

1989-2014: WWW, de una propuesta de gestión de información a un universo de información por descubrir.

Autores: Martínez Méndez, F. J.¹; Pastor Sánchez, J.A.²; López Carreño, R.³ y Rodríguez Muñoz, J.V.⁴ Departamento de Información y Documentación de la Universidad de Murcia.

Resumen

Este documento es una reflexión sobre la interrelación entre la gestión de la información en las organizaciones y los sistemas de hipertexto, un vínculo cada vez más conocido desde la propuesta de creación de la WWW realizada hace casi treinta años por Berners-Lee. Se lleva a cabo una revisión de los postulados teóricos del hipertexto, proporcionados principalmente por Bush y Nelson, como una posible solución a la crisis de la gestión de la información, comparándose con la situación de los primeros años noventa en las organizaciones. Se estudian los problemas detectados por Berners-Lee en su entorno de trabajo que le llevaron a desarrollar el WWW, la evolución que ha tenido esta tecnología y la forma en que todavía se necesita para una gestión adecuada de la información. El grado actual de desarrollo de la tecnología permite augurar una buena gestión de los contenidos de las organizaciones. Los problemas ahora están más relacionados con la actitud de los responsables de las organizaciones. Es necesario intentar que las organizaciones apuesten decisivamente por estas tecnologías para aumentar su efectividad en el momento de la gestión de la información. Además, la WWW es un sistema en continua evolución en el que debemos continuar investigando, aún no hemos llegado a vislumbrar el final de esta tecnología.

Palabras clave

Gestión de la Información, World Wide Web, Hipertexto

Abstract

This work is a reflection on the interrelation between the management of information in organizations and in hypertext systems, in the sense that it is a creation of the WWW. A review of the theoretical postulates of hypertext, proposed mainly by Bush and Nelson, is presented as a possible solution to the crisis of information management, comparing with the situation of the early nineties in organizations. We study the problems detected by Berners-Lee in his work environment that led him to develop the WWW, the evolution that this technology has had and how it is still needed for an adequate management of information. The current degree of technology development allows good management of the contents of organizations. The problems are now more related to the attitude of those responsible for the organizations. It is necessary to try that the organizations bet decisively for these technologies to improve their functioning at the moment of the management of the information. In addition, the WWW is a system in continuous evolution in which we must continue to investigate, we have not yet reached a final vision of this technology.

Keywords

Information Management, World Wide Web, Hypertext

1 javima@um.es

2 pastor@um.es

3 rosanalc@um.es

4 jovi@um.es

1. Introducción.

La preocupación por mejorar la gestión de la información ha sido, es y seguirá siendo una de las áreas de mayor interés en la evolución de nuestra disciplina, formando parte del núcleo central de la misma desde su génesis. De no haber sido así, ¿cómo hubiera podido el bibliotecario Calímaco de Cirene⁵ gestionar los miles de rollos de papiro que poblaban las estanterías y armarios de la *Biblioteca de Alejandría*? Obviamente, si no hubiera habido algún proceso de gestión de información la única respuesta posible es la magia, algo sobre lo cual estaban muy versados en la antigüedad pero que ya no está de moda en la actualidad.

Como disciplina científica, la Gestión de Información es la aplicación de los principios de gestión a la adquisición, organización, control, difusión y uso de información relevante para el funcionamiento eficaz de las organizaciones de todo tipo (Wilson, 2002). El término "información" hace referencia a todos los tipos de información de valor para una organización, tanto si tiene su origen dentro o fuera de la misma, incluyendo los recursos de datos (datos de producción o de facturación como ejemplo) y archivos relacionados (la organización del personal, los datos de investigación de mercado o de inteligencia competitiva) de entre una amplia gama de fuentes). La Gestión de Información, así entendida, se refiere al valor, la calidad, la propiedad, el uso y la seguridad de la información en el contexto del desempeño de la organización (Wilson, 2002).

Se puede decir que si bien la confluencia entre la Gestión de Información y la *World Wide Web* es un fenómeno contemporáneo (algo lógico tras ver la idea de Berners-Lee y su tremendo éxito), sus orígenes son algo más lejanos en el tiempo. Por un lado está el cambio de paradigma en la gestión documental auspiciado por los profesionales de los archivos federales norteamericanos a partir de la teoría del 'Records Management' de principios de los años cincuenta (Llansó i Sanjuan, 1993, p.73), algo que derivó en la necesaria reforma en los métodos de trabajo y, por otro lado, encontramos la corriente de pensamiento que surge alrededor de la idea del hipertexto concebida por Bush (1945), justo al final de la *II Guerra Mundial* y que desarrollan posteriormente otros investigadores, destacando entre todos ellos Nelson (1981) con su proyecto de sistema de gestión de la literatura universal *Xanadú* y Engelbart con el desarrollo de los primeros interfaces gráficos de usuario y los dispositivos que los hacían posible, como el primer *mouse* o ratón (Cantos et. al., 1994)

El embrión de la Gestión de Información como disciplina está en el desarrollo continuado en el tiempo de la idea del 'Records Management', proposición que está basada en una simple premisa: almacenar sólo la información verdaderamente necesaria para la gestión de las instituciones en su actividad cotidiana. Esta gestión se asienta en cuatro pilares: (1) el concepto de ciclo de vida; (2) los

5 Poeta y erudito griego que recibió de Ptolomeo II el encargo de ordenar la Biblioteca de Alejandría. Es tal la envergadura de su tarea que es considerado el padre de los bibliotecarios (o, por lo menos, de los catalogadores). Sus *Pinakes* (tablas) contienen el catálogo completo de la biblioteca ordenado cronológicamente y fueron de enorme valor para los posteriores estudios bibliográficos y literarios realizados sobre el período clásico.

programas de atención continuada; (3) el control de la gestión y (4) la especialización de quienes la practican (Llansó i Sanjuan, 1993, p.73). Bajo esta perspectiva, la gestión documental no se centra en conservar toda la documentación que día a día se genera (en la espera de que un siglo futuro un investigador tome una serie de datos para incluirlos luego en su tesis doctoral o en una novela histórica), sino se preocupa en analizar la información, describirla adecuadamente según estándares aceptados internacionalmente, identificar el ciclo de vida de la misma una vez ha entrado en nuestro sistema de información, procurar copias de seguridad para aquellos documentos de valor histórico perenne, decidir cuándo estos documentos pueden ser accesibles al público y de qué manera y, finalmente, proceder a la destrucción de los mismos cuando ya no resulten necesarios.

Esta disciplina permite poner en valor la información almacenada (incluso cuando ese documento parece ya haber perdido valor administrativo que no histórico), y propicia su reutilización en repositorios, portales y sitios web especializados, otorgando un valor añadido a documentos que, de otra manera, estarían apilados en estanterías inútiles sin poder ofrecer información alguna tras su recepción y/o creación. Así, la Gestión de Información devuelve (de alguna manera) parte de vida a documentos que de otra forma, serían olvidados con el paso del tiempo. Tal como indicaba Wilson (1989), estas ideas no pueden llevarse a cabo sin un adecuado soporte tecnológico que permita diseñar los sistemas de flujo de trabajo y los niveles de acceso a estas colecciones documentales. Una realidad fácilmente constatable es que resulta imposible ser gestor de información sin ser un tecnólogo de la información.

La aplicación de estas ideas y tecnologías asociadas en la Gestión de la Información de cualquier tipo de organización (una empresa de transportes, una fábrica de muebles, una concesionaria de venta de automóviles o una agencia de publicidad, por ejemplo) y sobre cualquier tipo de documento (estadísticas de clientes que usan el autobús a una hora determinada, los planos para el diseño de mesas y sillas, los datos de venta de un modelo de coche por determinadas áreas urbanas y sectores de edad o la gestión documental de los proyectos desarrollados previamente para uno de nuestros clientes), constituye la tercera pata del banco sobre el que se asienta la Gestión de Información actual, actividad que, en contextos y situaciones de negocio específicas, puede convertirse en estratégica.

