

---

## RESEÑA

---

**MOREIRO GONZÁLEZ, JOSÉ ANTONIO.** *Linguagens documentárias e vocabulários semânticos para a web: elementos conceituais.* Salvador: EDUFBA, 2011, 128 págs. ISBN: 978-85-232-9-0824-0.

Publicado en portugués, el libro de José Antonio González Moreiro ofrece un panorama sobre lo que permanece y lo que ha cambiado en los sistemas de organización del conocimiento, en especial con la presencia de la web semántica. En 11 capítulos, el autor repasa la evolución histórica de los instrumentos de tratamiento de la información para recuperación, desde la aparición de los sistemas de clasificación bibliográfica, hasta los topic maps, siempre teniendo como contexto la web y la búsqueda del mayor rigor e interactividad.

El texto está organizado de modo esquemático a partir de la experiencia docente. La estructura adoptada permite alcanzar una visión sintética del asunto, de forma que facilita el análisis introductorio a los alumnos y a los profesores, al tiempo que sirve de referencia consistente sobre la gama de recursos a utilizar ya sea desde el punto de vista teórico o desde el práctico. Su alcance lo determina como importante inventario para ser tomado como punto de partida hacia exploraciones más profundas en el terreno de los sistemas de organización del conocimiento.

Introduciendo el tema, el autor destaca la importancia de las operaciones del análisis documental, como medio para trabajar con los documentos, mostrando que las opciones varían desde un tratamiento morfológico -palabras-clave- a un tratamiento semántico -enfocado hacia los conceptos. Desde una perspectiva

del léxico y de las relaciones entre las palabras, atiende a los KOS (*Knowledge Organization Systems*), que incluyen listas de palabras (entre ellas, las folksonomías), glosarios, listas de nombres y diccionarios ligados por anillos semánticos (como, en parte, *Wordnet*); categorizaciones y clasificaciones (taxonomías, esquemas); e instrumentos que privilegian las relaciones asociativas entre conceptos (tesauros, redes semánticas, tesauros conceptuales, *topic maps* y ontologías).

En el primer capítulo, '*Bases epistemológicas da representação da informação*', Moreiro recupera los orígenes de la representación conceptual -las categorías de Aristóteles, Leibniz y Kant- y de la identificación de las relaciones entre conceptos: el árbol de Porfirio (primera red semántica), Ramón Llull (árbol semántico y mecanismo universal del hecho comunicativo), el movimiento de Port-Royal (conceptos lógicos de definición y de división del género por las diferencias), Descartes, los predicados clásicos (causalidad, finalidad, origen o modo), o los que revisaron los empiristas (David Hume), tanto los naturales (causa/efecto o semejanza) y los basados en relaciones filosóficas (causa/efecto, atribución, semejanza, contrariedad, cantidad y cualidad). El autor muestra, así, que cualquier propuesta de categorización de los conceptos siempre consideró la presencia de las relaciones entre ellos como forma fundamental del pensamiento (inducción). Lo que explica el hecho de que los diferentes

instrumentos de organización y representación del conocimiento siempre recurren a la asociación entre conceptos, destacando que las relaciones no-jerárquicas solo aparezcan cuando se consideran experiencias o campos específicos del conocimiento. El origen de las relaciones proviene de los predicados clásicos (causalidad, finalidad, origen o modo), con la identificación de sus tipos: naturales (de causa/efecto o semejanza) y por relaciones filosóficas (de causa/efecto, atribución, semejanza, contrariedad, cantidad, calidad). Se explican, por esa vía, los fundamentos de la constitución de los lenguajes documentales, basados sobre los mecanismos generales del razonamiento para la estructuración de los conceptos. Moreiro llama la atención también sobre el hecho de que la relación unívoca entre un concepto y un término pretende superar la condición inherente al signo lingüístico que es repetirse (teoría de la desconstrucción de Derrida), asegurando, por ese medio, la disminución de la ambigüedad. Desde el punto de vista de la visualización, el autor recuerda que la representación esquemática ya estaba considerada en Porfirio, alcanzando hoy, como arquetipo, los *topic maps*. En relación a las referencias para la selección de información a representar, destaca la importancia de las macroestructuras textuales.

