Por otro lado, los resultados se han contextualizado en el ámbito de la literatura, no solo documental, hecho que les confiere mayor credibilidad.

Aunque la invitación a participar en la investigación se dirigía indiferentemente a hombres y mujeres, todos los participantes fueron del sexo femenino. Nueve de las participantes tienen 1 hijo, tres de ellas 2 hijos y tres 3 hijos. La edad de los niños va del mes de vida hasta los siete años. La mayoría de las participantes estaba trabajando en el momento de las entrevistas pues solo dos de ellas no trabajaban fuera del hogar. En cuanto a nivel de educación, 8 de las participantes son licenciadas o graduadas, 1 post-graduada, 4 tienen una formación profesional de grado superior o medio, mientras que 2 tienen el título de bachillerato. La mayoría, 11, residen en la Comunidad de Madrid, 2 en Aragón, 1 en Galicia y 1 en Andalucía. El objetivo de las entrevistas era documentar el comportamiento informacional de las participantes de la manera más global posible, sin un foco especial en la web, como en mucha literatura actual sobre el tema. Las participantes hablaron sobre los temas que suelen buscar, las circunstancias que suelen desencadenar la necesidad de buscar información, la forma en la que valoran la información que reciben de diversos canales, incluidos los profesionales de la educación y los profesionales de la salud, sobre sus patrones de colaboración y puesta en común de la información, y finalmente sobre los criterios de evalua-

ción de la información que encuentran. En los extractos de las entrevistas los nombres se han cambiado intencionalmente.

### 3. Resultados y Discusión

Las participantes de este estudio confirmaron que efectivamente la búsqueda de información representa una constante en su experiencia de crianza y parentalidad, pues es la propia condición de madre que plantea todo un abanico de necesidades informativas. En el momento de convertirse en madres, las participantes experimentan un cambio en cuanto a la percepción de sí mismas. Por un lado la maternidad real es algo complementemente diferente a lo esperado, y por otro, el seguimiento del niño supone adaptarse a sus diferentes etapas de crecimiento. Alicia (madre de 2 niños de nueve meses y tres años) cuenta así su experiencia:

“[...] yo antes no me paraba a mirar nada prácticamente de tema niños, pero pensaba que era otra cosa completamente diferente y luego tuve a D. y me di cuenta de que no es como te lo dicen, como sale en la TV, como te dice la gente, tienes que buscar mucha información porque los profesionales hoy en día no te dan toda la necesaria. [...] desde que los he tenido he cambiado completamente mi vida entera... Sí, desde el momento parto no me esperaba nada de cómo iba a ser, ni con el niño ni con la niña. O sea, tú vas con una idea preconcebida y al final te sale todo al revés, en mi caso.”

Igualmente, la adaptación a las distintas fases de desarrollo del niño pone en cuestión constantemente el rol materno:

“Rocío (madre de un niño de 5 años): No sé... el desarrollo de los niños no es algo lineal, ¿no?, como que van avanzando, retrocediendo, como que van dos pasos para adelante, dos pasos para atrás, y en ese tipo de circunstancias en las que quizás dudas de si tu niño está teniendo un desarrollo normal o de si tú estás actuando bien de acuerdo a, cómo decirlo, las necesidades que van cambiando de tu niño, pues... sí que busco información para reafirmar o si me siento un poco perdida para leer otras experiencias, sí en este tipo de situaciones.”

Esta característica aparece también en el trabajo de Walker (2012) sobre un conjunto de padres y madres reclutados en varias escuelas de Leeds, en el cual el autor considera que “ser padres” de por sí supone la motivación principal para buscar información. Se trataría de una característica inherente a la condición de padre o madre. Papen (2013) habla de una “disposición para conocer” como características de las madres embarazadas de su estudio. Para el

conjunto de madres entrevistadas, la novedad y unicidad de la situación tiene como consecuencia que la percepción de uno mismo cambia y las participantes se definen en muchas ocasiones como “ignorantes”, “perdidas”, “no preparadas”, “vulnerables” o “inseguras”, especialmente al principio de la maternidad. Camila (madre de un niño de 14 meses) habla de pasar de la arrogancia por no querer ir a un curso de preparación al parto, a la ignorancia, después de tener al bebé, subrayando el cambio radical en cuanto a la percepción de uno mismo. Esta nueva percepción de uno mismo puede reforzarse en la interacción con los profesionales sanitarios, y no limitarse a los primeros meses de la maternidad, sino prolongarse, como confirma la historia de Inés (madre de 2 niños de tres y siete años), que subraya asimismo la relación entre esta nueva percepción de una misma y la información:

“Inés: [...] muchas veces sucede que en los centros de salud te tratan de ignorante o... es que es verdad que a veces eres ignorante ante muchas circunstancias y la facilidad que tiene ya internet, internet en los móviles que accedes inmediatamente a la información pues facilita esas búsquedas ¿no?”

Normalmente, estando a la narración de las participantes, los primeros problemas informativos tienen que ver fundamentalmente con salud, siendo la lactancia y el sueño del bebé los más comunes; de ahí que el pediatra sea a menudo el primer referente informativo. Más allá de eso, la necesidad de buscar información se sitúa más bien en unas situaciones conflictivas o de crisis no solo con los profesionales de la salud, sino también con los familiares, con el propio niño o con una misma. Puede tratarse de “conflictos existenciales”, como los define Génesis (madre de un niño de 7 años), hablando sobre una experiencia reciente en la que la información le ha permitido gestionar mejor las consecuencias emocionales de una mudanza, o de problemas que no es posible solucionar solo con las recomendaciones del pediatra, como una “crisis de lactancia”:

“Carol (madre de tres niños de 15 años, 4 años y 16 meses): [...] el caso es que yo al mayor le di pecho solamente un mes y medio - tiene quince años -, entonces era cada tres horas le das el pecho, veinte minutos en cada pecho, y entonces, claro, mi hijo a los dos meses dijo que tenía hambre, era una crisis de lactancia, ahora lo sé, entonces mi hijo tenía hambre, entonces yo dije, porque se llevan muchos años, entonces yo dije ‘con J. - que ha hecho 4 años ahora - yo le voy a dar el pecho’ y es más, sigue lactando, ahí es cuando empecé a buscar mucha in-

formación porque los pediatras no me ayudaban, los médicos no me ayudaban, entonces ahí comenzó mi curiosidad por la información por saber, por conocer porque no me aportaban nada sinceramente.”

Puede tratarse de una insatisfacción con respecto a una misma en la gestión del comportamiento del niño, temática que suele acaparar la atención de todas las madres una vez superados los problemas iniciales de alimentación y sueño:

“Michela: ¿Buscas más sobre la condición de madre que sobre los niños?”

Antonia (madres de tres de 2 meses, 3 y 5 años): Bueno, sobre maneras diferentes de tratar los conflictos con los niños... me veo épocas que soy super-gritona, que me paso el día gritándoles y me digo ‘esto no puede ser así, no es esto cómo quiero tratar a mis niños, no es como quiero que mis niños aprendan’. Entonces investigo modos alternativos de tratar el mismo problema, las cosas que... el día a día. O cuando está muy rebelde la pequeña por su edad y tal y a mí solo se me ocurre ‘que te dejo sin dibujos’, amenazas, y me digo ‘no, no, yo no quiero educar así, cómo puedo hacer para... cuál es la alternativa para llevar esto’. Sí, bueno a lo mejor sí que estoy buscando para modificar mi conducta y no la de los niños, ahora que lo dices, ahora que me estoy escuchando...”

El conflicto es implícito en la adopción de la crianza con apego como estilo de crianza, pues, según comentan las propias madres entrevistadas, se trata de una opción menos popular con respecto al estilo de crianza dominante y más tradicional, el conductismo, hecho que da lugar a discrepancias más o menos pronunciadas con personas de su entorno, incluidos los familiares y amigos de toda la vida. Es destacable que, en el estudio de Walker (2012), los familiares y amigos más allegados representan las primeras fuentes de información para los padres entrevistados por el autor británico. Las razones de la diferencia entre este estudio y el de Walker pueden encontrarse en el estilo de crianza específico adoptado por las madres del presente estudio, pero también en diferencias sociales y de educación más pronunciadas en la muestra de Walker. La historia de Brenda (madre de una niña de 5 años) es un ejemplo extremo pero significativo de conflictos con personas del entorno, aunque otras madres también denuncian ser a veces objeto de comentarios despreciativos por parte de familiares o amigos, que no comparten su concepción de la crianza.

“Brenda (madre de una niña de cinco años): De hecho cuando [los suegros] venían a casa y

veían los libros de crianza se reían de nosotros, no sé es como mucha falta de respeto, intenté dar una fiesta y casi todas mis amigas daban lactancia materna y mi suegra diciendo que 'qué asco', entonces hubo un momento en el que era tan marcada la diferencia que yo decidí un día enfrentarme, porque cogió S. a la fuerza y la llevó a lavar las manos con la niña gritando así, entonces le dije '¿Tú crees que es la manera de tratar a una niña? A lo mejor si se lo pides, se lava las manos' y se enfureció, me levantó la mano como para pegarme y yo dije... pues... ya está..."

La literatura sobre comportamiento informacional destaca la incertidumbre como estado emocional mayormente asociado con la búsqueda de información, mientras que aquí, como vemos, son más bien el conflicto o la crisis. La incertidumbre también puede manifestarse como una condición previa a la búsqueda de información, especialmente cuando las madres sienten la necesidad de reconfirmar que están actuando bien y que sus hijos se desarrollan en el ámbito de la normalidad, como confirma la experiencia de Aurora (madre de dos niñas de 2 y 4 años):

"Aurora: [...] yo me cuestiono mucho, no sé el resto de las madres, sí, sí lo sé que muchas madres hacen lo mismo que yo, yo me cuestiono mucho, cuestiono muchas veces si lo que estoy haciendo es correcto o no, si es correcto o si puedo hacerlo mejor, cuando ves que determinadas etapas evolutivas, yo por ejemplo... lo que te digo, yo tengo una hija de dos años y otra con cuatro. Ahora con dos años, estamos con todo el tema de... mi hija ha dejado el pañal muy pronto, ha dejado el pañal... empezó con año y medio, entonces bueno... hasta qué punto, yo me lo he planteado, ¿hasta qué punto estoy forzando yo la situación? ¿Es una cosa suya? ¿La tengo que dejar? ¿Realmente podemos esperar un poco más? Bueno, este tipo de cosas realmente te las preguntas porque no es lo habitual, cuando tienes hijas que no hacen cosas habituales, pues..."