Para ello, además de un profesional formado en técnicas de gestión de documentos y en TICs, hace falta alguien que conozca la organización y sepa interpretar, al menos de forma suficiente, la naturaleza y el objeto de los tipos documentales que se emplean en la misma, todo ello sin olvidar que este gestor debe adoptar una actitud proactiva hacia la organización, interesándose por analizar las necesidades y los flujos de información en su seno, identificando problemas y competencias (tanto internas como externas) y procurando estar ahí siempre ayudando a sus compañeros, superando la clásica actitud reactiva del paradigma de suministro de información bajo demanda y

adoptando el paradigma orientado al usuario (Dervin y Nilan, 1986), ayudando a configurar un entorno donde la información y su gestión contribuyan a la consecución de los objetivos de la organización (Kirk, 1999). Entonces, y sólo entonces, estaremos hablando de una adecuada Gestión de la Información en las Organizaciones

2. El hipertexto como solución de gestión de información.

El final de la *II Guerra Mundial* trajo consigo la constatación de que los sistemas de información existentes en esa época no permitían clasificar y organizar debidamente las ingentes cantidades de información que ya entonces comenzaban a manipularse, grandes conjuntos de datos que, de forma incipiente, no estaban siempre dispuestos en formato textual, sino que incorporaban otros tipos de media: gráficos, sonidos, planos de diseños, mapas, fotografías, etc⁶. Se habla de millones de documentos microfilmados por las tropas norteamericanas en archivos e industrias alemanas, documentos que fueron puestos a disposición de los gestores de información, quienes al ir a aplicar sobre ellos los sistemas de clasificación e indización de la época comprobaron su inutilidad. De esta situación de crisis surge la idea de disponer de otros sistemas que permitieran organizar esos fondos documentales, sistemas que colaboraran además, en una mejor comprensión de información de la misma, facilitando la asociación de ideas y conceptos, y permitiendo adquirir el conocimiento de forma no estrictamente secuencial.

Bush (1945), asesor científico del Presidente Franklin Delano Roosevelt, llegó a pensar en la construcción de una máquina llamada 'Memex'⁷ concebida como una gran base de datos donde se almacenarían todo tipo de documentos. Esta máquina constaría de una mesa con un teclado y unas palancas que permitirían la consulta de datos almacenados en microfilms que serían proyectados en unas pantallas translúcidas. El aparato incluiría también una opción para que el usuario pudiera tomar anotaciones en los márgenes, de manera que este usuario pudiera convertirse a su vez en autor, algo verdaderamente innovador y que se pretendía conseguir casi setenta años antes que se fijaran las características de la Web 2.0 (O'Reilly, 2005), que tanta importancia le confieren al doble sentido de la publicación en la Web. 'Memex' fue un proyecto teórico que nunca llegó a materializarse, la tecnología de la época no lo permitía. Lo que más subyace de los pensamientos de Bush es su idea de que algo nuevo había que hacer (de ahí el título de su artículo 'As we may think'), porque seguir gestionando la información de la misma manera, difícilmente podría llevar a buen término esta tarea.

6 Un hecho puntual, al que el cine y la literatura han hecho mención en alguna vez, especialmente en la interesante novela de Joseph Kanon "El buen alemán", fue el inicio primario de la llamada "guerra fría" entre las dos grandes potencias de entonces (Estados Unidos y la Unión Soviética) por la posesión de los preciados proyectos de ingeniería y de ciencia que se habían desarrollado en la Alemania de principios de siglo XX.

7 Para algunos autores, 'Memex' es el acrónimo de **Memory - Index**, probablemente por similitud a 'Rolodex' (**Rolling - Index**), un dispositivo para almacenamiento y búsqueda de información en fichas muy popular en esa época.

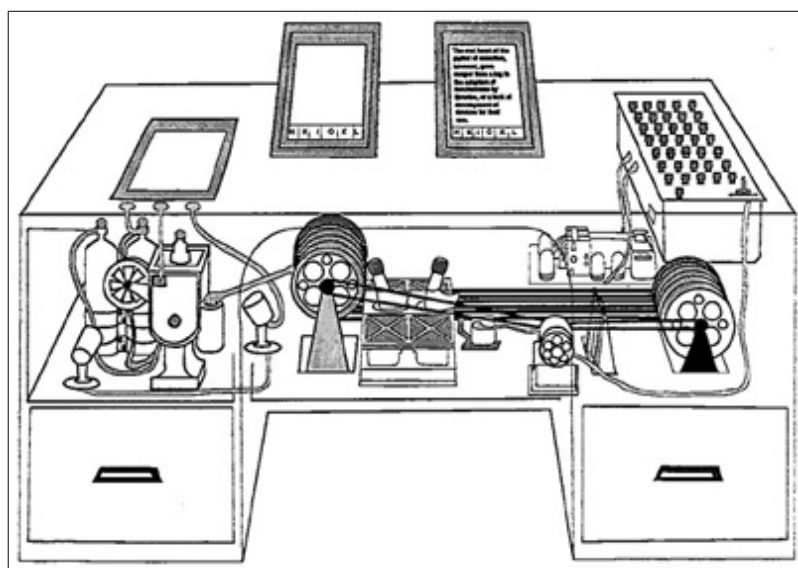


Fig. 1 Dibujo que representa la máquina 'Memex' concebida por Vannevar Bush.

Fuente: 'As we may think' <http://u-tx.net/ccritics/as-we-may-think.html>

De forma contemporánea a estos planteamientos y circunstancias, se impone paulatinamente la Arquitectura de Von Neumann (1945) para el diseño y construcción de las computadoras (una máquina basada en esta arquitectura, al igual que los ordenadores modernos, debía disponer de cuatro componentes esenciales: un dispositivo de operaciones, una unidad de control, memoria y dispositivos de entrada y salida, lo mismo que un ordenador actual). Han de pasar varios años para que estas máquinas primigenias pudieran llevar a cabo estas tareas en la forma deseada. Uno de los autores más destacados de este campo, el científico norteamericano Theodor Nelson, ha asistido prácticamente durante toda su vida al desarrollo de un sistema de información que él denominó “hipertexto” (Nelson, 1981), (Conklin, 1981). El hipertexto de Nelson, es un sistema capaz de gestionar piezas (fragmentos) de información de forma que permitiera un aprendizaje de sus contenidos y una gestión de los mismos de forma no estrictamente secuencial, permitiendo al mismo tiempo a sus usuarios la integración de cualquier pieza informativa en sus propios documentos, sin tener que copiar y pegar esa pieza informativa, sino estableciendo entre ellos un vínculo pasando los documentos de los usuarios a formar parte de una red universal (un sistema imaginario donde residiría toda la información científica: *Xanadú*⁸). Este proyecto tenía como objetivo principal la construcción de un servidor de hipertexto para almacenar y enlazar toda la literatura mundial, accesible desde cualquier ordenador. La idea era reunir toda la producción escrita existente y conectar unos textos con otros, estando esos documentos almacenados en ordenadores particulares y disponibles al mismo tiempo para el resto de los usuarios por medio de

8 El nombre **Xanadú** fue tomado al parecer de un poema de Samuel Taylor Coleridge. Nelson interpretó la palabra como "ese mágico sitio de la memoria literal donde nada se pierde nunca" (en el poema era el palacio de Kublai Khan).

una dirección única para cada uno de ellos, del mismo modo que ahora un objeto publicado en la Web tiene asociado un identificador único de documento URI/URL y los ficheros almacenados en ordenadores conectados a redes P2P resultan accesibles a través de aplicaciones de búsqueda y compartición de ficheros.

