

José A. Morcero González

---

**LA INDUSTRIA DE LOS CONTENIDOS EN LA SOCIEDAD  
DE LA INFORMACIÓN: VENTAJAS E INCONVENIENTES  
PARA LA ACTUACIÓN DOCUMENTAL**



## INTRODUCCIÓN.

El profesor Martín Fuertes, cuya memoria académica queremos honrar hoy aquí reunidos, desplegó sus quehaceres profesionales primero dentro del área de las Ciencias y Técnicas Historiográficas, para más tarde darles continuación en la de Biblioteconomía y Documentación. Si cambió de área de conocimientos, los conocimientos a los que se dedicó no variaron en demasía, pues en ambos casos cultivó el estudio para servirnos en bandeja lo mucho que guardan los archivos, poniendo al alcance de los demás informaciones inalcanzables. Desde su perspectiva existía una clara coherencia entre estos dos ámbitos científicos, pues él los consideraba auxiliares de la historia. Si la Archivística centra su interés en la agrupación documental, la Paleografía posibilita el acceso a los contenidos de esos mismos documentos, convirtiéndose en un medio necesariamente utilizable por todo historiador que se acerque a descodificar las fuentes documentales de los tiempos antiguos, el feudalismo o la modernidad. Entonces ¿por qué no hablar de él como documentalista de los escritos? Así parece serlo, tanto desde la perspectiva manual como impresa, debido a su vocación paleográfica nacida sin duda de su proximidad a los archivos. Intermediario informativo, pues, entre los escribanos que conformaron los originales y los historiadores que sistematizan causalmente los hechos. En este propósito suyo de hacer disponible la información contenida en los manuscritos se advierte una clara intención documental, pues no cabe duda de que estamos hablando de manejo de información y de técnicas que permiten el acceso a unos contenidos existentes, para facilitar el camino de las consultas a documentos de todo tipo guardados en archivos y bibliotecas. Si aún dudásemos de su vocación por el acceso y difusión de los contenidos, vendrá a sacarnos de dudas el impulso que dio a los estudios reglados de Biblioteconomía y Documentación en esta Universidad.

En respetuosa evocación de su figura universitaria me atrevo a disertar ante ustedes sobre los problemas que causa el acceso y el control de la información contenida en documentos más actuales, cuya generación y distribución se hace electrónicamente. Hemos entrado en un nuevo mi-

lenio donde la información ya no se fija principalmente mediante la tecnología de la imprenta. El hipertexto y el desarrollo de la web están causando en los sistemas de información un cambio comparable al que provocara Gutemberg en el siglo XV. Se han impuesto nuevas representaciones de la información que deberán llevar el determinante inexcusable de "legible por máquina". La vida se ha digitalizado y debemos entender esta nueva situación, mirándola desde nuestros intereses académicos en Biblioteconomía y Documentación, como de consecuencias determinantes a la hora de crearse, difundirse y usarse los documentos. Si hemos buscado un nombre para los tiempos que corren y definimos a la sociedad actual como de lo digital o de la información es porque estamos convencidos de que mucho han variado sus estructuras respecto a la sociedad industrial precedente.

Es nuestro propósito acercarnos a los cambios que para el mundo de la información y documentación ha supuesto esta etapa recién iniciada. Para hacerlo debemos considerar los parámetros que la explican, de tan hondo calado social que han llegado a dar nombre al momento por el que pasamos, estableciéndose así como referente indudable de cuanto afirmemos a lo largo de esta exposición. Los trazos esquemáticos fueron descritos por Juan Luis Cebrián como certezas de la nueva humanidad en el libro que coordinó la profesora Caridad sobre la Sociedad de la Información<sup>1</sup>:

- Desde *el efecto global*: por el que las divisiones económicas y geográficas entre los países de la tierra han ido desapareciendo, ya que cualquier actuación en la red informática que nos une tiene dimensión planetaria.
- Pasando por *la velocidad* no solo en la transmisión de información, si no especialmente en el ritmo de crecimiento y en los cambios que se producen vertiginosamente.
- Sin olvidarse de *la interactividad* o posibilidad intrínseca de lo digital de conceder participación activa a cuantos intervienen en los actos comunicativos.
- Ni de *la virtualidad*: la nueva realidad aparecida ahora, un campo de actividad antes inexistente.

<sup>1</sup> Comenta la descripción de CEBRIÁN, J. L. «Algunas certezas sobre la Sociedad digital», en CARIDAD SEBASTIÁN, M (Coord). *La sociedad de la información: Política, Tecnología e Industria de los contenidos*. Madrid: Ceura, 1999: XVII-XVIII.

- Hasta alcanzar al *descontrol* con que lo digital ha emergido, fuera de los cauces de organización precedentes, y que podemos describir, por similitud, como que la red no tiene guardias que ordenen lo que por ella circula.

Llegados aquí no escapa una constatación evidente: en la Sociedad digital ni la información ni el documento son estables y estáticos. La información y su manifestación expresiva, el documento, se han vuelto una realidad dinámica, simultánea y transversal<sup>2</sup>. Si la imprenta colocó al texto en el ombligo de la comunicación informativa, como trasunto de la oralidad, ahora la digitalización ha generado una *nueva escritura icónica* que intenta cumplir las mismas funciones que la alfabética, fragmentando el *continuum* de una representación, esquematizando hasta evidenciar la estructura de esa representación y permitiendo la abstracción de la idea a partir de la imagen<sup>3</sup>. Es el origen de una *cultura mosaico* desde el punto de vista sociológico<sup>4</sup> y de la *cibercultura* desde la perspectiva técnica, un nuevo escenario tecnológico que exige una nueva competencia comunicativa, capaz de referirse a diferentes semióticas<sup>5</sup>. Al momento que atravesamos se le ha querido determinar como *cultura de la pantalla* para simbolizar la comunicación informática y audiovisual, en contraste con la anterior *cultura de la escritura*. Pero también hablamos de *ingeniería* o *industrias de la información*, queriendo abarcar los conceptos y productos generados por el desarrollo tecnológico en relación con la comunicación, el conocimiento y la inteligencia. La intervención tecnológica en la capacidad y en la cantidad de cosas a aprender ha generado también la denominación de *Era del aprendizaje*, es decir, de la cesión y asimilación de contenidos por las personas.