Sin embargo, los resultados del presente estudio apuntan a que las situaciones que hacen de la búsqueda de información algo consustancial a la crianza, casi un hábito, son más bien las conflictivas o críticas. En un estudio con mujeres en otra época de transición existencial, la menopausia, Genuis (2012) destaca una vez más la incertidumbre como condición de base de la búsqueda de información, aunque menciona que las mujeres de su estudio se sienten responsabilizadas con respecto a la búsqueda de información por la existencia de información conflictiva y cambiante sobre su patología. El

tema de la información conflictiva es un tema emergente en la literatura médica, especialmente desde cuando los ciudadanos están accediendo a múltiples fuentes de información médica en la web. Hämeen-Anttila et al. (2014) estudian precisamente el tema de la multiplicidad de fuentes de información sobre medicación en el embarazo en un conjunto de mujeres embarazadas o con niños pequeños. Como consecuencia del conflicto entre fuentes de información la opción de muchas madres (el 43.6%) es la de no utilizar medicamentos.

Una consecuencia del conflicto, que podemos entender también como encubierto, es que la información sobre crianza normalmente se comparte solo dentro de un círculo de personas cuyas creencias y estilo de crianza son afines. Las madres entrevistadas, de esta manera, procuran evitar los comentarios y las críticas y, a la vez, intentan respetar otros puntos de vista. En la narración de Camila (madre de un niño de 14 meses), vemos la sorpresa de que las amigas más partidarias de un estilo de crianza tradicional, centrado en la madre más que en el niño, no hayan hablado nunca de su experiencia de maternidad con ella:

"Camila: Todas mis amigas en general han tenido niños antes, pero es muy curioso que - es una cosa que he comentado con mi marido - nadie habla después. Es cierto que todas mis amigas, de lo que es el parto... yo creo que hay una tendencia a... como hay un momento en el que solo hablamos de nuestros niños, nuestro parque y nuestras cosas, a las que en aquel momento éramos solteras o no teníamos niños, no nos cuentan ese tipo de cosas, 'pues he tenido un niño, he tenido un parto difícil, punto y ya está'. No te cuentan nada más. 'Me han dado tantos puntos'... también es cierto que todo mi grupo de amigas que han tenido niños antes eran... son de una crianza digamos tradicional, sabes, o por lo menos no hacen ver, que yo sepa, no colechan, han dado el pecho relativamente poco tiempo, o sea un poco como lo habitual."

Poner en común la información es algo que las participantes hacen constantemente tanto online como offline, pero solo con personas afines en cuanto a creencias y estilo de crianza. En este sentido, las comunidades virtuales o los grupos de crianza a los cuales muchas de las madres entrevistadas acuden permiten añadir una nueva dimensión al concepto de *information grounds* que Fisher y Naumer (2006) definen como cualquier lugar de interacción social en el cual la información se intercambie de manera espontánea. Se confirman, por otro lado, las connotaciones afectivas de la búsqueda de in-

formación colaborativa que Veinot (2009) encontraba estudiando un grupo de enfermos de SIDA/HIV. Esta autora además hace hincapié en que la “exposición a la información” es un fenómeno mediado por redes sociales, tanto online como offline, y que en estas redes la información se da y se recibe a la vez, fenómeno destacado también por las madres entrevistadas en este estudio.

Al hilo de estas cuestiones, cabe destacar que el conflicto del que hablan las participantes en este estudio se manifiesta muchas veces en el ámbito del sentir más que de la racionalidad, pues perciben una disconformidad entre lo que sienten y determinadas fuentes de información sean éstas el pediatra, los familiares, las propias experiencias o en general “las referencias que tienes [que] es lo que has vivido tú o lo que ves en otros”, como lo expresa Aitana (madre de una niña de 7 años y educadora infantil). Faircloth (2011, 2013) refiere que las participantes en su estudio comentan lo inexplicable que les resulta su experiencia de lactancia y define este “sentir” como “afecto”. Según la autora, el afecto más que las hormonas, el instinto y la intuición - que no sería sino una forma menos consciente de experiencia - puede justificar las prácticas no convencionales de lactancia de madres que amamantan a sus bebés más allá del año de edad. Refiriéndose al trabajo del neurocientífico Damasio (Faircloth, 2013: pp. 164-65), explica que el afecto, a diferencia de las emociones, no implica la clásica dicotomía entre mente y cuerpo o corazón y cabeza, que implican éstas, sino que se sitúa en un nivel previo de pre-consciencia: representa una reacción a un estímulo, al igual que las emociones, pero sin ser completamente consciente.

El tema del afecto es muy presente en la literatura sobre comportamiento informacional, aunque, estando a Fisher y Landry (2007) se limita más bien al ámbito de lo teórico. Éstas autoras analizan el tema del afecto en un conjunto de madres amas de casa, aunque lo estudian más bien como el entramado de emociones que coexisten con la búsqueda de información de estas mujeres, es decir que lo ven más bien como algo que se da contextualmente a la información. En el presente estudio el afecto emerge más bien como una fuente de información propiamente dicha o, en todo caso, como herramienta de evaluación y ponderación de la información. En este sentido, en la literatura informacional está emergiendo el concepto de “conocimiento intrapersonal”. Tidline (2005) comenta que la teoría del *sensemaking* de Dervin, citada para comprender la búsqueda de información sobre crianza (Jang et al., 2015), tam-

bién se ha utilizado para comprender el conocimiento intrapersonal, aunque no da referencias específicas. Genuis (2012) emplea la expresión “conocimiento intrapersonal” para definir un conocimiento comparable con las experiencias personales y que entra en juego en los procesos de evaluación de la información. De forma parecida, Papen (2013) subraya el papel esta vez de las emociones, como algo interior, en determinar la utilidad y validez de la información recibida por un grupo de mujeres embarazadas, concluyendo que “...their information practices are not purely rational processes” (sin número de página). En el presente estudio también, el afecto como conocimiento entra en juego especialmente a la hora de evaluar la información que procede de fuentes externas. Permite, por ejemplo, captar indicadores de fiabilidad de la información en niveles diferentes del racional como a nivel de percepción sensorial y de comportamiento, según cuenta Antonia (madres de tres de 2 meses, 3 y 5 años):

“Michela: ¿Y cómo valoras la información que te dan otras personas de tu entorno? Antonia: Pues, depende de la persona, de cómo gestiona sus propios problemas, sus propios... al final también es instinto, si el consejo que te están dando sobre una cosa concreta va acorde con lo que tú pensabas antes, te lo crees, te lo tomas. Y si de esa persona te fías mucho, porque crees que es una persona sensata, lo que sea, pues te fías, y si no te fías de esta persona ya de entrada, de la información que te den vas a escuchar poco o nada, al final yo supongo que esto lo hacemos todos como humanos clasificar primero a la fuente de información antes de escuchar la información. Si una madre que me parece que tiene a su hijo totalmente descontrolado y que le grita todo el tiempo y que no sé qué, no sé cuántos me está diciendo ‘huy, tienes que ir a este pediatra que es maravilloso’, seguramente a ese pediatra no me apetezca ir, directamente, porque lo que le parece maravilloso a esta señora que habla así a su hijo, que tiene esas ideas, seguramente no me parece maravilloso a mí.”

De manera parecida, la información de un pediatra que no se molesta en saludar a un niño que entra a la consulta carece de fiabilidad, como explica Marcela (madre de una niña de 14 meses), pues es muy probable que la información que vaya a dar no tenga en cuenta la peculiaridad y unicidad del paciente.

En la narración de Inés lo afectivo y la subjetividad como criterio de evaluación se justifica como una herramienta que compensa por la ignorancia en temas científicos:

“Inés (madre de 2 niños de 3 y 7 años): Sí, es un poco prejuicioso pero sí, la intuición de si me cae bien o no quien me lo está dando, si esa persona me ha agradado o me infunde confianza es más fácil que acepte esa información con mayor facilidad que aquella que no me infunde confianza, evidentemente eso es subjetivo y sí, es un poco prejuicioso, pero sí... [...] no tenemos toda la información, no tenemos todos los conocimientos, entonces muchas veces el descartar es en cuestiones de si me caes bien o no me caes bien, lo que te decía si una persona tiende más a tu ideología que otra, y a lo mejor la información que te está dando es verídica, es objetiva y eso, pero tú simplemente lo estás descartando porque no va concorde a tu línea de pensamiento.” El afecto conecta con la experiencia y las vivencias en el dominio de lo intrapersonal, según emerge de la narración de Aitana (madre de una niña de 7 años y educadora infantil):

“Aitana: Me suelo apuntar a que las cosas me las cuenten de viva voz y que yo pueda experimentarlas, pueda vivirlas, o sea, un artículo de arteterapia jamás te dice tanto como si tú lo experimentas, entonces lo que yo hago es ir a muchos cursos, ahí es donde busco más información, que alguien te lo cuente que haya tenido su propia experiencia...”

Michela: ¿Qué criterios de evaluación pones en juego para decidir si la información te sirve o no te sirve, es valiosa o no para ti?

Aitana: Sí, claro, lo que he venido diciendo, que sea... que la fuente sea fiable, me importa saber quién cuenta ciertas cosas, nunca me han importado los títulos, no busco si el que lo cuenta es médico, psicólogo, pedagogo, me da un poco igual eso... pero siempre hay otra historia, y ahora vuelvo al tema, pero yo, me puede gustar más leer, me puedo fiar más, alguien que por ejemplo diga ‘creó el proyecto de La Violeta, no sé dónde... después de tener un montón de niños, entonces quisieron hacer primaria, o se aventuraron a hacer escuela en el bosque, y llevan así 14 años...’, desde luego ya me atrae muchísimo más a que digan ‘licenciado, hizo pedagogía en Cambridge...’, te lo juro, a mí personalmente me llama muchísimo más una persona que tenga *recorrido vivencial, por lo que veo y oigo*, lo otro me da lo mismo, eso me importa, pero claro según lo que busque...”