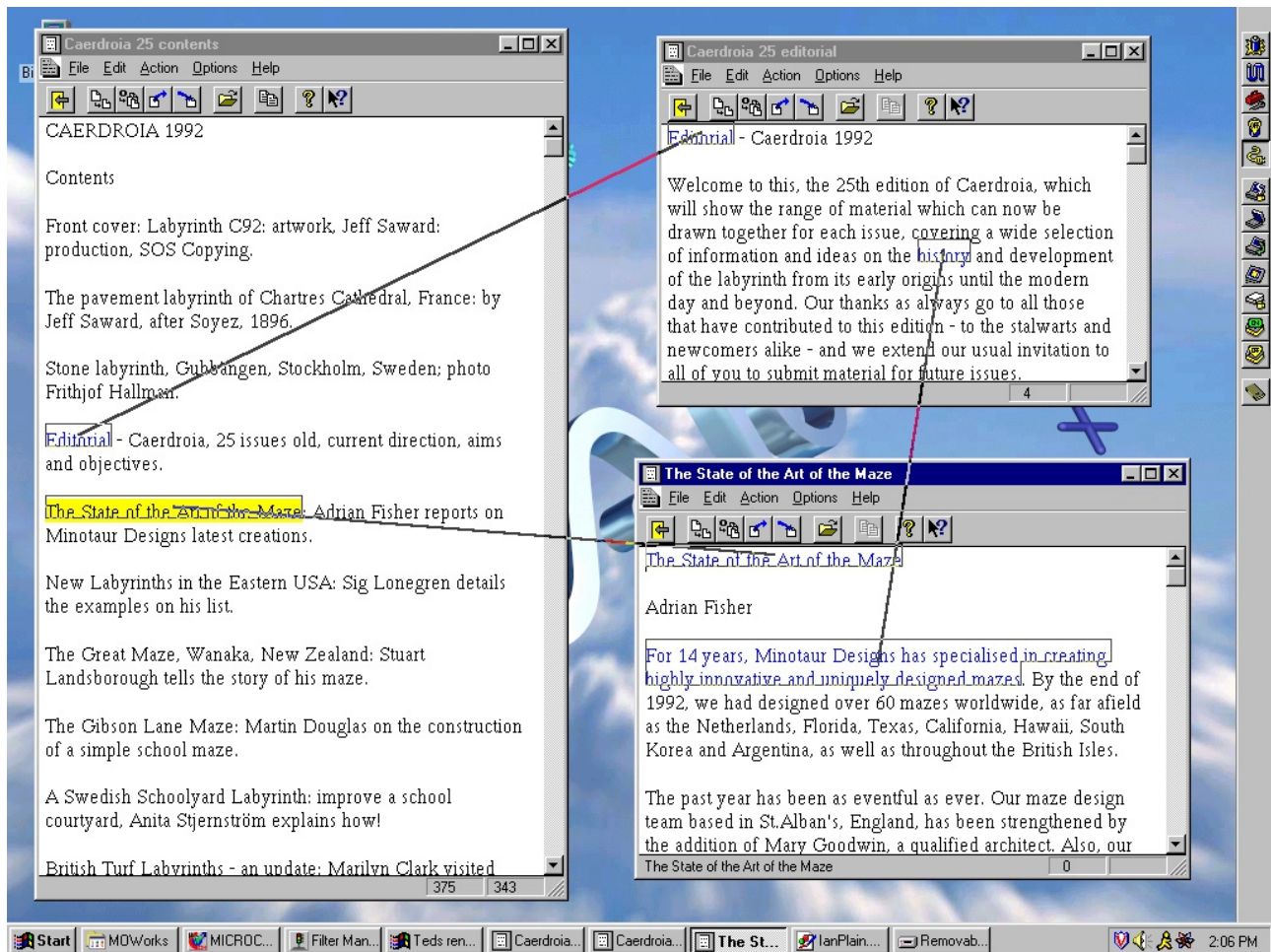


Fig. 2 Dibujo que representa el almacenamiento “xanalógico” de T.H. Nelson

Fuente: 'As we may think' <http://u-tx.net/ccritics/as-we-may-think.html>

Esta idea también fue irrealizable en los años sesenta, pero ciertamente es muy parecida a lo que disponemos hoy en día en la Web, aunque se presenta una importante diferencia: los documentos del hipertexto de Nelson se construirían en el momento de la consulta, en la edición el autor introduciría nuevos textos y enlaces a fragmentos de texto ya escritos por otros autores, la reconstrucción del documento para su lectura estaría garantizada por la técnica del almacenamiento “xanalógico” (Nelson, s.f.). Con esta técnica, el autor pretendía hacer viable su red y resulta curioso que esta idea no haya sido implementada en la Web actual, donde sabemos que abunda (mucho más de lo deseable) la copia de textos desde otros documentos originales. Así que, desde un punto de vista conceptual, el hipertexto de Nelson aún no se ha desarrollado del todo y no es de

extrañar la cierta frustración de este autor por tener que ceder el protagonismo y el reconocimiento de su idea a un joven (entonces) investigador británico que trabajaba becado en el Centro Europeo de Investigación Nuclear (CERN, hoy Organización Europea de la Energía). De hecho, hacia 1995 cuando comienza a popularizarse la tecnología de la Web, **Roy Rada** la presentaba en unas breves palabras: '*WWW, Xanadu at least*', homenajeando de alguna manera la aportación de Nelson a todo estos desarrollos, homenaje que no parece haber satisfecho del todo al autor quien ha reconocido en alguna entrevista que si bien no puede negar el éxito y trascendencia de la Web, este sistema no es (aún) un hipertexto completo en el sentido conceptual que él lo había imaginado sino una “brillante simplificación” (Whitehead, 1996).

Resulta verdaderamente curiosa la coincidencia, más si cabe en un intervalo de medio siglo, que la tecnología existente no permita una gestión efectiva de los documentos en las organizaciones (especialmente aquellos relacionados con la investigación científica donde se maneja una amplia variedad de fuentes, tipos y proyectos). Esta circunstancia termina uniendo en el tiempo a Bush, Nelson y Berners-Lee, investigador que ya asistía incrédulo diez años antes de su propuesta a la paradoja de comprobar día a día cómo el CERN perdía información y/o tenía problemas para localizar proyectos desarrollados por científicos de muy alto nivel tras costosísimas horas de trabajo. A Berners-Lee le llamaba poderosamente la atención que una “maravillosa organización”⁹ adoleciera de este problema, especialmente cuando en ella trabajaban varios miles de personas de alta cualificación intelectual (muchas de ellas muy creativas), colaborando unos con otros en la consecución de metas comunes. A pesar de que nominalmente estaban organizados en una estructura de gestión jerárquica, esto no limitaba la manera en la que se comunica y comparte información, equipo y software en todos los grupos. En realidad, más que de una jerarquía, la estructura de trabajo real del CERN era una red conectada que además aumentaba su tamaño con el paso del tiempo. En este entorno, una persona que se incorporaba a esta institución (o alguien que asumía una nueva tarea), usualmente recibía algunas pistas sobre quiénes serían contactos útiles para recabar información verbal sobre lo disponible para su proyecto en las instalaciones y poco más, el resto era un proceso de autoaprendizaje organizacional por parte de este nuevo trabajador. Como decía Berners-Lee (1989), a fin de cuentas, habitualmente los resultados de los proyectos del CERN alcanzan un éxito notable, a pesar de los malentendidos ocasionales y de la cierta duplicación de esfuerzos en la adquisición de la información y, por tanto, en la transmisión interna del conocimiento), sin olvidar ocasionales pérdidas de información. Los detalles técnicos de los proyectos anteriores a veces se perdían para siempre o sólo se recuperaban tras llevar a cabo una investigación típica de detective en una emergencia. A menudo, la información que incluso ya había sido registrada, simplemente no se encontraba.