Volviendo a nuestro asunto, esta época diferente en la que vivimos y que nos esforzamos por definir, ha tenido a los ojos de Pérez Tornero

<sup>2</sup> MCHALE, B. *Postmodernist fiction*. New York; London: Methuen, 1987, y *Constructing postmodernism*. London: Routledge, 1992.

<sup>3</sup> PÉREZ TORNERO, J. M.- «Los nuevos procesos de mediación: del texto al hipermedia», en *Comunicación y educación en la sociedad de la información*. Barcelona: Paidós, 2000: 78-80.

<sup>4</sup> MOLES, A.- *Teoría estructural de la comunicación y sociedad*. México: Trillas, 1983.

<sup>5</sup> Muy atractivo estudio de DERY, M. *Velocidad escape: la cibercultura en el final de siglo*. Madrid: Siruela, 1998.

unos efectos documentales evidentes:

1. En los contenidos documentales interactúan varias semióticas de forma sincrética y simultánea.
2. Los distintos sistemas de signos pueden establecer relaciones jerárquicas entre diferentes tipos de semiótica, pero son siempre relaciones inestables, variables, estratégicas.
3. El sentido del contenido documental deriva de la conjunción y complementariedad de diversos tipos de signos<sup>6</sup>.

No albergamos duda alguna de que los nuevos servicios y aplicaciones infomedias han supuesto una modificación radical en las maneras de acceder, crear y controlar la información<sup>7</sup>. Y, por ello, no haremos reserva alguna a la hora de describir el nuevo panorama tecnológico que ha causado cambios en la producción de los documentos, en su tipología, y en las posibilidades de acceso:

1. Así, el proceso de creación y difusión se ha digitalizado y se ha vuelto multimedia, igual que el tratamiento y el manejo de los documentos. La digitalización ha informatizado las diferencias precedentes y ha hecho que los documentos electrónicos puedan ser impresos, audiovisuales, sonoros y gráficos al mismo tiempo. Al homogeneizarse la vía de expresión se ha facilitado el acceso hipotético a cuantos documentos existen<sup>8</sup>.
2. Los rápidos medios tecnológicos han permitido que el momento de la creación coincida en el tiempo con el de la comunicación, que se ha conseguido establecer en tiempo real, y con precios más bajos.
3. Por otra parte, la facilidad para manipular y recomponer los mensajes ha fomentado la multiplicidad de formas y significados. Con el consiguiente aprovechamiento de esta ventaja tanto en el campo editorial como en el individual. Si bien han aparecido problemas

<sup>6</sup> PÉREZ TORNERO, J. M. «La nueva competencia comunicativa en un contexto mediático», en *op. cit.*: 85-103.

<sup>7</sup> FEATHER, J. *The Information Society*. 2<sup>nd</sup> ed. London: Library Association Publishing Ltd., 1998

<sup>8</sup> LINDQUIST, M. G. «Long term strategies for electronic documents», en *Journal of the International Association of Sound Archives*, (1995), 6: 33-39.

de control físico y de integridad de la información.

4. Mientras que la distribución se ha vuelto más flexible y puede hacerse mediante Internet, en línea, en CD-ROM o de forma impresa (incluso con impresión bajo demanda, o por autoedición). La distribución se hace sin manifestarse como tal las copias, lo que dificulta el reconocimiento de los derechos de autor. La permanencia de la información es imprevisible, pues depende fundamentalmente de la tecnología utilizada para su creación, y ésta envejece con rapidez. Además, hoy la publicación se ha hecho dinámica, mientras que antes se grababa y era difícil de borrar o reformar. Ahora los documentos pueden someterse a revisión cuantas veces se quiera.
5. No existen frenos técnicos para hacer copias del original, que además se realizan sin pérdida de calidad alguna. El concepto de garantía de autenticidad ha quedado descontextualizado. Desde esta perspectiva se han igualado acceso y posesión, préstamo y venta.
6. Nos hallamos, por tanto, frente a una información documental móvil, no sólo debido al incremento de los documentos audiovisuales que exigen un análisis documental que refleje el movimiento, ni tan siquiera al desarrollo de la documentación hipertextual e hipermedia, que al mismo tiempo reclama una identificación conceptual en mapas de navegación, sino sobre todo a la interacción sincrética de diversas semióticas con sus propias escrituras. Los hipertextos establecen asociaciones entre diferentes ítems de información imitando el funcionamiento del pensamiento humano. Permiten desplazarse con facilidad a través de toda la información contenida en las bases de datos. Usando estas estructuras lógicas los documentos digitales pueden enlazarse con otras informaciones que no pertenezcan físicamente a su unidad. Desde un documento puede saltarse a otros documentos custodiados a cientos de kilómetros de distancia. Cada pequeña representación de información se incrusta en algún sector de meta-información. Esto facilita el diseño y presentación del documento, los enlaces hipertextuales y las relaciones lingüísticas que permiten búsquedas semánticas o facilitan la traducción.

## **1. DE LA EDICIÓN IMPRESA A LA EDICIÓN DIGITAL.**

De la reflexión precedente se deriva que en la expresión del conocimiento se ha pasado de una edición impresa de aplicación

mayoritaria a una edición digital cada vez más frecuente.

