

UNIVERSIDAD NACIONAL
AUTÓNOMA DE MÉXICO

FACULTAD DE FILOSOFÍA Y LETRAS
DIVISIÓN DE ESTUDIOS DE POSGRADO

**“ORGANIZACIÓN DOCUMENTAL EN
LA SOCIEDAD DE LA INFORMACIÓN”**

TESIS

QUE PARA OBTENER EL GRADO DE:

**MAESTRA EN BIBLIOTECOLOGÍA Y ESTUDIOS
DE LA
INFORMACIÓN**

PRESENTA:

BEATRIZ JUÁREZ SANTAMARÍA

DIRECTOR DE TESIS:

DRA. GEORGINA ARACELI TORRES VARGAS



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

AGRADECIMIENTOS

Agradezco infinitamente todo el apoyo brindado durante el desarrollo de esta investigación, a mi asesora la Dra. Araceli Torres.

Quiero agradecer también al jurado conformado por el Dr. Filiberto Martínez, Dra. Catalina Naumis, Dr. Ariel Rodríguez y la Dra. Brenda Cabral por sus valiosos comentarios que hicieron a este trabajo.

Dedico esta tesis a mi padre, pero sobretodo a mi esposo Ricardo y a mi hija Dafne por todo el tiempo que no les dediqué durante la elaboración de esta investigación.

TABLA DE CONTENIDO

Introducción	i
---------------------------	---

Capítulo 1. La sociedad de la información en los albores del siglo XXI

1.1 Sociedad de la información	1
1.2 Brecha digital	10
1.3 El mercado de la información en la sociedad de la información	13

Capítulo 2. Organización documental en la sociedad de la información.

Estado actual

2.1 Situación actual de la organización documental	22
2.1.1 Documento digital.....	22
2.1.2 Objeto digital... ..	24
2.1.3 Repositorios.....	26
2.1.4 Creación de colecciones digitales mediante software libre.....	31
2.1.5 Organizar Internet <i>vs.</i> Organizar colecciones digitales	36
2.1.6 Sistemas para la organización de colecciones digitales	44
2.1.7 Lenguajes de marcado.....	50
2.1.8 Metadatos	55

Capítulo 3. Organización documental en la sociedad de la información.

Tendencias

3.1 Tendencias en la organización documental.....	65
3.1.1 La web 2.0 en Internet y en la biblioteca.....	69
3.1.2 La biblioteca 2.0.....	72
3.1.3 Las folksonomías en Internet y en la biblioteca	74
3.1.4 La web semántica en Internet y en la biblioteca	80
3.1.5 Redes académicas avanzadas	83
3.1.6 La biblioteca híbrida, Information Commons y Centro de Recursos para el Aprendizaje y la Investigación (CRAI).....	87

3.2 Alternativas para la organización documental en la sociedad de la información.....	92
Reflexiones finales.....	98
Referencias.....	100

INTRODUCCIÓN

En la sociedad actual han aparecido nuevas estructuras documentales, como bases de datos, hipertexto, multimedia y publicaciones digitales, por lo que ha surgido también la necesidad de organizar dichas estructuras.

Desde el punto de vista bibliotecológico, la organización documental, tiene entre sus finalidades preservar y difundir el conocimiento que se registra a través de los documentos digitales y orienta en el diseño y uso de los productos tecnológicos para recuperar los contenidos, por lo que podemos señalar que ésta se refiere a:

1. Determinar qué recursos existen y seleccionar los que sean revelantes para los usuarios.
2. Describir los recursos seleccionados.
3. Proporcionar puntos de acceso.
4. Analizar los contenidos de los recursos seleccionados.
5. Proporcionar la información necesaria para la localización de esos recursos.

Existen una serie de investigaciones y proyectos en los cuales se describe la importancia del empleo de las tecnologías para el procesamiento de la información. Al respecto Ariel Sánchez señala que la interrelación y convergencia entre las comunicaciones y la computación son fundamentales, pues la tecnología es cada vez más accesible para el procesamiento de la información.¹

Es por ello que el propósito de la presente investigación está orientado a analizar las tendencias que han emergido para establecer la organización documental, a partir de la influencia de las tecnologías de la información y la

¹ Ariel Sánchez Espinoza, *Análisis de las interacciones entre las tecnologías de la información y la sociedad de la información del siglo XXI en México*, p. 20.

comunicación (TIC), con el fin de poder definir cuál puede ser la mejor opción para ordenar las colecciones digitales.

Las preguntas que a continuación se mencionan se desprenden de la siguiente hipótesis: *los metadatos pueden sustentar una organización documental apropiada para la sociedad de la información.*

De acuerdo con lo señalado anteriormente se desprenden los siguientes planteamientos:

- ¿Cuáles pueden ser las alternativas más adecuadas para el establecimiento de una organización documental en el contexto de la sociedad de la información?
- ¿Los metadatos pueden ser la mejor alternativa?

A partir de las preguntas anteriores, se realizó lo siguiente:

- 1) La contextualización de la organización documental en el marco de la sociedad de la información, ya que los ejes en los que se desarrolla esta sociedad han determinado en gran medida las pautas para el ordenamiento de los documentos, sobre todo de los digitales. Esta tendencia por ver más hacia los documentos digitales que a los impresos, también es consecuencia de las propuestas que devienen de la sociedad de la información.
- 2) Una revisión bibliográfica, hemerográfica y de fuentes electrónicas de la situación en la que se encuentra la organización documental en la sociedad de la información.
- 3) El análisis de las tendencias que perfilan el desarrollo inmediato en la organización documental.
- 4) La propuesta de alternativas idóneas para lograr una organización documental adecuada en la sociedad de la información.

La tesis se compone de tres capítulos. En el capítulo uno, se desglosan las características de la llamada sociedad de la información y se describen los planos que la integran: económico, social y personal, así como el papel que juega la información dentro de las empresas.

En el capítulo dos, se habla de las herramientas que permiten organizar las nuevas representaciones de la información (materiales) en colecciones digitales, como son los sistemas de organización del conocimiento, los lenguajes de marcado, los metadatos, etc., con la finalidad de evaluar cuáles nos permitirán identificar, seleccionar y obtener dichos recursos de información.

En el capítulo tres, se analizan las tendencias que surgen en la sociedad de la información mediante el uso de las TIC, como son las llamadas: web 2.0, biblioteca 2.0, folksonomías y web semántica, en donde se mencionan las ventajas que ofrecen dichas tendencias para poder crear, preservar, utilizar y difundir la información y el conocimiento. Además, se proponen algunas alternativas que se consideran las más indicadas para organizar las colecciones digitales en la sociedad actual.

Por último, se incluyen las reflexiones finales en donde se confirma la hipótesis planteada en la Introducción de esta investigación, así como las obras consultadas.

CAPÍTULO 1

La sociedad de la información en los albores del siglo XXI

En este capítulo se hablará de la “sociedad de la información”, cuyo elemento principal, como su nombre lo indica, es la información conjugada con el uso de las tecnologías. Además, se presentarán algunas definiciones y características de ésta, así como las ventajas y desventajas que poseen los planos que la integran, como son el económico, social y personal. Asimismo, se describe y analiza la importancia que tiene la información y cuál es el papel que desempeña dentro de la sociedad de la información.

1.1 Sociedad de la información

A finales de los años sesenta aparece la denominada sociedad de la información, la cual se apoya en las tecnologías de la información y la comunicación (TIC), con lo cual se legitima este concepto como sociedad global. Uno de los elementos fundamentales de ésta son las redes de comunicación. Aquí cabe precisar que la computadora es sólo un medio de acceso a la sociedad de la información, no es su centro vital y que su función es facilitar el acceso a la información.

La sociedad de la información se refiere “al proceso de transformaciones sociales, económicas y políticas que se están produciendo como consecuencia del desarrollo de las tecnologías de información y las comunicaciones y a la capacidad prácticamente ilimitada que su uso aporta para obtener, almacenar, procesar y compartir información”.¹

Masuda define a esta sociedad de la información “como aquella que crece y se desarrolla alrededor de la información y que aporta un florecimiento general de la creatividad intelectual humana, en lugar de un aumento del consumo material”.²

A continuación se mencionan algunas características de la sociedad de la información que destacan Naisbitt y Toffler:

Naisbitt	Toffler
“La sociedad de la información es una realidad económica no una abstracción intelectual”.	La información pasa a ser el recurso central de la economía avanzada.
Las innovaciones en las telecomunicaciones y en las tecnologías de la computación van a acelerar el paso del cambio al producir el derrumbe de	El software, la información y las telecomunicaciones pueden ser elevados a nivel global.

¹ Plan Euskadi en la sociedad de la información. p.14, en <<http://www.euskadi.net/eeuskadi/datos/docs/plan>>, consultado el 25 de septiembre de 2008.

² MASUDA, Y. *La sociedad informatizada como sociedad postindustrial*. Madrid: Tecnos, 1994.

barreras de la información.	
Las nuevas tecnologías de la información al principio se aplicarán a antiguas tareas industriales; después y gradualmente, darán origen a actividades, procesos y productos nuevos.	Las tecnologías de información y comunicación activan las cosas, nos llevan hacia una economía instantánea en tiempo real.
En esta sociedad de intensiva alfabetización, el sistema educativo producirá individuos mejor preparados.	El conocimiento, a veces nada más como información y datos, puede utilizarse como sustituto de otros recursos.
La tecnología de la nueva era de información no es absoluta, tendrá éxito o fracasará de acuerdo con el principio de alta tecnología.	La información reduce la necesidad de materias primas, trabajo, tiempo, espacio y capital". ³

Además de las características mencionadas en la tabla anterior, es importante señalar que la educación en la sociedad de la información también juega un papel primordial, pues incrementa la capacidad económica, social, política y cultural de los ciudadanos. Al respecto Chomsky señala que para conseguir esa capacidad en los ciudadanos, es importante que coincidan los objetivos de los educadores y las

³ ESTUDILLO García, Joel. "Surgimiento de la sociedad de la información", *Biblioteca Universitaria*, nueva época, vol. 4, no. 2 (julio-diciembre): p.82.

demandas de los empresarios, políticos y otros grupos sociales para convertir a la educación en uno de los factores claves del progreso y crecimiento de los países.⁴

La educación debe considerarse como el componente esencial del desarrollo nacional, ya que proporciona a los ciudadanos, la oportunidad de volverse más productivos, escapar de la pobreza y mejorar así su calidad de vida,⁵ y, para ello, se requiere que los gobiernos inviertan en educación, capacitación, ciencia y tecnología.

En la sociedad de la información la calidad de vida, el sistema educativo y las formas de trabajo dependen cada vez más de la incorporación de las innovaciones tecnológicas, ya que permiten la creación, intercambio, recuperación, difusión y comercialización de la gran cantidad de información generada en la sociedad actual.

La importancia de la información se debe a que:

1. No es consumible, es decir, no se extingue con el uso.
2. No es transferible, pues quien la suministra no se deshace de ella y puede estar al mismo tiempo en todas partes.
3. Es indivisible, sólo puede utilizarse como conjunto.
4. Es acumulativa, cada gota de información posibilita conseguir más información.
5. No pertenece a ninguna industria ni a ningún negocio en particular.

⁴ CHOMSKY, Noam, Dietrich Heinz. *La sociedad global: educación, mercado y democracia*. p. 82.

⁵ *Ibidem*, p. 83.

6. La información se distingue de los demás productos básicos porque se sitúa en el marco de la abundancia y no de la escasez.

Debido a lo anterior, la información en esta sociedad es cada vez más importante en la vida económica, social y política. Un ejemplo de ello es el crecimiento de las profesiones especializadas en actividades de la información, las cuales ocupan un número mayor de empleos en muchas sociedades industrializadas. La categoría de más alcance es la de los procesadores de la información (que se refiere a la generación, almacenamiento, transmisión y manipulación de ésta), seguida por la de los productores de información, distribuidores y trabajadores de infraestructura.

La aceptación de las nuevas tecnologías de información ha sido más rápida que la de las otras tecnologías del pasado como son las máquinas de vapor o el motor eléctrico.

Detrás de la sociedad de la información se encuentran dos fenómenos: la globalización y las tecnologías de la información. Como se señaló anteriormente, las transformaciones que estamos viviendo abren grandes oportunidades de desarrollo en los planos: económico, social e individual, como se explica a continuación.

1.1.1. Plano económico

Existe la globalización de la oferta y la demanda en donde los consumidores tienen acceso través de Internet a una inmensa gama de productos y servicios. Asimismo, se ha dado una reducción de precios, ya que no existen barreras geográficas, y la innovación y mejora en la calidad de los servicios.

En un sector primario se incluyen todas las industrias que utilizan su información en mercados establecidos o en cualquier parte donde se le pueda otorgar un valor económico, como son los medios masivos de comunicación y la educación, entre otros, y en el sector secundario, se incluye la investigación y el desarrollo.

Estados Unidos es ahora una economía basada en la información, ahí las principales áreas de actividad económica son los productores de bienes y servicios de información.

1.1.2 Plano social

La democratización del acceso a la información permite avanzar hacia una mayor igualdad de oportunidades y participación de los individuos, es decir, verán incrementadas las oportunidades de empleo, educación y ocio.

Los gobiernos deben estar convencidos de que la educación será una de las características principales de la sociedad de la información y que la calidad de vida dependerá del acceso a la información.

En este rubro surgen nuevas formas de trabajo como es el teletrabajo.

1.1.3 Plano personal

La nueva sociedad se articula sobre la información y el conocimiento, lo que hace que las personas vuelvan a ocupar el centro de todo el proceso. La nueva era de la información es también la era de la sociedad civil y debe ser ella su principal protagonista.

Por ello, en esta sociedad

los estándares de vida, los patrones laborales, el sistema educativo y el mercado están marcadamente influenciados por los avances de la información y el conocimiento; esto es evidenciado por un incremento en los productos y servicios de información, comunicados a través de un amplio rango de medios, muchos de ellos de naturaleza electrónicos.⁶

La sociedad de la información se estructura en torno a redes en las que cada individuo es una fuente de información que interactúa con los demás, por lo que ésta debe constituirse como factor de integración social y de mejora en la calidad de vida de los individuos.

En este contexto, los individuos deberán ser:

- protagonistas, pues el futuro depende de todos;
- integradores en una sociedad de la información que pertenece a todos, y

⁶ William J. Martin. *The Global Information Society*. London: ASLIB-Gower, 1997, p. 47.

- formadores en la utilización de las nuevas técnicas y herramientas en un mundo donde los elementos de éxito son el nivel de conocimientos, el continuo aprendizaje y la adaptabilidad.

En la sociedad de la información existe la llamada *n-economía* que se basa principalmente en el conocimiento; ésta es dirigida por los trabajadores del conocimiento: profesionales, técnicos y consumidores del conocimiento. Los individuos que se mantengan informados y actualizados continuamente serán los que tendrán una participación activa en la *n-economía* y en la *e-conocimiento*.

Los símbolos de la sociedad de la información son las computadoras y las tecnologías de la información. Las tecnologías de la información son el centro de la sociedad informativa y los principales recursos de la economía que determinan la competitividad de las corporaciones y economías.

Por otro lado, se cree que las tecnologías ponen a la disposición de todos, la información, lo cierto es que ésta tiene un alto costo y no podrá ponerse al servicio de todos sin la intervención del gobierno. Por lo que actualmente se empieza a hablar de pobreza informativa y los países se agrupan de acuerdo con su acceso a infraestructura y telecomunicaciones, es decir, ha aparecido una nueva marginación llamada “brecha digital” para aquellos ciudadanos sin acceso a la información.

México, por ejemplo “se encuentra entre los países que tienen un bajo porcentaje de personas conectadas a Internet. Los países desarrollados tienen cerca del 70% de su población con conectividad, es decir entre el 12% y el 14%”.⁷

No todos los países cuentan con dicho recurso, es decir, los que están en desarrollo poseen escasos recursos económicos para poner en marcha acciones relacionadas con políticas y estrategias de información. En un país pobre el gobierno le resta importancia a dichas acciones porque sus prioridades son otras, como la producción agrícola y la industrial, ya que su desarrollo se fundamenta en éstas.

Los requisitos básicos de un país en desarrollo para entrar a la sociedad de la información son las tecnologías inalámbricas, satélites de uso múltiple y sistemas celulares.

Los países que han comenzado a construir una infraestructura para la sociedad de la información son Brasil, Rusia y Sudáfrica, pero por razones militares, no civiles.

Zaldívar señala que la formulación de políticas y estrategias de información debe hacerse de forma coordinada entre las diversas instituciones gubernamentales y privadas teniendo en cuenta sus necesidades y prioridades. El

⁷ Adolfo Rodríguez Gallardo. *Tecnologías de la información en México 2001-2005*. México: UNAM, Centro Universitario de Investigaciones Bibliotecológicas, 2005. p. 69.

papel del gobierno como ente catalizador entre los diversos actores sociales, es fundamental para el desarrollo exitoso de éstas.⁸

Estados Unidos es la primera sociedad global y es el único país que está en condiciones de llevar a cabo la revolución de la información; ya que cuenta con una vasta infraestructura en telecomunicaciones.

Cabe señalar que la sociedad de la información facilita la comunicación entre especialistas que intercambian información sobre sus disciplinas.

Actualmente está iniciando una nueva era en donde utilizar la información es decisivo e importante, de hecho ésta ha llegado a los hogares en donde existen nuevos sistemas domésticos, como realizar compras a través de Internet (telecompra).

El resultado de la sociedad de la información dependerá de las acciones y decisiones sociales, no sólo de los avances tecnológicos.

1.2 Brecha digital

En la sociedad de la información existe la llamada “brecha digital” que se refiere a los que no están incluidos en ella, es decir, los que no tienen contacto con las TIC. Es una línea que separa a las personas que ya se comunican y realizan sus actividades mediante redes digitales de quienes no han alcanzado este desarrollo.

⁸ Modesto Zaldívar Collazo. “Políticas y estrategias en la Sociedad de la Información”, en *Ciencias de la Información*, vol. 29, no. 3 (septiembre de 1988): p. 39.

La versión moderna de la brecha tecnológica es la brecha digital, expresión que sirve para describir las limitaciones de acceso a las nuevas tecnologías de la información: telecomunicaciones e informática.

Hablar de brecha digital es hablar de algo más que la falta de computadoras, ya que la tecnología no significa nada si no es usada, la gente que no utiliza las TIC no sólo se debe a que no saben hacerlo, sino también a que los gobiernos no han implementado la tecnología adecuada para que sea accesible.

En esta sociedad las diferencias de los individuos no se establecerán entre ricos y pobres, sino los que están conectados a la red y los que no lo están. Las desigualdades educativas, económicas y de infraestructura se pueden ver acrecentadas si no se toman las medidas pertinentes en la utilización de las tecnologías en beneficio de todos los grupos de la sociedad.

Por otro lado, el sector privado debe tener como prioridad el objetivo de reducir la brecha digital mediante la realización de inversiones y el suministro de servicios de comunicaciones electrónicas.

En América Latina existe una Asociación llamada AHCINET que es la que apoya el desarrollo de la sociedad de la información en Iberoamérica contribuyendo a la disminución de la brecha digital existente. Ésta es una asociación de empresas de telecomunicaciones no lucrativa, constituida en 1982 con los siguientes objetivos:

Ser punto de encuentro de las telecomunicaciones iberoamericanas para la cooperación, el intercambio de experiencias y la identificación de sinergias.

Fomentar el diálogo con organizaciones internacionales, reguladores y gobiernos para estimular el desarrollo sostenible de las telecomunicaciones y la sociedad de la información.⁹

AHCIET impulsa el desarrollo de la sociedad de la información involucrando a todos los sectores que tengan que intervenir, es decir, trata de eliminar la brecha digital que existe en dicha sociedad. Desarrolla también la capacitación en el uso de las tecnologías en la sociedad y trata de atraer inversión para acelerar el desarrollo de la sociedad de la información.

En Latinoamérica lo más importante para la sociedad de la información es la promoción de lo digital. Esto permitirá promover el desarrollo económico y social en la población de Latinoamérica.

Cabe señalar que el e-gobierno es la clave en el desarrollo de la sociedad de la información, ya que permite tener relaciones eficientes entre los ciudadanos y sus gobiernos.

⁹ Francisco Gómez Alamillo. "Red Iberoamericana de ciudades digitales", *Jornada de ciudades digitales*, Madrid, 4 de abril de 2003, p.1, en <www.iberomunicipios.org>.

1.3 El mercado de la información en la sociedad de la información

La información es un recurso relevante en la sociedad moderna, es el vehículo que transmite el conocimiento, por lo tanto es el elemento clave para el bienestar y el progreso.

Entre todos los mercados que existen en el mundo hay uno que controla a todos los demás y es el de las telecomunicaciones. Existen sólo tres empresas gigantes de telecomunicaciones que son Nippon Telegraph & Telephone (NT&T), American Telegraph & Telephone (AT&T) y Concert. Este mercado controla el flujo de la información a escala mundial.

En el mercado europeo de la información dentro del programa comunitario llamado Impact, se realizó el siguiente comentario respecto al mercado de la información

las actividades del mercado de la información no explican por sí solas los grandes cambios que están produciéndose en el sector de la información en Europa. Para poder comprender dichos cambios en toda su extensión hay que considerar la evolución del mercado bajo una perspectiva general y en el contexto de una sociedad en la que los servicios electrónicos de

información y su empleo desempeñan una función cada vez más importante en la vida económica, social y política.¹⁰

Por otro lado, la tasa anual de crecimiento de producción de información ha sido estimada alrededor del 10%, mientras que la tasa anual de consumo de información por la sociedad se estima entre el 3 y 4%. Obviamente se encuentra descompensado el balance entre producción y consumo de información, por lo que se puede identificar la siguiente característica fundamental de la sociedad de la información: se produce más información de la que se consume.

Debido a este crecimiento acelerado de la información en las últimas décadas, las organizaciones necesitan de una gestión adecuada para mantener sus procesos productivos y comerciales.

Actualmente la información en las empresas tiene las siguientes características:

1. *Relevancia.* Los usuarios deben tener la autoridad para determinar cuál es la información relevante y en la forma que ha de ser comunicada para su total aprovechamiento.
2. *Calidad.* La información debe ser precisa para tener calidad y su costo aumenta a medida que la calidad sube.
3. *Tiempo.* La rapidez es un factor importante a la hora de utilizar la información, las telecomunicaciones han contribuido de manera excepcional a la velocidad de transmisión de la información.

¹⁰ Impact *Líneas de actuación en el mercado de la información. Programa de trabajo de 1993.* Luxemburgo: Impact Central Office, DG. XIII, 1994, p.13-41.