9 “El CERN es un lugar donde todos los días se llevan a cabo pequeños milagros”, escucha el imaginario historiador Robert Langdon de boca de un también imaginario director del CERN en la novela “Ángeles y demonios”, escrita por Dan Brown.

La organización asumía este esquema como un imponderable del contexto, algo verdaderamente preocupante para Berners-Lee (1989), quien hacía mención especial a un problema: la alta rotación de personas en la institución aumentaba las pérdidas de información. La duración normal de una estancia investigadora era de dos años y este proceso de renovación de los investigadores desembocaba en una pérdida de información continuada, algo ya no tan sencillo de asumir por la dirección de la organización. La inserción de las nuevas personas en la organización requería de una buena cantidad de su tiempo y también del tiempo de los compañeros antes de que los nuevos trabajadores llegaran a tener alguna idea de lo que pasaba (y más al avanzado nivel de trabajo de esta institución). Si el ritmo de incorporación de trabajadores aumentaba, crecía paralelamente la falta de control de la información. También detectó otro problema que había pasado desapercibido hasta ese momento: el modo de registrar la documentación asociada a un proyecto. Si un experimento analizaba un fenómeno estático y particular, toda la información se podía registrar en un libro para posteriores consultas.

Pero esto no era lo frecuente, tanto el propio CERN como sus focos de interés cambian constantemente a medida que se producen nuevas ideas y se incorporan nuevas tecnologías al trabajo científico o se adquieren nuevos dispositivos que dan lugar a nuevos experimentos que no siempre tienen éxito. Cuando había que introducir un cambio en un proyecto que normalmente sólo afectaba a una pequeña parte de la organización (cambiar una parte del experimento o un sencillo detector de señales), esto implicaba que el investigador averiguara qué otras partes de la organización y otros proyectos se podrían ver afectados. El tipo de libro de registro utilizado era prácticamente imposible de mantener actualizado y no ofrecía respuestas a cuestiones tan simples como “¿cuándo se utiliza este módulo?” o “¿quién escribió este fragmento de código y dónde trabaja?”, interrogantes que quedaban sin una respuesta debidamente documentada, dependiendo su resolución casi absolutamente de la transmisión verbal de conocimiento dentro de la organización. Así, los trabajadores del CERN se convierten en personas que transmiten su conocimiento tácito (o partes del mismo) a otras personas, lo que convierte ese conocimiento en tácito en esas nuevas personas pero sólo unas pocas veces se convierte en conocimiento explícito, lo que limita enormemente el conocimiento organizacional como combinación de ambos tipos de conocimiento y se convierte en una clara desventaja competitiva (Maynez Guaderrama y Cavazos Arroyo, 2011, p.13-14), que no puede asumir una institución de investigación de alto nivel. Esto se complica bastante si además estas personas entran y salen de la organización en plazos pequeños de tiempo, lo que deja muchas veces la organización al margen (o casi) de esa transmisión de conocimiento. Con el paso del tiempo esto se hace insostenible, es un problema a resolver que no puede ser visto como un hecho aislado. Estos problemas de pérdida de información podían ser particularmente agudos en el CERN, pero en este caso (como en algunos otros), la organización de investigación avanzada era un modelo en miniatura del resto del mundo. Berners-Lee (1989) pensaba que si bien diez años después podría haber muchas soluciones comerciales a estos problemas, en ese preciso momento se precisaba disponer de alguna que permitiera continuar la tarea de la organización: un

nuevo sistema de gestión de la información que creciera y evolucionara con la misma y con sus proyectos. La supervivencia de una organización de investigación está íntegramente ligada a su capacidad para mejorar su gestión de información.

Para que esto fuera posible, el método de almacenamiento no debería poner sus propias restricciones a la información. Una "red" de notas con enlaces (referencias) entre los documentos era una solución mucho más útil que un sistema jerárquico fijo (típico de las carpetas de un administrador de ficheros clásico de un sistema operativo). Para describir un sistema complejo, muchas personas recurren a diagramas con círculos y flechas, esto permite describir relaciones entre los objetos de una manera que las tablas o directorios no pueden. Si llamamos a los círculos "nodos" y "enlaces" a las flechas e imaginamos cada nodo como una pequeña nota o pieza de información (da igual que sea un artículo, un resumen o un comentario), se puede construir un sistema vinculado de información entre personas y piezas informativas en constante evolución. A partir de ahora, la información de un proyecto no va a residir sólo en una carpeta de documentos que difícilmente un nuevo investigador va a poder reutilizar, ahora va a formar parte de la red informativa organizacional en la que se van a establecer vínculos entre otras personas y departamentos que permitirán la supervivencia de la información. Este sistema de almacenamiento va a conseguir implantar, al fin, la idea del hipertexto como sistema de información documental: "algo más que una simple herramienta de estructuración de información para su posterior recuperación, se debe considerar al hipertexto como una estructura activa que permite satisfacer necesidades de información mediante la elaboración de hiperdocumentos en los que se condensa la información relacionada en la red de hipertexto" (Pastor Sánchez y Saorín Pérez, 1995).

3. La solución de la Web.

En aquella época los científicos hacían un uso masivo de la difusión de la información por medio del protocolo de comunicaciones *uucp* (sistemas conocidos como 'newsgroups' que posteriormente fueron evolucionando en la Web hacia los "grupos de discusión" del tipo Yahoo Groups o similares). Un tema a debatir en esos sistemas daba lugar a una sucesión de temas paralelos, a su vez entrelazados entre sí de tal forma que una estructura tipo árbol como la que sustenta este protocolo no podía manejar (además del problema que representaba localizar un comentario en particular sin disponer de un sistema de búsqueda para ello), por tanto hacía falta diseñar algo que permitiera una mejor relectura de estos documentos. Entonces existía una aplicación informática basada en la idea de hipertexto y desarrollada por la empresa *Apple* que permitía resolver algunos de estos problemas en entornos locales: *Hypercard*¹⁰. Ya en su primera época en el CERN (al principio de la década de los ochenta), Berners-Lee había escrito el código de un programa – '*Enquire*' - de

10 Aplicación informática de Apple Computer que se contó entre los primeros sistemas hipermedia con éxito anteriores a la WWW. Conceptualmente se acerca a una base de datos, pues almacena información, siendo además gráfico, flexible y capaz de crear ficheros fáciles de modificar. También incluía *HyperTalk*, un lenguaje de programación potente y relativamente fácil de usar, para manipular los datos y la interfaz de usuario.

características muy similares a la citada solución comercial y que empleaba para documentar el seguimiento de los proyectos de desarrollo de software (en sustitución de los recuerdos de su memoria personal). Este programa permitía “llenar un documento con palabras que, cuando se hiciera clic sobre ellas, darían lugar a otros documentos para su elaboración” (Wright, 1997).