Es un hecho que a la hora de evaluar la información, la mayoría de las madres entrevistadas coinciden: los criterios clave son de quién procede y qué concepción de la crianza aboga y, según veíamos más arriba, práctica. Algunas madres hablan de “sentido común”, o, con pala-

bras de Sofía (madre de una niña de 11 meses) “hacer lo que tu corazón te dice que es correcto si además encaja con tu sistema de creencias anterior”. Otro criterio menos mencionado sería la referencia a otras fuentes, a través de bibliografías, por ejemplo, o el contraste de fuentes de información, incluidas las menos afines a las creencias de la persona. Asimismo, muchas madres insisten en la necesidad de que la información tenga una base de evidencia científica. La propia filosofía de la crianza con apego se basaría en evidencias científicas aunque, según Faircloth (2010, 2013: p. 144 y siguientes), éstas serían cuestionables por no tener en cuenta los contextos socio-culturales en las que se han producido. Por ejemplo, los supuestos beneficios de la lactancia prolongada respecto a la alimentación a través de biberón en los países desarrollados, no tienen en cuenta que las madres que amamantan por más tiempo suelen pertenecer a grupos socio-económicos más privilegiados. Según Faircloth, la dependencia de la evidencia científica en la crianza con apego permitiría estudiar un aspecto de la vida social de la ciencia. Los datos recabados del presente estudio, sin embargo, no son suficientes para poder decir algo sobre este tema, excepto que para muchas participantes la ciencia y la experiencia representan dos áreas de conocimiento igualmente respetables, y que para algunas la experiencia puede ser incluso más valiosa que la ciencia. Por otro lado, la subjetividad de los criterios de evaluación es algo que el conjunto de madres entrevistadas en este estudio comparte con los participantes del estudio de Walker (2012). Para los padres y madres entrevistados en (Walker, 2012), la fiabilidad de la información se articula en varias dimensiones, entre las cuales están la confianza y la credibilidad, y la afinidad de valores y creencias. Walker destaca asimismo la trayectoria y experiencia de quienes estén dando información – características que reúne bajo “cualificaciones” -, hecho muy reiterado también por las madres de este estudio. De manera parecida, el estudio de Fisher y Landry (2007) sobre el comportamiento informacional de madres amas de casa subraya la importancia de la fiabilidad seguida por la comunidad de experiencias a la hora de recibir la información por otras madres en espacios comunes como escuelas y parques. Mucha literatura sobre búsqueda de información para la crianza subraya el papel importantísimo de las experiencias, pues conocer las experiencias de otras madres representa una de las razones para acudir a foros y grupos de discusión en la web. En la literatura, a menudo feminista, se habla cada vez más de “conocimiento experiencial” o “evidencia experiencial”. Según Doyle

(2013), que estudia un foro de crianza, el “conocimiento experiencial” puede suponer una forma de retomar control sobre el conocimiento en la vida diaria, y para muchas madres representaría “a reaction to the specialization and expropriation of knowledge in modern scientific systems” (p. 25). Zaslów (2012), estudiando una comunidad virtual de madres de niños con trastornos sensoriales, destaca cómo estas madres interactuando y a través de sus historias producen conocimiento con base en una comprensión de sus hijos mucho más profunda y global que la de cualquier profesional médico, muchas veces apoyada por lo que la autora denomina el instinto. Este conocimiento no pretende sustituir de ninguna manera a los profesionales de la salud y la evidencia médica, pero es un hecho que para muchas madres es sin duda alguna un valioso complemento. Según Papen (2013), mezclar y utilizar conocimiento científico u obtenido de profesionales con conocimiento informal proveniente de otras madres o de la experiencia de otros sería una estrategia con la cual las madres desarrollan o defienden su “autoridad cognitiva”. Con respecto a la importancia del conocimiento experiencial en este estudio y en otros, cabe mencionar que la web tiene un papel determinante. Las experiencias que valoran las madres a la hora de buscar información proceden de personas afines en cuanto a creencias y estilo de crianza, que muchas veces es imposible encontrar en el entorno más próximo. Se trata de algo que subraya prácticamente la totalidad de madres entrevistadas. La web en este sentido ofrece como una ventana al mundo que permite seleccionar exclusivamente la experiencia de aquellas personas que las participantes consideran más afines.

Los aspectos que hemos venido apuntando como salientes en el comportamiento informacional de las madres entrevistadas, de alguna manera pueden interpretarse como, en términos de Faircloth (2010, 2013), “trabajo de identidad”. Faircloth emplea la expresión en un estudio de las prácticas de amamantamiento en un conjunto de madres vinculadas a la Liga de la Leche y a los ideales de la crianza con apego. Si el concepto de identidad a secas no está claramente definido en la literatura, el de “trabajo de identidad” apuntaría, según Faircloth, a un proceso activo de construcción de la identidad en el cual intervienen asimismo factores sociales, configurándose como “[...] the range of activities in which individuals engage to create, present and sustain personal identities, with particular reference to the constitution of relatedness” (2013: 32). Los estudios de Faircloth tienen en cuenta la información en el trabajo de identidad maternal, especialmente la información científica. La

información fortalece a las madres estudiadas por la antropóloga ofreciéndoles una elección informada que les permite defender sus estilos de crianza y amamantamiento no convencionales, como por ejemplo la lactancia materna más allá del año de edad del niño, ante la sociedad. En cierto sentido, podríamos decir que la búsqueda de información tanto experiencial como científica representa una de las actividades finalizadas al trabajo de identidad. En este sentido podríamos además encontrar referentes en la literatura sociológica sobre la condición de padre y madre. Las repercusiones negativas de la paternidad y maternidad en la calidad de vida que viene poniendo de manifiesto cierta literatura de corte economicista, como por ejemplo (Stanca, 2012), chocan, según Hansen (2012) con la opinión popular de que los hijos aportan beneficios emocionales y felicidad en la existencia de sus padres. Según el autor noruego, esta aparente contradicción se resuelve cuando en lugar de hablar de felicidad hablamos de “significado”. Los hijos aportarían significado en la existencia de sus padres, entendiendo que: “Meaning can be defined as having a sense of purpose and direction in life, and that one’s activities and efforts make sense and is part of something larger than oneself.” (Hansen, 2012: p. 50).

Desde el punto documental también, encontramos confirmaciones de que la búsqueda de información en la crianza puede conformar una de las actividades del trabajo de identidad, pues las madres entrevistadas mencionan a menudo los libros como herramienta informativa. Las menciones a los libros destacan especialmente frente a la práctica ausencia de referencias a revistas como fuente de información o a la prensa generalista que solo aparecen en la narración de las participantes con connotaciones negativas. En este sentido se confirman los resultados de Papen (2013), que estudia a un conjunto de mujeres con niveles educativos muy alto, destacando asimismo una peculiaridad de este estilo de crianza con respecto a los más convencionales, pues otros estudios apuntan a la importancia de la revista como fuente de información. Concretamente, Castelló Egea et al. (2011) en un estudio sobre fuentes de información de los padres con respecto a la salud de los hijos realizado en un Centro de Atención Primaria de Barcelona, mencionan las revistas de crianza como segundo recurso más popular tras internet. La gran mayoría de las madres entrevistadas en este estudio opta más bien por la lectura de artículos difundidos en la web o publicados en blogs. La revista por su vocación comercial y publicitaria no suele gustar a las madres entrevistadas. Aunque las motivaciones

por optar por los libros son varias, entre otras la fiabilidad y amplitud de la información, el libro marca a menudo la iniciación en la filosofía del apego – muchas la conocen a través de *Un regalo para toda la vida* de Carlos González - y representa un lugar a donde volver en caso de dudas. El libro se ha definido en la literatura documental, entre otras cosas, como un espacio de definición de la identidad personal. Ross (1999) habla de una relación recíproca entre la interpretación del texto y la auto-interpretación, y de cómo la comprensión del texto se puede convertir en una fuente de consuelo, confirmación de identidad y apoyo para el cambio.

#### 4. Conclusiones

Entrevistando a un conjunto de madres adheridas al estilo de crianza con apego con respecto a su comportamiento informacional, hemos analizado una serie de temas que se relacionan con la literatura existente y que plantean asimismo nuevas áreas de investigación. Se ha puesto de relieve cómo la condición de madre, en el conjunto de participantes estudiadas, supone una nueva percepción de una misma como mayormente necesitada de información que se refuerza en situaciones “conflicto” tanto con los profesionales médicos, como con familiares y personas del entorno, el propio niño o una misma. Este conflicto se manifiesta como una disconformidad con respecto a la base informativa que sustenta formas más convencionales de criar a los hijos. En las historias de las participantes, la forma tradicional de criar a los hijos, incluso cuando la patrocinan profesionales de la salud, choca con lo que las madres participantes sienten, con la experiencia de otras madres en situaciones parecidas o con la evidencia científica. En ausencia de respuestas completamente claras, estas madres defienden el subjetivismo de sus decisiones, conscientes de que sus tomas de decisiones siempre cuentan con un margen de desconocimiento. El subjetivismo coexiste con la evidencia científica, e implica que a la hora de evaluar la información que reciben de diversas fuentes entran en juego factores distintos a los racionales, que en este estudio, teniendo en cuenta la literatura previa, hemos definido afectivos. Sin duda, la investigación futura tendrá que comprender mejor el conocimiento experiencial y el afecto como fuente de información, especialmente en su relación con el conocimiento científico. La recepción de la ciencia por parte de actores sociales, en este caso madres (y padres) también es un tema que merece mayor atención en la investigación futura.