4. *Accesibilidad*. Es una de las características más importantes de la información ya que es determinante para el usuario la accesibilidad a la fuente de información.

Como puede observarse, la información es un recurso estratégico de la empresa, ya que revela alternativas, reduce incertidumbres y aporta soluciones entre otras cosas.

La información debe considerarse como uno de los principales recursos de la empresa. Es un recurso que debe manejarse con el mismo cuidado que los tradicionales recursos de la empresa.

La estrategia de administrar recursos de información debe estar ligada a la estrategia de negocios de la empresa.

Para que la información sea gestionada como producto debe seguir los siguientes pasos:

- Conocer las necesidades de información de los empleados
- Administrarla como el resultado de un proceso de producción que tiene un ciclo de vida
- Designar a un responsable que administre los procesos de información

El enfoque de información como producto tiene la finalidad de proporcionar al usuario información de calidad.

Los usuarios califican una información de calidad cuando reúne las siguientes características: calidad intrínseca (precisión, objetividad, credibilidad, reputación), calidad de accesibilidad a la información (accesibilidad, facilidad de tratamiento, seguridad), calidad contextual de la información (relevancia, valor añadido, puntualidad, carácter completo, cantidad de información), y calidad representativa de la información

(interpretabilidad, facilidad de comprensión, representación concisa y coherente).¹¹

Se espera que con el uso de las tecnologías de información se proporcione esa información de calidad, aunque cabe aclarar que tecnología no es sinónimo de calidad. Asimismo, la información como producto mejora la comunicación interna, ya que las actividades son más eficientes y repercute en la mejora de la rentabilidad, competitividad y la posición en el mercado.

Por otro lado, el principio de acceso a la información como bien común está siendo suplantado por otro principio: la información como recurso institucional. Es importante señalar la diferencia entre la información común (como el aire o el agua) y la que funge como recurso institucional (como el dinero). Esta última es la que justifica el cobro por su acceso. En conclusión, la información deja de ser un bien común y se vuelve un ingrediente necesario para la competitividad.

Dado que los servicios para administrar la información forman parte integral de los procesos de producción y comercialización de la organización, el gestor de la información debe ser competente en técnicas de comunicación y relaciones públicas, debe convencer a altos ejecutivos empresariales de los beneficios reales de los servicios de información y debe justificar la razón de su continua existencia.

Como se señaló anteriormente, la información es un recurso relevante en la sociedad actual ya que es el vehículo que transmite el conocimiento.

¹¹ Richard Wang, Leo Pipino Lee y Diane Strong. "Gestionar la información como producto", *Harvard Deusto Business Review*, no. 90 (noviembre-diciembre de 1999): p. 53.

En la sociedad de la información se pueden encontrar recursos digitales en Internet que puede ser desde cartas, artículos, reportes, conferencias, ponencias, libros e informes; también se encuentran pinturas, archivos de audio, videos y multimedios. De forma muy general se puede encontrar en Internet los siguientes tipos de recursos:

- 1) Páginas amarillas: denominadas así porque ofrecen información de empresas. Los datos que se pueden obtener pueden ser dirección, teléfono o correo electrónico sobre las mismas. Son muy parecidas a las guías telefónicas impresas.
- 2) Páginas blancas: tienen muchas similitudes con las páginas amarillas, con la única diferencia de que se encuentra información de personas y no de empresas.
- 3) Grupos de noticias: son por lo general una fuente de información útil. Es posible recoger información de los mensajes enviados a estos grupos.
- 4) Listas de distribución o redes académicas: éstas tratan por lo general temas especializados.
- 5) Imágenes, gráficos y mapas: encontramos en Internet, fotografías, mapas, animaciones o imágenes de cualquier tipo.
- 6) Archivos de audio: el sonido también lo encontramos en Internet en distintos formatos, como los que acompañan a páginas, los que transmiten información o los que son para descargarse.
- 7) Los chats o las conversaciones mediante texto o voz también son considerados fuentes de información.
- 8) Documentos hipertexto: el mayor porcentaje de la información que se encuentra en Internet está en formato HTML o página web. La calidad de

estos documentos han hecho que la información en este formato sea la principal.

En la llamada sociedad de la información se han colocado en la Red una gran cantidad y variedad de recursos de información que pueden ser utilizados para resolver todo tipo de necesidades que van desde vacaciones, salud, educación, economía etc., en México por ejemplo se diseñó la página de *e-México*, en la cual se pueden encontrar un amplio tipo de servicios de información de esta naturaleza. Estos servicios de información disponibles en Internet pueden ser públicos o privados. En el caso de *e-México* podría considerarse como un servicio de información público.

Por lo tanto, la creación, desarrollo y mantenimiento de estos servicios de información son considerados como una disciplina debido a que instituciones del sector público y privado desean tener presencia en Internet a través de interfaces web con el objetivo de satisfacer las necesidades de sus usuarios principalmente.

Ahora bien, el valor real de los portales se basa en ofrecer a los usuarios la posibilidad de acceder a la información y la capacidad para interactuar. Por ello, la gestión de los recursos de información tiene un valor estratégico para todas las organizaciones, ya que la información deja de ser un bien común para volverse una parte necesaria para la competitividad.

Como puede observarse, en la sociedad de la información todo se encuentra relacionado con la generación, almacenamiento y procesamiento de la información y, a la vez, todos los sectores relacionados con las TIC desempeñan un papel importante, pues son considerados los motores del desarrollo y progreso en la economía.

Pero también en la sociedad de la información existe la llamada brecha digital, considerada como uno de los obstáculos en el desarrollo de la misma, ya que los que forman parte de ésta se encuentran muy alejados de los beneficios y las ventajas de las TIC. En la actualidad se encuentran trabajando en conjunto los gobiernos nacionales, sector privado y la Organización de las Naciones Unidas para eliminar la pobreza del mundo y la brecha digital.

Como se ha visto anteriormente, la sociedad de la información no está limitada al uso de la Internet, aunque ésta sea un medio a través del cual se tiene acceso e intercambio de información. Las personas generalmente utilizan Internet porque la ven como una gran biblioteca, en la que casi siempre encuentran la información que están buscando. Internet ha venido a revolucionar el sector de la industria de la información, el de los productores y distribuidores de bases de datos.

Por ello, el especialista del manejo de la información tiene que convencer a los usuarios de que Internet no es la mejor opción para la búsqueda, localización y recuperación de la información sobretodo de tipo académico, pues la información que se localiza no se sabe si fue elaborada por y para fines académicos, además de que muchos recursos no son estables debido a que desaparecen constantemente.

Debido a lo anterior, en la sociedad actual existe una necesidad cada vez más creciente por organizar los diversos materiales digitales que se encuentran en Internet, a través del uso de las tecnologías que permitan procesar, almacenar, recuperar y transmitir información en cualquiera de sus formas (voz, textos, imágenes) sin importar la distancia, el tiempo o la cantidad.

Por ello, es necesario que las personas que trabajan en todas las áreas del conocimiento deban estar informadas y actualizadas para poder generar conocimiento, pues la sociedad de la información debe centrarse en que todos puedan crear, consultar, utilizar y compartir tanto la información como el conocimiento.

CAPÍTULO 2

Organización documental en la sociedad de la información. Estado actual

En este capítulo se hablará de las opciones que han surgido con el uso de las TIC para organizar los documentos digitales, por lo que se requiere definir los términos de documento y objeto digital, ya que ambos son el soporte de la sociedad de la información. Además, se describen los lenguajes de marcado y metadatos, así como el rol que juegan éstos en la biblioteca digital y repositorios, proyectos que han surgido principalmente para albergar, organizar y difundir la información digital principalmente de tipo académico.

Pero antes de comenzar a definir qué son documento y objeto digital, se expondrán algunas definiciones del término “digital”, con la finalidad de explicar por qué se optó por denominar a ambos términos como digitales y no electrónicos.

En la Sociedad de la Información han surgido nuevas representaciones de la información al momento de crearse, difundirse y usarse, es decir, la vida se ha digitalizado, y por ello también se define a la sociedad actual como “digital”, de ahí que dicho término sea utilizado durante el desarrollo de este trabajo.

A continuación se mencionan algunas definiciones del término *digital*.

Digital se refiere a la descripción de datos, la cual es almacenada o transmitida en una secuencia de símbolos discretos provenientes de un conjunto finito; normalmente para este conjunto se usa información "binaria" representada mediante señales electrónicas, electromagnéticas, etc.¹

Digital se denomina asimismo el objeto representado por dígitos, casi siempre se utiliza el sistema numérico binario.

Una vez expuestas las definiciones anteriores, se puede comenzar a hablar del papel que tienen ambos términos (documento y objeto) en la organización documental, así como de la necesidad e importancia que representa organizar por colecciones a la información digital contenida en la red.

2.1 Situación actual de la organización documental

Al hablar de la organización documental en la sociedad de la información, es necesario hacer referencia al documento digital. Si bien, el soporte tradicional dominante es el papel, así como la escritura manuscrita o impresa, el documento digital aparece como un medio distinto que otorga características novedosas en su creación y uso, por lo que resulta interesante acercarse a este elemento imprescindible.

2.1.1 Documento digital

De acuerdo con la definición del término *digital* señalada al inicio de este capítulo, se puede concluir que *documento digital* es aquel que ha sido creado de origen o

¹ Free On-Line Dictionary of Computing (FOLDOC), disponible en: <<http://foldoc.doc.ic.ac.uk/foldoc/index.html>>, consultado el 25 septiembre de 2009.

convertido a una forma de representación basada en números bajo un cierto patrón, con objeto de ser almacenado o transmitido. ²

Por su parte Rogelio Pedauque³ dice que cuando el documento se convierte en digital se puede representar con la siguiente ecuación:

$$\text{Documento digital} = \text{estructura} + \text{datos}.$$

En donde la estructura varía en función del tipo de documento, es decir, el documento tradicional al convertirse en documento digital evoluciona de una ecuación a otra, de soporte + inscripción a estructura + datos. Esta operación se hace a partir del primer documento que se digitaliza (*documento digital*= *estructura* + *datos*). Lo que significa que en los documentos digitales convivirá el texto con la imagen y el sonido, su carácter multimedia se acentuará, pero el lenguaje escrito seguirá siendo el principal vehículo para expresar ideas.⁴

Un documento digital tiene que ser fácil de leer en cualquier computadora y descifrado por distintas aplicaciones, por lo que la normalización es esencial. Un ejemplo de normalización es el lenguaje XML, ya que a través de éste se captura la estructura y el contenido del documento.

Así, se puede definir al documento digital como “la unidad flexible y dinámica, consistente de contenido no lineal, representado como conjunto de ítems de información, enlazados, almacenados en uno o más medios físicos o en red, creados y utilizados por uno o más individuos en el desarrollo de algún proceso o proyecto”.⁵

² VOUTSSÁS Márquez, Juan. *Bibliotecas y publicaciones digitales*. p. 61.

³ PEDAUQUE, Rogelio T. “Documento: forma, signo y medio, re-formulaciones de lo digital”, p. 8.

⁴ RODRÍGUEZ Bravo, Blanca. *El documento: entre la tradición y la renovación*, p. 164.

⁵ SCHAMBER, L. “What's a document? : Rethinking the Concept in Uneasy Times”, *Journal of the American Society for Information Science*, vol. 47, no. 9 (1996): 669-671.

El documento digital es un producto de información y de comunicación que se crea y se utiliza en un entorno digital. Se trata de un documento dinámico en cuanto que puede variar espacial, temporal e informativamente y es interactivo en cuanto el usuario/lector posee capacidad de decisión en su lectura/exploración/navegación.⁶

Rodríguez señala que documento digital “sería aquel documento electrónico que está codificado sobre la base de la codificación binaria, y que precisa de un ordenador para leerlo.”⁷ El ejemplo más claro de documento digital son las páginas que encontramos en la web, así como las aplicaciones hipertextuales.

A partir del ejemplo anterior, se puede deducir que el documento digital es un objeto codificado que, a través de diferentes tecnologías, se puede leer, ver, oír, etc. Los documentos digitales son unidades que se almacenan, procesan y transfieren por vía digital.

También se suele referir al documento digital con el nombre de objeto digital, aunque se hace la distinción entre ambos, por lo que a continuación se explica el término de objeto digital.

2.1.2 Objeto digital

El objeto digital es cualquier unidad de valor o descripción de contenido que puede ser texto, sonido, video, mundo VRML (lenguaje de modelación de realidad virtual), imagen, bases de datos, software, archivos tridimensionales. Consiste en una secuencia de bits que bajo una correcta codificación representan todo tipo de información.⁸

⁶ “Diseño y arquitectura”, en <<http://tramullas.com/ai/productos/htm>>, [Consultado el 19 de junio de 2007].

⁷ RODRÍGUEZ, BLANCA; Santos de Paz, Lourdes. “Del documento digital a la biblioteca virtual”, en *Scire*, vol. 8, no. 2 (julio-diciembre de 2002), p. 3, disponible en <<http://ibersid.eu/ojs/index.php/scire/article/viewFile/1170/1152>>, [Consultado el 1o de julio de 2009].

⁸ MIRELES Cárdenas, Celia. “Del documento al objeto digital”, *Revista de Investigación Bibliotecológica* vol. 17, no. 34 (enero-junio de 2003): p. 96.

Desde el punto de vista informático, un objeto digital puede considerarse como una secuencia de bits, de ceros y unos, ya que con ellos y una adecuada codificación se puede representar cualquier tipo de información, incluyendo los contenidos de los materiales tradicionales.⁹

Además, puede decirse que el objeto digital es cualquier entidad documental: texto, imagen, sonido, etc., que ha sido codificada numéricamente en algún formato y ensamblada junto con algún conjunto de metadatos de tal forma que puede ser almacenada, buscada, encontrada y usada a partir de una colección dentro de un sistema de cómputo.¹⁰

Una vez definido el término de objeto digital, es necesario señalar los principios que deben considerarse en la construcción de los objetos digitales con la finalidad de regir el desarrollo de una colección digital.

1. Un buen objeto digital debe ser producido de tal manera que apoye las prioridades de la colección al tiempo que mantiene las cualidades que contribuyen a su interoperabilidad y reusabilidad.
2. Un objeto digital es permanente. Debe existir la intención por parte de una persona o institución para identificar que ese objeto permanecerá accesible a lo largo del tiempo.
3. Un objeto digital está representado en un formato digital que soporta el uso actual para el que fue diseñado, así como usos futuros, es decir, debe poder copiarse a otros formatos que sí lo soporten. Por lo tanto, un objeto digital es intercambiable entre plataformas.
4. Un objeto digital será nombrado con un identificador único y permanente de acuerdo con un esquema bien documentado (como es el caso de su Localizador Uniforme de Recurso [URL]) ya que estos tienden a cambiar con el tiempo.

⁹ GARCÍA Camarero, Ernesto. La biblioteca digital p. 144.

¹⁰ VOUTSSÁS Márquez, Juan. *Bibliotecas y publicaciones digitales*, p. 47.

5. Un objeto digital debe poder ser auténtico en al menos tres aspectos:

- El usuario debe poder ser capaz de determinar el origen, estructura, versión y porción del objeto.
- El usuario debe poder determinar que el objeto es lo que pretende ser y,
- El usuario debe poder determinar que el objeto no está corrompido o alterado de forma intencional.¹¹

En un objeto digital, además de la información presentada al usuario, se incluyen una serie de datos adicionales que nos expresan algunas propiedades técnicas o descriptivas de los mismos objetos.

Por sus características estos contenidos constituyen las colecciones de la biblioteca digital, que actualmente están teniendo un desarrollo acelerado. Es por ello que, para referirse a los materiales que conforman las colecciones digitales, es más adecuado hablar de objetos digitales que de documentos digitales.

Es importante señalar que algunas de las repercusiones que presentan el documento y objeto digital en la bibliotecología se refieren a la producción, conservación, catalogación, clasificación y recuperación de éstos, por ello, son objeto de estudio de la disciplina.

2.1.3 Repositorios

Cuando se comenzaron a observar los nuevos fenómenos que trajo consigo el creciente uso de las TIC, se dieron acercamientos a los escenarios futuros que se presentarían en el manejo y uso de la información contenida en formato digital. Uno de esos escenarios fue que paulatinamente podría irse organizando toda la información contenida en Internet y tendríamos entonces acceso a ese gran universo de información.

¹¹ *Ibidem*, p. 59.

Existen dos puntos de vista con respecto al uso de la información digital:

1) El primero se refiere a que las bibliotecas no dejarán de estar presentes en el medio digital y que Internet no las suplantarán.

2) El segundo se refiere a que uno de los elementos importantes que forman parte de la biblioteca digital son las publicaciones en formato digital, ya que en la actualidad el conocimiento se encuentra en textos digitales, los cuales tienen la característica de actualizarse fácilmente, además de que pueden reproducirse varias versiones de estos mismos.

De esta manera, el objetivo principal de la biblioteca digital es el almacenamiento y organización de la información en formato digital, además de hacerla accesible a través de las redes de comunicación. Ésta es una de las principales características de la biblioteca digital.

El acceso universal a los objetos digitales que se atribuye a la biblioteca digital, se podrá concebir a partir de la cooperación de ésta con temáticas e intereses afines, es decir, de la constitución de sistemas de redes de información que ofrezcan acceso a los documentos que cada biblioteca tiene en sus servidores. Es así que se habla de repositorios, en el sentido de que hay diversos sitios (o servidores) en donde se albergan documentos digitales que pueden ser consultados por más de una biblioteca.¹²

Un repositorio es un lugar en los que las organizaciones almacenan y organizan los resultados de sus investigaciones. Por ello, a partir de éstos se pueden construir sistemas que contengan publicaciones de revistas digitales o de tesis doctorales. Algunas de las características de los repositorios son: 1) el contenido es académico y/o científico, 2) es acumulativo y 3) es interoperable y de acceso abierto.

¹² TORRES Vargas, Georgina Araceli. "Un modelo integral de biblioteca digital". p.7.

Una de las políticas de los repositorios es que los autores pueden archivar sus investigaciones ahí, lo que significa que los resultados de éstas no tienen que esperar hasta que salga la versión impresa de los *papers* en los que fueron publicadas

Como puede observarse, para hacer frente al escenario digital, los repositorios son una buena alternativa y como ejemplo podemos mencionar la red Scielo, en la cual se encuentran colecciones nacionales de revistas científicas de calidad. Dicha red ofrece los avances de la investigación científica de los países en desarrollo de América Latina y el Caribe, aunque también se encuentran países como España y Portugal.

El repositorio Scielo es un modelo para la publicación electrónica cooperativa de revistas científicas en Internet, en donde se asegura la visibilidad y el acceso a la literatura científica. Además, contiene procedimientos que permiten medir el uso y el impacto de las revistas. Este proyecto surgió en Brasil como resultado de la cooperación entre las instituciones de la Fundación de Apoyo a la Investigación del Estado de São Paulo (Fapesp) y del Centro Latinoamericano y del Caribe de Información en Ciencias de la Salud (Bireme), con la finalidad de difundir las revistas científicas médicas. El proyecto inició en 1997 y desde julio de 1998 opera regularmente incorporando nuevos títulos, mientras se expande a otros países. En la actualidad se están incorporando revistas de todas las áreas del conocimiento, ya que se pretende que sea una colección multidisciplinaria.

La red Scielo se encuentra conformada por tres componentes:

- 1) El primero se refiere a la publicación electrónica completa de las revistas científicas y la recuperación de los textos por sus contenidos.
- 2) El segundo componente concierne a que el proyecto Scielo favorece el desarrollo de sitios nacionales con la operación de sus aplicaciones, como puede observarse en los portales de Chile, Cuba, México, etc.

3) Y el tercer componente tiene que ver con las alianzas creadas entre los autores, editores, instituciones científicas, universidades, bibliotecas, etc., con la finalidad de perfeccionar y mantener la red de Scielo.

Se espera que en los próximos años Scielo contribuya a que la información científica generada en cada uno de los países, esté disponible rápidamente para su consulta.

Los países que forman parte de la red son Argentina, Brasil, Chile, Colombia, Cuba, España, Portugal, Venezuela, México, Costa Rica, Paraguay, Perú y Uruguay. Los últimos cinco países aún no se encuentran en el rubro de colecciones desarrolladas, pues no cuentan con la certificación por parte de Brasil; México se encuentra entre éstos ya que no posee aún el comité consultivo encargado de definir los criterios de selección y permanencia de las revistas en el repositorio.

La dirección para acceder al repositorio regional es a través de la siguiente dirección: <<http://www.scielo.org/php/index.php?lang=es>>, en la cual se pueden realizar búsquedas en las colecciones de todos los países o en alguno en particular, como puede observarse en la siguiente imagen:

The image shows a screenshot of the SciELO website interface. At the top right, there are language options: "portugués | english" and a "Contacto" link. The main header features the SciELO logo and the text "Scientific Electronic Library Online". The page is divided into several sections:

- Red SciELO:** Includes links for "Sobre SciELO", "Indicadores Bibliométricos", "Acceda via OAJ y RSS", and "Portal SciELO.org antiguo".
- Busca artículos:** A search section with a "método" dropdown set to "por palabra", a search input field, and a "donde" dropdown set to "Regional". It also includes a "Buscar" button and a list of "índices (Regional)" such as país, autor, palabra clave, título, organización, año de publicación, and tipo del documento.
- Lista revistas:** A section for "Búsqueda por Revistas" with a search input and "Buscar" button. It also offers "Por lista alfabética - todos" (A-Z) and "Por tema - todos" (Arquitectura, Ciencias Agrícolas, etc.).
- SciELO en números:** A statistics section showing "613 Revistas", "13.239 Números", "197.250 Artículos", and "3.891.531 Citas".
- Nuevos:** A section for "Última actualización - 20/mar/2009" showing "96 Números" and options to view by "tema" or "colección".
- Servicios personalizados:** Includes a "Login" form with "Login" and "Login" buttons, and a "Regístrese para servicios personalizados" link.
- Noticias:** A news section with several articles, including "Estudio destaca SciELO y J-Stage como plataformas de revistas en Acceso Abierto" and "Cooperación técnica de la OPS en Brasil fortalecerá la gestión de información y conocimiento".
- Enlaces:** A list of external links including "Biomed Central", "DOAJ", "HighWirePress", "OpenJGate", "Periódicos Capes", and "PubMed Central".

México ingresa a la red Scielo en marzo de 2005 e inicia su prueba piloto en la UNAM con siete revistas basándose en el índice de Revistas Mexicanas de Investigación Científica y Tecnológica de Conacyt, como criterio de selección. En noviembre del 2005 se establecen compromisos de trabajo entre el Instituto Nacional de Salud Pública y la UNAM en donde se seleccionan 8 revistas del área médica basándose en Artemisa.