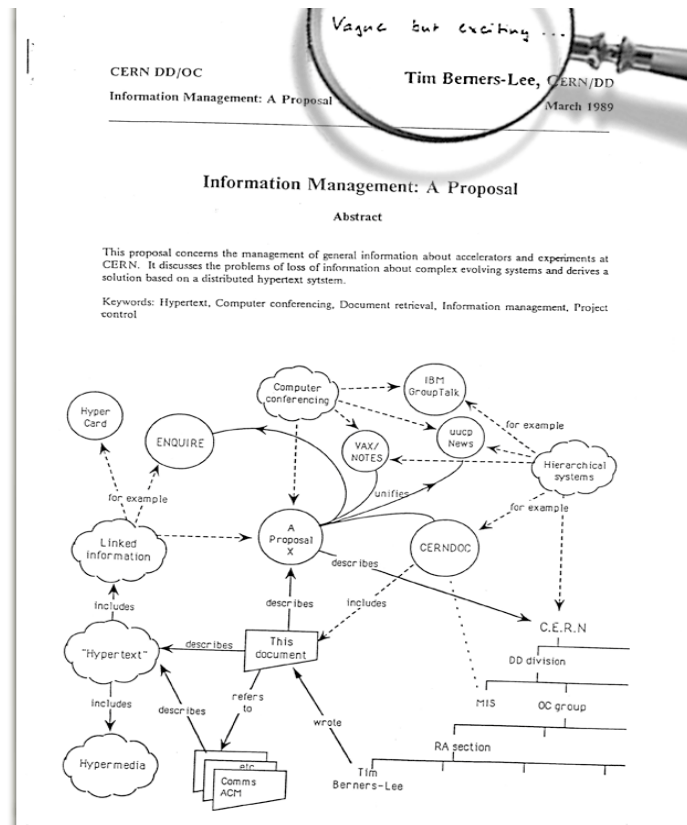


Fig. 3 Imagen de la primera página del memorándum con el que **Berners-Lee** presentó su propuesta al CERN. Fuente: <http://www.w3.org/History/1989/proposal>

Cuando Berners-Lee volvió al CERN casi una década después todo había cambiado: la propia organización, la aplicación *Hypercard* que estaba en pleno auge y también el uso de la propia red Internet había cambiado: “el mundo había cambiado. La Internet, aunque todavía desconocida para gran parte del público, estaba ahora firmemente arraigada. Era esencialmente una escueta infraestructura, un enrejado de tubos vacíos. Había formas de recuperar los datos, pero no existían maneras muy fáciles para llevarlo a cabo, y ciertamente, no había nada y nada con la estructura intuitiva y neuronal estructura del hipertexto” (Wright, 1997). Había llegado la hora de proponer un solución para este problema y la misma conllevó la creación de la Web. Realmente no se anduvo con medias tintas planteando propuestas teóricas pendientes de implantación posterior, todo lo contrario, buscó una solución a partir de nuevos desarrollos y adaptaciones de aportaciones

anteriores. De esta forma desarrolló el protocolo ligero de comunicaciones que permite llevar a cabo la navegación por la web – *http* - y desarrolló el lenguaje de marcado que permite la construcción de las páginas web: el HTML. A estas dos propuestas unió las soluciones informáticas para el soporte (servidor Web), para la navegación (clientes Web o navegadores), y la edición de textos en el lenguaje de marcado HTML. Con ello se puso en marcha el sistema de información que todo lo revolucionó.

Siendo posiblemente la mayor creación tecnológica jamás desarrollada en nuestro campo, su importancia queda algo reducida si la comparamos con el resultado obtenido una vez transcurridos algunos años: un verdadero orbe de información prácticamente ilimitado, de acceso libre y universal (sin limitaciones en lo geográfico como en las posibles discapacidades personales o tecnológicas), a través del verdadero invento que todo lo transformó: la red *Internet*. Si alguna vez Berners-Lee fue consciente de la grandiosidad de su invento sólo él lo sabe, y sólo a él, en justicia cabe preguntarle, los demás sólo podemos elucubrar al respecto. Lo que sí es cierto es el hecho irrefutable de que la Web y sus tecnologías asociadas forman parte del modo de vida cotidiano de gran parte de los ciudadanos de este mundo y muchos de ellos, los llamados “nativos digitales” no conciben la vida sin su existencia, especialmente desde la popularización del uso de aplicaciones como el correo electrónico *hotmail* o los motores de búsqueda *Google* o *Yahoo!* (por citar algunas de las tecnologías más clásicas), o la mensajería por *Whatsapp* o el intercambio de fotos, opiniones y comentarios en la red de microblogging *Twitter* (aplicaciones algo más contemporáneas).

La implantación de la Web fue meteórica, en pocos meses había ocupado el espacio como sistema de información distribuido que habitaba *Gopher*, un sistema de información de campus en formato de tablero electrónico que había popularizado su uso entre las comunidades universitarias y comenzaba a plasmar en la red dos ideas importantes: la interconexión de servidores *Gopher* (a modo de enlace entre dos sitios web) y la interoperabilidad con aplicaciones a las que se accedía por acceso remoto o *telnet*, como era el caso de los primeros módulos de consulta en línea de catálogos de bibliotecas (OPACs). Se comentaba que hacia 1995 existían en Internet más de dos millones de servidores basados en esta tecnología¹¹, si bien la experiencia de usuario y la multitud de ventajas aportadas por la Web convirtió el uso de esta aplicación en un recuerdo casi romántico en la mente de aquellos que la llegaron a manipular. Leonard Kleinrock, uno de los inventores de Internet¹², comentó en una entrevista hace unos años que lo más le gustaba de su invento era el hecho de “haber estado allí mientras todo pasaba”. A un nivel infinitamente menor de importancia y desde el punto de vista de un humilde profesor, muchos profesores de Tecnologías de la Información y Documentación Automatizada en los estudios de Biblioteconomía y Documentación en España

11 Algo que parece exagerado, si bien no le resta importancia alguna a este sistema.

12 Leonard Kleinrock es un científico de la Computación y profesor en UCLA. Es autor de diversas contribuciones extremadamente importantes en el campo teórico de las redes de ordenadores y desempeñó un papel importante en el desarrollo de la red ARPANET en UCLA, germen de la actual Internet.

tenían que utilizar diapositivas o versiones de demostración para simular a los alumnos la recuperación de información online en bases de datos bibliográficas (Medline, ERIC o Lisa por ejemplo) porque no se podían asumir los costes de la conexión telefónica. La conexión a Internet de las universidades y la creación de la Web acercó a los profesores y estudiantes a una industria de la información hasta entonces inaccesible y pudieron, de una manera humilde obviamente, colaborar en el desarrollo de este nuevo entorno que además ayudó a aumentar el conocimiento de la tecnología que lo sustentaba. De hecho, la creación de la Web coincidió con la mejora y liberalización del acceso a la Internet en muchos lugares del mundo, por lo que en el imaginario colectivo de buena parte de los ciudadanos reside la idea de que la Web trajo la Internet a nuestra vida, algo que tiene algo de cierto, incluso puede ser que mucho.

Y por una vez, que no la única afortunadamente, acompañando a profesores e investigadores en el descubrimiento de los posibles usos y aplicaciones de la Web, ahí estaban también los profesionales de la información participando en su desarrollo, no dejándolo (como ocurre con otras tecnologías) en las manos exclusivas de los informáticos. Y con ello comenzaron a publicarse las primeras páginas Web - la mayoría con un diseño manifiestamente mejorable - y se comenzó a dar forma a los primeros sitios Web para, poco a poco, conseguir la integración de la información con servicios y aplicaciones en el formato de portal Web (López Carreño, 2004). Todo esto no fue flor de un día, sino que precisó de algunos años para su consolidación y desarrollo, período de tiempo que, obviamente, no resultó igualitario entre países y organizaciones. En esa primera etapa se trataba de una Web de un sólo sentido: desde el editor (autor) al usuario (lector). Al principio no había retroalimentación ni interactividad alguna, algo que hoy en día parece imposible de concebir para muchos. En estos primigenios sitios web - la ahora llamada por algunos autores "Web 1.0" - solía incluirse una página con enlaces a un conjunto de otras páginas que el autor consideraba interesantes para sus lectores. Esta acción no era otra cosa que la traslación del muy tradicional servicio de referencia que desde tiempos inmemoriales llevan a cabo los profesionales de la información en las bibliotecas y constituyó el germen para el desarrollo de los primeros sistemas de recuperación de información en la Web: los índices o directorios, sistemas de los cuales *Yahoo!* fue durante un tiempo el mejor ejemplo. Los directorios, como todos recordarán son un producto documental considerado una fuente de información de carácter secundario porque dirige a la fuente original, justo lo que hacían y actualmente hacen estos sistemas de recuperación. Una actividad de gestión de información vuelve a confluir con la tecnología de la Web.