La escritura ha sido la vía común por la que históricamente se ha transmitido el conocimiento. Ahora, los nuevos medios tecnológicos, lejos de actuar en su contra, han potenciado su utilidad. Internet se está convirtiendo en la vía preferida para publicar. Por encima de todo, permite llegar a los destinatarios de inmediato. Y favorece publicar cosas hasta ahora inimaginables: asuntos de todo tipo, del interés más disperso. Cualquiera puede publicar, comentar, juzgar o avisar de algo. Y aunque la WEB no es un proceso de publicación, *estimula la tradición académica de la publicación abierta de ideas y resultados*<sup>9</sup>. Es un medio recién llegado con el que se ha iniciado un comportamiento diferente en la difusión de la información y han aparecido nuevas oportunidades de edición, especialmente para la escritura que se ha visto fomentada por:

- La posibilidad de publicar ilimitadamente: al ofrecer libertad, universalidad de temas y soportes, y comunicación instantánea, que además, por si fuera poco, es más barata.
- El fácil intercambio de ideas entre las fuentes y los destinatarios, de modo especial mediante las aplicaciones del correo electrónico, que permiten intercambiar mensajes y ficheros.
- La oferta para buscar datos e información en texto completo.
- La existencia de hiperenlaces entre diferentes textos y documentos, así como la de variados métodos de recuperación con potentes herramientas de búsqueda.
- La sencilla anexión de componentes multimedia y de cualquier tipo de anotaciones.
- El marketing y la promoción de los productos que efectúa la propia red.

La contrapartida está en que mucho de lo que se publica no aporta casi nada, es inútil o está falto de la mínima calidad científica, empresarial, o social. Además, el conocimiento distribuido por la red no guarda un orden paralelo al mundo real que representa. La tarea de organizarlo afecta sobre todo a los editores y a quienes trabajan en el ámbito académico, pero más que a nadie a quienes siempre cuidaron del acceso y control de la información, que de manera especial están obligados a ofer-

<sup>9</sup> LEINER, B. ; CERF, V. et AL. «Una breve historia de Internet», en *Novática*, (1999), 1: 8.



tar al usuario diversos caminos para orientarse sobre los recursos existentes y la calidad de los mismos. Aparece así destacada la intervención de los profesionales de la Biblioteconomía y Documentación para organizar lo que ha nacido desordenadamente.

Los modos de hacer de las artes gráficas pueden servir de referencia. Son un negocio de larga tradición, con mucha experiencia acumulada, cuya actividad exige un proceso administrativo y técnico complejo. Hasta ahora, los editores han financiado, gestionado y vendido productos impresos. Para sacar adelante su inversión han tenido que asegurarse la calidad de los originales. Por ello los evalúan y filtran. Nunca los aceptan sin una discusión previa. Si se quiere introducir un poco de seriedad en el proceso, es obligado que las funciones plenamente establecidas de la edición tradicional se mantengan en las publicaciones electrónicas<sup>10</sup>.

## **2. LA PRODUCCIÓN DIGITAL EN LOS MEDIOS DE COMUNICACIÓN Y DE INFORMACIÓN CORPORATIVA E INSTITUCIONAL.**

Si bien la digitalización ha llevado de inmediato sus efectos más allá del mundo científico y de la cultura, alcanzando a la producción y gestión de todo tipo de empresas e instituciones, ha tenido especial incidencia en los medios de comunicación social. La difusión en línea, usando páginas web, ha resultado ser un medio tecnológico muy apropiado para los diarios y revistas de información general. En su confección se emplean las mismas planas que sirvieron para elaborar la tirada en prensa<sup>11</sup>. La principal utilidad de la edición digital es que la información corriente está disponible casi desde el momento de su composición. A lo que se añade que con un sistema de menús se facilita la localización de la información deseada. Y que su distribución es, además y hasta ahora, gratuita.

Desde luego, en el sector audiovisual, la cultura y el ocio están interpretando el papel principal, fomentadas por lo digital. No cabe duda de que entre las aplicaciones multimedia ofrecen una destacada presencia los

<sup>10</sup> COHEN, A. «A system for electronic peer review», en BUTTERWORTH, I (ed.). *The Impact of Electronic Publishing on the Academic Community*. London: Portland Press, 1998: 69-71.

<sup>11</sup> ABADAL, E. «El futuro de la edición, ¿es electrónico?: tecnologías de creación y empleo de la información», en *Jornadas Españolas de Documentación Automatizada (4as: Gijón, 1994)*. IV *Jornadas Españolas de Documentación Automatizada*. Gijón: FESABID, 1994. p. 378

contenidos culturales, educativos y lúdicos, junto a los generados por el cine, el vídeo, y la música, además de las noticias. La televisión por cable ofrece acceso a Internet y trata de emitir por su medio. Es una nueva concurrente en la producción y distribución de programas. Su llegada a este mercado ha hecho que las audiencias se disgreguen aún más. La liberalización del sector de las telecomunicaciones junto a la posibilidad de desarrollo tanto de sus redes tradicionales como de la televisión por cable ha favorecido el establecimiento de alianzas empresariales por las que han intercambiado sus competencias específicas. Además, la televisión por cable es interactiva. Ahí reside su principal ventaja. Y a su lado crece constantemente el mercado de los canales temáticos, a gusto del espectador: programas de carácter científico, deportivo, artístico, cultural, infantil, de aventuras, de viajes, de la naturaleza, musicales, o de noticias. Pero también el desarrollo de juegos interactivos en red, de aplicaciones a la teleenseñanza, o a la televenta ofertada mediante catálogos en línea. Cada uno de estos aspectos supone una oportunidad de desarrollar contenidos para la industria audiovisual.