El capítulo segundo, *Para melhor entender a recuperação léxica da informação*, es bastante breve. En él busca demostrar que, de una u otra forma, el léxico y los contextos de referencia han sido utilizados desde la antigüedad para constituir, en cierto modo, redes semánticas: desde la referencia a las ciencias naturales al predominio de la función comunicativa

del lenguaje; con referencia a las obras que abrieron la trayectoria de los lenguajes documentales en el siglo XX: la cobertura enciclopédica y universal con la CDD, por una parte, y las normas para la organización de las listas de encabezamientos de materia, de Cutter, por otra; el enriquecimiento de los vocabularios y lenguajes documentales con el establecimiento de relaciones entre los términos (ítem que se expone en otro capítulo); y la propuesta de superación de las jerarquías a partir de la ventaja concedida a las asociaciones entre conceptos (Vannevar Bush).

En el capítulo tercero, *Requerimentos fundamentais da nova web*, el texto discurre sobre los requisitos que caracterizan a la representación en ese ambiente que, a falta de parámetros definitivos, viene haciéndose a partir de tentativas de mejora de las estructuras y de los sistemas de organización del conocimiento. Moreiro González muestra cómo en la *web 2.0* (web social) dominan las folksonomías, una arquitectura colaborativa, y la usabilidad, mientras que en la web semántica, lo hacen los vocabularios controlados a partir de las ontologías, y el resto de KOS y vocabularios de metadatos que privilegian la interoperabilidad en detrimento de la usabilidad. Frente a esa dicotomía, el autor apunta la posibilidad de que una evolucione hacia la otra, que las dos converjan en la web 3.0, o que la web 2.0 y a web semántica coexistan manteniendo sus diferencias.

En el capítulo cuarto, *Léxico e Linguagens documentárias*, el autor sugiere la mejora de los instrumentos de organización y representación mediante las normas terminológicas, aspecto que viene siendo apuntado por la literatura,

inclusive en la brasileña. Distingue entre, por una parte, los *KOS* (palabras-clave independientes, glosarios, listas de nombres, diccionarios; categorizaciones y clasificaciones; tesauros, *topic maps* y ontologías) que se diferencian principalmente por su grado de estructuración y por sus dificultades de implantación; de otro, de los *SKOS* (*Simple Knowledge Organization for the Web*) que permiten establecer un puente entre las tendencias de la web semántica y las prácticas tradicionales incluidas en los *KOS*, con codificación de vocabularios en XML basados en RDF, y donde los conceptos y relaciones de un lenguaje de indización corresponden a instancias de clases y declaraciones sobre esas instancias. Los dos tipos de sistemas son analizados brevemente en este capítulo mediante la comparación de sus características y de las ventajas que ofrece cada uno de ellos.

El quinto capítulo está a las *Folksonomias*, que son brevemente analizadas a partir de sus características, tipos, funciones, ventajas y debilidades, incluyéndose una comparación entre ellas y los tesauros.

En el capítulo 6, *Taxonomia*, Moreiro González se remonta a los orígenes del concepto (en biología, botánica y zoología) para pasar a considerar las taxonomías sistemáticas en su uso contemporáneo. Atiende a sus componentes principales, discrimina sus funciones, destaca los usos actuales, así como sus características más relevantes, dedicando buena parte del capítulo al proceso de elaboración de esos instrumentos, con especial atención a las aproximaciones facetadas y facilitando ejemplos.

El *Tesauro* es atendido en el capítulo 7, mediante la definición de su concepto, sus objetivos, la estructura en

forma de red de relaciones conceptuales, sintaxis, ventajas e inconvenientes de su uso, y sugerencias para la superación de sus problemas, de manera destacada mediante las aportaciones de las nuevas normas, BS 8723 (2005-2007) y ANSI/NISO Z39.19 (2005). Luego, el autor compara el tesauro con otros *KOS* y aporta sugerencias para el incremento pragmático de relaciones, en especial las de asociación entre conceptos.