El portal de Scielo México (<<http://www.scielo.org.mx/scielo.php>>) se encuentra conformado por una colección de 37 títulos de revistas entre las que se encuentran principalmente del área médica, aunque también comprende otras áreas del conocimiento. Contiene 373 fascículos y más de 3,619 artículos, como puede observarse en la siguiente ilustración:



Colección de la biblioteca

Lista alfabética - 31 seriadas listadas

- [Agricultura técnica en México](#) - 5 números
- [Agrociencia](#) - 5 números
- [Andamios](#) - 3 números
- [Archivos de cardiología de México](#) - 20 números
- [Archivos de neurociencias \(México, D.F.\)](#) - 8 números
- [Argumentos \(México, D.F.\)](#) - 6 números
- [Atmósfera](#) - 18 números

- [Boletín médico del Hospital Infantil de México](#) - 16 números
- [Boletín mexicano de derecho comparado](#) - 4 números

- [Computación y Sistemas](#) - 8 números
- [Contaduría y administración](#) - 4 números
- [Cuestiones constitucionales](#) - 2 números

- [Estudios sociales \(Hermosillo, Son.\)](#) - 5 números

- [Gaceta médica de México](#) - 16 números
- [Geofísica internacional](#) - 7 números

- [Ingeniería, investigación y tecnología](#) - 12 números

- [Perfiles educativos](#) - 13 números
- [Perinatología y reproducción humana](#) - 9 números
- [Política y cultura](#) - 8 números

- [Región y sociedad](#) - 3 números
- [Revista de investigación clínica](#) - 15 números
- [Revista del Instituto Nacional de Enfermedades Respiratorias](#) - 13 números
- [Revista internacional de contaminación ambiental](#) - 5 números
- [Revista latinoamericana de investigación en matemática educativa](#) - 8 números
- [Revista mexicana de astronomía y astrofísica](#) - 3 números
- [Revista mexicana de biodiversidad](#) - 7 números
- [Revista mexicana de ciencias geológicas](#) - 3 números
- [Revista mexicana de física](#) - 21 números
- [Revista mexicana de física E](#) - 2 números

- [Salud Pública de México](#) - 45 números

Cabe señalar que en la actualidad existe la preocupación de difundir los resultados de las investigaciones científicas de manera rápida y eficaz a través de los repositorios o depósitos institucionales, con la finalidad de apoyar a las futuras investigaciones de estudiantes, profesores e investigadores, por ello los repositorios han pasado a formar parte primordial como proyectos para difundir las aportaciones de tipo científico, como es el caso de Scielo.

Los investigadores y los estudiantes de posgrado ven con muy buenos ojos a los repositorios digitales, ya que los consideran fuentes serias que difunden los resultados de las investigaciones científicas actuales que ayudan a fortalecer la cooperación y el aprovechamiento de los recursos.

Se ha visto también que el papel de las bibliotecas en este contexto ha evolucionado en el sentido de que los recursos que se encuentran en éstas no son sólo para que los usuarios trabajen con la información, sino para que también puedan crearla con el apoyo de todas las herramientas que ofrecen las nuevas tecnologías de la información, es decir, pueden consultar todos los recursos que tiene la biblioteca, como son catálogos en línea y bases de datos, pero además puedan tener acceso a las bibliotecas digitales y a los repositorios. Aunado a esto, podemos observar que el software libre es considerado como una de las herramientas principales para el diseño de repositorios o depósitos digitales, ya que permite que los usuarios que hacen uso de éstos puedan hacer adaptaciones en el software según sus necesidades.

La mayoría de los repositorios digitales se encuentran diseñados con software libre como es el caso del repositorio de la biblioteca de MIT que utilizó Dspace o el proyecto llamado *Securing a Hybrid Environment for Research Preservation and Access* (SHERPA), el cual está creando repositorios institucionales de acceso abierto en las universidades del Reino Unido con el software libre Eprints para facilitar su rápida difusión para la investigación.

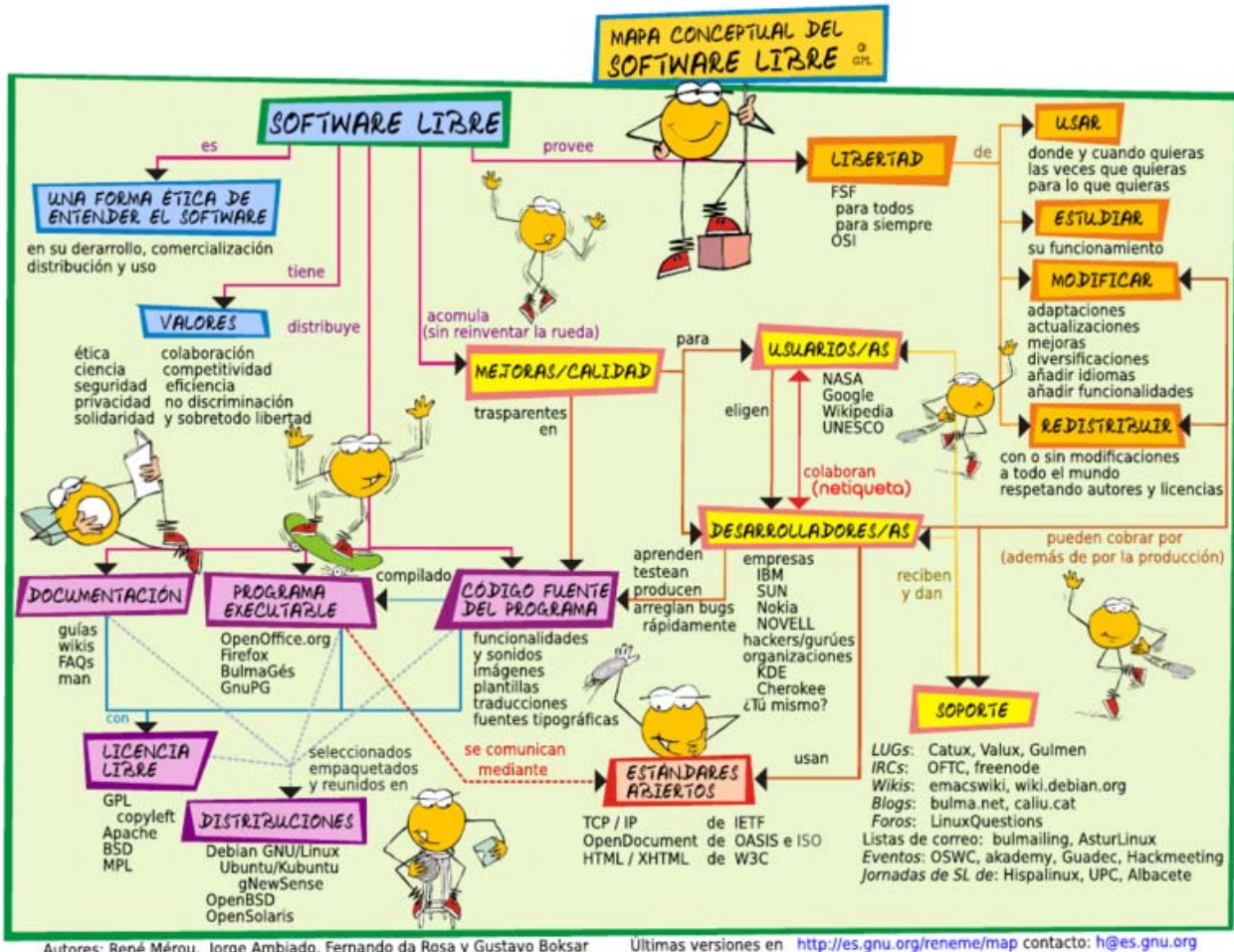
2.1.4 Creación de colecciones digitales mediante software libre

Existe una amplia proliferación de herramientas para crear colecciones digitales y esto hace que su selección requiera de un proceso de análisis antes de escoger la que se utilizará en cada institución, según sus necesidades y las bondades que ofrezca dicha herramienta. La selección debe centrarse principalmente en los siguientes aspectos:

- *Interfaz*: la forma de presentación al usuario final, así como la presentación a la persona que se ocupa del procesamiento.
- *Flexibilidad*: adaptación de la herramienta, según las características institucionales.
- *Lenguaje*: idiomas del ambiente de procesamiento y de la interfaz de recuperación.
- *Contenidos*: formato de los documentos que acepta en sus colecciones.
- *Procesamiento*: facilidades para procesar los documentos para una recuperación efectiva.
- *Recuperación*: formas que tiene el usuario de acceder a los documentos.
- *Requerimientos de sistema*: características de las computadoras que soportarán la herramienta y de las que harán uso de las colecciones.
- *Servidor Web*: requerimientos de los servidores en los que se soportará la herramienta.
- *Licencia*: si es libre o privada.
- *Costo*: gratis o pago.

El software libre nace de la Fundación para el Software Libre (Free Software Foundation), creada por Richard Matthew Stallman en el año de 1985 con la finalidad de eliminar las restricciones de la copia, redistribución y modificación de los programas de computadoras.

A partir de los años noventa muchas compañías y autores escriben sobre qué es el software libre, pero para entender mejor la filosofía por la cual surge, se presenta a continuación un mapa conceptual de éste:



Se le denomina software libre ¹³ por la libertad con que cuentan los usuarios para ejecutar, copiar, distribuir, estudiar, cambiar y mejorar el software. De modo más preciso, se refiere a cuatro aspectos:

¹³ Principios del Software Libre, disponible en <http://doc.ubuntu-es.org/Software_Libre>, [Consultado el 22 de junio de 2008]

- La libertad de usar el programa, con cualquier propósito.
- La libertad de estudiar cómo funciona el programa y adaptarlo a tus necesidades.
- La libertad de distribuir copias, con lo que puedes ayudar a tu vecino.
- La idea de mejorar el programa y hacer públicas las mejoras a los demás, de modo que toda la comunidad se beneficie. El acceso al código fuente es un requisito previo para esto. Un programa es software libre si los usuarios tienen todas estas libertades.

El software libre y el acceso libre suponen notables beneficios para los sistemas bibliotecarios gracias a la flexibilidad que presentan los sistemas y a las posibilidades de implementar programas de intercambio de información. En este sentido, el software libre ayuda en el ahorro de costos a la biblioteca, pero también aporta algunas características a la colección digital que la hace diferente de aquellas colecciones de documentos impresos. Por ello, es que se hace imprescindible mostrar qué es lo que ocurre en cuanto al uso de programas basados en software libre para la creación de colecciones digitales.

Actualmente se muestra una fuerte tendencia por el uso de software libre, por ejemplo en la creación y administración de colecciones digitales, ya que la filosofía de éste establece que la gente pueda usar el software de todas las formas que considere útiles, pues el código está disponible para poder descargarlo, modificarlo y usarlo del modo que más le convenga.

Pero, ¿cómo saber que software nos conviene para administrar nuestras colecciones digitales? Algunos señalan que a la hora de evaluar el software libre, es necesario tomar en cuenta las siguientes funciones:

- 1) Proceso de instalación, ya que no siempre es fácil y se requiere de ayuda por parte del personal de cómputo.
- 2) Incorporación de documentos al repositorio y la gestión de éstos, es decir, que puedan incorporarse toda clase de documentos sin importar si es texto, imagen, video, etc.
- 3) Gestión de usuarios.
- 4) Recuperación de la información.

Hoy en día, no existen modelos de evaluación de herramientas de software, las que se han trabajado han sido evaluadas para decidir cuál plataforma de software libre es la más conveniente para llevar a cabo el proyecto de repositorio digital. También se pueden considerar las siguientes funcionalidades básicas del software para poder evaluarlo:

- Especificaciones técnicas.
- Interoperabilidad.
- Estándares de metadatos.
- Administración del sistema.
- Configuración del sistema.
- Soporte técnico (dentro de este rubro se puede manejar también el mantenimiento del sistema).
- Difusión (interfaz del usuario y funcionalidad de recuperación de información).¹⁴

¹⁴ Las funcionalidades del software libre para su evaluación fueron proporcionadas por el Mtro. Dante Ortiz en el curso "Construcción de Repositorios Institucionales con Fedora" llevado a cabo en la Dirección General de Bibliotecas de la Universidad Nacional Autónoma de México del 19 al 23 de mayo de 2008.

Puede observarse que la creación de depósitos o repositorios institucionales (Institutional Repositories) está muy relacionada con el uso de software libre, ya que facilita la gestión de los depósitos y la asignación de los metadatos conforme al estándar OAI-PMH, algunos ejemplos son Dspace, Fedora, Contentdm, Greenstone, etc. En la actualidad existen algunos proyectos que trabajan con dos tipos de software libre, pues lo que le hace falta a uno se puede encontrar en otro, como es el caso del Proyecto de Daedalus desarrollado por la Universidad de Glasgow, en dicho proyecto se optó por combinar el software Eprints y Dspace. El ejemplo anterior nos lleva a que debemos comparar las funcionalidades que interesan al proyecto que se está trabajando sin entrar en cuestiones solamente tecnológicas.

2.1.5 Organizar la Internet vs. organizar por colecciones digitales

Internet es una herramienta real que es usada por miles de usuarios y exige ideas para organizar constantemente su sistema documental tan complejo. El volumen de información contenida en la red y los mecanismos de acceso a ella ofrecen, sin duda, muchas posibilidades, pero falta organización y adquisición de conocimiento. La World Wide Web comprende documentos en formato hipertexto, como son documentos textuales, gráficos, multimedia, acceso a bases de datos y aplicaciones complementarias tales como el correo electrónico, ftp, Telnet, Archie, etc. La WWW asume todas las funcionalidades generales de Internet y constituye un sistema complejo.

Es cierto que en muchos de los artículos y en literatura especializada se habla de la necesidad de ver hacia la organización de Internet. Pero estos ideales no pueden ser fructíferos en razón de que se tiene que organizar para el gran universo de usuarios, que parten de contextos y necesidades variadas, por ello, es más conveniente organizar la información para comunidades de usuarios que toda la información contenida en Internet.

Por tales motivos en este estudio se prefiere hablar de colecciones pertenecientes a las bibliotecas del medio digital (llámense virtuales, digitales, 2.0, etcétera). Aunque es necesario hablar brevemente de la situación en que se encuentra la información disponible en Internet y de su organización.

En primer lugar hay que señalar que no existen normas para el lenguaje de etiquetado de los documentos que se encuentran en la red, por lo que cada documento sigue patrones individuales en su presentación, distribución, descripción de enlaces, diseño, etc.

Otro problema en Internet es la falta de calidad de la información que se introduce en la red. El uso efectivo de la información en la red exige conocer la fuente, el mantenimiento de esa información, los criterios de elaboración, etc. No olvidemos que en Internet encontramos desde documentos de calidad, de prueba o hasta las páginas personales que cualquier usuario puede poner.

Al navegar en Internet hallamos herramientas generales de búsqueda en la red, que dan la sensación de una caja negra, en la que los documentos recuperados son un conjunto heterogéneo que al azar nos atrae, siendo muchos de ellos prescindibles desde el momento mismo de su creación.¹⁵

Ante un volumen de información como el de Internet no resultan válidas las herramientas de selección y navegación. Un sistema ideal sería el que ofreciera al usuario un resumen de un conjunto de documentos. El modelo de hipertexto integra, por ejemplo, además de las herramientas de búsqueda y navegación, las de elaboración y condensación de documentos, convirtiéndose en un poderoso entorno de organización del conocimiento.¹⁶

¹⁵ PASTOR Sánchez, Juan A. "Exposición conceptual y aplicación de un modelo para la organización del conocimiento en Internet: el hipertexto documental", en Encuentro de ISKO-España 1993: Organización del Conocimiento en Sistemas de Información y Documentación 2, p. 249.

¹⁶ *Ibidem.*

Sin embargo, han surgido dos vías para localizar recursos en la red. La primera se refiere al desarrollo de motores de búsqueda que son sistemas de búsqueda por palabras clave o con árboles jerárquicos por temas e incorporan automáticamente páginas web mediante robots de búsqueda en la red, es decir, los motores de búsqueda emplean el uso de arañas para compendiar la información de los sitios web, recopilando los datos y enlaces que se encuentran en las páginas. Las arañas rastrean Internet recolectando las páginas nuevas y actualizadas de la web para mantener sus resultados al día. Aunque son útiles, tienden a recuperar demasiada información irrelevante.

Algunos ejemplos de los motores de búsqueda o robots son Altavista o Infoseek que indizan automáticamente la información a texto completo de los distintos servidores. La segunda vía consiste en la organización por índices o repertorios temáticos que ayuden a los usuarios a visualizar los recursos. Esto condujo a la utilización de esquemas de clasificación que proporcionan una estructura semántica, es decir, son concebidos como el conjunto de recursos organizados jerárquicamente en función del esquema de clases establecido por el administrador del sistema.

Sin embargo, Gorman señala que existen los siguientes problemas al organizar los documentos digitales:

- Existen muchos documentos digitales.
- Muchos de los documentos y sitios nunca han sido y nunca serán de interés para las bibliotecas y sus usuarios.
- La inmensa mayoría de documentos digitales son de uso temporal, uso local o no tienen uso.
- Se tiene poca o ninguna garantía de que cualquier documento digital sea lo que dice que es. Ante tanta disponibilidad de información surge la pregunta ¿Qué tan confiable es esa información? La información que se localiza en

Internet no tiene control, no existe una centralización, cada uno es responsable de lo que aporta. Cuando una empresa dice que publica algo en Internet, es la responsable de lo que dice, aunque también existe la información anónima. Siempre que haya alguien que se responsabilice de los contenidos habrá alguien que responda por ellos. La mayoría de los contenidos en Internet tienen dueño.

- Algunos desaparecen después de ser registrados.
- Tenemos poca o ninguna seguridad de que cualquier documento electrónico o sitio sea el mismo en su próxima ubicación o si existirá.¹⁷

Para Gorman no tiene sentido la catalogación de los documentos digitales que se encuentran en Internet, considera que la catalogación de la web es inalcanzable e indeseable, ya que la mayoría de los documentos son:

- Efímeros (como web personales, pronósticos meteorológicos, anuncios publicitarios, comunicados de prensa).
- Lugares comerciales (desgraciadamente las únicas empresas con éxito son las de pornografía).
- Publicaciones seriadas electrónicas independientes, es decir, que no proceden de las impresas.
- Archivos digitalizados (textuales, sonoros y visuales).

¹⁷ GORMAN, Michael. "Metadata: Hype and Glory", en *Cataloging the Web: Metadata, AACR and MARC21*. p. 180.

Por otro lado, Méndez apoya la utilización de metadatos para organizar la información en Internet siempre y cuando sea parte de entornos finitos abarcables y susceptibles de planificación para determinar su aplicación en relación con la calidad de recursos.¹⁸ Señala que es necesario plantearse las siguientes preguntas antes de utilizar los metadatos:

- ¿Realmente es necesario que todos los documentos web lleven metadatos?
- ¿Sólo deben tener metadatos los recursos de calidad?
- ¿Qué se entiende por recursos de calidad?
- ¿Quién selecciona la calidad de la red?

Al describir los recursos digitales nos tenemos que hacer las siguientes preguntas:

1. ¿Cómo voy a definir el objeto de la descripción?
2. ¿Cómo lo describo, es decir, uso Dublin Core o adapto MARC?
3. ¿Cuál va a ser el uso o cuáles serán los usos de mis registros?

Méndez apoya el uso de los quince elementos de Dublin Core, pero sugiere que se tienen que hacer las siguientes preguntas:

- ¿Cuál es el objeto de la descripción? En la catalogación bibliográfica trabajamos materiales físicos, pero en los materiales digitales catalogar no es tangible, ya que para poder acceder al contenido de una obra se necesita un *browser* o *buscador* para recuperar la página web y también una conexión vía Internet al servidor donde la página se encuentra almacenada.

¹⁸ MÉNDEZ Rodríguez, Eva. *Metadatos y recuperación de información*, p. 104.

- ¿Cuál es la función del registro? Los bibliotecarios siempre nos hemos preocupado por la creación de catálogos que permitan a los usuarios navegar, localizar, identificar, seleccionar y obtener información. El catálogo web debe llevar los recursos directamente al espacio virtual de trabajo del usuario y las herramientas que usa para manipular el recurso (texto, imágenes, sonidos, datos, etc.).¹⁹

Ante esta situación descrita, puede verse que en el escenario tecnológico actual es preferible hablar de la organización de colecciones digitales, más que de Internet y, al hablar de colecciones, se hace referencia a las bibliotecas. Es en este sentido, que se puede ver a la biblioteca digital como la imagen más o menos aproximada de lo que podría ser una biblioteca dedicada al manejo de la información digital. Para entender mejor el concepto y su funcionalidad en el contexto de la organización de colecciones digitales, se presentan algunas definiciones en la siguiente tabla:

AUTOR	DEFINICIÓN
David Garza Salazar	La biblioteca digital es una colección organizada de documentos almacenados en formato digital que a su vez ofrece los servicios de búsqueda y recuperación de información. ²⁰
Ben Gross	El concepto de biblioteca digital no es sólo equivalente a una colección digitalizada con herramientas que gestionan información sino el conjunto de colecciones, servicios y personas que participan en el ciclo vital

¹⁹ BEACOM, Matthew. "Crossing a Digital Divide: AACR2 and Unaddressed Problems of Networked Resources", presentado en: la conferencia *Bibliographic Control for the New Millenium*, llevada a cabo en la Biblioteca del Congreso, Washington, D.C., noviembre de 2000.

²⁰ GARZA Salazar, David; Sordia Salinas, Martha; Martínez Treviño, Yolanda. "Biblioteca digital Phronesis", disponible en <<http://www.mty.itesm.mx/dgi/transferecia/Transparencia45/eli-03.htm>>, [Consultado el 5 de enero de 2008].

	completo para la creación, difusión, uso y conservación de datos, información y conocimiento. ²¹
Ernesto García Camarero	<ul style="list-style-type: none"> • Es la colección organizada de documentos digitales para cuya consulta se precisa de una computadora, unos programas informáticos y, en algunos casos, de un sistema de telecomunicaciones compuesto por un módem, una línea telefónica, una empresa que facilite el acceso a las redes telemáticas y unos programas de comunicaciones. • -Organismo o parte de él cuya función principal consiste en construir bibliotecas digitales, mantenerlas, actualizarlas y facilitar el uso de los documentos digitales que precisen los usuarios para satisfacer sus necesidades de información, investigación, educación o de esparcimiento, contando para ello con personal especializado.²²
American Digital Library Federation	Son organizaciones que proveen recursos, incluyendo equipo especializado, para seleccionar, estructurar, controlar, interpretar, distribuir, preservar la integridad y asegurar la persistencia de colecciones de materiales digitales que aporten conocimientos, a fin de tenerlos fácil y económicamente disponibles para una comunidad o varias comunidades definidas. ²³
Digital Library Project	El concepto de biblioteca digital no es únicamente el equivalente de repertorios digitalizados con métodos de

²¹ GROSS, Ben. "Digital library Information and Resources", disponible en <<http://interspace.grainger.uiuc.edu/~bgross/digital-libraries>> consultado el 5 de enero de 2008.