Es pujante, de modo especial, la creación de contenidos relativa a programas para la educación y el entretenimiento. La utilización de la realidad virtual y el apoyo a los métodos didácticos con el ordenador y las redes permite tanto al que enseña como al que aprende tener gran libertad de actuación. Los límites del aula se han visto ampliamente superados. No se trata de apoyar mecánicamente la docencia, si no de utilizar el ciberespacio para favorecer la comprensión de conceptos abstractos, el aprendizaje de lenguas, la actuación de los hechos físicos y geográficos, las reconstrucciones históricas, la recreación de ambientes biológicos, etc.

La mayor facilidad de generación y difusión de los contenidos favorecida por las web ha alcanzado también a la información creada por las instituciones, empresas y sociedades. En la creación de información están implicados, en primer lugar, todos los niveles de las administraciones públicas y las instituciones oficiales, organizaciones no gubernamentales y todo tipo de asociaciones. Pero asimismo, la mayoría de las empresas se vuelcan por estar presentes en la red para acercar su imagen, actividades, servicios y productos a quien desee conocerlas.

En todo el mundo hay millones de páginas web creadas por organismos oficiales. Desde ellas se puede acceder a datos de las administraciones locales, regionales, estatales, supranacionales y de los organismos internacionales. Suelen tratarse de textos (legales, informativos, explicativos de servicios, directorios, etc.), pero hay también lugar para gestiones

y tramitaciones administrativas que se pueden hacer desde casa, o para la descarga de formularios de interés para los ciudadanos. Son destacables, por la importancia de los contenidos que crean y distribuyen, otros centros oficiales como Academias, Institutos de Normalización, Bibliotecas Nacionales y redes de bibliotecas, o Bases de datos de la acción gubernamental sobre la Sociedad de la Información<sup>12</sup>.

Finalmente, consideramos en este apartado, de forma somera, la implicación de lo digital en la gestión integrada del conocimiento en las empresas. Las magnitudes alcanzadas por la creación y el manejo de información electrónica dentro de la actividad de las empresas ha fomentado la actualización de los procedimientos con los que tratar las grandes cantidades de datos almacenados. La existencia de nuevos parámetros a la hora de crear, gestionar y difundir el conocimiento está delimitada *por la complejidad, la incertidumbre y la necesidad de respuesta en tiempo real que requiere cualquier organización*<sup>13</sup>. Así las cosas, el objetivo principal de cualquier sistema de gestión consiste en mantener el control si se quiere tener capacidad de predicción sobre la organización para la que se trabaja. Las empresas se mueven ahora en contextos cada vez más competitivos, donde es muy importante administrar bien la información, cuyo valor se aprecia a la hora de tomar decisiones o de marcar las estrategias a seguir. El problema está en que muchas han pensado que el reto radica en propiciar la potencia del procesamiento de datos, cuando donde realmente se sitúa es en comprender el potencial de la tecnología para desarrollar y difundir información como fuente de ventaja competitiva<sup>14</sup>.

Los *Data Warehouse* son potentes bases de datos cuyos recursos provienen de los diferentes sistemas operativos de una empresa (facturación, compras, ventas, producción, personal, finanzas, marketing, etc.) que, orientados por temas, integrados, variables en el tiempo y no volátiles se

<sup>12</sup> En <http://www.internet.gouv.fr/francais/index.html> está programada toda la acción gubernamental francesa sobre la Sociedad de la Información. Internet.gouv.fr, no es más que otra tentativa gubernamental vía web de potenciar la nueva sociedad del conocimiento.

<sup>13</sup> NITIN, J. B. «An Action Perspective: The Crux of the New Management», en *California Mananagement Review*, (1994), 2: 54.

<sup>14</sup> BARTLETT, C. y GHOSHAL, S. «Changing the Role of Top Management: Beyond Systems to People», en *High Business Review*,(1995), 3: 46.

emplean como apoyo a la toma de decisiones administrativas<sup>15</sup>. Un Data Warehouse posibilita:

- El acceso universal a los datos disponibles.
- La puesta en marcha de un sistema por el que fluyen recursos informativos externos e internos.
- La selección de los datos de acuerdo con el contenido de su información y con su importancia para tomar decisiones.
- El almacenamiento diferenciado de las bases de datos que conforman el *Data warehouse* y los metadatos de aquellas otras con datos de los diferentes sistemas que operan en la empresa<sup>16</sup>.
- El establecimiento de herramientas sobre las que establecer las consultas del usuario final.

Hasta aquí hemos revisado someramente las posibilidades que para el desarrollo de los contenidos ha traído lo digital a diferentes sectores de la vida social, vamos a entretenernos ahora con aquello que está más próximo a la actividad científica.

### 3. LAS REVISTAS CIENTÍFICAS DIGITALES.

De manera especial, dados nuestros propósitos académicos, nos interesa atender al impacto que ha tenido lo digital sobre las revistas de carácter científico y técnico. Quienes las editan saben que manejan un negocio delimitado por la peculiaridad de los destinatarios y del producto<sup>17</sup>, ya que

- en cada tirada lanzan un número escaso de ejemplares, pues se dirigen a un restringido sector de usuarios como son los especialistas de los campos científicos;
- manejan unos contenidos escogidos y particularizados cuyo envejecimiento es rápido a causa de la constante evolución de las investigaciones;

<sup>15</sup> INMON, B. *Building the Data Warehouse*. New York: John Wiley & Sons, 1992.

<sup>16</sup> CORNELLÁ, A. *Información digital para la empresa: una introducción a los servicios de información electrónica*. Barcelona: Macombo-Boixareu, 1996.

<sup>17</sup> COLLIER, H. *The electronic publishing maze. Strategies in the electronic publishing industry*. Tetbury: Infonortics, 1998: 55.