#### Bibliografía

- Attachment Parenting International. *Bienvenido a API*. Recuperado el 8-07-2015 de: <http://www.attachmentparenting.org/espanol>
- Cooper, L., Rogers, C. (2015). Mothering and 'insider' Dilemmas: Feminist Sociologists in the Research Process. *Sociological Research Online*, 20(2), 5.
- Coronado Ferrer, S., Peset Mancebo, F., Ferrer Sapena, A., González de Dios, J., Aleixandre-Benavent, R. (2011). Web 2.0 en medicina y pediatría (y II). *Acta Pediátrica Española*, 69(2), 79-87.
- Doyle, E. (2013). Seeking advice about children's health in an online parenting forum. *Medical Sociology Online*, 7(3), 17-28.
- Egea, A. C., del Riego, M. R., García, I. M., García, A. G., Casteliví, M. M., Bordes, L. G. (2011). Fuentes de información de los padres respecto a la salud de sus hijos. *Acta Pediátrica Española*, 69(10), 450.
- Faircloth, C. (2010). 'What Science Says is Best': Parenting Practices, Scientific Authority and Maternal Identity. *Sociological Research Online*, 15(4), 4.
- Faircloth, C. (2011). 'It feels right in my heart': affective accountability in narratives of attachment. *The Sociological Review*, 59(2), 283-302.
- Faircloth, C. (2013). *Militant lactivism? Attachment parenting and intensive motherhood in the UK and France*. New York/Oxford: Berghahn Books, 266 p., ISBN 978-0-85745-758-5.
- Fisher, K. E., Naumer, C. M. (2006). Information grounds: Theoretical basis and empirical findings on information flow in social settings. En Spink, A., Cole, C. (eds.) *New directions in human information behavior*. Springer: Netherlands, p. 93-111.
- Fisher, K. E., Landry, C. F. (2007). Understanding the information behavior of stay-at-home mothers through affect. En: Nahl, D., Bilal, D., *Information and emotion. The emergent affective paradigm in information research and theory*. Medford, New Jersey: Information Today Inc. 978-1-57387-310-9.
- Fox, S., Duggan, M. (2013), *Health Online 2013*. Pew Research Center Report. Recuperado el 9/07/2015 de <http://www.pewinternet.org/2013/01/15/health-online-2013/>
- Genuis, S. K. (2012). Constructing “sense” from evolving health information: a qualitative investigation of information seeking and sense making across sources. *Journal of the American Society for Information Science and Technology*, 63(8), 1553-1566.
- Gibson, L., Hanson, V. L. (2013). 'Digital motherhood': how does technology support new mothers. En: *CHI '13: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. (pp. 313-322). New York: Association for Computing Machinery. 10.1145/2470654.2470700
- González-Teruel, A., Abad-García, M. F. (2012). Grounded theory for generating theory in the study of behavior. *Library & Information Science Research*, 34(1), 31-36.
- Hämeen-Anttila, K., Nordeng, H., Kokki, E., Jyrkkä, J., Lupatelli, A., Vainio, K., & Enlund, H. (2014). Multiple information sources and consequences of conflicting information about medicine use during pregnancy: a multinational Internet-based survey. *Journal of Medical Internet Research*, 16(2).
- Hansen, T. (2012). Parenthood and happiness: A review of folk theories versus empirical evidence. *Social Indicators Research*, 108(1), 29-64.

- Hays, S. (1996). *The cultural contradictions of motherhood*. New Heaven, CT and London: Yale University Press.
- Jang, J., Dworkin, J., Hessel, H. (2015). Mothers' Use of Information and Communication Technologies for Information Seeking. *Cyberpsychology, Behavior, and Social Networking*, 18(4), 221-227.
- Lima-Pereira, P., Bermúdez-Tamayo, C., Jasienska, G. (2012). Use of the Internet as a source of health information amongst participants of antenatal classes. *Journal of Clinical Nursing*, 21(3-4), 322-330.
- Lopez, L. K. (2009). The radical act of 'mommy blogging': redefining motherhood through the blogosphere. *New Media & Society*, 11(5), 729-747.
- McDaniel, B. T., Coyne, S. M., Holmes, E. K. (2012). New mothers and media use: Associations between blogging, social networking, and maternal well-being. *Maternal and Child Health Journal*, 16(7), 1509-1517.
- Medina, S., Magnuson, S. (2009). Motherhood in the 21st century: Implications for counselors. *Journal of Counseling & Development*, 87(1), 90-96.
- Niela-Vilén, H., Axelin, A., Salanterä, S., Melender, H. L. (2014). Internet-based peer support for parents: A systematic integrative review. *International Journal of Nursing Studies*, 51(11), 1524-1537.
- Papen, U. (2013). Conceptualising information literacy as social practice: a study of pregnant women's information practices. *Information Research*, 18(2).
- Pedersen, S., Smithson, J. (2013). Mothers with attitude—How the Mumsnet parenting forum offers space for new forms of femininity to emerge online. *Women's Studies International Forum*, 38 (May-June), 97-106.
- Peset Mancebo, M. F., Ferrer Sapena, A., Coronado Ferrer, S., González de Dios, J., Aleixandre-Benavent, R. (2011). Web 2.0 en Medicina y Pediatría (y II). *Acta pediátrica española*, 69(2), 79-87.
- Plantin, L., Daneback, K. (2009). Parenthood, information and support on the internet. A literature review of research on parents and professionals online. *BMC Family Practice*, 10(1), 34.
- Reddy, M. C., Jansen, B. J. (2008). A model for understanding collaborative information behavior in context: A study of two healthcare teams. *Information Processing & Management*, 44(1), 256-273.
- Ross, C. S. (1999). Finding without seeking: the information encounter in the context of reading for pleasure. *Information Processing & Management*, 35(6), 783-799.
- Schoenebeck, S. Y. (2013, June). The Secret Life of Online Moms: Anonymity and Disinhibition on YouBeMom.com. EN: *Proceedings of the 7th International AAAI Conference on Weblogs and Social Media*. Cambridge, Massachusetts, USA, July 8–11, 2013: AAAI Press, Palo Alto, California.
- Shah, C. (2014). Collaborative information seeking. *Journal of the Association for Information Science and Technology*, 65(2), 215-236.
- Stanca, L. (2012). Suffer the little children: Measuring the effects of parenthood on well-being worldwide. *Journal of Economic Behavior & Organization*, 81(3), 742-750.
- Tidline, T. J. (2005). Dervin's sense-making (pp. 113-117). Fisher KE, Erdelez S. et McKechnie LEF (éds.), *Theories of information behaviour*, Medford, NJ, Information Today, xxii-431 p.
- Torres, A., Suárez, A., & Rodrigo, M. J. Educar en Positivo: Primeros resultados y retos de futuro. *Revista Iberoamericana de Sistemas, Cibernética e Informática*, 11(2), 1-13.
- Veinot, T. C. (2009). Interactive acquisition and sharing: Understanding the dynamics of HIV/AIDS information networks. *Journal of the American Society for Information Science and Technology*, 60(11), 2313-2332.
- Webb, L. M., & Lee, B. S. (2011). Mommy blogs: The centrality of community in the performance of online maternity. En Moravec, M. (ed.) (2011) *Motherhood online*. Cambridge: Cambridge Scholar Publishing, 244-257.
- Zaslow, E. (2012). Revalorizing feminine ways of knowing: the challenge to biomedical epistemology in an online mothers' health community. *Information, Communication & Society*, 15(9), 1352-1372.

---

# Formação discursiva, sistemas de classificação e o discurso estrangulado de Charlie Hebdo: impactos, desdobramentos e implicações na representação da informação da ciência da religião no século XXI

*Discursive formation, classification systems and the strangling discourse of Charlie Hebdo: impacts, consequences and implications for the representation of information on the science of religion in the twenty-first century*

---

Rosana Matos da Silva Trivelato (1), Maria Aparecida Moura (2)

(1) Escola de Ciência da Informação, rosanatrivelato@gmail.com

(2) mamoura@ufmg.br

## Resumo

Os processos de globalização colocaram em evidência o multiculturalismo, doutrina que reconhece a diversidade cultural como uma parte das sociedades políticas, e também as marcas das visões de mundo pautadas por clivagens religiosas, étnicas e socioeconômicas. Como estes diferentes discursos são evidenciados em nossos sistemas de representação da informação? Seria oportuno então, investigar se os sistemas de classificação da informação dão conta das formas de representar a religião no alvorecer do século XXI. Este artigo tem o objetivo de compreender a pertinência da concepção de formação discursiva na área de ciência da religião em interface com as dimensões sócio-históricas e seus impactos e implicações em sistemas de classificação. Os objetivos específicos são: (1) situar inicialmente as superfícies de emergência, onde surge o objeto do discurso, fazê-lo nomeável e descritível; (2) descrever as instâncias de delimitação e regulamentação; quem define o objeto do discurso; (3) analisar as grades de especificação, especificidade. O percurso metodológico empregado é um estudo de caráter exploratório, utilizando o método teórico-investigativo. O corpus empírico é a classe 2, religião, na Classificação Decimal de Dewey (CDD) e na Classificação Decimal Universal (CDU). O propósito deste estudo foi explorar inicialmente questões para análise futura.

**Palavras chave:** Formação Discursiva. Sistemas de classificação bibliográfica. Classificação Decimal Universal. Classificação Decimal de Dewey. Ciência da Informação. Ciências da Religião.

## 1. Introdução

A violência política que culminou em recente ataque e assassinato de profissionais do jornal satírico francês Charlie Hebdo, de policiais e

## Abstract

The processes of globalization have in evidence the multiculturalism doctrine that recognizes cultural diversity as a part of political societies, and also the prints of worldviews guided by cleavages religious, ethnic and socioeconomic. As these different discourses are evidenced in our representation of information systems? Then it would be appropriate, to investigate whether the classification of information systems realize the ways of representing religion at the dawn of the twenty-first century. This article aims to understand the relevance of the concept of discursive formation in the science of religion area of linking with the socio-historical dimensions and its impact and implications for classification systems. The specific objectives are: (1) initially identify the emergency content where the discourse object arises, do it identifiable and describable; (2) describe instances of delimitation and regulation; who defines the address of the object; (3) to analyze the specification grids, specificity. The methodology approach is a exploratory study, using the theoretical and investigative method. The empirical corpus is the class 2, religion, Dewey Decimal Classification (DDC) and the Universal Decimal Classification (UDC). The purpose of this study was initially explore issues for future analysis.

**Keywords:** Discursive formation. Systems of bibliographic classification. Universal Decimal Classification. Dewey Decimal Classification (DDC). Information Science. Religion Science.

judeus na França, inundou a mídia de protestos que vão desde as marchas e declarações reivindicando "Je suis Charlie Hebdo" a movimentos de repulsa ao que consideram que a livre

expressão do veículo de comunicação representa um deboche agressivo aos muçulmanos. O jornal assume o discurso de defensor da liberdade de expressão e denunciante de atraso religioso. Em contrapartida muitos acusam o seu discurso como sendo uma provocação irresponsável em tempos de violência e instabilidade no mundo islâmico.