²² GARCÍA Camarero, Ernesto; García Menero, Luis Ángel. La biblioteca digital, p. 27.

²³ American Digital Library Federation, disponible en <<http://www.diglib.org/about/dldefinition.htm>>, [consultado el 30 de enero de 2000]

	gestión de la información. Es más bien, un entorno donde se reúnen colecciones, servicios y personal que favorece el ciclo completo de la creación, difusión, uso y preservación de los datos, para la información y el conocimiento. ²⁴
Juan Voutssás Márquez	Es un ente complejo, el cual conlleva un concepto social que va mucho más allá de la simple computarización de los acervos y servicios de una biblioteca tradicional. ²⁵

Bruce Schatz y Hsinchun Chen señalan que las bibliotecas digitales almacenan principalmente materiales en formato digital y gestionan eficazmente grandes cantidades de dichos materiales.²⁶ Así, investigar en bibliotecas digitales significa realmente hacerlo en sistemas de redes de información. Los aspectos clave de la tecnología se centran en cómo buscar y mostrar las selecciones solicitadas desde y a través de las grandes colecciones. En la práctica, las bibliotecas digitales deben enfocarse hacia temas relacionados con los costos de acceso y la tecnología de la digitalización, la investigación se centra en cómo desarrollar la infraestructura necesaria para la gestión masiva y eficaz de información en la red.

²⁴ Digital Library Project. University of California Berkeley. "Digital Library Using Next Generation Internet", *IEEE Communications Magazine*, vol. 37, no. 1 (enero de 1999), disponible en <<http://elib.cs.Berkeley.edu>>.

²⁵ VOUTSSÁS Márquez, Juan. "Un modelo de bibliotecas digitales para México", p. 30.

²⁶, SCHATZ, Bruce and Chen, Hsinchun. "Building Large-Scale Digital Libraries", *Computer. Monográfico: US Digital Library Initiative* (mayo de 1996), disponible en www.computer.org/computer/dli [Consultado el 15 de febrero de 2008]

La biblioteca digital contará no sólo con su acervo, sino con colecciones de otros lugares, además deberá tener acceso a la información como si se hiciera físicamente para asegurar su disponibilidad permanentemente. También puede contener diversas fuentes de información: revistas y libros electrónicos, bases de datos referenciales y de texto completo, así como programas de multimedia para apoyar la educación.

Existen otras ideas de lo que podría ser la biblioteca en este medio digital, por lo que independientemente de los conceptos, estaríamos hablando de un repositorio en donde se encuentra información de naturaleza digital, disponible para un grupo determinado de usuarios y el objetivo principal de las colecciones digitales sería compartir los conocimientos y resultados obtenidos en el proceso de investigación, así como evitar la duplicidad de esfuerzos en este sentido.

2.1.6 Sistemas para la organización de colecciones digitales

Si bien en la actualidad todavía está la pregunta de cómo se han de organizar las colecciones digitales, ya se puede hablar de algunas opciones, pues varios autores afirman que los sistemas de organización del conocimiento (SOC) pueden ser una buena alternativa.

En la actualidad los SOC son usados para organizar materiales con el propósito de recuperar y manejar una colección. Un SOC sirve de puente entre la necesidad de información del usuario y el material en la colección, con éste el usuario puede identificar un objeto de interés sin conocimiento previo de su existencia., por ello, los SOC son mecanismos que están en el corazón de cada biblioteca, museo y archivo, ya que tratan de abarcar todos los tipos de esquemas que permitan organizar la información para promover el conocimiento.

A pesar de su diversidad, los SOC tienen la siguiente característica al organizar bibliotecas digitales: imponen una vista particular del mundo en una colección y los artículos dentro de ella.

Los SOC, también conocidos como Knowledge Organization Systems (KOS, por las siglas en inglés) son sistemas que se enfocan en todos los tipos de proyectos para organizar información digital y para promover el manejo del conocimiento; además incluyen esquemas de clasificación para organizar los materiales a un nivel general, con encabezamientos de materia que otorgan un acceso más detallado y archivos de autoridad que controlan varias versiones de información clave (nombres geográficos y personales). También incluyen proyectos menos tradicionales, como ontologías y redes semánticas.

Las formas más simples de un SOC vienen siendo las tablas de contenido y los índices de los libros de texto. El conocimiento se halla en el texto; los SOC son una herramienta complementaria que ayuda al lector a transitar a lo largo del texto y han comenzado a ejercer funciones más amplias, ya que han requerido denominaciones más notables, como lenguajes de recuperación, taxonomías, categorizaciones, léxicos, tesauros, u ontologías. Son vistos hoy como esquemas que organizan, gestionan y recuperan información".²⁷

Para describir a los SOC, se tiene que tomar en cuenta las características de estructura y complejidad, las relaciones entre términos y su función. Por ello, los sistemas se agrupan dentro de tres categorías generales²⁸: *listas de términos*, las cuáles ponen énfasis en listados de definiciones; *clasificaciones y categorías* que ponen énfasis en la creación de grupos de sujetos; y *listas de relaciones*; las cuales ponen énfasis en las conexiones entre términos y conceptos.

²⁷ VICKERY, Brian. On 'knowledge organization', 2008. Disponible en: <http://www.lucis.me.uk/knowlorg.htm> [Consultado: 30 de julio de 2009].

²⁸ Sistemas de organización del conocimiento, en <<http://es.geocities.com/ontologia2004/sko.htm>>, [Consultado el 10 de julio de 2009]

Listas de términos

Entre las listas de términos se encuentran los archivos de autoridad, los glosarios, los diccionarios y los diccionarios geográficos.

Archivos de autoridad. Son listas de términos que son utilizadas para controlar los nombres variantes para una entidad o un valor de dominio para un campo en particular, como por ejemplo, nombres de países, individuos y organizaciones. Este tipo de SOC generalmente no incluye una estructura de organización muy profunda o compleja. La presentación puede ser alfabética u organizada por un esquema de clasificación superficial.

Glosarios. Es una lista de términos, normalmente con definiciones. Éstos se definen dentro de un ambiente específico y rara vez incluyen significados variantes, como por ejemplo, los términos de “Ambiente”, de la Agencia de Protección del Ambiente (The Environment Protection Agency Terms of Environment).

Diccionarios. Los diccionarios son listas alfabetizadas de palabras y sus definiciones. Un diccionario puede proveer sinónimos y antónimos, mediante las definiciones, palabras relacionadas, pero no existe una estructura jerárquica explícita.

Diccionarios geográficos. Es una lista de nombres de lugares. Los diccionarios geográficos han sido publicados como libros o han aparecido como índices o atlas. Cada entrada puede ser identificada por su tipo de rasgo, como un río, ciudad o escuela. Un ejemplo de diccionario geográfico es el Código Estadounidense de nombres geográficos (U.S. Code of Geographic Names).

Clasificaciones

Las clasificaciones y las categorías persiguen la misma idea de categorizar en clases y subclases, es decir, clasificar en entidades separadas un dominio del conocimiento. Se dividen de la siguiente manera:

Encabezamientos de materia. Proveen un grupo de términos controlados para representar los temas de artículos en una colección. Las listas de encabezamientos de materia pueden ser extensivas y cubrir un rango muy amplio de temas, pero su estructura es generalmente muy superficial, con jerarquías limitadas. Como ejemplo tenemos los Encabezamientos de materia de la Biblioteca del Congreso (The Library of Congress Subject Headings <http://www.loc.gov/cds/lcsh.html>).

Esquemas de clasificación, taxonomías y esquemas de categorización:

Estos términos se utilizan a menudo alternativamente. Son un tipo de SOC que proveen tópicos de nivel amplio.

Esquemas de clasificación. Entre éstos encontramos los Esquemas de Clasificación de la Biblioteca del Congreso (Library of Congress Classification Schedules), un sistema extensible y abierto; la Clasificación Decimal de Dewey (Dewey Decimal Classification), un sistema cerrado de diez secciones numéricas con extensiones decimales, y la Clasificación Universal de Decimales (Universal Decimal Classification, basada en Dewey pero extendida para incluir facetas o aspectos particulares de un tema).

Taxonomía. Una taxonomía es una estructura jerárquica de elementos o grupos de elementos, por lo que es normal que este término sea utilizado en relación con los tesauros, ontologías, sistemas de clasificación, etc. En una taxonomía se crean categorías jerárquicas para organizar las cosas en mapas simples del mundo.

El término taxonomía etimológicamente procede de los términos griegos “taxis”, ordenación, y “nomos”, norma. Aristóteles fue uno de los primeros en utilizar este término en el año 300 a. C., para designar esquemas jerárquicos orientados a la clasificación de objetos científicos.

En los años noventa, el concepto de taxonomía se incorpora a áreas del conocimiento, como la psicología, las ciencias sociales y la informática para designar sistemas de acceso a la información. El término se utiliza para designar uno de los métodos existentes para organizar los contenidos en Internet. En el entorno digital, ésta tiene como finalidad mejorar la navegación y desarrollar sistemas de búsqueda.

La idea básica de la taxonomía es proveer un vocabulario controlado a los atributos de metadatos y especificar las relaciones entre términos en el vocabulario controlado. La taxonomía incorpora las relaciones de equivalencia y jerarquía.

En tanto vocabulario controlado, todos los términos están conectados mediante algún modelo estructural jerárquico, arbóreo o facetado, orientado especialmente a los sistemas de navegación, organización y búsqueda de contenidos en los sitios web.²⁹

Una vez definido el término de taxonomía, hay que señalar que existe otro tipo, llamada corporativa; ésta surge cuando la empresa requiere una organización adecuada de la información basada en el uso de las taxonomías; esto es, al plantearse la cuestión de organización y recuperación de contenidos digitales en entornos corporativos surge el concepto de taxonomía.

Al emplear taxonomías corporativas se tienen que tomar en cuenta cuatro procesos que consisten en:

²⁹ CENTELLES, Miguel. Taxonomías para la categorización y la organización de la información en sitios web. “Hipertext.net”, No.3, 2005 p.5. <http://www.hipertext.net> [Consultado el día 28-agosto-2006]

1. Delimitación de la realidad (por ejemplo, área del conocimiento o sector industrial) que será representada por la taxonomía.
2. Extracción del conjunto de términos o categorías que representan dicha realidad. Para llevar a cabo este proceso se tienen que establecer las fuentes prioritarias y los mecanismos de extracción adecuados para cada una de ellas. Existen tres tipos de fuentes:
 - a) Fuentes personales integradas por usuarios del web.
 - b) Fuentes documentales, que son los documentos representativos de los tipos de contenidos.
 - c) Las taxonomías o instrumentos de representación del conocimiento.
3. Control terminológico de los términos o categorías, en donde se identifican los diferentes términos que designan un mismo concepto. Durante este proceso se establece la relación de equivalencia entre todos los términos de la taxonomía.
4. Establecimiento de los esquemas y la estructura de organización de los términos o categorías. Se incluye el criterio utilizado para dividir y agrupar las categorías, los ejemplos más utilizados como criterios son temas, materias, disciplinas, entidades, destinatarios, procesos, tipos de documentos.³⁰

En lo que se refiere a la automatización de los procesos de construcción de la taxonomía, podemos señalar que actualmente difícilmente se aplican sistemas manuales para la creación de taxonomías, en caso de que exista un nivel mínimo de automatización se encuentran dos tipos de soluciones:

- a) Taxonomías preelaboradas que deben ser adaptadas a las condiciones específicas de una determinada organización.

³⁰ CENTELLES, Miguel. Op Cit. p.6-7.

- b) Las herramientas de edición de taxonomías que ofrece a los administradores un entorno amigable para el establecimiento de relaciones entre los términos y diferentes modalidades de presentación y visualización de los resultados.³¹

Listas de relaciones

Entre las listas de relaciones se encuentran los tesauros. Éstos se basan en conceptos que muestran relaciones entre términos. Las relaciones que se expresan en un tesoro incluyen jerarquía, equivalencia (sinónimo) y asociación o relación. Estas relaciones son generalmente representadas mediante la noción término amplio (TA), término estrecho (TE), sinónimo (SI) y término asociativo o término relativo (TR).

La mayoría de los tesauros se desarrollan para una disciplina específica, muchos son muy largos y pueden incluir más de 50,000 términos.

2.1.7 Lenguajes de marcado

Un *lenguaje de marcado* o lenguaje de marcas es un conjunto de reglas que codifican un documento, reúne etiquetas y marcas que contienen información para definir la estructura del texto, su presentación y la semántica del documento. El *lenguaje de marcado* más popular es el HTML (Hypertext Markup Language) y es la base de la WWW (World Wide Web).

Los formatos utilizados en estos momentos para la creación de bibliotecas digitales pueden ser de tipo abierto o cerrado. Son formatos de tipo abierto los que pueden manipularse con facilidad como son SGML (Standard Generalized Markup Language), HTML, XML (Extensible Markup Language), etc. Son formatos de tipo cerrado los que presentan dificultad para ser manipulados, como son Acrobat pdf, entre otros.

³¹ *Ibidem.*

Los lenguajes HTML y SGML surgen como propuestas para generar la sintaxis en la que puede ser representada la información electrónica, por lo que es importante definir lo que es un formato digital desde el punto de vista de cómputo.

Otros términos usados son formato, que es el arreglo predeterminado de elementos que permiten definir una sintaxis en la que puede ser representada la información, y digital, que es la forma o representación de elementos del mundo real, en la que se utilizan valores discretos de 0 y 1, con el propósito de realizar operaciones con un mayor control y precisión.

Pero, para fines de esta investigación, se entiende por formato digital a los diversos lenguajes que sirven para la representación de los documentos digitales.³²

A continuación se detallan las características y estructura de los lenguajes de marcado: HTML, SGML, XML, así como la función e importancia que juegan los metadatos en la descripción de los documentos digitales.

HMTL

El HTML es un lenguaje cuya finalidad es organizar la estructura de un documento digital en forma de hipertexto, con objeto de que pueda ser fácilmente procesado por un programa visualizador, independientemente de la máquina o el sistema operativo que se emplee.³³

Uno de los formatos digitales para representar la información electrónica es el formato HTML, que es un lenguaje de marcas que define la representación gráfica de los documentos, es decir, el resaltado, separación de párrafos, negrita, subrayado, etc. El HTML da forma al contenido de los documentos que se quieren publicar en la WWW. Sirve para representar tanto elementos de contenido

³² JUÁREZ Santamaría, Beatriz; Martínez Ortega, Martha Patricia. "Uso de los metadatos en el orden documental", en *XXXII Jornadas Mexicanas de Biblioteconomía*, 2001 p. 4.

³³ LAFUENTE López, Ramiro; Garduño Vera, Roberto. *Lenguajes de marcado de documentos digitales*. México: UNAM, CUIB, 2001 p. 105.

(marcado estructural) como de presentación (marcado de procedimiento) relativos a un documento.

El lenguaje de marcas es el sistema formal por el cual se añade información codificada a la forma electrónica de un documento, de tal manera que represente su significado y permita así, controlar su procesamiento.³⁴

El formato HTML es superficial en lo que a contenido se refiere, ya que indica a los navegadores cómo deben presentar la información en vez de lo que ésta significa. Es sólo una tecnología de presentación. El HTML tiene dos funciones básicas: 1) formalizar a través de etiquetas la presentación de documentos digitales para facilitar su recuperación a través de Internet; 2) orientar al navegador a partir de las etiquetas para que interprete, recupere y realice la presentación del documento digital en pantalla.³⁵

El HTML ofrece a los autores los medios para:

- Publicar documentos en línea con cabeceras, textos, tablas, listas, fotografías.
- Encontrar información en línea mediante los enlaces de hipertextos, simplemente haciendo clic sobre un botón.
- Diseñar formularios para realizar transacciones con servicios remotos, para usarlos en la búsqueda de información para hacer reservas, pedidos de productos, etc.
- Incluir videoclips, clips de sonido y otras aplicaciones directamente en los documentos.³⁶

³⁴ BURNARD, Lou; Light, Richard. "Three SGML Metadata Formats : TEI, EAD and CIMI", en *Work Package 1 of Telematics for libraries project BIBLINK*, UKOLN Metadata Group, diciembre de 1996, disponible en <<http://ukoln.ac.uk/metadata/BIBLINK/wp1/sgml>>, [Consultado el 15 de febrero de 2008].

³⁵ *Ibidem*, p. 107.

³⁶ GARCÍA Camarero, Ernesto y Luis Ángel García Menero. "La biblioteca digital", p. 70.

SGML (Estándar Generalizad Markup Language)

El formato SGML es un metalenguaje que define lenguajes de anotación para procesar y almacenar un texto, además logra una independencia de la plataforma y procesador de texto que utiliza una organización, es decir, es un sistema para la organización y etiquetado de documentos que facilita el intercambio y conservación de recursos electrónicos ampliamente estructurados. Se ha convertido en el estándar de la industria, ya que productos como Word han incluido ideas y propuestas de SGML.

Tiene una notación formal para el texto a través de etiquetas enmarcadas con paréntesis en ángulo <>.

XML (Extensible Markup Language)

La primera versión de XML apareció en noviembre de 1998. Es un subconjunto simplificado del estándar SGML de ISO que contempla parámetros para hacer operar aplicaciones Intranet de manera ágil y con un alto grado de eficiencia, lograr que los datos de Intranets sean fáciles de gestionar y desarrollar aplicaciones de alta velocidad a través de mecanismos que permitan compartir procesos en diversos equipos.³⁷

El formato XML es un subconjunto simplificado del estándar SGML. En una Intranet XML, permite que los datos sean más fáciles de gestionar y permitirá el desarrollo de aplicaciones más rápidas, pasando parte del procesamiento de los servidores a las PC. Algunas de las características más importantes de XML son:

1. Una arquitectura más abierta y extensible. No se necesitan versiones para que pueda funcionar en futuros navegadores.
2. Se puede hacer el intercambio de documentos entre las aplicaciones tanto en la propia PC como en una red local o extensa.

³⁷ MÉNDEZ Rodríguez, Eva. *Metadatos y recuperación de información*, p. 113.

3. Permite agrupar una amplia variedad de aplicaciones (desde páginas web hasta bases de datos), debido a su extensibilidad y flexibilidad.
4. Los motores de búsqueda devolverán respuestas más adecuadas y precisas.

Una de las aplicaciones que se ha desarrollado con XML es la Text Encoding Initiative (TEI) que partió de diversas asociaciones profesionales en los campos de humanidades. Trata de establecer etiquetas que propicien la descripción de textos científicos y literarios, buscando así una normalización de la estructura y el contenido. XML es más flexible que HTML y menos complejo que SGML.

XML es una tecnología sencilla que tiene a su alrededor otras que la complementan y le brindan posibilidades mucho mayores. Tiene un papel muy importante en la actualidad, ya que permite la compatibilidad entre sistemas para compartir la información de una manera segura, fiable y fácil.³⁸

XML se ve como el único estándar para editar, manipular y almacenar datos, de la misma forma que el formato MARC se convirtió en un formato de intercambio en el mundo de la información bibliográfica.³⁹ Por ello, todos los documentos codificados en XML pueden ser editados, guardados y enviados sin importar el programa analizador. Los documentos XML proveen hojas de estilo que permiten a los navegadores convertir los documentos para presentarlos en HTML en la pantalla.

En conclusión se puede decir, que los lenguajes de marcado (HTML, SGML y XML) contienen un esquema para que la computadora entienda y lea los documentos digitales, además de que permiten la compatibilidad entre los sistemas con la finalidad de poder compartir e intercambiar información.

³⁸ Extensible Markup Language. En: Wikipedia. http://es.wikipedia.org/wiki/XML#Ventajas_del_XML [Consultado el 16-julio-2009]

³⁹ *Ibidem*, p. 45.

2.1.8 Metadatos

El término de metadata surge en los años sesenta utilizado por Jack E. Myers para la descripción de conjuntos de datos. Myers escogió este término para que fuese fácil de recordar y posteriormente para llamar así a su empresa: Metadata Information Partners.

Los metadatos son un conjunto de elementos que se utilizan para ayudar a la identificación, descripción y localización de recursos electrónicos por medio de una representación de la descripción bibliográfica de los mismos, además permiten indizar y clasificar grandes cantidades de información de diversos tipos, por ello son imprescindibles en la recuperación global de la información en Internet.

Un metadato es un elemento que describe el contenido, condiciones, características, etc., de un documento, con el fin de definir, identificar, organizar, indicar, filtrar, colocar, preservar, recuperar y administrar ese documento como una parte de un conjunto ordenado de recursos de información electrónicos o no.⁴⁰

Los metadatos tienen como regla primordial que los documentos introducidos en la red incluyan todos los datos necesarios para su búsqueda, localización y recuperación.

“Los metadatos responden a las necesidades de proporcionar una opción para la catalogación y de proporcionar una alternativa entre los costos de la catalogación tradicional y la facilidad de la búsqueda por palabras clave”.⁴¹

No sólo identifican y describen un objeto informático, sino que también documentan su comportamiento, función, uso y su relación con otros objetos informáticos, así como la forma en que debería gestionarse. Son datos que nos proporcionan los atributos con los que se describe un objeto digital. No sólo son claves para la búsqueda, sino también sirven para facilitar el uso de los recursos,

⁴⁰ VOUTSSÁS, Bibliotecas y publicaciones digitales. México: p. 155.

⁴¹ GORMAN, Michael. “¿Metadatos o catalogación? Un cuestionamiento erróneo”, en *Internet, metadatos y acceso a la información en biblioteca y redes en la era electrónica*, p. 1.

por ejemplo, indican las condiciones técnicas o comerciales para que ese recurso pueda ser utilizado.

Son ejemplos de metadatos:

- El encabezamiento de un fichero multimedia (imagen, vídeo o audio).
- El resumen de un documento.
- El catálogo de una base de datos.
- Los términos asignados haciendo uso de un tesoro.
- Las palabras extraídas de un texto.
- Las fichas catalográficas en cualquier formato (ISBD, MARC, etc.).
- Las páginas amarillas.

Una de las características generales de los metadatos es su capacidad de relación y de establecer enlaces. Además tienen como regla primordial que los documentos introducidos en la red, incluyan todos los datos necesarios para su búsqueda, localización y recuperación tanto por personas como por programas.