- además tienen una audiencia, un sector de la comunidad científica, muy dispersa internacionalmente;
- es común que los autores no cobren. Basta con saber que contribuyen al progreso del conocimiento, junto al prestigio profesional que esto conlleva;
- hasta ahora su periodo de composición ha sido demasiado dilatado y formal. La meticulosidad de las artes gráficas alargaba en exceso el distanciamiento entre la entrega de los originales y la edición, con el consiguiente desgaste de los contenidos;
- la publicación final era muy restringida y controlada al detalle. Es bien sabido que muchas revistas devuelven más del 50% de los originales recibidos.

Contexto este desde el cual nos interesa considerar a Internet como vía de publicación regulada y crítica. Son muchas las posibilidades que la red ofrece a la edición y distribución de las publicaciones. Inicialmente se duplicaban, en impreso y electrónicas, las ediciones aprovechando la digitalización de texto e imágenes. Por lo que ambas ediciones ofrecían la misma calidad y control. Después, ha ido creciendo el número de publicaciones distribuidas tan solo digitalmente. Incluso han ido apareciendo revistas electrónicas que nunca tuvieron precedentes impresas. Su configuración es similar a la de las publicaciones impresas: sirven a unos propósitos y un campo de aplicación; cuentan con editores, director, consejo de redacción y colaboradores; tienen fijada su aparición periódica; presentan sumario de contenidos y unas secciones<sup>18</sup>.

La facilidad con que se han alistado las revistas científico-técnicas a la edición digital se ha debido a que:

- De entrada, el público al que se dirigen está especialmente preparado: posee habilidades e infraestructuras informáticas y puede comunicarse en varias lenguas y, desde luego, en inglés.
- A lo que se añade que es muy sencillo editar en versión electrónica. Cualquier grupo de investigación o centro universitario puede actuar como editor, pues tan solo precisa de una infraestructura ofimática común y contar con programas de autoedición.
- Suponiendo, de esta manera, unos costes escasos, con la consi-

<sup>18</sup> ABADAL, E. «Els serveis d'informació electrònica: què són i per a què serveixen?» Barcelona: Edicions de la UB, 1997.

guiente adaptación a las obligadas tiradas cortas en que se difunden los resultados de sus investigaciones.

- Si añadimos que, como hemos observado antes, son de distribución inmediata en línea, vemos que se evitan las demoras y los costes producidos por el transporte y el proceso de mercado de los impresos.
- No cabe duda de que favorecen la circulación de los artículos (incluso antes de su terminación: los conocidos *working papers*).
- Ofertan mayor eficacia y versatilidad, con facilidad de movimientos por las tablas de contenidos y los resúmenes de los artículos.
- En fin, permiten el establecimiento de enlaces más fáciles y ricos, impensables en los impresos. Desde un artículo se puede enlazar con fuentes externas: otro artículo que se cita en las notas, una referencia bibliográfica, y añadir comentarios personales con lo que se facilita la creación de archivos propios.

Pero la publicación de revistas digitales de calidad plantea aún interrogantes radicales tales como<sup>19</sup>: ¿de qué manera fiarnos de la calidad de los contenidos ofertados por autoedición? ¿quién sostiene la publicación? O ¿quién evaluará los contenidos?, ¿con qué rigor?, ¿quién corre con los gastos que supone la valoración?

#### **4. LA CREACIÓN Y EL SUMINISTRO INTEGRADOS DE CONOCIMIENTO CIENTÍFICO Y TÉCNICO.**

Un paso más y consideramos Internet como vía de suministro del conocimiento científico y técnico. Su implantación exigió a los centros implicados en la industria del conocimiento, que se habían anticipado a la difusión de sus productos en línea, a que adaptasen su forma de conexión. A la vez que la atención primera que el sector de la información electrónica prestó a las referencias (información secundaria) se fue ampliando a los originales que hasta hace poco se editaban solo en papel. Ahora existen muchas compañías con la doble función de producir y distribuir información científico-técnica. Comúnmente facilitan enlaces

<sup>19</sup> SCUPOLA, A. «The impact of electronic commerce on the publishing industry: Towards a business value complementarity framework of electronic publishing», en *Journal of Information Science*, (1999), 25, nº 2: 135. También ALTBACH, P. La edición, en *Informe mundial sobre la información*. Madrid: UNESCO-CINDOC, 1997: 336-343.

que partiendo de los registros bibliográficos en línea, dirigen la navegación hacia el texto íntegro.

Dado que la publicación electrónica no solo compete a Internet, si no también a los servicios en línea tradicionales, a los fabricantes de discos ópticos y a cualquiera que venda información desde un sistema autónomo, es lógico que el servicio ofertado por las grandes compañías que comercian con la generación y distribución de información sea considerado cada vez más como merecedor de retribución económica. De alguna manera, la inversión que han hecho debe retornar. Entre los servicios en línea tradicionales, perduran las actividades precedentes. Pensemos en DIALOG, ORBIT, ESA-IRS, LEXIS-NEXIS, STN, grandes servicios centralizados que se adelantaron a los ordenadores personales en dar servicios de información en línea<sup>20</sup>. La nueva situación delimitada por la presencia de la web, los enlaces entre contenidos y los motores de búsqueda se gestó fuera de ellos. Hoy tienen su futuro comprometido. Su tamaño y la dificultad de gestión operan en su contra.