El crecimiento acaecido en la edición y publicación de contenidos en la Web (páginas fundamentalmente pero también otros tipos de documentos), hizo muy pronto imposible la gestión manual de estos sistemas de recuperación de información que sólo alcanzaban a realizar una revisión muy superficial de lo que se publicaba. Esto propició el desarrollo de los motores de búsqueda, sistemas que alimentan sus bases de datos a partir de la ejecución de unos programas de

rastreo ('crawlers') que recopilan direcciones de páginas a partir de los enlaces insertos en las mismas y las indexan de forma automática llevando a cabo una revisión mucho más profunda de lo publicado y de los cambios producidos en los documentos ya recopilados anteriormente. *Altavista*, *Lycos* y *AlltheWeb* fueron algunos de estos sistemas y representaron una innovación de gran impacto en su momento, por primera vez se podía acceder a grandes cantidades de documentos con sólo introducir unas pocas palabras en la ecuación de búsqueda sin necesidad de tener apenas que estudiar el lenguaje de recuperación de información. Si bien su aceptación fue espectacular, estos sistemas no estaban al margen de críticas ya que se basaban en algoritmos de rastreo e indexación desarrollados en entornos previos a la creación de la Web, y por esta causa, factores como el refresco de los índices elaborados con los documentos recopilados no estaban del todo bien implementados y tampoco tenían la capacidad de indizar la página Web completa. La suma de estos (y otros) inconvenientes, propició una amplia serie de dudas sobre la efectividad de la recuperación de información de estos sistemas, desarrollándose diversos estudios para proponer medidas que evaluaran esta efectividad¹³ (Martínez Méndez, 2001), (Martínez Méndez y Rodríguez Muñoz, 2003).

La desconfianza que se pudiera tener sobre el éxito de estos ingenios de búsqueda desapareció rápidamente cuando dos “niños prodigio” de la Universidad de Stanford crearon el motor de búsqueda *Google*, sistema que ya, casi desde el principio, superaba al resto en número de documentos indexados y en efectividad en sus operaciones de recuperación de información. ¿Cuál era la base de sustento del nuevo ingenio?, una vez más, sus autores se inspiraron en nuestro campo (la Información y Documentación) cuando diseñaron *Pagerank* su algoritmo de alineamiento de los resultados (Page et al., 1999) porque el mismo es una réplica del *factor de impacto* de las publicaciones periódicas, introducido por Gardfield (2006) y utilizado desde hace muchos años para establecer rankings de estas publicaciones científicas con base en el cómputo de citas recibidas por los artículos en ellas publicados. Este “nuevo” criterio para presentar los resultados, basado en el número de enlaces que una página Web recibe (y en la calidad de los mismos obtenida a partir del número de enlaces que entran y salen de las páginas que enlazan con la página analizada), fue aceptado en muy poco tiempo por los usuarios de los motores de búsqueda que prácticamente dejaron de utilizar el resto de sistemas hasta ese momento desarrolladas y llevaron a este nuevo motor a la posición predominante de uso que actualmente ocupa. Pero no ha sido sólo el modo de presentar los resultados lo que propició este éxito, otras razones fueron, sin duda alguna, la mejora en la revisión de los índices de páginas recopiladas, la mayor capacidad de indexación (de un pequeño conjunto de caracteres que se indexaba al principio se pasó en muy poco tiempo a indexar la totalidad de la página), la presentación agrupada de resultados de un mismo sitio Web, las

13 Algunos autores en tono algo catastrofista llegaron hablar de crisis en la recuperación de información en la WWW, postura algo radical pero con ciertas dosis de verosimilitud. En qué corto espacio de tiempo el ser humano pudo llegar a olvidar la gran cantidad de ventajas que estos sistemas trajeron consigo y el tiempo que se necesitaba antes de su existencia para localizar un documento sin ellos.

búsquedas sugeridas, etc. Todo ello ha estado acompañado de una serie de mejoras permanentes en las búsquedas y el lanzamiento de otros muchos servicios alrededor del motor que han convertido a *Google* en una de las principales empresas informáticas del planeta, si no la que más.

De esta forma, la puesta en marcha de un sistema para la Gestión de Información, diseñado originalmente para facilitar el seguimiento de proyectos de investigación y el aprendizaje organizacional, en muy poco tiempo engarzó con el campo de la recuperación de información y propició el desarrollo de los sistemas de búsqueda más avanzados desarrollados hasta ahora. Dos campos que parecían ir por separado convergieron de esta manera con muy interesantes resultados. Si bien no debería de extrañarnos porque para recuperar documentos es necesario que los mismos hayan sido almacenados previamente (es casi una ley de obligado cumplimiento), y ese proceso de almacenamiento forma parte del *ciclo de vida de la información*, es decir, es un proceso nuclear de la Gestión de Información (Wilson, 2002). Por tanto, la recuperación de información, de alguna manera “siempre ha estado ahí” aunque no se le había prestado mucha atención quizá y se había estudiado desde otros enfoques. A lo largo de estos veinticinco años han sido muchas las mejoras introducidas en estos sistemas. Lo cierto es que muchas de ellas pasan totalmente inadvertidas para los usuarios como son los constantes añadidos y/o sustituciones a los algoritmos de búsqueda, indexación y alineamiento, y otras no tanto, como la presentación de búsquedas sugeridas y/o personalizadas a partir de nuestro historial previo de búsquedas. El entorno donde se desenvuelven ahora estos motores no es el mismo del año 1995, el mismo ha cambiado sustancialmente y la edición y publicación de contenidos se ha multiplicado exponencialmente. La Web es ahora de doble sentido (Bacallao Pino, 2010), es decir, un usuario de la web (lector) puede convertirse con muy esfuerzo en un editor de contenidos (autor).

4. La Web 2.0: prosigue la convergencia.

El fenómeno del doble sentido de la publicación en la Web es una de las bases del actual paradigma llamado Web 2.0 (O'Reilly, 2005). El mismo tiene su origen en la popularización del intercambio de información vía *blogs*, es decir, una idea que comenzó para proporcionar un soporte documental a departamentos de desarrollo de tecnología (generalmente software), fue utilizada por científicos y escritores interesados en difundir sus conocimientos sin las rigideces impuestas por las editoriales científicas tradicionales propiciando el desarrollo de blogs científicos, algunos de ellos imprescindibles para el estudio de determinadas materias (como por ejemplo, la propia recuperación de información).