Nesse contexto, revela-se mais atual do que nunca tecer pontes entre a concepção da formação discursiva em sistemas de representação da informação e o campo religioso. Assim, este artigo tem por objetivo verificar se os sistemas de classificação bibliográfica corroboram intencionalmente para promover abordagens de tolerância entre credos, etnias e se dão conta das formas de representar a religião no alvorecer do século XXI.



Figura 1. Fonte: [http://www.playbacpresse.fr/documents/charlie/mq\\_charlie\\_5.pdf](http://www.playbacpresse.fr/documents/charlie/mq_charlie_5.pdf)

A fim de atualizar o debate e aproximá-lo dos desafios no âmbito da informação escolhemos o domínio cultural consolidado: classe de Religião, e, a sua acomodação na Classificação Decimal de Dewey (CCD) e na Classificação Decimal Universal (CDU), bem como as clivagens religiosas, étnicas e socioeconômicas com diferentes discursos estão evidenciados. Por fim, realiza-se uma avaliação sobre a pertinência do estudo em relação aos objetivos traçados e às reflexões construídas para seu desenvolvimento.

## 2. Formação discursiva

As obras de Michel Foucault "A ordem do discurso", de 1971 e Arqueologia do saber, 1972 sugerem um exercício de deslocamento metodológico para dar-se conta do discurso.

[...] "não se pode reconstruir um sistema de pensamento a partir de um conjunto definido de discursos. Mas esse conjunto é tratado de tal maneira que se tenta encontrar, além dos próprios enunciados, a intenção do sujeito falante, sua atividade consciente, o que ele quis dizer, ou ainda o jogo inconsciente que emergiu involuntariamente do que disse ou da quase imperceptível fratura de suas palavras manifestas; de qualquer forma, trata-se de reconstruir um outro discurso [...] restabelecer o texto miúdo e invisível que percorre o interstício das linhas escritas e às vezes, as desarruma." (Foucault, 2012, p.33-34)

Para Foucault, os membros de uma sociedade estão imersos no discurso e nos regimes de regulação discursivas que lhes dão um direcionamento e não há como ficar imune a isso. O discurso pode ser regido um conjunto de regras anônimas e ser usado para designar os modos de pensamento que caracterizam as instituições, os domínios da cultura, distinguir diferentes áreas de estudo, para identificar a linguagem de diferentes grupos sociais ou ocasiões. Essas práticas discursivas estarão sempre determinadas no tempo e no espaço, em uma dada época e destinadas a uma área social, econômica, geográfica ou linguística.

Foucault (2010) assinala que o caráter desconcertante do discurso está na descontinuidade, na necessidade de assumir recortes provisórios e de escolher domínios de análise em que as relações corram o risco de serem numerosas e densas.

Nesse contexto as formações discursivas são enunciados diferentes em sua forma, dispersos no tempo que formam um conjunto quando se referem a um único e mesmo objeto.

Segundo Foucault, [...] "a produção do discurso é ao mesmo tempo controlada, selecionada, organizada e redistribuída por certo número de procedimentos que têm por função conjurar seus poderes e perigos, dominar seu acontecimento aleatório, esquivar sua pesada materialidade". O poder se manifesta como resultado da vontade de estruturar o campo possível da ação dos outros. Assim, os discursos dão conteúdo e forma a um conjunto de relações de poder entre uma variedade de indivíduos e instituições. (FOUCAULT, 2013: 8-9)

Na análise do discurso considera-se as maneiras em que objetos e ideias são expostos, ao invés da ontologia dos objetos e as próprias ideias. Pode suscitar perguntas como: o que torna alguns discursos dominante, e deixa outros permanecerem marginalizados? Quais são as instituições através das quais esses discursos são organizados, divulgadas e legitimados?

A instituição da biblioteca tem um papel fundamental a desempenhar na manutenção e legitimação de discursos particulares em detrimento dos outros (Radford, 1992). Um exemplo disso é o ato de classificar um item informacional, processo desde a escolha da classificação bibliográfica institucionalmente constituída a escola no número notacional, usar uma classificação bibliográfica permite que a biblioteca enquadre discursivamente itens informacionais.

Frohmann defende a pertinência do uso da análise do discurso como um método multidisciplinar capaz de fornecer uma contextualização, uma das principais vantagens da pesquisa qualitativa. No âmbito da Biblioteconomia e Ciência da Informação (library and information science - LIS), a formação discursiva de Michel Foucault é um método de pesquisa útil.

"Pelo menos desde 1876 até os dias atuais, os discursos de LIS são completamente interligados com formas institucionais específicas, através do qual o poder sobre a informação, os seus usuários, e seus usos é, foi, e continuará a ser exercida. Estes discursos incluem fala especializada sobre a informação, a sua organização, quem a usa e quem não, o que seus usos são, foram ou poderiam ser, os papéis sociais e culturais das organizações responsáveis por isso, as análises introspectivas do profissional e até mesmo pessoais, identidades dos seus detentores, e os pronunciamentos programáticos dos seus teóricos que falam sobre como essas coisas devem ser faladas" (FROHMANN, 1994 : p.121).

Segundo Foucault, para entender a formação discursiva é necessário elucidar as regras de formação do objeto, seu regime de existência enquanto objetos do discurso. Nesse sentido ele aponta três direções: a) demarcar inicialmente as superfícies de emergência, onde surge, "a possibilidade de limitar seu domínio, de definir aquilo de que se fala, de dar-lhe o status de objeto - ou seja, de fazê-lo aparecer, de torná-lo nomeável e descritível"; b) a descrição das instâncias de delimitação, quem define o objeto do discurso; c) analisar as grades de especificação, especificidade, "trata-se dos sistemas segundo os quais opomos, associamos, reagrupamos, classificamos, derivamos" (FOUCAULT, 2012, p.50-51).

### 3. Classificar

Un resumen proporciona información sobre O ato de classificar é uma preocupação que surge com a própria evolução do homem, que anseia por criar mecanismos de organização e representação do pensamento. Como aborda POMBO:

[...] "nada nos parece mais "natural", óbvio e indiscutível que as classificações dos entes, dos factos e dos acontecimentos que constituem os quadros mentais em que estamos inseridos. Elas constituem os pontos estáveis que nos impedem de rodopiar sem solo, perdidos no desconforto do inominável, da ausência de "idades" ou "geografias" (POMBO, 1998: p.19)

A nossa visão de mundo, as tentativas de organizar e representar o conhecimento são impregnadas de valores em contexto determinados. Portanto, quando classificamos estamos categorizando pelo nosso viés, clivados pelo nosso pensamento e pelo nosso contexto socio-histórico.

Barbosa (1969) define a classificação como um processo que reúne o livro em grupos segundo o seu assunto e enquadrando-os em um sistema pré-estabelecido. Dando desta forma "um lugar certo na coleção".

Do ponto de vista religioso, Eco (2000, p. 15) assinala que interpretação hermenêutica pode introduzir viés à compreensão religiosa.

"Historicamente, o fundamentalismo religioso é um princípio hermenêutico, ligado à interpretação de um livro sagrado. O fundamentalismo ocidental moderno nasce nos meios protestantes dos Estados Unidos, no século XIX, e caracteriza-se pela vontade de interpretar literalmente as Escrituras, em particular no que concerne às observações sobre cosmologia, quando a ciência da época parecia negar a veracidade do conto bíblico. Onde, é claro, a recusa frequentemente intolerante a qualquer interpretação alegórica e a qualquer forma de educação que ponha em dúvida as Escrituras - como se verificou por ocasião da polêmica sobre a teoria de Darwin."

De acordo com Apostel (apud POMBO, 1998, p. 19-33), a classificação apresenta as seguintes características,

1. cada classificação tem por detrás um determinado mecanismo classificador que executa, melhor ou pior, as operações necessárias à classificação,
2. cada classificação persegue uma mais ou menos sistemática multiplicidade de fins que, em última análise, vão determinar a sua estrutura,
3. cada classificação exerce-se sobre um domínio da realidade cujas estruturas internas tornam mais ou menos fácil as operações necessárias à classificação,
4. cada classificação constrói-se no contexto das classificações precedentes do mesmo domínio, ou seja, há uma inexorável historicidade das classificações ao longo da qual os domínios classificados podem ser modificados, as divisões podem ser completadas, novos critérios de classificação podem ser acrescentados,

5. para cada classificação existe um produto externo da atividade classificadora que se apresenta como uma árvore genealógica mais ou menos regular, isto é, toda a classificação supõe uma dupla operação: o estabelecimento de equivalências entre classes do espaço classificatório global; o estabelecimento de hierarquias entre subclasses no interior das classes previamente estabelecidas.

A classificação bibliográfica, também constitui-se de um instrumento verbal de representação da informação e a forma de organizar os registros do conhecimento. As questões sobre organização desses registros, e a própria organização do conhecimento, são abordadas pela Biblioteconomia e Ciência da Informação, campos de estudos para resolver os problemas de ordenação e acesso ao conhecimento

### 3.1. Superfície de emergência das classificações

O processo de classificação dos saberes e dos seres de Aristóteles embasou os sistemas de classificação bibliográficos, isso objetivando simplificar e facilitar o livre acesso a documentos. As classificações bibliográficas não deixam de ser classificações do conhecimento e utilizam de termos originalmente filosóficos resignificando e adaptando as especificidades da classificação pretendida. (ANJOS, 2008)

Historicamente, conhecemos o trabalho precursor de Ranganathan que introduziu o termo faceta e a Classificação de dois pontos e Dewey com o sistema numérico decimal.

Contudo, Esteibar (1984) analisa o manuscrito "Idée, d'une bibliothèque" escrito por Leibniz provavelmente no final do século XVII e o trabalho de Ott Christoph Hilgenberg intitulado "Zur Entstehung", e sugere que Dewey conheceu os intentos de Leibniz.

Segundo Esteibar, Leibniz enfoca a importância da do intercâmbio universal do conhecimento para o progresso da ciência e postula a biblioteca como um lugar de uso. A conjunção entre ciência, técnica e entretenimento resultando em uma biblioteca enciclopédica.

A proposta de Leibniz, ainda segundo Esteibar, utilizava uma notação numérica universal para possibilitar o intercâmbio de conhecimento entre sábios de diferentes países.

Dewey enfatizou meios eficientes para alcançar a missão cultural da biblioteca.