Sin duda es más llamativa la adaptación que han sufrido las grandes compañías editoras de información científica. Son los casos de Springer-Verlag, a través de su LINKService, o de IDEAL con su Academic Press, y del servicio ScienceDirect™ de Elsevier<sup>21</sup>, también de SWETS, del OCLC FirstSearch Service<sup>22</sup>; así como de EBSCO Information Services, y de SilverPlatter<sup>23</sup> con sus sistemas de recuperación *Silver Linker* y *KnowledgeCite Library*. Todos ellos permiten consultar libremente las tablas de contenido y re-

<sup>20</sup> COLLIER, H. *The electronic publishing maze. Strategies in the electronic publishing industry*. Tetbury: Infonortics, 1998: 123

<sup>21</sup> Los usuarios pueden imprimir o guardar partes de los textos para su uso personal. Véase SHOFF, P. L.- The traditional journal, en BUTTERWORTH, I. (ed.). *The Impact of Electronic Publishing on the Academic Community*. London: Portland Press, 1998: 66-69.

<sup>22</sup> Se incluyen las bases de datos de OCLC: WorldCat, OCLC FirstSearch Electronic Collections Online, OCLC NetFirst, OCLC ArticleFirst, OCLC ContentsFirst, OCLC FastDoc, OCLC PapersFirst, OCLC ProceedingsFirst, y OCLC Union Lists of Periodicals. HomePage de *Online and Computer Library Center*, accesible en: <http://www.oclc.org/oclc/menu/home1.htm>

<sup>23</sup> El nacimiento de la industria del CD-ROM tuvo lugar precisamente en 1985, cuando SilverPlatter presentó bases de datos en este soporte dentro del Congreso de la *American Library Association*. Véase *CD-ROM: a practical guide for information professionals*. BURTON, P. F., y MOORE, C. eds. 2<sup>nd</sup> ed. London: UK Online User Group y Library Information Technology Centre (LITC), 1994.

súmenes de los artículos de las revistas que distribuyen. Mientras que el acceso a los textos completos solo está permitido para aquellos usuarios que cuenten con una licencia. Se esfuerzan por integrar en línea los servicios de acceso y suministro de documentos, con objeto de fomentar la utilización de las colecciones bibliográficas<sup>24</sup>. Algunos de ellos, como *Kluwer Academic Publishers*<sup>25</sup> presentan tal adaptación a las necesidades del usuario que imprimen un libro o revista solo cuando lo precisa un usuario (*Printing-on-demand*). Los contenidos permanecen almacenados en discos ópticos y se pasan al papel tan solo cuando existe una petición.

## 5. LOS CONTENIDOS ESTRUCTURADOS: LA BIBLIOTECA HÍBRIDA.

Pero, si algún organismo se ha adaptado a la situación digital, no cabe duda de que ha sido la biblioteca junto al resto de sistemas de información que podemos acoger bajo este nombre generalizador. Con sus procesos, perfectamente estructurados durante siglos, han sabido sacar ventaja de lo favorable y superar los inconvenientes que la red plantea. Podemos sentir que hemos pasado del *todo está en los libros* al *todo está en la red*, pues mediante su concurso estamos alcanzando el viejo mito de los bibliotecarios de llevar hasta los usuarios todo el conocimiento documentado<sup>26</sup>. Se está haciendo realidad la aspiración manifestada por Theodore Nelson en Xanadú: disponer de una ilimitada red de información, accesible al instante, donde se guarden y desde donde se difundan todos los

<sup>24</sup> FEINBERG, D. «Full-Text Online and Abstracts on CDROM: Nexis, Dialog and Abi/Inform 175, en *Online&CDROM Review*, (1995, 19, n° 3: 143-148.

<sup>25</sup> About *Kluwer Academic Publishers* [documento www]. Dordrecht: Kluwer Academic Publishers, 1999. Disponible en: <http://www.wkap.nl/kapis/CGI-BIN/WORLD/kaphtml.htm?ABOUTKAP>

<sup>26</sup> Desde las utopías renacentistas, la aspiración a un acceso ilimitado al conocimiento ha vivido sucesivas renovaciones. Véase MILLARES CARLO, A. (ed.). *Utopías del Renacimiento*. México: FCE, 1941. Aquellas teorías visionarias, pese a lo que el tiempo ha caminado, han mantenido su vigencia como referencias a lo ideal. Podemos recordar a Leibniz, a Otlet, a Bush, a Wells. Desde luego, la más conocida de ellas es la *Bibliotheca universalis* de La Fontaine y Otlet, cuyo empeño se centraba en eliminar las dificultades de acceso a la información causadas por la distancia y la concurrencia de usuarios. LA FONTAINE, H. y OTLET, P. «L'état actuel des questions bibliographiques et de l'organisation internationale de la documentation», en *Institut International de Bibliographie, Bulletin*, (1908), 13: 167.



contenidos del saber humano<sup>27</sup>.

Como recurso informativo, Internet es similar a una biblioteca sin libros que recoge la más variopinta información de todo el mundo. Es la nueva biblioteca de Alejandría. Estamos en el camino de coleccionar digitalmente todo el conocimiento humano y de hacerlo llegar a todo tipo de personas. En esta biblioteca universal caben también los documentos preexistentes. Para poder usar esta mezcla de recursos las bibliotecas deberán olvidar la dependencia de cualquier soporte, pues lo que les interesa es satisfacer las demandas y esto se logra con la selección y difusión adecuadas de los recursos de acuerdo con sus contenidos. Se alcanza así la auténtica cooperación internacional, a la hora de crear y de intercambiar tanto los registros bibliográficos como los originales multimedia. La apertura de los depósitos de información a la web ha facilitado el acceso a los documentos completos y a todo tipo de información.

Volvemos a valorar que Internet haya traído nuevos presupuestos de actuación:

- Las puertas de la ciencia son mayores para los especialistas, pero también para la gente de la calle.
- El conocimiento y los datos se comunican con mayor rapidez.
- El intercambio de ideas y el flujo de información entre los especialistas es más fácil.