Los metadatos son estructuras de organización de la información, legibles por máquina, cuya finalidad es hacer útiles los datos, de distintas formas, según las necesidades concretas de cada servicio de información digital y según la aplicación que se les otorgue.⁴²

Los metadatos son estándares o modelos, como SGML. Todas las iniciativas de metadatos basadas originalmente en SGML están evolucionando al mercado flexible XML.

Existen diversos tipos de metadatos que corresponden con diferentes tipos de usuarios con tipos de información distinta, por ello, los metadatos se pueden clasificar según su origen, forma, funcionalidad, uso, etc.

⁴² *Ibidem.*

Baca⁴³ señala que al desarrollar sistemas de redes de información digital es importante dividir a los metadatos en las siguientes categorías:

TIPOS DE METADATOS Y SUS FUNCIONES		
TIPO	DEFINICIÓN	EJEMPLOS
Administrativo	Metadatos usados en la gestión de los contenidos y administración de recursos de información	-Adquisición de información -Control de derechos y reproducciones -Información sobre localización -Criterios de selección para la digitalización -Institución en cuyo servidor se alberga
Descriptivos	Metadatos usados para describir o identificar información sobre recursos	-Catalogación de registros -Ayudas para la búsqueda o URN -Índices especializados -Relaciones hiperenlazadas entre recursos
Conservación	Metadatos relacionados con la conservación de recursos de información	-Documentación de recursos sobre la condición física
Técnico	Metadatos relacionados con el funcionamiento de los sistemas o el comportamiento de los metadatos	-Documentación de hardware o software -Digitalización de la información -Autenticación y seguridad de los datos
Uso	Metadatos relacionados con el nivel y el tipo de uso de los recursos de información	-Registro de exhibiciones -Seguimiento de usos y de usuarios -Uso repetido del contenido e información sobre versiones múltiples

⁴³ GILLILAND Swetlad, Anne J. "La definición de los metadatos", en *Introducción a los metadatos: vías de la información digital*. p. 9.

Para Méndez⁴⁴ los metadatos se clasifican de la siguiente manera tomando en cuenta su aplicación:

- Para describir los recursos que se encuentran en la Web, asociados con la biblioteca digital.
- Que definen los elementos de los registros catalográficos en bibliotecas.
- Metadatos asociados a la información localizada en museos.
- Para describir recursos geográficos y geoespaciales como el modelo Content Standard for Digital Geospatial Metadata (CSDGM, 2001) y el Federal Data Geographic Committee (FGDC).
- Para la descripción de información gubernamental como el modelo Government Information Locator Service (GILS).

Un catálogo de biblioteca son metadatos, es decir, cada ficha de un libro es un metadato, ya que proporciona diferentes puntos de acceso, como por ejemplo las obras de un autor que se relacionan con otras obras del mismo autor.

Los metadatos son utilizados en tres categorías: descriptivas (identifican recursos de información como son el título, resumen, las palabras clave, etc.), administrativa (facilitan la gestión y el procesamiento de colecciones digitales como son el lenguaje, formato, la fecha de creación, el tamaño, etc.) y estructural (facilitan la navegación y presentación de los recursos).

El proceso para asignar metadatos se refiere a añadir la cabecera que describe el contenido, la propiedad intelectual, la tipología, la materia del documento y se requiere conocimiento de SGML, XML y HTML.

El número y tipo de metadatos que se van a aplicar dependerá de la naturaleza de los contenidos. Si por ejemplo lo que se está catalogando es una fotografía, los metadatos que se van a usar serían autor, formato, tamaño, peso, resolución, color y las palabras clave.

⁴⁴ MÉNDEZ Rodríguez, Eva. *Metadatos y recuperación de información*, p. 56.

Cabe señalar que los tipos de metadatos más útiles son los descriptivos para el usuario y la recuperación de la información, ya que indican explícitamente sobre qué trata la unidad de contenido.

Los metadatos asociados a un documento u objeto digital se pueden agrupar, en cuanto a su función, en las siguientes categorías:

- **Asidero (identificación o nombre)** Es la categoría compuesta por los metadatos que se usan para la identificación unívoca del objeto y para su localización. Facilitan la gestión de los objetos digitales.
- **Propiedades.** Esta clase está formada por los metadatos utilizados para describir los objetos digitales.
- **Firma.** Es la categoría que sirve para realizar la autenticación del documento, es decir: para certificar que el objeto no ha sido alterado.
- **Transacciones.** Es la información que nos indica cómo ha sido usado el objeto, por ejemplo: nos informa sobre quién, cuándo y cómo se ha utilizado o distribuido el objeto.⁴⁵

Los metadatos son importantes en las bibliotecas digitales porque permiten la localización y uso de cualquier documento, ya que describen el contenido de éstos.

Las bibliotecas digitales cuentan con normas establecidas para crear metadatos y uno de los esquemas que se propone para su creación es el estándar Dublin Core (DC), el cual fue desarrollado por OCLC en Dublin, Ohio, de ahí el nombre de Dublin Core.

El estándar se encuentra conformado por quince elementos o etiquetas que son más simples que los utilizados para la catalogación tradicional.

⁴⁵ Ernesto García Camarero y Luis Ángel García Menero. "La biblioteca digital", pp. 146-147.

En el estándar Dublin Core cada elemento es opcional, puede repetirse y aparecer en cualquier orden. Si los elementos de DC se lograran aceptar internacionalmente, todos los robots que indizan documentos en Internet encontrarían en la cabecera de los mismos todos los datos necesarios para su indización. La eficacia de los robots como Altavista y Yahoo mejoraría notablemente. Si el DC lograra ser operativo, los metadatos de la cabecera de los documentos facilitarían su indización automática y mejorarían la efectividad de los motores de búsqueda.

El Dublin Core ha ido evolucionando hacia un formato de registro para el intercambio y recuperación de información en el espacio digital.

Existen tres grupos de estándares para metadatos estrechamente vinculados entre sí y son de particular importancia en el ámbito de la Web: las etiquetas <META> de Altavista, el Dublin Core y el Resource Description Framework.

- Etiquetas <META> de Altavista

Altavista sirve de apoyo a dos elementos que son palabras clave y descripción que pueden insertarse en recursos de la Web usando la etiqueta HTML<META>. La información suministrada por las palabras clave contribuye a la clasificación de la información, mientras que la información que proporciona la descripción se usa al visualizar los resultados para obtener un resumen más preciso de un recurso en particular.

El desarrollo del Dublin Core está determinado por las siguientes directrices:

- Los elementos tienen que ser fáciles de entender y de usar para que cualquier persona que introduzca recursos en la Web los describa sin necesidad de capacitarlo. Cada elemento es opcional y repetible.

Resource Description Framework (RDF). Es un lenguaje para representar información sobre los recursos de la World Wide Web y las relaciones entre ellos. Está particularmente diseñado para representar metadatos tales como título, autor, fechas de las páginas web, copyright.⁴⁶

El RDF es una aplicación de metadatos del XML considerado por la mayoría de los expertos como el sucesor del HTML y el futuro lenguaje de la Web, además facilitará una arquitectura flexible para la utilización de diversos paquetes de metadatos.

RDF es una estructura flexible para la recuperación de información basada en metadatos que propone un modelo de datos coherente y un marco sintáctico para los distintos modelos como MARC, DC, TEI, GILS, etc.

RDF proporciona soluciones que habilitarán un mayor grado de fiabilidad, relevancia y exactitud de las aplicaciones y servicios orientados al descubrimiento del recurso y a la gestión de sitios Web y otros recursos de Internet.⁴⁷

Para la creación de metadatos se deben considerar los siguientes puntos:

1. Seleccionar el estándar metadatos que mejor se ajuste a nuestras necesidades.
2. Conocer y entender el estándar metadatos seleccionado.
3. Recopilar toda la información sobre los datos.
4. Analizar los datos y decidir como se codificará la información referente a los mismos.
5. Capturar los metadatos.
6. Verificar y modificar la estructura sintáctica de acuerdo al estándar metadatos seleccionado.

⁴⁶ María Jesús Lamarca Lapuente. "Lenguajes hipertextuales", en *Hipertexto: el nuevo concepto de documento en la cultura de la imagen*, disponible en <http://www.hipertexto.info/documentos/lenguajes_h.htm>, [Consultado el 16 de julio de 2009].

⁴⁷ Pedro Manuel Díaz Ortuño. "Problemática y tendencias en la arquitectura de metadatos web", en *Anales de documentación*, no. 6 (2003): p. 55.

7. Revisar el contenido de los metadatos para comprobar que la información contenida en ellos, describe los datos de manera completa y correcta.

Cuando se implemente el uso de metadatos en la biblioteca digital también debe considerarse lo siguiente:

1. No inventar su propio estándar metadato.
2. Ajustarse a los estándares existentes para poder tener intercambio de información con el resto del mundo.
3. Familiarizarse con el estándar que se elija.
4. Determinar la prioridad de los datos.
5. Documentar hasta un nivel en el que se preserve el valor de los datos.
6. No limitar la participación de profesionales.
7. Contar con una adecuada infraestructura de cómputo y comunicaciones.

Por otro lado, existen diversos software y aplicaciones para la edición de metadatos:

Herramientas para
la creación de metaetiquetas

TIPO	SOFTWARE
Plantillas	Metatag Builder Position Site
	Owner Metabuilder Vancouver
Software Cliente	9X metaMagic 2.07
	Legal y Metatag Builder 2.0
	Meta Master v.5.0
	Meta Demon 2.1.1
	Meta Former 2000
	Meta Tag Maximizer
	Metty 1.21

Herramientas para la creación de metadatos

Plantillas (Templates)	Klority Nordic Metadata Project Template DC-Dot (UKOLN) Reggie (DSTC) Reg (DSTC)
Software Cliente	Note Tab Light 4.86+Dcmeta.clb TagGen (Hi Software) Meta Contents Builder (Hi Software) Metabrowser 1.5 Metamanage 2.5.15
SIG-DLO	Wachfire WebXM MetaStar 7.0 (Blue Angel) MetaMatrix Meta Tagger (Interwoven) Ex-Libris DigiToo Library/Metalib

Por último, considerar qué software nos conviene utilizar para la creación de metadatos dependerá de

- 1) si somos los creadores de un recurso web y necesitamos metaetiquetas mínimas que entiendan los buscadores generales, simplemente para aumentar su visibilidad en los sistemas de búsqueda.
- 2) si queremos crear metainformación para construir una biblioteca digital muy bien organizada y un servicio de información de calidad.

En conclusión los metadatos se utilizarán:

- a) Desde el enfoque del control bibliográfico para resumir, buscar, evaluar, seleccionar, recuperar y especificar el uso e interpretación de

los datos a los que se refieren, así como para la descripción de los mismos.

- b) Desde el punto de vista de la gestión de los datos, sirven para administrar y gestionar un DLO dentro de un servidor, para definir las relaciones entre los distintos objetos de datos y para especificar la estructura y los componentes lógicos de los objetos de información y el modo de acceso a ellos.⁴⁸

Podemos observar que existen una gran variedad de formatos que nos permiten organizar la estructura de un documento digital en forma de hipertexto con la finalidad de que pueda ser fácilmente procesado por un programa. Hoy en día, el formato de XML es el más utilizado, pues permite que los datos sean más fáciles de gestionar y permite el desarrollo de aplicaciones más rápidas, además de ser uno de los componentes claves del desarrollo web, ya que está siendo usado para definir esquemas de datos.

Como se ha señalado a lo largo de los capítulos anteriores, el surgimiento de los documentos digitales ha traído consigo el diseño de nuevas herramientas que permitan su descripción, organización y recuperación. Se puede señalar que los metadatos son considerados como una de las herramientas más completas, pues permiten describir de una manera completa a los documentos digitales, lo que nos llevará a que la localización y recuperación de éstos sea de manera rápida y sencilla.

⁴⁸ Méndez Rodríguez. *Metadatos y recuperación de información*, p. 55.

CAPÍTULO 3

Organización documental en la sociedad de la información. Tendencias

En este capítulo se habla de cuáles son las tendencias que han surgido en la organización de los documentos digitales con la aplicación de las nuevas tecnologías, por lo que se señalarán algunas definiciones del término *tendencia*, para entender la aplicación de este término en el campo de la bibliotecología.

3.1 Tendencias en la organización documental

El término *tendencia* se refiere a la inclinación o disposición de que algo suceda, es decir, designa el curso que toman los acontecimientos. Una segunda definición señala que tendencia es la idea religiosa, económica, política, artística, etc., que se orienta en determinada dirección.¹

En este contexto, el término tendencia se aplica en el presente capítulo para saber cuál es el rumbo que ha tomado la organización documental con la aplicación de las TIC en la llamada sociedad de la información; por ello se hablará de las tendencias que apoyan la organización documental, como son la web 2.0, la biblioteca 2.0, las folksonomías, la web semántica, la red académica, etcétera.

¹ Diccionario de la Lengua Española. 22^a ed., en <http://buscon.rae.es/draeI/SrvltConsulta?TIPO_BUS=3&LEMA=tendencia> [consultado el 4 de agosto de 2009].

Hoy en día existe la necesidad de contar con un rápido y adecuado acceso a la información, por lo que debe organizarse a través de sistemas de clasificación o lenguajes documentales, como se apreciará a lo largo de este capítulo. Por ejemplo, en los materiales impresos, se organiza la información aplicando términos y conceptos para recuperarlos, pero esto no es tan sencillo de aplicar en los documentos digitales, pues existe una cantidad ingente de documentos en la Internet, además de ser un medio en el que existe una gran cantidad de tipos de usuarios que recurren a aquélla para buscar información, es decir, se dificulta la organización de la información digital, debido a que no se dirige a un tipo de usuarios con necesidades particulares, sino a un público bastante amplio. Si a esto añadimos, como ya se dijo anteriormente, que la información en la Internet se incluye, modifica o desaparece diariamente, tenemos como resultado que resulta difícil la elaboración de herramientas que nos permitan indizar la información contenida en este medio.

Cabe señalar que la navegación es un aspecto importante a considerar en cuanto al uso de las nuevas tecnologías aplicadas a las tendencias que se analizarán posteriormente. Al utilizar Internet, la información se transmite, busca y recupera de forma diferente a como se realizaba en medios impresos (a través de los catálogos electrónicos en las bibliotecas), por ello, deben crearse herramientas que acerquen al usuario a la información que requiere, mediante el uso de motores de búsqueda² o agentes inteligentes.

El papel que desempeña el hipertexto dentro de la navegación es fundamental, puesto que constituye una red de relaciones semánticas, al igual que el texto impreso, pues el hipertexto se encuentra constituido por palabras, frases y párrafos. El hipertexto es una nueva forma de manejar y representar la información; la lectura y escritura son totalmente distintas a las de los medios impresos, puesto que no hay que seguir una secuencia para leer el hipertexto,

² Un motor de búsqueda es un sistema informático que indexa archivos almacenados en servidores web. Un ejemplo son los buscadores en la Internet. Definición disponible en: <http://es.wikipedia.org/wiki/Motor_de_b%C3%BAqueda>, [Consultado el 28 de septiembre de 2009]

debido a los enlaces que se indican por medio de palabras resaltadas, gráficas o íconos.

Debido a esta nueva forma de representación de la información contenida en Internet, se han creado distintas formas de búsqueda y recuperación, comenzando porque los términos (búsqueda y recuperación) se denominan *navegación* y *exploración*, pero lo importante no es cómo se les llame, sino la creación de otras modalidades de acceso a los documentos digitales.

La navegación se refiere al movimiento que se realiza a través de la información digital dispuesta en hipertexto o multimedia, mediante botones e interfaces.³ Al navegar, el lector no sigue un orden, es decir, cada usuario recorre un camino propio, a diferencia de lo que sucede con el libro impreso.

Ahora bien, a lo largo de esta investigación se ha observado que, al hablar del entorno digital, en el que la información contenida en Internet tiene características diferentes a la información impresa, la manera de buscar también es diferente, por lo que, cuando el usuario busca información en la red, se le llama exploración.

Para encontrar información útil, el usuario la busca a través de medios automatizados presentes en la red, que puede ser a través de los motores de búsqueda, mejor conocidos como buscadores de Internet, entre cuyos ejemplos se hallan Altavista, Google, Excite, Infoseek, Yahoo y otros. En éstos, las búsquedas se realizan mediante palabras clave o por medio de árboles jerárquicos por temas, por ello, los buscadores se clasifican en dos tipos: 1) índices o directorios por temas: la búsqueda se realiza por temas, cuya creación la realizan personas; 2) motores de búsqueda: la búsqueda se realiza por medio de palabras clave, a través de robots (bases de datos) que se encargan de incorporar las páginas web.

³ TORRES Vargas, Araceli Georgina. "La navegación y exploración en la información digital del World Wide Web", en *La información en el inicio de la era electrónica*. p. 166.

Con el uso de los robots, los motores de búsqueda contienen más información que los directorios. Algunos robots indizan a partir de los títulos o los primeros párrafos del texto; otros, en el texto completo, aunque la desventaja es que no todos los documentos contienen título y, por lo tanto, no es analizado dicho documento. Entonces, como la indización solamente se aplica al título y a los primeros párrafos, resulta poco confiable porque se excluye mucha información de la web del análisis que realiza el robot.

Actualmente, Google es el motor de búsqueda más utilizado, debido a que las estrategias de búsqueda ofrecidas al usuario permiten que aquélla sea por palabras (búsqueda en lenguaje natural), en la que el usuario suele escribir las palabras que recuerda sobre la información requerida, creando así sus propias estrategias de búsqueda.

Por lo anterior, diversas bibliotecas han diseñado catálogos en los que los usuarios realicen búsquedas de este tipo, es decir, en lenguaje natural; por ejemplo, como en los casos de la Biblioteca del Congreso de Washington y la Biblioteca "Daniel Cosío Villegas" del Colegio de México, han optado por contar con ambos catálogos: por un lado, con el catálogo tradicional OPAC y con el nuevo catálogo en el que los usuarios realizan las búsquedas en lenguaje natural, parecido a Google, lo que permitirá acercar al usuario al material que busca, aun cuando no recuerde el nombre completo del autor o el título exacto de una obra.

Es un hecho que los motores de búsqueda ayudan al usuario a localizar la información que realmente necesita, pero, debido a que la información contenida en Internet cambia y desaparece vertiginosamente, además de que existe información de muy diversa naturaleza (artículos, anuncios, páginas informativas, noticias y demás), es casi imposible la creación de herramientas que indiquen para siempre esta información. Pero también es cierto que los usuarios actuales realizan sus búsquedas en dichos motores de búsqueda utilizando solamente lenguaje natural, lo que nos lleva a los bibliotecarios a tomar la iniciativa de incorporar en nuestras bibliotecas, centros de

documentación e información estos nuevos tipos de catálogos, a una nueva opción de búsqueda para los usuarios de hoy que crecieron y desarrollaron sus formas de búsqueda con el uso de las nuevas tecnologías.

3.1.1 La web 2.0 en Internet y en la biblioteca

La web 2.0 se refiere al Internet convertido en un espacio social que forma parte de la sociedad de la información, por lo que se le ha llegado a llamar *web social*. Cualquier aplicación o servicio que gire en torno a las características propuestas como propias de la web 2.0, han hecho que se le califique como web social. Todo el universo de esta web se sustenta en una diversidad considerable de aplicaciones y servicios, agrupados todos bajo el concepto de software social.

La web 2.0 surgió como una forma innovadora para introducir la inteligencia necesaria en la red, es decir, trata de involucrar directamente al usuario, dando lugar a una web semántica.

La web 2.0 incorpora una o varias tecnologías, teniendo como prioridad al usuario. También en ésta aparecen los blogs, considerados como una herramienta personal para la comunicación y la expresión individual en la red. Esta red supone un cambio de rumbo, ya que ésta se acerca al usuario y no el usuario a la red, pues éste es quien elige la información y la selecciona para que, automáticamente y de acuerdo a sus necesidades, la información llegue a su computadora.

El éxito de los grandes líderes de la web 2.0 se debe a su poder para explotar la inteligencia colectiva, es decir, los hipervínculos son la base de esta red y, a medida que los usuarios agregan contenidos y sitios web nuevos, se enlazan con la estructura de la web gracias a otros usuarios que descubren el contenido y se enlazan con él, así es como la red de conexiones crece como resultado de la actividad colectiva de todos los usuarios de la web. Entre los ejemplos más relevantes se hallan los siguientes:

Yahoo: nació como un catálogo o directorio de enlaces, pero su valor actual reside en su papel como portal del trabajo colectivo de los usuarios.

Google: su innovación en la búsqueda se debe al método llamado PageRank, que utiliza la propia estructura de enlaces en la web para proporcionar mejores resultados de búsqueda, en lugar de usar sólo las características de los documentos.

Amazon: actualmente cuenta con invitaciones para participar de diversas maneras en casi todas sus páginas; utiliza la actividad de los usuarios para producir mejores resultados de búsqueda. Con la participación del usuario, las ventas de Amazon han llegado a ser mayores que las de sus competidores. Las compañías que adoptan la idea de la participación del usuario dejan su marca en la web, como se aprecia en las antes mencionadas. Hoy, al hablar de la web 2.0, se hace referencia al uso de determinadas tecnologías, como Ajax,⁴ software social, etc., es decir, dichas tecnologías están presentes en los sitios web 2.0.

Ajax: una combinación de XML con Javascript que permite crear aplicaciones web ejecutables en el cliente, reduciendo el tráfico de datos y la carga de trabajo del servidor. Como ejemplo tenemos Google Maps o Gmail.

Software social: un elemento clave en todo el desarrollo de la web 2.0, especialmente en lo que compete al aprovechamiento de la inteligencia colectiva, referente al conjunto de aplicaciones que permiten a los individuos comunicarse unos con otros y seguir las conversaciones a través de la web. En este rubro encontramos tres componentes que influyen en la web 2.0: los blogs, los foros y las wikis:

- ❖ Los blogs son páginas web que contienen mensajes ordenados cronológicamente y enviados por uno o varios autores, lo que permite al resto de usuarios leerlos y escribir sus comentarios. Algunos periódicos tienen sus propios blogs y han desarrollado una nueva forma de columnas de opinión, por ejemplo, los periódicos Elpais.com, El mundo.es.
- ❖ Los foros funcionan de manera similar a los blogs: permiten a cualquier usuario registrado plantear un tema de discusión y el resto participar en ésta. Surgieron en la web

⁴ Acrónimo de Asynchronous JavaScript and XML (JavaScript asíncrono y XML). Es una técnica de desarrollo web para crear aplicaciones interactivas, las cuales se ejecutan en el cliente, es decir, en el navegador de los usuarios mientras se mantiene la comunicación asíncrona con el servidor en segundo plano. De esta forma es posible realizar cambios sobre las páginas sin necesidad de recargarlas, lo que significa aumentar la interactividad, velocidad y usabilidad en las aplicaciones, en <<http://es.wikipedia.org/wiki/AJAX>>, [Consultado el 17 de julio de 2009]

1.0, cuyos predecesores son los grupos de noticias o las listas de correos.