[...] 'Durante meses imaginei, noite e dia, que deveria haver, em alguma parte, uma solução satisfatória. No futuro teríamos milhares de bibliotecas, a maioria das quais aos cuidados de pessoas com pequena capacidade ou escasso adiestramento. O

primeiro requisito da solução há de ser a maior simplicidade possível. O provérbio dizia 'simples como a, b, c', mas ainda mais simples do que isso era 1, 2, 3. Após meses de estudo, um domingo, durante um longo sermão do pastor Stearns, enquanto o encarava com firmeza sem lhe ouvir uma palavra, e absorvida a mente no problema vital, a solução coruscou-me ante os olhos, a ponto de fazer-me saltar da minha cadeira e de quase levar-me a gritar 'Eureka!' Consistia ela em alcançar a simplicidade absoluta empregando os mais simples símbolos conhecidos, os algarismos arábicos na qualidade de decimais, com o zero revestido da significação usual, a fim de numerar uma classificação de todo o conhecimento humano impresso; completando tal expediente pelos símbolos que, depois dos números, eram os mais simples, ou seja, a, b, c, e indexando todos os títulos das tabelas, de forma que seria mais fácil usar a classificação com 1.000 títulos assim ajustados, do que os 30 ou 40 títulos que reclamam cuidadoso estudo antes de serem utilizados." (SCHELLENBERG, 1980, p. 11) apud SOUSA, 2007, p. 12-13)

Segundo Frohmann seus procedimentos tecnoburocráticos correspondem uma ordem cultural em que a informação desejada por leitores é construída pelos elementos de uma classificação burocratizada, mecanizada e racionalizada de assuntos do livro que não são controladas por um guardião da alta cultura tradicional, como almejado por ele, a autoridade cultural passou a gestores profissionais de capital intelectual.

### 3.2. Instâncias de delimitação

A Classificação Decimal de Dewey (CDD), foi idealizada por Melvil Dewey, bibliotecário norte-americano de visão pragmática e funcionalista. O intento de Dewey era criar uma ferramenta prática de fácil aplicação de forma responder as questões pragmáticas de organização da informação com vistas rapidez de na localização do item e o livre acesso as estantes. Essencialmente a CDD é uma classificação com base no conhecimento, estruturada hierarquicamente, constituindo de uma série de dez classes principais que correspondem a disciplinas tradicionais ou área de estudo, subdivididas sucessivamente. O desenvolvimento é ilimitado. Quanto aos princípios de divisão parecem provir da lógica, da tradição e da prática (ANJOS, 2008, p. 170).

A Classificação Decimal Universal (CDU), concebida por Paul Otlet e Henry La Fontaine como intuito de organizar um repertório bibliográfico universal de trabalhos publicados em qualquer lugar do mundo desde a invenção da imprensa. Otlet não pensava na ordenação de livros na estante e sim no estabelecimento de relações entre dados informacionais visando o uso da

informação. Adaptado do sistema de Dewey, a CDU retoma as classes principais da CDD (reflete disciplinas tradicionais) e a sua notação decimal. A classificação de estrutura hierárquica segue uma ordem sistemática que parte do geral para o particular, do todo para a parte, do gênero para a espécie etc., porém se afasta do modelo original (ANJOS, 2008, p. 178).

Anjos apresenta uma sumarização em forma de tabela que nos possibilita visualizar as principais características dos sistemas de classificação: Classificação Decimal de Dewey (CDD), Classificação Decimal Universal (CDU) e a Classificação da Biblioteca Library of Congress (LCC).

| Data      | Classificação | Características   |
|-----------|---------------|---|
| 1876      | CDD           | Hierarquia geral;<br>Racionalismo rege a divisão;<br>Notação decimal pura;<br>Divisão do conhecimento em 9 classes principais.  |
| 1905-1907 | CDU           | Hierarquia universal;<br>Notação não rigorosamente decimal, mas mista, praticada em função da necessidade;<br>Divisão do conhecimento em 9 classes principais;<br>Uso de procedimentos sintáticos que multiplicam as possibilidades de uso. |

Tabela 1: Principais características dos sistemas de classificação. Fonte: ANJOS, 2008, p. 226.

### 3.3. Grades de especificação

A aplicação do conceito de formação do discurso pode fornecer novos elementos de análise de sistemas de classificação. Imaginado livros arranjados em prateleiras de uma biblioteca por um profissional qualificado, pode-se legitimamente perguntar o por que dos itens estarem organizados de uma forma e não outra. Certamente o bibliotecário responderia que os livros são organizados de acordo com a proximidade do assunto e de acordo com o sistema de classificação adotado.

Coleções em bibliotecas podem, por exemplo, ser organizados dentro dos dez programações do sistema de Classificação Decimal de Dewey, ou seja, classe 000-900. A ideia de uma formação discursiva encena o mesmo princípio que o arranjo de livros em uma prateleira. Há algo além dos próprios documentos que permite que o catalogador reúna ou disperse um grupo de documentos.

Plácida faz uma breve abordagem sobre a cobertura das religiões modernas e religiões

étnicas não são cobertas pela LCC nos 2 primeiros níveis da classificação hierárquica. A CDU se refere às religiões modernas e étnicas em uma subclasse (290), já a CDU sob o título estranho de “religiões não providas para outro lugar” (299). Aparentemente, esta é uma cobertura incompleta e não sistemática. A CDU se refere à religião moderna na classe dos movimentos espirituais modernos. Religiões étnicas são cobertas pela CDU nos cultos menores e religiões (25), mas a cobertura é

Um estudo feito por Eisenhart (1960), aborda sobre as dificuldades das bibliotecas teológicas americanas devido a inadequação da classe de Teologia aos sistemas de classificação CDU, CDD e LC. Nesse estudo, ela aponta que mesmo as questões do passado remoto podem causar incomodo ao classificador, como por exemplo, a descoberta dos Manuscritos do Mar Morto que gerou um grande volume de estudos a seu respeito, exigindo novas interpretações quanto a classificação bibliográfica.

Outra questão que causa embaraço ao classificador é a orientação / tradição religiosa por trás do sistema de classificação. Como por exemplo, Os de religiosidade popular que ficam classificados em Folclore, e a literatura que trata Deus sem religião? Como achar enunciados válidos aos temas que abrangem as atuais investigações do campo de Ciência da Religião? Como pensar um sistema um sistema de representação informação em Ciência da Religião a partir do conceito de formação discursiva?

Ao colocarmos lado a lado a classe de Religião dos sistemas CDD e CDU e Dewey, percebe-se o viés religioso predominante cristão que a CDD ainda mantém na estrutura da classe principal da classe, que de alguma forma privilegia o discurso religioso cristão.

| CDD   | CDU  |
|---|--|
| <b>200 - Religião</b>                                 | 2 - Religião. Teologia   |
| <b>210 - Filosofia e Teoria da religião</b>           | 21 - Religiões pré-históricas e primitivas                                     |
| <b>220 - Bíblia</b>                                   | 22 - Religiões originárias do Extremo Oriente                                  |
| <b>230 - Cristianismo</b>                             | 23 - Religiões originárias do subcontinente indiano. Hinduísmo em sentido lato |
| <b>240 - Ética cristã e Teologia devocional</b>       | 24 - Budismo   |
| <b>250 - Congregações cristãs e Teologia pastoral</b> | 25 - Religiões da antiguidade. Cultos e religiões menores                      |

|   |   |
|---|---|
| <b>260 - Sociologia cristã e Teologia social cristã</b> | 26 - Judaísmo                                     |
| <b>270 - História do cristianismo</b>                   | 27 - Cristianismo. Igrejas e denominações cristãs |
| <b>280 - Denominações cristã</b>                        | 28 - Islamismo                                    |
| <b>290 - Outras religiões</b>                           | 29 - Movimentos espirituais modernos              |

Tabela 2: Classe principais

Acredita-se que pensar os sistemas de classificação, à luz do conceito de formação discursiva, implica assumir que os dispositivos de representação da informação não se referem a instrumentos neutros e alheios ao tempo histórico no qual foram concebidos. Além disso, é preciso compreender que os arranjos biblioteconômicos, decorrentes de tais dispositivos, efetivam formas de classificação e recuperação da informação demarcadas por hierarquizações e contextualizações pensadas na origem dos sistemas e reiteradas em novas edições. Dessa perspectiva, a transformação da produção do conhecimento e a movimentação sócio-histórica pode, nesses contextos, ser categorizada por um instrumento classificatório que, por vezes, reflete e dialoga pouco com as transformações em curso.

Refletir, sobre a dimensão discursiva dos sistemas de classificação bibliográficas implica, assim, em ampliar o escopo formativo dos profissionais da informação para que possam compreender as implicações do seu gesto classificatório e os limites do enquadramento classificatório proposto nos processos de oferta e circulação do saber.

#### 4. Considerações finais

A proposta da classificação bibliográfica como forma de ordenação de livros nos suscita uma certa "rigidez" na ação do bibliotecário. Conforme aborda Radford, pode ser interpretado como uma imagem de flagelação, semelhante ao tapa na palma da mão com um bastão dado por um pai autoritário ou o professor, no usuário que não devolver o texto para o seu devido lugar pelo tempo designado. A imagem da biblioteca perfeita, o resultado final dos esforços do bibliotecário, é a de um lugar onde tudo é em última instância responsável por, de volumes fechados. (RADFORD: 1998:619)

Radford ao explorar o ensaio "A biblioteca fantástica" de Foucault evidencia que o filósofo desenvolve uma noção em que esses opostos se conjugam para formar uma nova noção, a

incorporação de ordem na biblioteca está em contraste direto com a noção de fantasia.

O coração da análise de Foucault é a alegação de que os reinos da biblioteca e do fantástico não podem mais ser mantidas separadas. Foucault (1967/1977) escreve que o "domínio de fantasmas não é mais a noite, o sono da razão, ou o vazio incerto que está diante de desejo, mas, pelo contrário, a vigília, a atenção incansável, erudição zelosa, e uma vigilância constante". (FOUCAULT citado por RADFORD, 1998: 628)

Liberto ou preso a uma formação discursiva o bibliotecário da Biblioteca fantástica de Foucault pode classifica e incorporar um novo livro a ser arquivado juntamente com tantos outros, mas na realidade amplia o espaço que os livros existentes podem ocupar. Ele recupera livros (mostra) em detrimento de outras (esconde), isso em um único movimento, "ele faz com que eles brilhem e desaparecem".