El fin de las bibliotecas continúa siendo el acceso referencial y el uso material de los documentos. Y así debe ser, por más que a la nueva situación se la haya buscado definir con nombres como *biblioteca digital*, *biblioteca electrónica*, *biblioteca sin paredes*, *biblioteca virtual*, e incluso con el más adecuado de *biblioteca híbrida*, en alusión a la colección mixta que manejan<sup>28</sup>. En ella encuentran paralelismo las rutinas propias de una biblioteca tradicional: la adquisición de los materiales previa a las demandas de los lectores; la custodia, el préstamo interbibliotecario; la creación de instrumentos que identifiquen y localicen las diferentes materias y temas. Por ello, se continúa precisando de la ayuda de un profesional y de herramientas que guíen a los usuarios hacia las existencias que buscan; así

<sup>27</sup> NELSON, T. *The Xanadu Paradigm*. San Antonio (Texas): Theodor H. Nelson, 1987.

<sup>28</sup> OPPENHEIM, C. y SMITHSON, D. «What is the hybrid library?», en *Journal of Information Science*, (1999), 25, nº 2: 97-112.

como de un orden lógico en la disposición de los fondos<sup>29</sup>. Los documentos en papel y las publicaciones digitales deberán cohabitar por largo tiempo, por lo que están llamados a seguir unos idénticos objetivos. Sin embargo, algo está cambiando desde hace más de una década. Las bibliotecas, valiéndose de las redes han roto amarras con la anterior tarea principal de conservar la información y, sin duda, se dirigen hacia la difusión y la cooperación. Los nuevos depósitos (servidores) facilitan la ubicuidad ilimitada en el acceso y la difusión, en función del lugar en el que esté el usuario.

Los trabajos y funciones bibliotecarias están cambiando con gran rapidez. Además, es evidente que las bibliotecas son causantes de la introducción de entornos digitales en las instituciones a las que dan servicio. Hasta ahora hemos creído que una biblioteca era un lugar donde se guardaba la información<sup>30</sup>. Sin embargo, vemos que su equivalente electrónico tiene que atender a informaciones dispersas, dentro de un marco novedoso<sup>31</sup>:

1. Aunque la responsabilidad sobre el archivo de los documentos electrónicos aparece aún poco determinada, el conocimiento expresado en los documentos digitales tiene también que preservarse. Los contenidos científicos y técnicos de grabación digital son escasos si lo comparamos con los que están impresos, lo que no impide que deban ser atendidos adecuadamente. El contenido es el factor dominante y hay que preservarlo con independencia del soporte en el que se dispuso originariamente.
2. Percibimos la biblioteca digital como una entidad compleja e inestable, que cuenta con escasas apoyaturas teóricas. Esta inestabilidad hace que la inversión necesaria para ponerla en marcha sea aún de alto riesgo. Nos movemos en una situación en la cual los aspectos económicos no se comprenden todavía muy bien.

<sup>29</sup> MEADOWS, J. «The development of digital libraries», en BUTTERWORTH, I. (ed.). *The Impact of Electronic Publishing on the Academic Community*. London: Portland Press, 1998: 118-125.

<sup>30</sup> MACKENZIE OWEN, J. «Preservation of digital materials for libraries», en *European research libraries cooperation; the LIBER Quarterly*, (1996), 6, nº 4: 440.

<sup>31</sup> Descripción ampliada de M. COLLIER. «A model for electronic university library», en HEAL, A. y WEISS, J (eds). *Towards a worldwide library: a ten year forecast*. Essen: Essen University Library, 1997: 180 – 190.

3. Opera en un entorno global: nuevos productos y servicios se hacen normales en corto espacio de tiempo. Así, se tiende al acceso a los recursos en red, más que a los materiales almacenados en el centro. El mismo catálogo de biblioteca se ha vuelto una herramienta para descubrir recursos en línea. En consecuencia, las referencias bibliográficas se están ampliando hacia recursos no bibliográficos: como referencias a personas, organizaciones y conjuntos de datos.
4. Para mantener la competitividad, la cooperación resulta primordial. Con un solo ejemplar se pueden satisfacer necesidades de información de origen disperso, gracias a que la red difunde los recursos por todas partes.
5. Finalmente, se observa que con frecuencia creciente la interactividad con el usuario se suele establecer a distancia (desde el ordenador), en preferencia sobre el acceso personal.

## 7. EL TRATAMIENTO DE LOS NUEVOS CONTENIDOS.

Vamos viendo cómo sólo aquellos sectores que han organizado y tratado la información, tanto tradicionalmente como recién llegados, son los que obtienen ventajas de la oferta informativa que Internet ha supuesto. En ellos, las técnicas de tratamiento de los contenidos documentales continúan pretendiendo establecer un puente conceptual entre los documentos y los usuarios<sup>32</sup>. Si de alguna manera hubiera que determinar

<sup>32</sup> Los modelos teóricos de tratamiento de los contenidos son de naturaleza lingüística: se centran en el estudio de los diversos niveles del lenguaje natural: fonológico, morfológico, léxico, sintáctico, semántico y pragmático. Podemos nombrar: la *Lingüística del texto*, en la que Van Dijk que explica los textos desde las: macro-, micro- y super-estructura (véase MOREIRO, J. *Aplicación de las Ciencias del texto al resumen documental*. Madrid: BOE – Universidad Carlos III, 1993); el *Modelo reticular* de Rumelhart: relaciona la información con los esquemas de conocimiento de emisores y receptores; Los *Modelos modulares* o integradores del procesamiento del lenguaje desde sus diversos niveles; El *Modelo estructural*, de Thorndyke basado en las estructuras documentales: natural, conceptual y física; y la *Gramática de casos*, de Fillmore que desde la situación concreta de los discursos que da significado a los términos que los componen. Pero también de carácter estadístico: como el modelo vectorial, y el de ponderación, así como métodos de agrupación en clases y algoritmos conexionistas. Finalmente, de índole cognitiva, que integra estructuras textuales y del conocimiento junto a las estrategias y técnicas de representación, los fines de los sistemas de información,