- ❖ Las wikis permiten a cualquier persona leer, escribir o modificar los mensajes escritos por otros usuarios. El ejemplo más famoso es la wikipedia.

En el contexto de la biblioteca digital, al utilizar la web 2.0, los usuarios no sólo recopilarán información como ocurría anteriormente, sino que además harán uso de ésta, es decir, la aprovecharán de la siguiente manera:

- Creando contenidos por parte de los usuarios, como es el caso de las wikis, en las que los usuarios registran su conocimiento y crean nueva información.
- Pueden compartir objetos digitales, ya sean vídeos, fotografías o documentos.
- Aportan comentarios a los objetos digitales compartidos, como en los videos contenidos en YouTube, o los materiales presentes en las colecciones digitales dentro de la biblioteca digital o en los libros referidos en Amazon.

Todas las formas anteriores de intercambio y aprovechamiento del conocimiento colectivo tienen como denominador común la mejoría del servicio, por ejemplo, entre más comentarios tenga un libro en Amazon o los servicios que ofrece la biblioteca, más útiles serán tales servicios para los mismos usuarios.

El nuevo reto para los profesionales de la información consistirá en diseñar servicios atractivos, útiles para los usuarios y que los inviten a participar y a aportar su conocimiento.⁵ Algunos de los servicios presentes en la web 2.0 son

Amazon. Uno de los modelos de librería virtual más conocidos y consolidados de Internet.

⁵ MARGAIX Arnal, Dídac. "Conceptos de web 2.0 y biblioteca 2.0: origen, definiciones y retos para las bibliotecas actuales", en *El profesional de la información* 16, núm. 2, marzo-abril de 2007 p. 101, en <<http://www.oei.es/tic/kx5j65q110j51203.pdf>>, [Consultado el 23 de noviembre de 2007]

Wikipedia. Enciclopedia libre en la que los usuarios consultan gratuitamente sus artículos y que también amplían o modifican la información localizada en esos artículos.

Google Maps. Sistema de búsqueda de localizaciones geográficas. Las claves de su éxito estriban en el uso de tecnología Ajax y en la creación de una “application programming interface” (API), la cual permite a los desarrolladores externos crear aplicaciones nuevas, utilizando gratuitamente la información de Google Maps.

Youtube. Sitio que permite compartir videos. Cualesquier persona puede subirlos, añadir comentarios, seleccionarlo como favorito, entre otras acciones.

3.1.2 La biblioteca 2.0

Los primeros esfuerzos en la biblioteca digital se centraron en almacenar, organizar y dar acceso a la información en formato digital. Ahora, con la aplicación a las bibliotecas de la web 2.0, se ha dado paso al desarrollo de servicios orientados al usuario, espacios de participación y colaboración, es decir, la biblioteca aprovechó las herramientas de la web 2.0, como el hecho de que los usuarios tomen decisiones de acuerdo a sus necesidades y en el diseño de los servicios y actividades que ofrece la biblioteca.

El término de biblioteca 2.0 va de la mano con el término web 2.0, debido a que la web se ha convertido en un fenómeno social, nutrido por la propia sociedad que decide cómo funciona. Esta activa participación del destinatario final ha cambiado las formas de trabajo en la web, de tal manera que hoy existen enciclopedias creadas por sus lectores (wikipedia y demás wikis), archivos de acceso abierto mantenidos por los usuarios, sitios donde los destinatarios colocan las etiquetas a los contenidos que leen (como YouTube) y enormes espacios en los que se comparten desde videos, audios u otro tipo de archivos.

La primera aplicación de la web 2.0 a las bibliotecas fue aprovechar estas nuevas herramientas para mejorar los servicios. Al aplicarlas, se observó que a través de éstas los usuarios podían, igual que en la web 2.0, decidir, de acuerdo a sus necesidades, participar activamente en distintas acciones, como el diseño de los servicios y actividades que se llevan a cabo en la biblioteca.

El modelo biblioteca 2.0 incluye principalmente un incremento en el flujo de información, desde el usuario hacia la biblioteca. Se intenta incorporar al usuario en el diseño e implementación de los servicios bibliotecarios, incluyendo la retroalimentación y participación activa. Así, los servicios bibliotecarios se actualizan constantemente y se reevalúan para atender mejor las necesidades de su comunidad.

Maness señala que la biblioteca 2.0 es la aplicación de tecnologías bibliotecarias basados en web.⁶ En cambio, Michael C. Habib define la biblioteca 2.0 como “la aplicación y la adaptación del modelo web 2.0 al entorno bibliotecario, tanto virtual como físico”.

Los servicios bibliotecarios se diseñan para cubrir las necesidades de los usuarios causadas directa o indirectamente por los efectos de la web 2.0.⁷ Con este concepto de biblioteca 2.0, se espera que los usuarios participen activamente. Los elementos que conforman la biblioteca 2.0 se aprecian en la siguiente gráfica:⁸

⁶ MANESS, Jack M. “Library 2.0 theory : web 2.0 and its implication for libraries”, en *Webology* 3, núm. 2 (junio de 2006), en <<http://webology.ir/2006/v3n2/a25.html>>, [Consultado el 26 de septiembre de 2008].

⁷ HABIB, MICHAEL C. Habib. Defining academic library 2.0: lessons learned from web 2.0, en <<http://www.unc.edu/~mchabib/inls342/DefiningAcademicLibrary20Presentation.pdf>>, [Consultado el 26 de septiembre de 2008].

⁸ La imagen fue tomada de Dídac Margaix Arnal. “Conceptos de web 2.0 y biblioteca 2.0: origen, definiciones y retos para las bibliotecas actuales”, en <<http://www.oei.es/tic/kx5j65q110j51203.pdf>>, [Consultado el 15 de enero de 2009].

Elementos de la Biblioteca 2.0



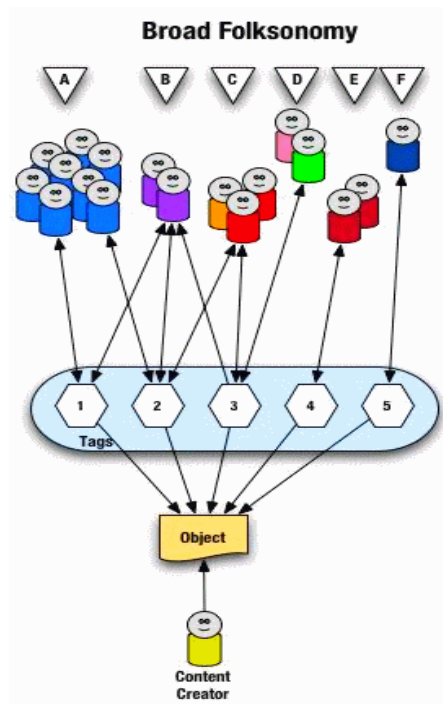
De dicha gráfica se concluye lo siguiente: la aplicación de las herramientas de la web 2.0 a la biblioteca 2.0 se centra básicamente en la mejora de los servicios que ofrece, es decir, los usuarios, de acuerdo a sus necesidades, participan activamente en el diseño e implementación de nuevos servicios, así el ambiente se vuelve más participativo y colaborativo entre biblioteca y usuarios, y los procesos de la biblioteca se centrarían en las necesidades de sus usuarios.

La biblioteca 2.0 establece que el espacio bibliotecario (virtual y real) sea más interactivo y colaborativo, de tal manera que se aprovechen los servicios gratuitos que ofrece la Internet (blogs, open access, open source y otros) para mantener una comunicación con los usuarios.

3.1.3 Las folksonomías en Internet y en la biblioteca

Las folksonomías surgen en Internet. Son un tipo de sistema de clasificación distribuida, generalmente creada y empleada en la web 2.0 por un grupo de individuos que por lo general son usuarios de esos recursos. Este nuevo método se usa en varios servicios web, como del.icio.us, en el que los usuarios añaden y clasifican las páginas web y fotografías, es decir, los usuarios agregan etiquetas

(tags)⁹ a los ítems en línea, que pueden ser imágenes, videos o texto, como se aprecia en la siguiente figura:



Las etiquetas son una especie de metadatos y no un reemplazo para los sistemas de clasificación formal, como Dublin Core; más bien se trata de un medio suplementario para organizar información y ordenar, a la vez, los resultados de las búsquedas.

La folksonomía muestra su utilidad cuando los miembros de la comunidad clasifican directamente un mismo elemento, generando etiquetas similares o iguales, o usan las mismas etiquetas para elementos similares. Con la clasificación folksonómica no se tiene que pensar o recordar una categoría determinada para clasificar, ya que se añaden etiquetas como uno quiera.

⁹ Son un conjunto, libremente elegido, de palabras clave textuales. Una posible definición sería que son palabras clave, nombres de categoría o metadatos. La definición fue consultada en "Folcsonomías y tags: como lograr un tagging efectivo", en <http://www.masternewmedia.org/es/2006/02/02/folcsonomias_y_tags_como_lograr.htm>, consultado el 13 de abril de 2008..

Pero, ¿qué es exactamente una etiqueta (o *tag*)? Son palabras clave, nombres de categoría o metadatos o un conjunto de palabras claves textuales. Sin embargo, cabe señalar que no las crean especialistas de la información; tampoco se sigue ninguna guía formal, lo que significa que los ítems se categorizan con cualquier palabra que defina una relación entre el recurso en línea y el concepto que surgió de la mente del usuario.

Las *tags* se utilizan en taxonomías dinámicas y flexibles, generadas automáticamente por la Internet para recursos en línea, como las páginas web, imágenes digitales, etc. Por ello, el concepto *tagging* se ha comenzado a asociar al movimiento web 2.0.

Un ejemplo del uso de las folksonomías en Internet se muestra desde el punto de vista de la organización personal de datos, por ejemplo, en el sitio flick existe la tag *me* (yo), con fotos personales de los usuarios o relaciones que son historias en las que un usuario comienza un tag y otros la pueden continuar. Al respecto, Catuxa Seoane señala que el éxito de las folksonomías se debe a que: somos creadores y consumidores: contribuimos y nos beneficiamos del *tagging*; reflejan el vocabulario de la gente; son un atajo (de abajo a arriba) hacia las categorías, y porque favorecen la navegación y proponen otras formas de búsqueda.¹⁰

Existen dos sistemas de folksonomías en la web: el primero es del.icio.us, que es un servicio de gestión de marcadores sociales en la web. Permite agregar los marcadores que se guardan en los navegadores y categorizarlos con folksonomías (*tags*); además de almacenarlos, permite compartirlos con otros usuarios, de tal manera que se guardan los enlaces preferidos. El segundo es flickr, un sistema de almacenaje en línea de fotos, el cual permite a los usuarios

¹⁰ SEOANE, Catuxa. "Flexibilidad de las folksonomías", en <<http://eprints.rclis.org/11558/1/Anuario-ThinkEPI-2007-Catuxa-Flex-Folkson.pdf>>, [Consultado el 22 de febrero de 2008].

identificar sus fotografías por medio de un conjunto de *tags*. Algunas de las cosas que del.icio.us permite hacer a sus usuarios son las siguientes:

- a) Guardar los enlaces preferidos y acceder a éstos desde cualquier computadora.
- b) Es posible utilizar etiquetas (*tags*) para organizar y recordar los enlaces preferidos de manera flexible y útil, a través de carpetas.
- c) Se descubren contenidos al navegar, es decir, se aprovecha el trabajo de los usuarios de la comunidad para después profundizar en lo que nos interesa.
- d) Se comparten favoritos con la comunidad de del.icio.us y fuera de ésta.
- e) Se etiquetan por tipo de archivo, por ejemplo, `<http://del.icio.us/tag/system:filetype:mp3>`, método con el que se podría realizar nuestra búsqueda diciendo que nos muestre sólo los mp3 de música jazz, combinando etiquetas.

Cabe señalar que hay que situarse en la mente de los usuarios y ver por qué envían ciertos *tags* para permitir que éstos sean más *buscables*. En del.icio.us, por ejemplo, se aprecian los top de 50 etiquetas agregadas por los usuarios, ya que existe una variedad de herramientas disponibles que ofrecen numerosos métodos de visualización, como

- 1) Tag.alicio.us, es un diseño experimental que opera como filtro de tags, recobrando enlaces de del.icio.us, de acuerdo con tags y restricciones de tiempo (por ejemplo, etiquetas de tal hora, día actual o de la semana).
- 2) Extisp.icio.us exhibe una distribución aleatoria de los tags de un usuario, a través del número de veces que aquél ha reutilizado cada tag.
- 3) Facetious reutiliza la base de datos de del.icio.us, la cual clasifica y agrupa los tags debajo de encabezados como “por lugar” (México, Australia, Japón), “por tecnología” (blogs, wiki, website), etcétera.

Entre las ventajas de las folksonomías destacan que la asignación de las etiquetas es flexible; su simplicidad en la gestión y utilización; que las consultas son específicas y adecuadas al vocabulario del usuario, asimismo que la creación es de bajo costo y poco esfuerzo.

Ahora bien, otros autores señalan que algunas de las desventajas de las folksonomías se deben a que las etiquetas son inexactas (por ejemplo, *del.icios.us*), ambiguas, imprecisas y subjetivas. Las tres últimas debido a que se trata de sinonimias y homonimias. Esto se debe a que muchos de los sitios permiten solamente una palabra, es decir, no se admiten términos compuestos. Por el momento no existe un control de los sinónimos y homónimos.

Los administradores de los sistemas no imponen control sobre las etiquetas escogidas por los usuarios, cuyo resultado obtenido es un conjunto caótico y sin control de términos marcados con *tags* que no son las adecuadas para la búsqueda, como sí lo son los vocabularios controlados.

Los usuarios no advierten lo anterior como problema, pues las etiquetas están ahí para ayudar al usuario final, quien envía esos *tags*, por lo que cada palabra seleccionada por el usuario tiene un significado único.

Podemos decir que el problema real no es el caos reinante en sus etiquetas, sino que se intenta cubrir las colecciones personales y colectivas. En la práctica, se promueve la inclusión de los *tags* más comúnmente utilizados por los de un solo uso, o bien por los que se usan de manera infrecuente, como los sinónimos.

Uno de los métodos para mejorar las etiquetas es que se organice mejor su alfabetización en las folksonomías, lo cual implica dos procesos: primero, que la comunidad necesita estar lista para definir reglas y coincidir en una serie de estándares de *tags* y, segundo, que los usuarios necesitan ser conscientes y aceptar seguir esas reglas. Por el momento no existen lineamientos estándar para una buena selección de *tags*; sin embargo, quienes pertenecen a la comunidad de la folksonomía han propuesto varias ideas, las cuáles se dan a conocer en blogs y sitios de discusión de folksonomía.

Algunas propuestas que se señalan en los blogs a la hora de asignar *tags* son los que se enuncian enseguida:

- ❖ Utilizar el plural en vez del singular.
- ❖ Uso de minúsculas.

- ❖ Continuar con la convención de las etiquetas iniciada por otros y agregar sinónimos.
- ❖ Usar tags personales, siempre y cuando se utilicen los más genéricos.

Podemos añadir que una folksonomía “mezcla, diverge y evoluciona de la misma manera que lo hace el idioma, a través del uso y la interacción”.¹¹ La folksonomía aplicada a un portal web sería un sistema de autoclasificación de contenidos, basado en las descripciones de los usuarios al contenido del portal con sus propias palabras (*tags*), creando así un lenguaje natural, sin jerarquías preestablecidas.

En las bibliotecas, por ejemplo, el uso de los catálogos (u OPAC), siempre han sido criticados por carecer de una interfaz fácil de usar. El usuario desconoce de qué manera se encuentra estructurada la CDU o las ISBD, y se sentirá incomodo buscando en un medio basado en estos términos. Dentro de este contexto, los bibliotecarios, documentalistas, arquitectos de la información y demás, deben ser expertos en la creación y manejo de clasificaciones, lenguajes controlados, tesauros y otros recursos, que de nada sirven si no se logra que el usuario, cuando busque información, no la encuentre, pues conviene recordar que el usuario, cuando realiza sus búsquedas lo hace, la mayoría de las veces, en lenguaje natural, por lo que para él resulta más fácil buscar en toda la Internet a través de Google, que en los catálogos de una biblioteca.

Se puede señalar, además, que la web 2.0, la biblioteca 2.0 y las folksonomías se caracterizan por tener un alto grado de colaboración, en donde el usuario decide qué poner y cómo usarlo, las contribuciones del usuario son la clave para el dominio del mercado en la era de la web 2.0.

¹¹ GUY, Marieke; Tonkin, Emma. “Folksonomía y tags : como lograr una tagging efectivo”, en <http://www.masternewmedia.org/es/2006/02/02/folcsonomias_y_tags_como_lograr.htm>, consultado el 29 de mayo de 2008.

Como se puede observar, las tendencias actuales en la web nos llevan a ser unos usuarios más participativos, por ejemplo: en las redes académicas para la investigación y la docencia, todos los usuarios intercambian sus conocimientos con otros colegas. También, en los sitios del.icio.us y flickr, los usuarios agregan etiquetas a todos los recursos que se agregan o que encuentran en el sitio.

3.1.4 La web semántica en Internet y en la biblioteca

Actualmente, la información en la web está estructurada mediante lenguajes de etiquetado que describen cómo dicha información debe presentarse al usuario por el navegador, pero no expresan nada sobre su significado, es decir, su semántica.

La web semántica es una aplicación que nace en Internet y con ésta se trata de construir una biblioteca multimedia, integrando herramientas de búsqueda perfeccionadas. Asimismo, los promotores de la web semántica crean una serie de lenguajes documentales que han sido llamados ontologías.

La web semántica sería una extensión de la web actual, dotada de significado, esto es, un espacio en el que la información tendría un significado bien definido, de manera que sea interpretada por personas y por computadoras, por ejemplo, podríamos buscar los medios de transporte disponibles en una ciudad determinada o conseguir un mapa de la ciudad.

La web semántica ha sido impulsada por el creador de la WWW, Tim Berners-Lee y otros especialistas relacionados con la 3WC. En el 2000, Berners-Lee ofreció una conferencia en la que propuso que la nueva información debía reunirse para que se comprendiera, en vez de ponerla en una lista. La web semántica sería una red de documentos inteligentes, es decir, dotan a las páginas web de contenido semántico, así la web semántica no sólo almacenará datos sino también entenderá e interpretará el sentido de dicha información.

La mayoría de los sitios web están contruidos en lenguaje HTML, pero el desarrollo de la web semántica requiere de otro tipo de lenguajes, como los lenguajes XML y el RDF, que pueden dotar a cada página de contenido y a cada recurso de una lógica y significado que permitan a las computadoras conocer el significado de la información que manejan.

Se necesita un lenguaje¹² común basado en la web, con suficiente capacidad expresiva y de razonamiento para representar la semántica de las ontologías. Por ello, la utilización de lenguajes como el OWL¹³ es un paso más en el desarrollo de la web semántica. En ésta es necesario crear vocabularios descriptivos/semánticos, definidos en formato RDF para determinar el significado contextual de cada palabra, por medio de la consulta a la ontología apropiada. Así, tanto los agentes inteligentes y programas autónomos buscarían en la web automáticamente, sólo en las páginas que contengan la palabra buscada con el significado con que interpretamos ese término.

En este sentido, los agentes inteligentes son software que recogen, filtran y procesan la información contenida en la web, algunas de cuyas características que deben tener los agentes inteligentes son las que se indican a continuación:

Comunicativo: debe entender las necesidades, objetivos y preferencias del usuario y debe poder comunicarse mediante representaciones compartidas de conocimiento llamadas ontologías.

Capaz: el agente debe proporcionar no sólo información sino también servicio.

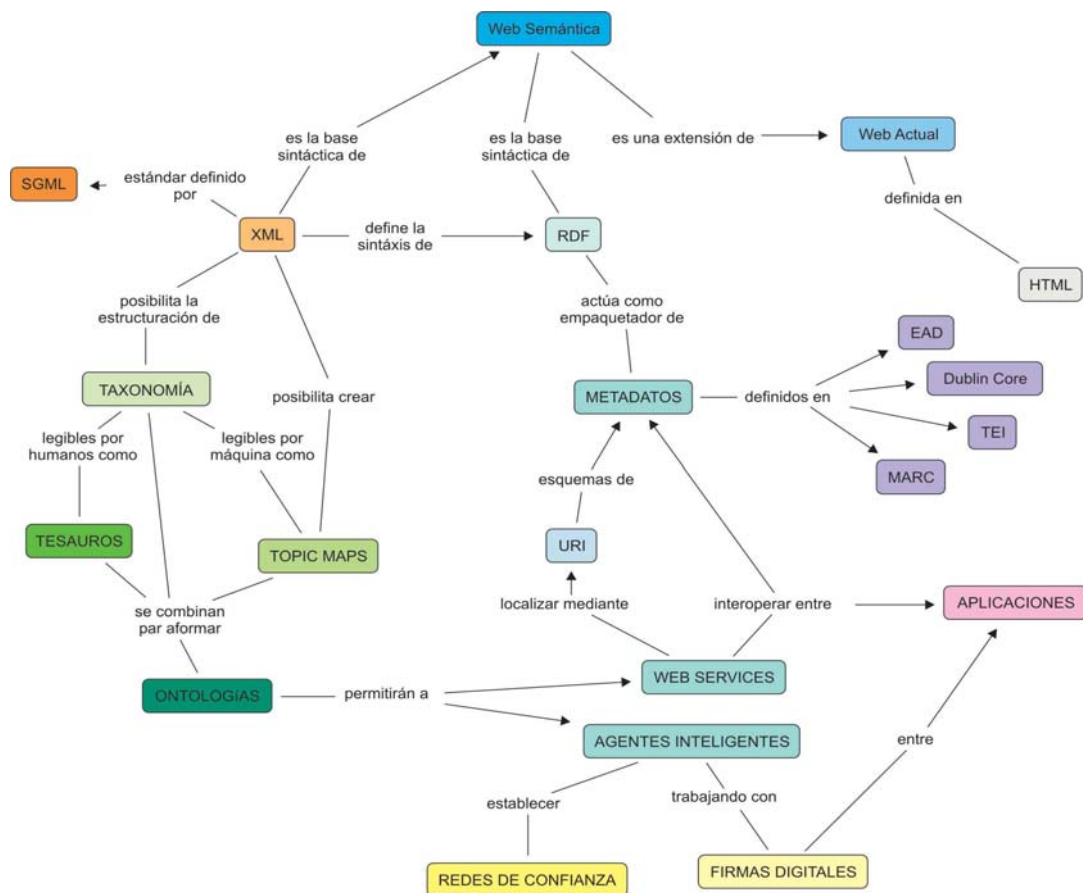
¹² Se refiere a lenguaje de marcas también llamado "lenguaje de marcado" o "lenguaje de descripción de documentos", con los que se construye un conjunto de reglas que definen todo lo que es parte de un documento digital, pero que no pertenece al texto del mismo. Los lenguajes de marcas no son lenguajes de formato similares a los lenguajes que se usan en la Internet, como los de descripción de páginas (archivos PostScript, archivos PDF, etc.), ni son lenguajes de programación (Java, Perl, C++...), sino que se trata de lenguajes orientados a definir la estructura y la semántica de un documento. María Jesús Lamarca Lapuente. "Lenguajes hipertextuales". En *Hipertexto: El nuevo concepto de documento en la cultura de la imagen*. Disponible en <http://www.hipertexto.info/documentos/lenguajes_h.htm>, consultado el 16 de mayo de 2008.