Nesse contexto, parece paradoxal que a liberdade religiosa seja um dos principais direitos do homem incluídos em tratados internacionais e, ao mesmo tempo representar aquele que é mais frequentemente ameaçado por formas diversas de intolerância.

A ação discursiva pressuposta nos aparatos informacionais que orientam a organização da informação, por vezes, parece subjugar ou ignorar o papel determinante que a mediação informacional pode aportar aos processos de produção do conhecimento. Se por um lado a leitura radical do sistema de crenças religiosas parece turvar os avançados do conhecimento, por outro lado, nota-se no campo da Organização da Informação, tendência dogmatizante em relação ao alcance e à neutralidade presumida dos instrumentos de representação da informação.

O propósito deste estudo foi explorar inicialmente questões para análise futura, sendo oportuno salientar que será objeto de estudo do discurso que os sistemas de representação da informação promovem dão conta das formas de pensar a religião no alvorecer do século XXI.

#### 5. Agradecimentos

Agradecimentos são devidos ao CNPq e a Fapemig no desenvolvimento do presente trabalho.

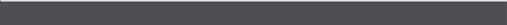
#### Referencias

Anjos, L. (2008). *Sistemas de classificação do conhecimento na Filosofia e na Biblioteconomia: uma visão histórico-conceitual crítica com enfoque nos conceitos de classe,*

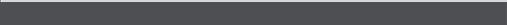
- de categoria e de faceta. 291 f. Tese (Doutorado em Ciência da Informação) – Escola de Comunicações e Artes, Universidade de São Paulo.
- Barbosa, Alice (1969). Teoria e prática dos sistemas de classificação bibliográfica. Rio de Janeiro: Instituto Brasileiro de Bibliografia e Documentação.
- Eco, Umberto (2000). *Definições léxicas*. IN: FORO INTERNACIONAL SOBRE INTOLERÂNCIA, 1997, Paris, França. A intolerância: Foro Internacional sobre a intolerância. Rio de Janeiro: Bertrand Brasil.
- Eisenhart, Ruth (1960). *The Classification of Theological Books*, Library Trends 9 (April 1960): 263.
- Esteibar, Belén Atuna (1-6 de octubre de 1984). Apuntes para historia de las ciencias de la documentación: la concepción leibniziana de una biblioteca. Actas del III Congreso de la Sociedad Española de Historia de las Ciencias, San Sebastian.
- FORO INTERNACIONAL SOBRE INTOLERÂNCIA, 1997, Paris, França. A intolerância: Foro Internacional sobre a intolerância. Rio de Janeiro: Bertrand Brasil, 2000.
- Foucault, Michel (2013). A ordem do discurso: aula inaugural no College de France, pronunciada em 2 de dezembro de 1970. Trad. Laura Fraga Sampaio. 23. ed. São Paulo: Loyola. 74 p.
- Foucault, Michel (2012). Arqueologia do saber. Trad. Luiz Felipe Baeta Neves. 8. ed. Rio de Janeiro : GEN : Fofense. 254 p.
- Frohmann, Bernd (1994). The Dewey Decimal Classification as technobureaucratic discourse. Home Page. Researchs. Recuperado 04-12-2014, de <http://www.firms.uwo.ca/people/faculty/frohmann.html/>.
- Pombo, Olga (1998). Da classificação dos seres à classificação dos saberes. *Revista da Biblioteca Nacional de Lisboa*, Lisboa, n. 2, 1998. p. 19-33.
- Radford, G.P. (1998), "Flaubert, Foucault, and the Bibliothèque Fantastique: toward a postmodern epistemology for library science", Library Trends, v. 46, n. 4, pp. 616-34.
- Sousa, Renato Tarciso Babosa de. Os princípios da teoria da classificação e o processo de organização de documentos de arquivo. *Arquivo & Administração*. Rio de Janeiro: Associação dos Arquivistas Brasileiros. v.6, n. 1, jan./jun. 2007, p. 5-26.



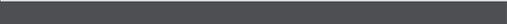
[Redacted text block]



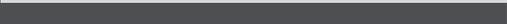
[Redacted text block]



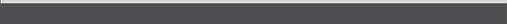
[Redacted text block]



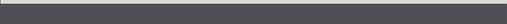
[Redacted text block]



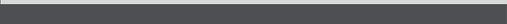
[Redacted text block]



[Redacted text block]



[Redacted text block]



[Redacted text block]



**Presentación comercial**

---

# TesaurVAI: software para la creación y gestión de tesauros

---

Laura Martín-Pérez González

Dail-Software, Campus Montegancedo, Universidad Politécnica de Madrid, Edificio CAIT, 28223 Pozuelo de Alarcón (Madrid), info@dail-software.com

## Resumen

Se presenta la herramienta TesaurVAI. Esta aplicación integra en un solo producto un potente extractor terminológico de términos desde 1 hasta 7 palabras y un gestor de tesauros multilingüe. TesaurVAI se adapta a la Norma ISO 25964.

**Palabras clave:** software, extractor terminológico, tesauro, gestor de tesauros.

## Abstract

Novel application that integrates in one product a powerful terminology extractor and a thesaurus manager. TesaurVAI complies with the Norm ISO 25964.

**Keywords:** software, terminology extractor, thesaurus, thesaurus manager.

## 1. Dail-Software y su herramienta TesaurVAI



Dail-Software es una empresa especializada en el desarrollo de aplicaciones de ingeniería lingüística. Nace desde el Grupo de Validación y Aplicaciones Industriales de la Universidad Politécnica de Madrid y pertenece

al holding tecnológico eGauss Business Holding I+T. Dail-Software se orienta a la creación de herramientas y a la prestación de servicios relacionados con el tratamiento multilingüe de textos de forma automática. Para ello, Dail-Software cuenta con profesionales destacados en sus áreas, formando un equipo multidisciplinar dispuesto a afrontar los retos marcados. Como resultado a uno de esos retos, nace TesaurVAI.

TesaurVAI es una herramienta para la creación y tratamiento de tesauros y lenguajes especializados que facilita el trabajo a documentalistas y gestores de la información. Está compuesto de dos módulos: un extractor terminológico y un gestor de tesauros.

Existen versiones para Windows, Mac y Linux.

El interfaz de TesaurVAI está disponible para los idiomas inglés y español.



Ilustración 3. Menú principal de TesaurVAI

## 2. Descripción de TesaurVAI: extractor terminológico

El extractor terminológico permite extraer términos de ficheros .txt, .doc, .docx, .html, .htm.

Posee un panel donde se pueden configurar distintos parámetros de la extracción. En primer lugar se indica el número máximo y mínimo de veces que se desea que aparezcan los términos simples (1 palabras). En segundo lugar, se indica si se desea o no extraer términos compuestos. En caso afirmativo se debe definir el número máximo de palabras (hasta 7) que deben tener los términos extraídos. Igualmente se debe indicar el número máximo y mínimo de veces que se desea que aparezcan estos términos.

Para la extracción se pueden utilizar ficheros de palabras vacías, pudiéndose elegir entre los predefinidos para las lenguas español, inglés, portugués y ruso proporcionados por el propio

programa o utilizar uno confeccionado a medida por el mismo usuario.

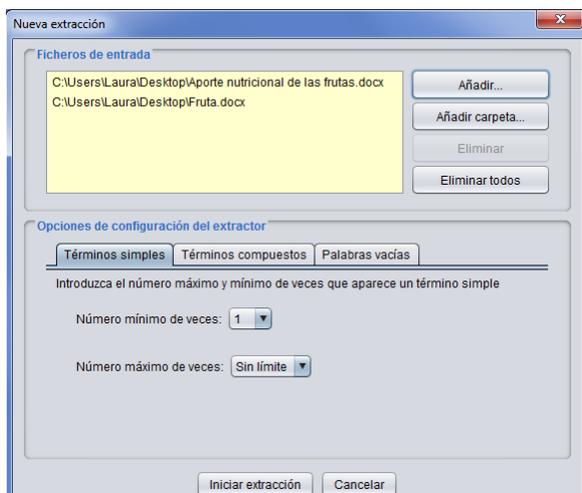


Ilustración 4. Menú del extractor terminológico

Una vez realizada la extracción, se muestra una tabla con los términos que cumplen los parámetros definidos acompañados de su frecuencia de aparición y de los contextos donde se hallan insertos. Se muestra también a qué documento pertenecen. Los términos pueden ser ordenados por frecuencia o por orden alfabético.



Ilustración 5. Tabla de la extracción de términos

En dicha tabla el usuario puede realizar una selección de términos para crear un tesaurus o insertarlos en uno ya existente. Así mismo, puede imprimir o guardar dicha tabla para ser recuperada con posterioridad. También puede generar informes en .pdf, .doc, .csv y .txt con los términos seleccionados.

El extractor contiene un potente buscador que permite localizar términos.

### 3. Descripción de TesaurVAI: gestor de tesaurus

TesaurVAI se adapta a la *norma ISO 25964. Thesauri and interoperability with others vocabularies*. Puede gestionar un número ilimitado de términos y relaciones de equivalencia, jerárquicas y asociativas. Además, posibilita la creación de equivalencias idiomáticas en inglés, francés, alemán, italiano, catalán, euskera y gallego, permitiendo la configuración de tesaurus multilingües.

| Término           | Relaciones        | Idiomas                 | Estado    |
|-------------------|-------------------|-------------------------|-----------|
| Citricos          | TE(5),TG(1)       | DE(1),EN(1),FR(1),ES(1) | Candidato |
| Frambueas         | TG(1),TR(2)       | DE(1),EN(1),FR(1),ES(1) | Candidato |
| Fresas            | TG(1),TR(2)       | DE(1),EN(1),FR(1),ES(1) | Aprobado  |
| Fresones          | TG(1),TR(2)       | EN(1),FR(1),ES(1)       | Candidato |
| Frutas            | TE(6)             | DE(2),EN(2),FR(2),ES(2) | Candidato |
| Frutas del bosque | USE(1)            | EN(1),FR(1),ES(1)       | Candidato |
| Frutas tropicales | TE(2),TG(1)       | DE(1),EN(1),FR(1),ES(1) | Aprobado  |
| Frutos rojos      | TE(3),TG(1),UP(1) | FR(1),ES(1)             | Candidato |

Ilustración 6. Tabla de términos del gestor de tesaurus

Los términos pueden ser agrupados por categorías, ser definidos o acompañarse de una nota personal.