El capítulo 8, *Ontologías, o predomínio dos conceitos*, contextualiza el uso de las ontologías en la web semántica dando énfasis a la capa ontológica como medio para inferir conocimiento. Para Moreiro González, las ontologías suponen un salto cualitativo en la recuperación, ya que en lugar de motores de búsqueda por palabras-clave se emplean agentes de software inteligentes que operan con conceptos. Contextualiza las ontologías en la web semántica a partir de la discriminación de sus capas (sintáctica, semántica, ontológica, lógica, de prueba y confianza), destacando la importancia de los componentes ontológicos como medio para inferir conocimiento. Argumenta que una ontología se diferencia de los *KOS* por permitir definir relaciones más complejas, siendo usada para describir *SKOS* con diferentes grados de estructura, desde simples taxonomías (como una jerarquía) y esquemas de datos (como el Dublin Core) hasta teorías lógicas más complejas. Especifica los componentes de una ontología, sus objetivos y características, los lenguajes para su creación, y aporta ejemplos de herramientas para su construcción. Buena parte del capítulo está dedicada al proceso de elaboración de ontologías, atendiendo a las diversas metodologías existentes, a los criterios y principales

requisitos para su diseño, así como a las etapas principales a cumplir, junto con ejemplos y programas para su construcción, además de una abundante bibliografía.

Ya en el capítulo 9, *Tesouros: progressão conceitual y redes semânticas*, el autor presenta la evolución seguida por el tesoro tradicional y las experiencias con él relacionadas, refiriéndose a los ‘tesoros conceptuales’, que pueden recurrir a las ontologías con el objetivo de alcanzar mayor precisión en la recuperación; a los ‘tesoros de descriptores verbales’ más adecuados a la representación y recuperación de imágenes en movimiento; la base de datos de referencia léxica Wordnet, que realiza la desambiguación mediante el uso de verbos; las redes semánticas de grafos, que más allá de las relaciones tradicionales de los tesauros introducen otras relaciones que permiten la representación declarativa de objetos; y los mapas conceptuales, que pueden representar el conocimiento por medio de grafos cognitivos o redes de conceptos sin ningún control. Se desataca así la posibilidad que ofrecen los tesauros para utilizar recursos de cara a la desambiguación conceptual, a la representación declarativa de objetos, así como para aprovechar los procesos cognitivos por asociación libre de conceptos.

El capítulo 10, *Topic Maps*, trata del documento o conjunto de documentos SGML o XML que, interrelacionados en un espacio multidimensional, facilitan el

acceso y la recuperación de información en la web. El capítulo atiende al concepto de *topic map*, explica el modelo desde sus componentes y características, sugiere herramientas apropiadas para construir *Topic Maps*, entre las que destaca ‘*demos*’ disponible en internet, presenta también las ventajas e inconvenientes de su uso, para concluir con una tabla comparativa de las características de los *Topic Maps* frente a las que ofrecen los Mapas conceptuales y los Tesoros.

Cierra el texto una *Reflexão Final* presentada por ítems (cap. 11) que tratan, en síntesis, de las inversiones necesarias en este ámbito para abordar los problemas de recuperación en la web, de la nueva terminología que va substituyendo a la tradicional a la hora de referirse a los instrumentos de tratamiento y recuperación, así como de la importancia de las ontologías como referencia por su mayor riqueza semántica, entre otros aspectos destacados.

El libro, en fin, es un instrumento didáctico de suma importancia para la enseñanza y el aprendizaje. Las incorrecciones que presenta la expresión portuguesa del texto no llegan a comprometer su comprensión y, considerando su importancia, merecería una nueva edición con más amplia difusión, así como una versión en lengua española.

Marilda Lopes Ginez de Lara  
Universidade de São Paulo