¹³ Ontology Web Language, un lenguaje de marcado para publicar y compartir datos, utilizando ontologías en la www. El OWL tiene como objetivo facilitar un modelo de marcado construido sobre RDF y codificado en XML, en <<http://es.wikipedia.org/wiki/OWL>>, consultado el 24 de agosto de 2009.

Por otro lado, las tecnologías de la web semántica desarrollan una red en la que sea más fácil localizar, compartir e integrar información y servicios para sacar mayor partido a los recursos disponibles en la web.

La web semántica rescata la noción de ontología del campo de la inteligencia artificial, como vehículo para cumplir este objetivo. La idea es que la web semántica esté formada por una red de nodos tipificados e interconectados mediante clases y relaciones definidas por una ontología compartida por sus distintos autores.¹⁴

Para comprender mejor cómo se estructura la web semántica, observemos la siguiente figura.¹⁵



¹⁴ CASTELLS, Pablo. La web semántica.

<<http://arantxa.ii.uam.es/~castells/publications/castells-uclm03.pdf>> [Consultado el 26/mayo/2008].

¹⁵ La gráfica fue tomada de Keylin Rodríguez Perojo y Rodrigo Ronda León. "Web semántica: un nuevo enfoque para la organización y recuperación de información en el web", en <http://bvs.sld.cu/revistas/aci/vol13_6_05/aci030605.htm>, [Consultado el 30 de mayo de 2008].

Para explotar la web semántica, se necesitan lenguajes semánticos más potentes, es decir, lenguajes de marcado capaces de representar el conocimiento, basándose en el uso de metadatos y ontologías. Dichos lenguajes tiene que ser estandarizados y formalizados para que su uso sea universal.

Al igual que las bibliotecas, la web semántica propone introducir descripciones de los recursos digitales mediante representaciones procesales o entendibles, para lograr que las máquinas utilicen lo que la web contiene. Esta web está integrada por agentes (software) capaces de navegar y realizar operaciones para ahorrarnos trabajo y optimizar resultados.

Los lenguajes semánticos que se señalan en la web semántica son los mismos que se utilizan en la biblioteca digital y en los repositorios, conocidos como metadatos, pues, como ya se señaló antes, son la herramienta que mejor permite representar la información digital, con la finalidad de localizarla y recuperarla fácilmente.

3.1.5 Las redes académicas avanzadas

La investigación académica sigue siendo el motivo más sólido que justifica el gasto y el esfuerzo invertidos en la red de redes. En la actualidad existen redes especializadas a través de las cuales se consigue que la creación del conocimiento se evalúe y discuta. En las redes académicas se incorporan conocimientos en la red para que los conozcan otros colegas. Para un académico, publicar sus conclusiones en una revista especializada requiere de un largo proceso de preparación, sin embargo, no hay trámite alguno para sumergirse en un foro de intercambio de colegas, como las redes académicas.

Una red académica o de conocimiento es un grupo de personas de diferentes instituciones que se asocian para investigar o desarrollar proyectos, para lo cual se apoyan en la información que aportan. La finalidad de estas redes es mejorar la calidad del trabajo científico, posibilitar que el flujo de la

información sea libre entre los grupos y fortalecer la cooperación y el aprovechamiento de los recursos.

Desde los noventa, en el mundo se desarrollan las redes académicas de vanguardia de alta velocidad, cuyo principal objetivo es facilitar el avance de las tecnologías y de nuevas aplicaciones de la Internet2.

Las redes académicas (también denominadas redes de investigación) permiten, a quienes trabajan en la investigación, colaborar y compartir información a través de las redes interconectadas, que posibilitan la transferencia de datos y aplicaciones derivadas de sus investigaciones.

Estas redes permiten que los investigadores dispersos geográficamente, puedan comunicarse con otros, es decir, sin las redes de alta velocidad, los proyectos de investigación no se darían a conocer. Entre las ventajas de las redes académicas avanzadas sobresale que son de uso exclusivamente académico, lo que favorece a los investigadores, académicos y estudiantes; que tienen calidad estable durante todo el día, es decir, no hay que acceder a éstas en horas con menos tráfico para transferir archivos o visualizar videos, asimismo, que están a la vanguardia en el desarrollo de protocolos y servicios del futuro.

En nuestro continente, las redes académicas de Argentina, México, Brasil y Chile ya se han integrado a la Internet2. A partir del 2003, en Uruguay se establece la Cooperación Latinoamericana de Redes Avanzadas (Clara), entre cuyos objetivos de dicha asociación civil destacan: 1) la coordinación entre las redes académicas nacionales de América Latina y con otros bloques; 2) la cooperación para la promoción del desarrollo científico y tecnológico, y 3) la planificación y aplicación de servicios de redes para la interconexión regional.

Se ha pensado en la Internet2 como base del desarrollo de las redes académicas avanzadas, porque el uso dado a la Internet hoy en día es para fines comerciales, más que académicos. Esto la vuelve inapropiada para los propósitos de investigación y experimentación.

Las redes avanzadas no reemplazan, por ahora, a la Internet actual. Su meta es dotar a las instituciones académicas nacionales, regionales e internacionales con los recursos necesarios para desarrollar nuevas tecnologías y aplicaciones.

En México, existe la llamada red denominada Corporación Universitaria para el Desarrollo de la Internet (CUDI), asociación civil sin fines de lucro que gestiona la Red Nacional de Educación e Investigación para promover el desarrollo de nuestro país y aumentar la sinergia entre sus integrantes, con el adelanto de una red de alta velocidad unida con la Internet2, la cual pretende ofrecer a la comunidad científica y universitaria de México una red de telecomunicaciones que le permita crear una nueva generación de investigadores, dotándolos de mejores herramientas que permitan incrementar las aplicaciones científicas y educativas de alta tecnología a nivel mundial.

La visión de la CUDI hacia el año 2030 es ser la Red Nacional para la Educación e Investigación líder mundial en el aprovechamiento, innovación, investigación, servicios y aplicaciones basadas en las tecnologías de la información y la comunicación.¹⁶

Esta red académica avanzada más conocida es la realizada en Estados Unidos por la Corporación Universitaria para el Desarrollo del Internet Avanzado (UCAID, por sus siglas en inglés), institución creada en 1997, que reúne a la academia y a empresas que desarrollan las tecnologías de información, con el fin de desarrollar la Internet2, es decir, una nueva Internet con fines educativos y de investigación, para lograr así un mejor intercambio de información en la sociedad del conocimiento.

El éxito tecnológico y académico de la Internet2 ha promovido el desarrollo de redes avanzadas en Argentina, Bolivia, Brasil, Colombia, Costa Rica, Cuba, Chile, Ecuador, El Salvador, Guatemala, Nicaragua, Panamá, Paraguay, Perú, Uruguay, Venezuela y México. Así surgió en América Latina el proyecto de Clara, para interconectarse regionalmente y compartir recursos y enlazarse, al

¹⁶ <<http://www.cudi.edu.mx/> [Consultado el 28-marzo-2009].

mismo tiempo, con las grandes redes de alta velocidad de todo el mundo, como Canarie (Canadá), Apan (Asia) y Dante (Europa).

Los cinco principales nodos IP de Clara se ubican en São Paulo, Buenos Aires, Santiago de Chile), la ciudad de Panamá y en Tijuana. Desde Brasil, la red Clara enlaza a las redes nacionales de América Latina a Géant2, la red avanzada paneuropea, y a Whren-Lila, en Estados Unidos; desde México, se establece el enlace directo a la Internet2.

Las principales redes académicas avanzadas del continente son las siguientes:

- Argentina, Red TeleInformática Académica (Retina) (www.retina.ar).
- Bolivia, Agencia para el Desarrollo de la Sociedad de la Información en Bolivia (ADSIB) (www.bolnet.bo).
- Brasil, Red Nacional de Enseñanza e Investigación (RNP) (www.rnp.br).
- Colombia, Red Nacional Académica de Tecnología Avanzada (Renata) (www.renata.edu.co).
- Costa Rica, Red Nacional de Investigación (CR2Net) (www.crnet.cr/cr2net).
- Cuba, RedUniv (www.mes.edu.cu).
- Chile, Red Universitaria Nacional (Reuna) (www.reuna.cl).
- Ecuador, Consorcio Ecuatoriano para el Desarrollo de Internet Avanzado (CEDIA) (www.cedia.org.ec).
- El Salvador, Red Avanzada de Investigación, Ciencia y Educación Salvadoreña (RAICES) (www.raices.org.sv).
- Guatemala, Red Avanzada Guatemalteca para la Investigación y Educación (RAGIE) (www.ragie.org.gt).
- Honduras, Universidad Tecnológica Centroamericana (Unitec) (www.unitec.edu);
- México, Corporación Universitaria para el Desarrollo de Internet (CUDI) (www.cudi.edu.mx).
- Nicaragua, Red Nicaragüense de Internet Avanzada (Renia) (www.renia.net.ni).
- Panamá, Red Científica y Tecnológica (Red CyT) (www.redcyt.org.pa).
- Paraguay, Arandu (www.arandu.net.py).
- Perú, Red Académica Peruana (RAAP) (www.raap.org.pe).
- Uruguay, Red Académica Uruguay (RAU) (www.rau.edu.uy/redavanzada/).
- Venezuela, Red Académica de Centros DE Investigación y Universidades Nacionales (Reacciun) (www.reacciun2.edu.ve).

La Red Hemisférica Interuniversitaria de Información Científica y Tecnológica (Red HUCyT) Regional, de la OEA (<http://www.redhucyt.oas.org/webesp>) tiene como objetivos conectar a las instituciones de los Estados miembros a la Internet para crear una red interoperante para el intercambio de información científica y tecnológica. Relaciona bibliotecas y unidades de información, indicadores, investigadores, listas de interés y promociona posgrados, maestrías, doctorados, proyectos, además de posibilitar el acceso a diferentes publicaciones. La Red HUCyT la integran México y América Central (Costa Rica, El Salvador, Guatemala, Honduras, Nicaragua, Panamá).

Las redes académicas o redes de investigación han sido posibles gracias a la rápida evolución de las tecnologías de telecomunicación y principalmente de las de comunicación de datos.

3.1.6. La biblioteca híbrida, Information Commons y Centro de Recursos para el Aprendizaje y la Investigación (CRAI)

La biblioteca híbrida

Ésta se concibió a finales de los años noventa como la integración de recursos electrónicos, con las funciones tradicionales de la biblioteca, es decir, delimita dos espacios: por un lado, a la biblioteca como lugar y, por el otro, al espacio o entorno de la información digital. Moreiro señala como características de la biblioteca híbrida las siguientes:

- Sus servicios se sustentan en recursos de información digitales distribuidos por la red.
- La interactividad con el usuario suele establecerse desde el teclado (a distancia).
- Se dirige el acceso a los recursos en red.

–Se manifiestan nuevos modelos organizativos y nuevas funciones distribuidas que se fundamentan en la cooperación.¹⁷

La Biblioteca Híbrida se diseñó para armonizar todo tipo de tecnologías y debe integrar el acceso a cuatro tipos de recursos usando tecnologías digitales:

1. *Recursos tradicionales*: cualesquier recurso no digital, no sólo impreso, sino también documentos que se encuentren en soporte magnético.
2. *Recursos de transición*: son los recursos que han sido digitalizados.
3. *Nuevos recursos*: se refiere a los recursos informativos que han sido creados para el entorno digital.
4. *Recursos futuros*: recursos digitales normalizados en los que los datos y los métodos de acceso forman una unidad.

De acuerdo con lo anterior, podemos señalar que la biblioteca híbrida existe desde el momento en que se incorporan las tecnologías digitales a la biblioteca tradicional.

Information Commons

Como mencionamos antes, una tendencia que se observa es la dirección hacia la conjunción de todo tipo de materiales, sean impresos o no. Un ejemplo es el Information Commons (IC), relativo a un espacio físico, generalmente en una biblioteca académica, donde todos los miembros de la comunidad participan en los procesos de investigación, recolección y producción de la información. El propósito del IC es mejorar los servicios, con miras a ofrecer un ambiente útil que apoye los trabajos de los usuarios.

Para ejemplificar esto, comento el caso de la biblioteca de la Universidad de Alberta, en California, la cual ofrece a sus estudiantes un lugar para estudiar, trabajar y colaborar. Este espacio se denomina Knowledge Commons (KC), y cuenta con 140 computadoras, impresoras y software para cubrir todas las necesidades del estudiante; además, cuenta con conexiones inalámbricas a la

¹⁷ GONZALEZ, Moreiro A. “La distribución de los contenidos en la nueva sociedad informacional”, en Sebastián M. Caridad (coord.), *La sociedad de la información: política, tecnología e industria de los contenidos*. p. 315.

Internet sin restricciones para las computadoras portátiles y áreas más grandes. También ofrece cubículos para trabajar en grupo, que tienen una computadora y, en algunos casos, una pantalla, así como conexión inalámbrica para computadoras portátiles.

En lo que se refiere a los recursos que integran las colecciones dentro del IC, el movimiento Open Access ha despertado gran interés en los investigadores, a través de los llamados repositorios o depósitos institucionales desarrollados gracias a este movimiento, ya que facilitan el acceso a las colecciones digitales, es decir, la información contenida en éstos es útil y de fácil acceso, además de que cuentan con varias entradas para localizar la información que requieren los estudiantes e investigadores de todos los niveles.

El CRAI

Éste se concibe como un centro dinamizador del nuevo aprendizaje dentro de una universidad. Es un nuevo diseño de espacios y servicios que reúnen elementos ubicados fuera de la biblioteca; espacios no sólo con mobiliario para libros, sino también para estaciones de trabajo que aseguren el acceso a la información digital. Entre los objetivos del CRAI destacan:

- 1) Facilitar a los usuarios el acceso a toda la información y documentación que necesite de la universidad y fuera de ésta, fácil, rápida y organizadamente.
- 2) Programa el crecimiento de las diversas colecciones bibliográficas e integra otros materiales (impresos y en formato electrónico y digital).
- 3) Integra los servicios de la universidad relacionados con el aprendizaje.
- 4) Dispone de diversos servicios generales, pero también personalizados, según las necesidades de los usuarios.

El CRAI también ofrece los siguientes servicios bibliotecarios para la investigación:

- De información y referencia especializada.
- De consulta a bases de datos y e-revistas.

- De obtención de documentos externos.
- De búsqueda documental en bases de datos de pago.
- De formación de usuarios en el uso de herramientas electrónicas de acceso a la información.
- De colecciones especializadas.

Aunque el CRAI también ofrece servicios bibliotecarios digitales, destinados a los usuarios virtuales, ya sean servicios

- De acceso a la biblioteca digital y a los repositorios institucionales.
- De acceso a los productos multimedia.
- Web y acceso a la Internet.
- De préstamo en línea.
- De adquisición y pedidos de libros en línea.
- De formación en búsquedas avanzadas en las colecciones digitales y en la Internet.
- De soporte documental y bibliográfico en línea.

Como conclusión de este capítulo, conviene subrayar que las tendencias de la web se enfilan hacia la colaboración, participación e intercambio de conocimientos e información por parte de los usuarios, por ejemplo, tenemos las redes académicas, sitios como Youtube, Amazon, las wikipedias y del.icio.us, en los que los usuarios agregan sus comentarios, fotos, etiquetas (*tags*) de las páginas web para compartir o recuperar información.

Además, existe una gran cantidad de sistemas para organizar la información contenida en la web, como los metadatos, las ontologías, folksonomías y otras. En la actualidad, el sistema que se utiliza en la web son las folksonomías, por ser un sistema de clasificación en el cual los usuarios asignan sus propias etiquetas (o *tags*) en varios sitios de la red para recuperar información. En esos sitios se dan a conocer las posiciones de los *tags* asignados para conocer cuáles son los de mayor número de asignaciones por los usuarios y cuáles las de menor número.

Entre los sitios que recurren a esta modalidad, encontramos a *del.icio.us*, que es el más conocido, pero también Elsevier, en su sistema de recuperación de información en revistas, cuenta con el sistema de clasificación basado en folksonomías, en el que se guarda la información cada vez que los usuarios teclean un término en sus búsquedas, de tal manera que después cualquier usuario puede consultarlo, sólo que en este sistema el término que aparece con letra de tamaño grande es el que ha sido asignado más veces por el usuario al realizar sus búsquedas y puede ser escogido para realizar otras búsquedas.

En la Biblioteca 2.0 ocurre lo mismo: el usuario es más participativo y, por ende, se intenta incluirlo en el diseño e implementación de los servicios bibliotecarios. Los usuarios deciden, de acuerdo a sus necesidades, participar en las tomas de decisiones, para lo cual se aprovechan las herramientas contenidas en la web 2.0.

Una de esas herramientas es el software libre que se encuentra en la web para la creación de repositorios, los cuales son manejados por bibliotecarios, ingenieros y usuarios, es decir, la participación del usuario radica en que puede opinar sobre la interfaz que se le presenta a la hora de recuperar la información. La mayoría del software libre tiene la característica de modificar su interfaz.

A lo largo de este capítulo, observamos que, en los últimos años, la web ha traído consigo una serie de cambios: se han creado herramientas que permiten la participación del usuario, lo cual ha cambiado las formas de trabajo, por ejemplo, existen los llamados weblogs, las enciclopedias creadas por los lectores llamadas wikis (wikipedia), sitios de acceso abierto que son creados y mantenidos por los usuarios, asimismo sitios en los que, además de compartir archivos de audio y video, los usuarios etiquetan esos contenidos, como es el caso de YouTube.

En suma, la web se ha convertido en una web social, pues permite que todos los usuarios participen de una u otra manera en su desarrollo. Este nuevo paradigma también ha cambiado la forma de trabajo en las bibliotecas, ya que, como se señaló anteriormente, la aplicación de las herramientas de la web 2.0 a las bibliotecas ha permitido que los usuarios también sean más participativos, como pasa en la llamada biblioteca 2.0, con la que se pretende que los usuarios participen y colaboren, es decir, que trabajen conjuntamente bibliotecarios y su comunidad de usuarios en la creación de nuevos servicios.

A continuación se presentan algunas alternativas que serían las más completas y fáciles de incorporar a nuestro quehacer cotidiano como bibliotecarios.

3.2. Alternativas para la organización documental en la sociedad de la información

Lo hasta aquí expuesto permite llevar a cabo una serie de reflexiones que delimiten algunas alternativas para la organización documental en la sociedad de la información. Como se recordará, en el primer capítulo se señaló que, en la sociedad de la información, las computadoras y las TIC son los principales recursos de la economía, ya que ponen a disposición de todos la información. Mas no puede generalizarse cuando existen personas que no tienen acceso a una computadora y porque la información ofrecida también tiene un costo que mucha gente no puede pagar, aun cuando pueda tener acceso a ésta. De hecho, la brecha digital es una consecuencia de esta inequidad en el acceso a la información, sea por motivos tecnológicos, económicos, o de cualquier otra índole.

Por ello, cuando se habla de la llamada sociedad de la información, se tiene que aludir la brecha digital, pues ésta refleja la diferencia entre quienes sí tienen acceso a la información y los que no. Aquí cabe aclarar que, hasta la fecha, la brecha digital se mide de acuerdo a parámetros de acceso a las TIC, ya que por lo general se contabiliza cuántas computadoras personales hay por

habitante, o bien, cuántos teléfonos celulares hay por habitante. De ahí que sea muy difícil determinar cuál es el nivel de acceso a la información.

En la sociedad de la información se prioriza la importancia de la información digital, siendo que la impresa no dejará de existir y, por lo tanto, también representa un bien para la sociedad. En este sentido, la importancia que se le da a lo digital hace parecer que lo impreso no constituye parte de esta sociedad, y que hay que dar un trato aparte a los documentos digitales. Pero resulta necesario y prioritario integrar la información impresa y la digital en un universo de documentos útiles en el contexto de la sociedad de la información.

Ahora bien, si se analiza qué pasa en el medio digital, se advierte claramente que la facilidad en el acceso no es del todo cierta. Además de la brecha digital mencionada, se carece también de documentos en texto completo que se utilicen a partir de su identificación de los catálogos en línea.

Existen muchas razones para que esto no sea posible: desde el problema de los derechos de autor, hasta el poco número de acciones que conduzcan a la creación de repositorios institucionales relacionados con proyectos de bibliotecas digitales. En la mayoría de los proyectos de biblioteca digital (BD), no encontramos repositorios digitales, pues éstos se consideran proyectos aparte, debido a que no se establece la relación entre ambos, ni se contempla la BD como un elemento a través del cual se organizan y recuperan las colecciones digitales que conforman los repositorios.

La tendencia actual para la administración de colecciones distribuidas (de los depósitos institucionales, es decir, los repositorios) y bibliotecas digitales, es la utilización del software libre, ya que las bibliotecas invierten menos recursos al llevar a cabo sus proyectos digitales, gracias a la flexibilidad que presentan los sistemas y la posibilidad de poner en marcha programas de intercambio de información y de acceso a otros sistemas. Además, el software libre cuenta con la característica de que se basa en el estándar de metadatos Dublin Core para la descripción o procesamiento de los documentos, lo que significa que siguen ciertos estándares conocidos y muchas veces usados por los bibliotecarios.