Ilustración 7. Menú Edición de un término

Desde la pantalla de Edición, los términos pueden ser localizados en Internet de forma directa.

TesaurVAI genera los siguientes informes de términos: alfabético, jerárquico, por categorías, KWIC y KWOC. Todos ellos en formatos JasperReports, PDF, RTF, ODT, DOCX, HTML, XLS y XML.

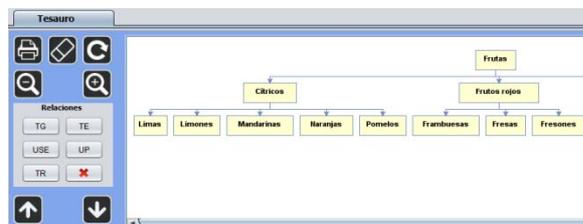


**Ilustración 8. Tipos de informes**

TesaurVAI permite la exportación e importación de tesauros en formatos TXT y XML.

### **Novedades**

En enero de 2016 estará disponible la versión gráfica de TesaurVAI que incorpora un módulo que permite visualizar el tesauro en forma arborescente. Además, desde la interfaz gráfica se podrán crear términos y establecer relaciones de forma directa y sencilla.



**Ilustración 9. Módulo gráfico**

### **Direcciones web**

TesaurVAI puede ser adquirido en la página “tienda” de la web de Dail-Software (<http://www.dail-software.com>). Se ofrece un periodo de prueba gratuito durante 2 meses.

## Índice de autores

---

**A**

|  |               |
|--|---------------|
| AGUSTÍN LACRUZ, M <sup>a</sup> Carmen..... | 470           |
| ALMEIDA, Carlos Cândido de .....           | 498, 556, 603 |
| ALVES, Bruno Henrique.....                 | 570           |
| ALVITE DÍEZ, María Luisa.....              | 266           |
| ARAÚJO, Eliane Pawlowski Oliveira .....    | 279           |
| ARBOIT, Aline Elis.....                    | 16            |
| ASSUMPÇÃO, Helena Campos .....             | 24            |

---

**B**

|                                   |         |
|-----------------------------------|---------|
| BAÑOS-MORENO, María-José.....     | 290     |
| BAPTISTA, Ana Alice.....          | 12      |
| BARBOSA, Marilene Lobo Abreu..... | 160     |
| BENTES PINTO, Virgínia.....       | 258     |
| BORGES, Jussara .....             | 234     |
| BUFREM, Leilah Santiago .....     | 16, 570 |

---

**C**

|   |     |
|---|-----|
| CÂNDIDO, Gilberto Gomes.....              | 304 |
| CARREÑO CORCHETE, Esther .....            | 83  |
| CARVALHO, Ângela Maria Grossi de .....    | 37  |
| CASTANHA, Renata Cristina Gutierrez ..... | 508 |
| CENDÓN, Beatriz Valadares .....           | 116 |
| COLOMBO, Stephanie.....                   | 598 |
| CORDEIRO, Rosa Inês de Novais .....       | 195 |
| COSTA, Luzia Sigoli Fernandes .....       | 438 |

---

**D**

|                                     |     |
|-------------------------------------|-----|
| DAL'EVEDOVE, Paula Regina .....     | 332 |
| DALESSANDRO, Rafael Cacciolari..... | 518 |
| DAVANZO, Luciana.....               | 187 |
| DE ASSIS RODRIGUES, Fernando.....   | 244 |
| DIAS, Célia da Consolação.....      | 96  |
| DIAS, Guilherme Ataíde .....        | 105 |
| DÓRIA, Maria Rosa Gomes.....        | 430 |
| DUARTE, Emeide Nóbrega .....        | 446 |

---

**E**

|                              |     |
|------------------------------|-----|
| EVANGELISTA, Victorino ..... | 527 |
|------------------------------|-----|

---

**F**

|   |     |
|---|-----|
| FADEL, Bárbara .....                          | 41  |
| FARIAS, Maria Giovanna Guedes; .....          | 160 |
| FARIAS, Mona Cleide Quirino da Silva .....    | 603 |
| FAUSTINO, Rene G.....                         | 16  |
| FELIPE, Eduardo R.....                        | 290 |
| FERNEDA, Edberto .....                        | 105 |
| FERRAREZI, Maria Amélia Duarte Oliveira ..... | 41  |
| FREITAS, Ângela Cristina Basílio de .....     | 41  |
| FREITAS, M. Cristina V. de.....               | 206 |
| FRÍAS, José Antonio.....                      | 206 |

FUJITA, Mariângela Spotti Lopes..... 319, 332, 345, 411, 470

---

## G

GALLOTTI, Monica Marques Carvalho..... 357  
GARCÍA MARCO, Francisco Javier..... 498  
GARRIDO PICAZO, Piedad ..... 152  
GESTEIRA, Rodrigo ..... 362  
GOMES, Gracielle M. R..... 116  
GONÇALVES SANT'ANA, Ricardo César ..... 244  
GRÁCIO, Maria Cláudia Cabrini ..... 508, 544  
GUIMARÃES, José Augusto Chaves ..... 461, 518, 527, 533, 544

---

## H

HOFFMANN, Wanda Aparecida Machado ..... 220

---

## K

KOBASHI, Nair Yumiko..... 387

---

## L

LIMA, Gercina Ângela Borém de Oliveira ..... 50, 290, 398  
LOURENÇO, Cíntia de Azevedo..... 96  
LÖW, Marieta Marks ..... 173

---

## M

MACULAN, Benildes C. M. S. .... 50  
MARSULO, Thabyta Giraldelelli..... 37  
MARTÍNEZ GONZÁLEZ, M. Mercedes ..... 266  
MARTÍNEZ-ÁVILA, Daniel..... 518, 533  
MARTÍNEZ-BÉJAR, Rodrigo ..... 290  
MARTÍN-PÉREZ GONZÁLEZ, Laura ..... 637  
MARZAL GARCÍA-QUISONDO, Miguel Ángel ..... 234  
MATOS, Daniela Fernanda de Oliveira ..... 544  
MEDEIROS, Filipa..... 130  
MÉNDEZ ZAPATA, Patricia Alejandra ..... 373  
MILANI, Suellen Oliveira..... 508  
MONTESI, Michela ..... 618  
MORAES, Isabela Santana de ..... 187  
MORAES, João Batista Ernesto de ..... 304, 461  
MOREIRA, Walter ..... 187, 461  
MOREIRO GONZÁLEZ, José Antonio ..... 130  
MOTA, Denysson ..... 387  
MOURA, Maria Aparecida ..... 362, 629

---

## N

NAUMIS PEÑA, Catalina ..... 373  
NETTO, Cristiane Mendes..... 398  
NEVES, Dulce Amélia de Brito ..... 411, 446

---

## O

OHLY, H. Peter..... 61  
OLIVEIRA, Jacqueline Pawlowski ..... 279  
OLIVEIRA, Leandro H. M. .... 50  
OLIVEIRA, Ely Francina Tannuri de ..... 16, 570

|                                      |     |
|--------------------------------------|-----|
| OLIVEIRA, Henry Pôncio Cruz de ..... | 258 |
|--------------------------------------|-----|

---

**P**

|   |         |
|---|---------|
| PALETTA, Francisco Carlos .....           | 138     |
| PANDO, Daniel Abraão .....                | 556     |
| PASTOR-SÁNCHEZ, Juan Antonio .....        | 290     |
| PAULA, Claudio Paixão Anastácio de .....  | 279     |
| PAVANELLI, Maria Aparecida .....          | 570     |
| PIEROZZI JR, Ivo .....                    | 50, 398 |
| PRADO, Marcos Aparecido Rodrigues do..... | 570     |
| PRATES, Raquel O. ....                    | 116     |
| PRET, Raquel Luise;.....                  | 195     |

---

**R**

|                                   |     |
|-----------------------------------|-----|
| RIBEIRO, Claudio Jose Silva ..... | 419 |
| RODRÍGUEZ BRAVO, Blanca .....     | 206 |
| RODRÍGUEZ YUNTA, Luis.....        | 485 |
| RUBI, Milena Polsinelli .....     | 345 |

---

**S**

|                                       |     |
|---------------------------------------|-----|
| SALDANHA, Gustavo Silva .....         | 584 |
| SALES, Odete Máyra Mesquita .....     | 258 |
| SÁNCHEZ CASABÓN, Ana I. ....          | 152 |
| SANTOS, Fernando Bittencourt.....     | 430 |
| SANTOS, Rodrigo Eduardo dos.....      | 67  |
| SANTOS, Cintia Almeida da Silva.....  | 220 |
| SANTOS, Marcel Pereira .....          | 438 |
| SATUR, Roberto Vilmar .....           | 446 |
| SCHMIDT, Clarissa .....               | 579 |
| SILVA, Andrieli Pachu da .....        | 461 |
| SILVEIRA, Naira Christofolletti ..... | 584 |
| SIMÕES, Maria da Graça Melo.....      | 206 |
| SMIT, Johanna.....                    | 579 |
| SMITH, Marinês Santana Justo .....    | 41  |
| SOBRINHO, Heliomar Cavati .....       | 319 |
| SOUZA, Jacqueline Aparecida de .....  | 357 |

---

**T**

|                                    |     |
|------------------------------------|-----|
| TERRA, Ana Lúcia .....             | 470 |
| TORRES, Simone.....                | 279 |
| TRAMULLAS, Jesús .....             | 152 |
| TRAVIESO RODRÍGUEZ, Crispulo ..... | 83  |
| TRIVELATO, Rosana.....             | 629 |

---

**V**

|                              |     |
|------------------------------|-----|
| VARELA, Aida Varela .....    | 160 |
| VIDAL LIY, José Ignacio..... | 485 |

---

**Z**

|                             |        |
|-----------------------------|--------|
| ZAFALON, Zaira Regina ..... | 24, 67 |
|-----------------------------|--------|

**XII Congreso ISKO España  
II Congreso ISKO España-Portugal**

**Facultad de Comunicación y Documentación.  
Universidad de Murcia.  
ISBN: 978-84-608-3558-5**

**UNIVERSIDAD DE  
MURCIA**