Rápidamente, cualquier usuario de la Web se convirtió en potencial autor y la proliferación de blogs de cualquier naturaleza y objetivo constituyó un éxito sin precedentes en el que se inspiraron, sin duda alguna, varios de los autores de servicios tan populares hoy en día como *Wikipedia* (fruto de la arquitectura de participación que “Jimbo” Wales había visto en los blogs y

otro ejemplo de cómo nuestras disciplinas forman parte del ADN de la Web), o *Facebook* (Zuckerberg percibió el interés que tenían muchas personas en compartir información de forma rápida y sencilla en cualquier momento y a cualquier hora). De alguna manera, la eclosión de estas aplicaciones sociales parece haber dejado en un segundo plano a las dos disciplinas objeto de esta charla, como si todo estuviera ya resuelto. Nada más lejano de la realidad, la proliferación de información presente en la Web y su gran ritmo de crecimiento viene a modelar un contexto muy parecido a cuando los técnicos de la *Biblioteca del Congreso* enrolados en el ejército norteamericano tuvieron que hacer frente a la gestión de esos millones de documentos microfilmados. Los motores de búsqueda deben ahora discernir entre el grano y la paja a la hora de identificar cuál es el verdadero blog autor de un texto frente a la cantidad de otros blogs y sitios Web que han copiado literalmente ese mismo texto (se habla de que sólo un 9% del contenido publicado en blogs es original).

Otro problema, no muy comentado y directamente implicado con esta cuestión, tiene que ver con la finalidad de la búsqueda, algo que resulta complicado de expresar en dos o tres palabras que suelen conformar la ecuación de búsqueda. Así, si un estudiante de Bachillerato necesita preparar una redacción sobre “sistemas de economía mixta de mercados”, además de la inevitable referencia a *Wikipedia* que el motor de búsqueda va a devolver en un lugar destacado de la respuesta (y que no le sirve porque la profesora ha dicho que no se use esa enciclopedia en línea como fuente), seguramente los motores de búsqueda van a devolver una amplia cantidad de programas de asignaturas de Economía impartidos en varias universidades, documentos muy escuetos para ayudar a este alumno en la elaboración de la redacción. En esta tesitura, la sobreabundancia de documentos juega en su contra y le obliga a hacer algo que ya habíamos olvidado (o casi), consultar la segunda o tercera página de resultados de búsqueda del motor. Una posible solución sería poder suprimir este tipo de documentos en la propia ecuación de búsqueda, o indicar en la misma que los documentos a recuperar sean fuentes de referencia para el estudio. Este tipo de ideas se satisfacen parcialmente con sistemas de búsqueda tipo *Google Académico* pero, dependiendo de la materia y del nivel, no son del todo satisfactorios.

La *Web Semántica* (Berners-Lee and Hendler, 2001), representa una nueva oportunidad de sinergia entre gestores de información y los desarrolladores de sistemas informáticos para la recuperación de información. La idea original de aportar una capa semántica a las páginas Web por medio de la inclusión de metadatos no ha sido un éxito en una Web abierta y generalista (Zhang and Dimitroff, 2004, 311), (Phelps, 2012, 328). Sólo en algunos sistemas de información más específicos (bibliotecas y otras colecciones digitales) se detecta un mayor uso de esquemas de metadatos para proporcionar más información sobre el contenido de los documentos a los motores de búsqueda y mejorar las capacidades de recuperación de información. Sin embargo, la idea es buena y positiva y poco a poco parece que comienza a consolidarse por la vía de los denominados

microformatos, especialmente los microdatos (Google, 2015), que sí parecen ser objeto de interés por parte de los motores porque ayudan en una presentación más dinámica y estética de los resultados de búsqueda y han llevado a cabo una alianza para el desarrollo de schema.org (García Marco, 2013), (Sulé, 2015). Llevar a cabo un proceso de documentación de las páginas Web con microdatos es una buena manera de garantizar su presencia entre los resultados destacados de una búsqueda y de promocionar el sitio Web en los motores de búsqueda sin tener que recurrir a costosísimas inversiones en marketing digital. Además, se trata de una actividad de Gestión de Información, el “matrimonio” entre estas disciplinas sigue vivo.

El campo del *Linked Open Data* (Bizer et al., 2008) y la disposición de datos abiertos para la transparencia y buen gobierno por parte de los estados (Concha y Naser, 2012), es otro nexo de unión entre ambas disciplinas. Son los documentalistas quienes deben preparar esos documentos que han de ser puestos a disposición de los ciudadanos en portales Web y también han de ser quienes se preocupen de velar por la actualización y la calidad de la información en ellos contenida.

5. Conclusiones.

En definitiva, son muchos los frentes abiertos en el campo de la recuperación de la información en la Web que deben ser resueltos. En la mayoría de casos, el éxito de la solución informática propuesta dependerá en gran medida del correcto tratamiento de la información. En otros casos, la solución informática podrá paliar fallos o deficiencias pero no terminará de ocultar una inadecuada gestión de información previa. Lo que está claro es que la Web y la Gestión de Información son disciplinas interdependientes y complementarias, ahora y en un futuro inmediato.

El propio Berners-Lee (2014) escribió un post en el blog oficial de Google en conmemoración de los 25 años transcurridos desde que presentó su propuesta de desarrollo de un sistema de gestión de información a sus jefes del CERN, con la idea de mejorar los flujos de información en la organización por medio de una: "web de notas con enlaces entre ellos". En este comunicado, recordaba que, como el CERN es un centro de investigación altamente especializado en Física, este desarrollo se aprobó casi como proyecto secundario porque no podían dedicar recursos a algo tan genérico y no estrictamente relacionado con la institución: la Gestión de Información. Esta decisión fue la que permitió a miles de personas comenzar a trabajar juntos en la construcción de la Web, compartiendo información y generando con el paso del tiempo miles de millones de euros de valor económico, transformando actividades tradicionales como la educación y la asistencia sanitaria, activando movimientos sociales y creando nuevos hábitos y necesidades, en aquella época posiblemente insospechados. Y para su inventor, Tim Berners-Lee, prácticamente aun estamos empezando. Eso sí, además de celebrar los primeros 25 años de vida de este sistema, también es una ocasión para pensar, hablar y hacer. Las decisiones clave sobre la gobernabilidad y

el futuro de Internet se avecinan, y es de vital importancia hablar del futuro de la Web y responder a esta serie de cuestiones:

1. ¿cómo podemos asegurarnos de que el “otro 60 por ciento” de personas de todo el mundo que aun no están conectados podrán acceder a Internet?
2. ¿cómo podemos asegurarnos de que la WWW es compatible con todos los idiomas y culturas, no sólo las dominantes?
3. ¿vamos a permitir que otros empaqueten y restrinjan nuestra experiencia en línea - el fin de la neutralidad de la Web que se han planteado algunos gobiernos y operadores telefónicos (Cullell March, 2012) -, o vamos a proteger a la magia de la Web abierta y el poder que nos da para decir, descubrir y crear libremente?

Estas son algunas de las cuestiones que plantea Berners-Lee (2014), quien pregunta a continuación: "¿y cuáles son las tuyas?", exigiéndonos que le ayudemos a imaginar y construir los futuros estándares de la Web y que presionemos en cada país para desarrollar un proyecto de ley de los derechos digitales que avance en la consolidación de una Web libre y abierta para todo el mundo.