Otra ventaja que ofrece el software libre es que se puede trabajar con todo tipo de formatos, trátase de texto, imágenes (JPEG, GIF, BMP, PNG, TIFF y otros), videos (MPEG, Quicktime), bases de datos, lenguaje XML, HTML. Debido a las ventajas que ofrece, los bibliotecarios debemos promover el uso del software libre.

Como hemos observado, la mayoría de las bibliotecas y repositorios digitales han optado por el uso de los metadatos para ordenar sus documentos, pero también algunos se han inclinado por el uso de los sistemas para la organización del conocimiento (SOC), debido a que se enfocan a todos los tipos de proyectos que requieren organizar información, aparte de que incluyen diversos esquemas de clasificación, encabezamientos de materia, listas de términos y de relaciones que proveen vocabularios controlados, por lo que el uso de los SOC también son una buena opción a la hora de crear nuestra biblioteca digital o repositorio.

Hoy en día, con el surgimiento de la llamada web 2.0, en la que los usuarios son cada vez más participativos, surge algo novedoso llamado folksonomía, un tipo de sistema de clasificación distribuida, creada por los usuarios mismos, sistema que se describe como un tipo de metadatos y como un medio suplementario para organizar la información y, a la vez, los resultados obtenidos de las búsquedas.

En la llamada folksonomía, los usuarios agregan los tags o etiquetas a los recursos digitales que, como ya se dijo antes, pueden ser textos, imágenes, videos u otro tipo de información. Estas etiquetas se denominarían como palabras clave, nombres o metadatos, y al proceso de asignación de *tags* se le conoce como *tagging*. Cabe señalar que, por muy novedoso que parezca, tiene algunas desventajas que, desde mi personal punto de vista, son importantes indicar cuando se organizan las colecciones digitales, algunas de éstas serían: a) que las etiquetas (o *tags*) son creadas por los usuarios, no por bibliotecarios o especialistas de la información; b) no existen reglas o lineamientos que delimiten la asignación de esos *tags*; c) no se tiene un control en los términos

asignados, lo que redundará en que las etiquetas son inexactas, inadecuadas, imprecisas y ambiguas, debido a su sinonimia y homonimia.

Resultante de lo anterior, las folksonomías no son por el momento una buena alternativa para la organización documental, ya que se orientan a ordenar las colecciones personales, principalmente, mas no las colectivas; además, como previamente se dijo, no cuentan con estándares o lineamientos para la asignación de *tags*, por lo que los administradores de los sistemas tienen como resultado un caos en los términos por la falta de control en su asignación.

En la actualidad existen algunas tendencias en cuanto a las formas en que se recompondrán los patrones de la organización documental, una de las cuales es la conjunción de documentos digitales e impresos, como recursos de toda biblioteca que participen en el escenario de la sociedad de la información.

Ofrecer materiales de todo tipo es algo deseable en la biblioteca. Para ello se requiere mantener un catálogo que relacione los universos de lo impreso y de lo digital, de tal forma que el usuario tenga la posibilidad de conocer qué existe en la temática de su interés y qué requiere. Pero también resulta importante subrayar que, aparte de los registros, es imprescindible dar el documento, porque de lo contrario la biblioteca será una especie de base de datos gigantesca, pero carente o vacía de documentos. Por ello, el repositorio se presenta como una de las alternativas en las que se conjugan los materiales impresos y digitales.

Cabe decir que los repositorios o depósitos digitales se diseñan con software libre en el que se localizan libros y revistas científicas en texto completo, pero también se tiene acceso a los recursos comunes de la biblioteca, como los catálogos, bases de datos y material impreso.

Como se mencionó en líneas precedentes, la sociedad de la información no está limitada al uso de Internet, aunque ésta sea un medio a través del cual se tiene acceso e intercambio de la información, por lo que un claro ejemplo de este factor es la aparición de la propuesta denominada Information Commons, referente a la mejoría de los servicios que ofrece la biblioteca tradicional y combinarlos con las nuevas tecnologías, con miras a ofrecer un ambiente útil en

el que los usuarios puedan acceder, manejar y producir información de toda naturaleza en un mismo sitio de trabajo.

De acuerdo con estas consideraciones, sería plausible proponer algunas alternativas para establecer una organización documental dentro del escenario actual de la sociedad de la información:

1. Es deseable que en la sociedad de la información el desarrollo de bibliotecas que respondan al ideal del IC, a través de las cuáles los usuarios localicen y consulten información en formato impreso o digital, respondiendo así al ideal de acceso a la información que se maneja en este entorno social. El IC también corresponde con la tendencia del trabajo en colaboración que cada vez se observa con mayor vigor y presencia en el ámbito académico.
2. En el seno del IC, es posible desarrollar repositorios que muestren a la comunidad académica la producción de la institución. En la creación de estos repositorios, es importante el uso de software libre como alternativa ante los elevados costos de los programas, así como de las opciones que ofrecen para la creación de colecciones digitales distribuidas.
3. Conviene tomar en cuenta el elemento “colaboración” emanado de la idea de la web y de la biblioteca 2.0, pues la red se ha convertido en un fenómeno social, ya que la sociedad decide cómo funciona. La colaboración entendida en la creación de contenidos, pero también en la propuesta –al bibliotecólogo– de los temas y relaciones temáticas que reflejen estos contenidos. Como dijimos antes, las etiquetas no son lo ideal, pero si se ve desde la óptica de la colaboración con el bibliotecólogo, sería un medio de enriquecimiento y retroalimentación, pues el usuario participaría en las decisiones del diseño y mantenimiento de los servicios y actividades de la biblioteca, según sean sus necesidades; asimismo los servicios se actualizarían constantemente para servir mejor a su comunidad. Además, los usuarios, al crear contenidos, registran su conocimiento y generan nueva información. Esto permitiría aproximarse a la personalización, manifiesta como elemento esencial en los CRAI.
4. Las nociones de colaboración y de colecciones digitales distribuidas están estrechamente relacionadas con la idea de redes académicas avanzadas, que son de uso exclusivamente académico. En este ambiente, será posible establecer tanto parámetros para la publicación y generación de repositorios, como para un ordenamiento adecuado a las áreas de interés de investigadores, profesores y alumnos de posgrado de las instituciones participantes. Esto tiene correspondencia con la web semántica, formada por una red de nodos tipificados e interconectados por medio de clases y

- relaciones definidas por una ontología compartida por sus distintos autores.
5. La organización documental en la sociedad de la información requerirá de herramientas (como metadatos, folksonomías, ontologías y otras) para el ordenamiento de documentos digitales (que conformen las colecciones digitales distribuidas), pero también de documentos electrónicos y digitales, así como de materiales docentes elaborados por profesores y estudiantes de la universidad que formen parte de los IC. Algunas alternativas serían la creación de listas de términos que complementen las búsquedas de documentos o los tesauros, que nos permiten organizar los documentos digitales, pues funcionan como un soporte de la recuperación de la información, basado en el conocimiento que facilitará el acceso a los diversos contenidos. Mediante estos sistemas (SOC) se organiza la información y, al mismo tiempo, se promueve el uso del conocimiento, ya que cuentan con esquemas de clasificación para organizar los materiales a nivel general, encabezamientos de materia y archivos de autoridad, por lo que se considerarían una herramienta completa para la organización documental.
 6. Los metadatos se consideran como la mejor opción para trabajar con colecciones digitales, pues permiten describir los objetos digitales, indizar grandes cantidades de información, así como crear relaciones y establecer enlaces, por lo que se conciben como una herramienta completa muy usual en la actualidad en todos los proyectos de bibliotecas y repositorios digitales, porque permiten representar a la información digital con la finalidad de localizarla y recuperarla fácilmente.

Reflexiones finales

De acuerdo con lo expuesto a través de esta investigación, se desprenden las siguientes consideraciones:

- En la llamada sociedad de la información, el elemento clave para tener éxito es el acceso a la información que en ella se genera, por lo que para que todos los individuos puedan lograr esto se deben desarrollar colecciones digitales para un grupo determinado de usuarios, pero también ha de darse importancia a la información impresa, convergiendo en un mismo universo de información.
- Las bibliotecas digitales y los repositorios deben orientarse principalmente a satisfacer las necesidades de información de *sus usuarios*, además de difundir y preservar los documentos de cada institución.
- El uso de las tecnologías de la información en las bibliotecas no sólo es para consultar los catálogos y bases de datos, sino para buscar información y poder crearla al mismo tiempo y en el mismo espacio.
- Las tendencias de las bibliotecas, que corresponden con una idea más completa de la sociedad de la información, son el Information Commons y el CRAI, en las que se busca que los documentos coexistan en una organización documental que permita identificarlos en un universo común. Esto nos daría una idea de todos los documentos que conforman las colecciones, con la ventaja de que se podrá elegir el material que más satisfaga las necesidades de información, independientemente del formato en que ésta se encuentre.

- Los bibliotecarios debemos defender el principio de equidad del acceso, promover el acceso libre a la información y al conocimiento y promover el uso del software libre por las ventajas que éste nos ofrece.

Para poder lograr una organización documental en la sociedad actual debe establecerse una relación estrecha entre materiales digitales e impresos, la cual puede conseguirse a través de proyectos de bibliotecas y repositorios digitales, con el uso de formatos digitales, como es el estándar de metadatos, así como la aplicación de algunos sistemas para organizar colecciones digitales (SOC), por ser las mejores opciones con la finalidad de representar, administrar y recuperar a las colecciones digitales.

Podemos señalar que la importancia de los formatos bibliográficos y digitales en la organización documental radica en el hecho de que permiten identificar, describir, representar y localizar los documentos digitales, además de que permiten el intercambio de información. Por ello, es importante señalar que el papel del bibliotecario sigue siendo primordial a la hora de administrar y crear colecciones, ya sea en formato impreso digital. Las herramientas que siempre ha utilizado para organizarlas siguen siendo indispensables en la era digital.

Como puede observarse en el desarrollo de este trabajo, los metadatos son la herramienta más completa pues permiten describir (o identificar los recursos como son título, autor, resumen, palabras clave, etc.) administrar (o facilitar la gestión y el proceso de colecciones digitales como son el idioma, la fecha, la creación, etc.) y además facilitan la navegación y presentación de los recursos. Por ello, primero fueron utilizados en los proyectos de bibliotecas digitales ya que permiten la localización y uso de cualquier documento, y hoy en día, son utilizados en la creación de repositorios. El estándar más utilizado es el estándar Dublin Core que se encuentra compuesto por 15 elementos simples y en donde cada uno de ellos es repetible y puede aparecer en cualquier orden,

además de que son fáciles de usar, para que cualquier persona que introduzca recursos en Internet los pueda describir con facilidad.

En este sentido, la hipótesis planteada en este estudio se pudo confirmar, debido a que los metadatos son una herramienta primordial que apoya y permite lograr la organización documental, ya que apoyan la administración y la creación de nuevas colecciones digitales, además de que permiten trabajar, describir, localizar e intercambiar los diversos tipos de formatos que se encuentran en las colecciones digitales y que pueden ser desde un archivo de texto, un PDF, una imagen, etc.

En relación con los objetivos planteados en esta investigación se analizaron los diversos formatos digitales demostrando que el uso de los metadatos es importante porque permiten realizar una descripción del contenido de las colecciones digitales siguiendo normas establecidas, lo que permite el intercambio y recuperación de la información en el espacio digital.

Debido a los avances que presentan las tecnologías de la información, surgirán nuevos formatos digitales, sistemas para la organización del conocimiento, así como software libre, que deberán estudiarse para conocer cuáles son sus características y beneficios que puedan brindar a la organización documental.

REFERENCIAS

American Digital Library Federation.

<http://www.diglib.org/about/dldefinition.htm>

BEACOM, Matthew. "Crossing a digital divide: AACR2 and Unaddressed problems of networked resources." Presentado durante la conferencia *Bibliographic Control for the New Millenium*, llevada a cabo en la Biblioteca del Congreso, Washington, D.C. Noviembre 2000.

BURNARD, Lou; Light, Richard. "Three SGML Metadata Formats : TEI, EAD and CIMI." En: Work Package 1 of Telematics for libraries project BIBLINK, UKOLN Metadata Group, diciembre 1996

<http://ukoln.ac.uk/metadata/BIBLINK/wp1/sgml>

CASTELLS, Pablo. La web semántica.

<http://arantxa.ii.uam.es/~castells/publications/castells-uclm03.pdf>

CENTELLES, Miguel. Taxonomías para la categorización y la organización de la información en sitios web." *Hipertext.net*", No.3, 2005. p.5

<http://www.hipertext.net>

Corporación Universitaria para el desarrollo de Internet, A.C.

<http://www.cudi.edu.mx/>

CHOMSKY, Noam, Dietrich Heinz. *La sociedad global: educación, mercado y democracia*. 2ª ed. México: Editorial Planeta Mexicana, 2004. 198 p.

DÍAZ Ortuño, Pedro Manuel. "Problemática y tendencias en la arquitectura de metadatos web." En: *Anales de documentación*. No.6, 2003. p. 35-58.

DICCIONARIO DE LA LENGUA ESPAÑOLA. 22ª ed. Disponible en:

http://buscon.rae.es/draeI/SrvltConsulta?TIPO_BUS=3&LEMA=tendencia

Digital Library Project. University of California Berkeley.

<http://elib.cs.Berkeley.edu> "Digital Library Using Next Generation Internet" *IEEE Communications Magazine* 37(1) (January 1999).

Diseño y arquitectura. En: <http://tramullas.com/ai/productos/htm>

ESTUDILLO García, Joel. "Elementos que conforman la sociedad de la información" En: *Investigación Bibliotecológica*, Vol .15, No.3, 2001. p. 163-194.

Extensible Markup Language. Disponible en wikipedia:
http://es.wikipedia.org/wiki/XML#Ventajas_del_XML

FERNÁNDEZ Aballi, Isidro. Ponencia magistral. 22 de noviembre del 2000.
<http://www.uco.mx/interfaces/interfaces2001/ponencias/ponencias/Pm01.pdf>

Folcsonomías y tags: como lograr un tagging efectivo.
http://www.masternewmedia.org/es/2006/02/02/folcsonomias_y_tags_como_lograr.htm

FOLDOC. Free Online Dictionary of Computing. Disponible en:
<http://foldoc.doc.ic.ac.uk/foldoc/index.html>

GARCÍA Camarero, Ernesto. La biblioteca digital. Madrid. Arco Libros, c2001, 380 p.

GARCÍA Camarero, Ernesto; García Menero, Luis Ángel. "La biblioteca digital". Madrid : Arco/Libros, 2001. 380 p.

GARZA Salazar, David; Sordia Salinas, Martha; Martínez Treviño, Yolanda. "Biblioteca digital Phronesis". 4 p.
<http://www.mty.itesm.mx/dgi/transferecia/Transparencia45/eli-03.htm>

GILLILAND-Swetlad, Anne J. "La definición de los metadatos". En: Introducción a los metadatos: vías de la información digital. Los Angeles: J. Paul Getty Trust, 1999. p. 1-8.

GÓMEZ Alamillo, Francisco. "Red Iberoamericana de ciudades digitales." En: Jornada de ciudades digitales, Madrid 4 de abril 2003.
www.iberomunicipios.org

GONZALEZ, Moreiro A. "La distribución de los contenidos en la nueva sociedad informacional", en Sebastián M. Caridad (coord.), *La sociedad de la información: política, tecnología e industria de los contenidos*. Madrid: Centro de Estudios "Ramón Areces", Universidad Carlos III, 1999, pp. 312-316.

GORMAN, Michael. "Metadata : hype and glory." En: Cataloging the web : metadata, AACR and MARC21. Lanham, Mayland : Scarecrow, 2002. 199 p.

GORMAN, Michael. "¿Metadatos o catalogación? Un cuestionamiento erróneo". En: Internet, metadatos y acceso a la información en biblioteca y redes en la era electrónica. comps. Filiberto Felipe Martínez Arellano, Lina Escalona Ríos ; tr. de los documentos en inglés Filiberto F. Martínez. México: UNAM, Centro Universitario de Investigaciones Bibliotecológicas, 2000. 112 p.

GROSS, Ben. "Digital library Information and Resources"

<http://interspace.grainger.uiuc.edu/~bgross/digital-libraries>

GUY, Marieke; Tonkin, Emma. Folksonomía y tags : como lograr una tagging efectivo. En:

http://www.masternewmedia.org/es/2006/02/02/folcsonomias_y_tags_como_lograr.htm

HABIB, MICHAEL C. Defining acadmic library 2.0: lessons learned from web 2.0 Consultado en:

<http://www.unc.edu/~mchabib/inls342/DefiningAcademicLibrary20Presentation.pdf>

IMPACT. Líneas de actuación en el mercado de la información. Programa de trabajo de 1993. Luxemburgo: IMPACT Central Office, DG.XIII, 1994. p.13-41.

JUÁREZ Santamaría, Beatriz; Martínez Ortega, Martha Patricia. "Uso de los metadatos en el orden documental". En: XXXII Jornadas Mexicanas de Biblioteconomía 2001.

KEYLIN Rodríguez Perojo; Rodrigo Ronda León. Web semántica: un nuevo enfoque para la organización y recuperación de información en el web. En: http://bvs.sld.cu/revistas/aci/vol13_6_05/aci030605.htm [Consultado el 30-mayo-2008]

LAFUENTE López, Ramiro; Garduño Vera, Roberto. Lenguajes de marcado de documentos digitales. México: UNAM, CUIB, 2001.

LAMARCA Lapuente, María Jesús. Hipertexto: El nuevo concepto de documento en la cultura de la imagen. Tesis doctoral. Facultad de Ciencias de la Información, Universidad Complutense de Madrid.

<http://www.hipertexto.info/>

MANESS, Jack M. "Library 2.0 theory : web 2.0 and its implication for libraries." En: Webology, 2006, June vol.3 no.2 Consultado en:

<http://webology.ir/2006/v3n2/a25.html>

MARGAIX Arnal, Dídac. "Conceptos de web 2.0 y biblioteca 2.0: origen, definiciones y retos para las bibliotecas actuales." En: El profesional de la información n v.16, n.2, marzo-abril, 207 p.101.

<http://www.oei.es/tic/kx5j65q110j51203.pdf>

MARTÍN, William J. The global information society. London : ASLIB : Gower, 1997, 233 p.

MASUDA, Y. La sociedad informatizada como sociedad postindustrial. Madrid: Tecnos, 1994. 197 p. ISBN:8430910379.

MÉNDEZ Rodríguez, Eva. Metadatos y recuperación de información. Gijón: Trea, 2002. 409 p.

MIRELES Cárdenas, Celia. "Del documento al objeto digital". En: Revista de Investigación Bibliotecológica, enero-junio, Vol.17, No.34, 2003. pp. 88-143.

MOREIRO González, J. A.: "La distribución de los contenidos en la nueva sociedad informacional" En: Caridad Sebastián, M. (Coord.) : *La sociedad de la información: política, tecnología e industria de los contenidos*. Madrid: Centro de Estudios Ramón Areces: Universidad de Carlos III, 1999, pp. 312-316.

PASTOR Sánchez, Juan A. "Exposición conceptual y aplicación de un modelo para la organización del conocimiento en Internet : el hipertexto documental." En: Encuentro de ISKO-España 1997: Getafe, España. Organización del conocimiento en sistemas de información y documentación 3. Zaragoza : F. J. García Marco, 1999. 286 p.

Plan Euskadi en la Sociedad de la Información. p.14

<http://www.euskadi.net/eeuskadi/datos/docs/plan>

Principios del Software Libre. http://doc.ubuntu-es.org/Software_Libre

PROTHMAN, Bill. "Meta Data: Managing Needles in the Proverbial Haystacks." En: IEEE Potentials , Vol. 19, No.1, February-march, 2000. (Documento PDF) p .20.

<http://www.ieee.org/membership/students/potentials/febmar2000/pdf/pages20.pdf>

RODRÍGUEZ, BLANCA; Santos de Paz, Lourdes. "Del documento digital a la biblioteca virtual". En: Scire. 8: 2 (jul.-dic. 2002) 43-52.

<http://ibersid.eu/ojs/index.php/scire/article/viewFile/1170/1152>

RODRÍGUEZ Bravo, Blanca. El documento: entre la tradición y la renovación. España: Ediciones Trea, 2002. 281 p.

RODRÍGUEZ Gallardo, Adolfo. Tecnologías de la Información en México 2001-2005. México: UNAM, Centro Universitario de Investigaciones Bibliotecológicas, 2005. 97 p.

SÁNCHEZ Espinoza, Ariel. Análisis de las interacciones entre las tecnologías de la información y la sociedad de la información del siglo XXI en México (Tesis de Maestría). México : El autor, UNAM, FFyL, 188p.

SCHATZ, Bruce y Hsinchun Chen. "Building Large-Scale Digital Libraries", *Computer. Monográfico: US Digital Library Initiative* (mayo de 1996), disponible en www.computer.org/computer/dli

SEOANE, Catuxa. Flexibilidad de las folksonomías. En: <http://eprints.rclis.org/11558/1/Anuario-ThinkEPI-2007-Catuxa-Flex-Folkson.pdf>.

SCHAMBER, L. (1996). What's a document?: Rethinking the concept in uneasy times. En: *Journal of the American Society for Information Science*. 47:9 (1996) 669-671.

Sistemas de organización del conocimiento. Disponible en: <http://es.geocities.com/ontologia2004/sko.htm>

TORRES Vargas, Araceli Georgina. La navegación y exploración en la información digital del World Wide Web. En: *La información en el inicio de la era electrónica*. México: UNAM, Centro Universitario de Investigaciones Bibliotecológicas, 1998. Vo.1, 297 p.

TORRES Vargas, Araceli. *Un modelo integral de biblioteca digital*. México, D. F. : UNAM, Centro Universitario de Investigaciones Bibliotecológicas, 2008. ix, 76 p. (Cuadernos de investigación ; 6)

VICKERY Brian. *On 'knowledge organization'*. 2008. Disponible en: <http://www.lucis.me.uk/knowlorg.htm>

VOUTSSÁS Márquez, Juan. *Bibliotecas y publicaciones digitales*. México: UNAM, Centro Universitario de Investigaciones Bibliotecológicas, 2006. 338 p.

VOUTSSÁS Márquez, Juan. "Un modelo de bibliotecas digitales para México". Tesis de doctorado. UNAM, Facultad de Filosofía y Letras, División de Estudios de Posgrado, 2005. 516 p.

WANG, Richard; Lee, Pipino, Leo y Strong, Diane. "Gestionar la información como producto." En: *Harvard Deusto Business Review*. No. 90 Noviembre-Diciembre, 1999.

ZALDÍVAR COLLAZO, Modesto. "Políticas y estrategias en la Sociedad de la Información". En: *Ciencias de la Información* Vol.29, No.3, Sept., 1988.