Referencias

- Bacallao Pino, L.M. (2010) Representaciones mediáticas de las redes sociales: un estudio de casos. *Revista Latina de Comunicación Social*, 65, 114-125. Disponible en: http://www.revistalatinacs.org/10/art/887_UZaragoza/09_Lazaro_Bacallao.html [21-06-2015]
- Berners-Lee, T. (1989) Information Management: A Proposal. <http://www.w3.org/History/1989/proposal.html> [18-11-2014]
- Berners-Lee, T. and Hendler, J. (2001) Publishing on the semantic web. *Nature* 410, 1023-1024 (26 April 2001) doi:10.1038/35074206
- Berners-Lee, T. (2014, 11th March) *On the 25th anniversary of the web, let's keep it free and open*. [Blog] Disponible en: <https://googleblog.blogspot.com.es/2014/03/on-25th-anniversary-of-web-lets-keep-it.html> [18-11-2014]
- Berners-Lee, T. (2014) *On the 25th anniversary of the web, let's keep it free and open* [blog]. <https://googleblog.blogspot.com.es/2014/03/on-25th-anniversary-of-web-lets-keep-it.html> [18-11-2014]
- Bush, V. (1945, July). As we may think. *Atlantic Monthly*, 176, 101-108 <http://www.theatlantic.com/magazine/archive/1945/07/as-we-may-think/303881/> [18-11-2014]
- Bizer, C., Heath, T., Idehen, K. & Berners-Lee, T. (2008). Linked data on the web (LDOW2008). *Proceedings of the 17th international conference on World Wide Web*, 1265-1266.
- Cantos Gómez, P., Martínez Méndez, F.J. y Moya Martínez, G. (1994) Hipertexto y Documentación. Murcia: Universidad, Servicio de Publicaciones. Disponible en

https://books.google.es/books/about/Hipertexto_y_documentaci%C3%B3n.html?id=_4GTzaNTIosC&hl=es [21-06-2015]

Conklin, J. (1987) Hypertext: an introduction and survey. *Computer*, 9 september 1987, vol. 20. p.17-41 doi:10.1109/MC.1987.1663693

Concha, G. y Naser, A. (2012) El desafío hacia el gobierno abierto en la hora de la igualdad. CEPAL, Naciones Unidas. Disponible en: <http://iis7-e2.cepal.org/ddpe/publicaciones/xml/9/46119/W465.pdf> [21-06-2015]

Cullell March, C. (2012). El futuro de la Web ante la neutralidad de la Red: estado de la cuestión en la Unión Europea. *El Profesional de la Información* 21, 1. Disponible en: <http://dx.doi.org/10.3145/epi.2012.ene.10>

Dervin, B. & Nilan, M.(1986) Information needs and uses, *Annual Review of Science and Technology* 21, 3-32

García-Marco, F.J. Schema.org: la catalogación revisitada. *Anuario ThinkEPI*, 2013, 7, 169-172. Disponible en: <http://www.thinkepi.net/schema-org-catalogacion-revisitada> [21-10-2015]

Gardfield, E. (2006). The Agony and the Ecstasy - The history and meaning of the journal impact factor. *Jama*, 295 (1), 90-93. Disponible en <http://www.garfield.library.upenn.edu/papers/jifchicago2005.pdf?wa=IPEMB114> [21-06-2015]

Google. (2015) *Promote Your Content with Structured Data Markup*. Disponible en: <https://developers.google.com/structured-data/> [21-10-2015]

Kirk, J. (1999) Information in organisations: directions for information management" *Information Research*, 4 (3). Disponible en: <http://informationr.net/ir/4-3/paper57.html> [21-06-2015]

Llançó i Sanjuan, J. (1993) *Gestión de documentos. Definición y análisis de modelos*. Vergara: Centro de Patrimonio Documental de Euskadi. Disponible en: <http://eah-ahe.org/pdf/ikerlanak7.pdf> [12-09-2015]

López Carreño, R. (2004) Análisis taxonómico de los portales periodísticos españoles. *Anales de Documentación*, 7, 123–140. Disponible en <http://revistas.um.es/analesdoc/article/view/1671/1721> [18-11-2014].

Martínez Méndez, F.J. (2001) Aproximación general a la evaluación de la recuperación de información mediante motores de búsqueda en Internet. *Scire: representación y organización del conocimiento*. 7, 1, 11-31 Disponible en: <http://ibersid.eu/ojs/index.php/scire/article/viewFile/1139/1121> [20-11-2015]

Martínez Méndez, F.J. y Rodríguez Muñoz, J.V. (2003) Síntesis y crítica de las evaluaciones de la efectividad de los motores de búsqueda en la Web. *Information Research: an international electronic journal*, 8 (2), paper 148. Disponible en: <http://www.informationr.net/ir/8-2/paper148.html> [20-11-2015]

Maynez Guaderrama, A. y Cavazos Arroyo, J. (2011) Conocimiento tácito: su transferencia dentro de la organización, como fuente de ventaja competitiva sostenible. *Administración y*

- organizaciones, Junio 2011, p.10-26 Disponible en: http://148.206.107.15/biblioteca_digital/articulos/9-578-8224ife.pdf [11-07-2015]
- Nelson, T.H. (1981) *Literary Machines: the report on, and of, Project Xanadu concerning word processing, electronic publishing, hypertext, thinkertoys, tomorrow's intellectual revolution, and certain other topics including knowledge, education and freedom*. Sausalito, CA: Mindful Press.
- Nelson, T.H.; (s.f.) *Xanalogical Structure, Needed Now More than Ever: Parallel Documents, Deep Links to Content, Deep Versioning and Deep Re-Use*. <http://www.xanadu.com.au/ted/XUsurvey/xuDation.html> [12-07-2015]
- O'Reilly, T. (2005) *What is Web 2.0: Design Patterns and Business Models for the Next Generation of Software*. <http://www.oreilly.com/pub/a/web2/archive/what-is-web-20.html> [18-11-2014]
- Page, L.; Brin, S.; Motwani, R. and Winograd, T. (1999) *The PageRank Citation Ranking: Bringing Order to the Web*. Technical Report. Stanford InfoLab. (1999) *The PageRank Citation Ranking: Bringing Order to the Web*. Technical Report. Stanford InfoLab. Disponible en: <http://ilpubs.stanford.edu:8090/422/1/1999-66.pdf> [21-06-2015]
- Pastor Sánchez, J.A. y Saorín Pérez, T. (1995) El hipertexto documental como solución a la crisis conceptual del hipertexto. El reto de los documentos cooperativos en redes. *Cuadernos de Documentación Multimedia*, número 4, junio de 1995. Disponible en: <http://pendientedemigracion.ucm.es/info/multidoc/multidoc/revista/cuadern4/hiperdoc.htm> [18-11-2014]
- Phelps, T.E. (2012) An Evaluation of Metadata and Dublin Core Use in Web-Based Resources. *Libri* 62, 4, 326-335.
- Schema.org (s.f.) Getting started with schema.org using Microdata. Disponible en: <https://schema.org/docs/gs.html> [8-11-2015]
- Sulé, A. Schema.org, la mejora de la visualización de los resultados en los buscadores y mucho más. *BID: textos universitarios de biblioteconomía y documentación*, 34 (junio 2015). Disponible en: <http://bid.ub.edu/es/34/sule.htm> [8-11-2015]
- Von Neumann, J. (1945) *First Draft of a Report on the EDVAC*. <https://web.archive.org/web/20130314123032/http://qss.stanford.edu/~godfrey/vonNeumann/vnedv.ac.pdf> [18-11-2014]
- Wilson, T.D. (1989). Towards an information management curriculum. *Journal of Information Science*, 15, 203-210 Disponible en: <http://www.informationr.net/tdw/publ/papers/infmagt89.html> [8-6-2014]
- Wilson, T.D. (2002) Information Management. *International Encyclopedia of Information and Library Science*, 2nd ed. http://informationr.net/tdw/publ/papers/encyclopedia_entry.html [18-11-2014]
- Witehead, J. *Orality and Hypertext: An Interview with Ted Nelson*. http://www.ics.uci.edu/~ejw/csr/nelson_pg.html [12-07-2015]

Wright, J. (1997) The man who invented the web. *Time*, May 19, Vol. 149 No. 20. Disponible en: http://faculty.washington.edu/mfan/ebiz509/download/readings/TheManWhoInventedWeb_Time97.pdf [18-11-2014]

Zhang, J. and Dimitroff, A. Internet search engines' response to metadata Dublin Core implementation. *Journal of Information Science*, 30, 4, 310-320.