(cont.)

en qué se han centrado los esfuerzos habría que responder que lo han hecho en automatizar, ya que así se puede explotar, estructurar y tratar directamente el enorme flujo de información que circula por Internet<sup>33</sup>. Entre los intentos de estructuración y control de los documentos electrónicos vamos a destacar estos:

### **7.1. Metadatos (datos sobre los datos).**

La catalogación entendida como estructuración de la información, desde la que se generalizó el formato MARC, había tendido a verse en algunos ámbitos como una actividad extraña que correspondía tan sólo al mundo de los bibliotecarios y archiveros. Apreciación que se está modificando desde que comenzó a participar en ella un sector más amplio de la comunidad científica. La trascendencia de los problemas que plantea la recuperación de información más allá de los límites de las bibliotecas y de los sistemas de información se ha consolidado con la generalización de la información electrónica. Una de las actividades que lo demuestra es el desarrollo de diversos modelos de metadatos que persiguen organizar los recursos informativos digitales, ayudando a identificarlos, describirlos y localizarlos. Dentro de estos formatos el que cuenta con mayor desarrollo y aceptación es el Dublin Core que, desde su aparición en 1995, ha ido evolucionando para representar la información electrónica a través de quince etiquetas de información básica. También los formatos de descripción bibliográfica más tradicionales han contemplado en su normalización los recursos electrónicos, como testimonian las ISBD-ER de 1997.

### **7.2. Indización automática de conceptos.**

De otra parte, las técnicas de análisis del contenido se han volcado en la indización automática de los conceptos, cuya identificación es imprescindible para recuperar documentos en la red y para valorar su pertinencia<sup>34</sup>. El procesamiento lingüístico de los documentos está ahora marca-

pretensiones de los usuarios, así como el contextos en que se produce el análisis y la recepción de información.

<sup>33</sup> LYNCH, D. Searching the Internet [documento www]. *Scientific American*, March 1997. Disponible en: <http://www.sciam.com/0397issue/0397lynch.html>

<sup>34</sup> GIL LEIVA, Isidoro y RODRÍGUEZ MUÑOZ, José Vicente. Tendencias en los sistemas de indización automática. Estudio evolutivo, en *Revista Española de Documentación Científica*, (1996), 19, nº 3: 273-291.

do por la utilización de las bases de conocimientos como instrumentos de análisis. En ellas se indican los tipos de relaciones que existen entre los conceptos y concretan el sentido de los enlaces sintácticos. Son una especie de tesauros enriquecidos con información morfológica, sintáctica y semántica, en los que se especifican las relaciones asociadas de los tesauros. Este modo de actuar sigue el conocimiento (semántica de mundo) que los especialistas han depositado en los documentos. En su aplicación intervienen la estadística, la informática, la lingüística y la Inteligencia Artificial<sup>35</sup>. Las bases de conocimientos buscan establecer relaciones semánticas desde un concepto con sus causas y consecuencias. Los tesauros ya tenían relaciones de asociación, pero limitaban mucho la aplicación de relaciones semánticas entre sus términos. Ahora, las bases de conocimientos especifican cómo es esa asociación, y la representan mediante estructuras arborescentes o en planos infográficos<sup>36</sup>.

### **7.3. La generación automática de los tesauros conceptuales.**

Otra herramienta terminológica es la generación automática de tesauros conceptuales: un conjunto estructurado de conceptos útiles para describir los conocimientos propios de un determinado sistema de información. La situación actual obliga a diseñarlos y construirlos de tal manera que sean capaces de almacenar, procesar, gestionar y recuperar cualquier tipo de documento, sin importar su presentación, soporte y forma de acceso. Y además lo deben hacer a través de un sistema auto-generable que permita que el propio sistema se actualice conforme vaya almacenando. Para lograrlo deben contar con la construcción automática de relaciones; con las aportaciones de términos hechas desde un tesoro de descriptores global (utilizando la superestructura de los tesauros); y, desde luego, con analizadores sintácticos, semánticos y morfológicos.

Para identificar y adquirir las palabras representativas de un dominio se realizan los siguientes subprocesos: análisis léxico, tratamiento de palabras vacías, tratamiento de términos flexionados, tratamiento de palabras compuestas y filtrado de términos.

Mientras que para realizar filtrados sobre los posibles términos representativos de un dominio se utiliza la Indización Estadística de Términos

<sup>35</sup> POLANCO, X.- *Infométrie et ingénierie de la connaissance*. Nancy: INIST-CNRS, 1995.

<sup>36</sup> MOREIRO, José A., y MÉNDEZ, Eva. «Lenguaje natural e Indización automática», en *Ciencia de la Información*, (1999), 30: 11-24.

por Frecuencias<sup>37</sup> o filtrado basado en la ley de Zipf<sup>38</sup>, así como el Método N-grams, algoritmo aplicado a la solución del tratamiento de palabras compuestas. Para la obtención de relaciones entre componentes se utilizan diferentes tipos de clasificadores<sup>39</sup>: los cuantitativos para el análisis de coocurrencia de las palabras; los estadísticos para la agrupación en clases<sup>40</sup>; y también los de Redes Neuronales<sup>41</sup>.

Con el fin de alcanzar a describir con exactitud los contenidos de los documentos, en especial cuantos se refieren a actividades, acciones y procesos, se están iniciando ensayos de construcción de Tesoros de verbos de estructuración paralela a los clásicos tesauros de descriptores, pero que presentan grandes variaciones en las características de extracción, clasificación, y asociación.

Siguiendo esta línea se llega a los Mapas Conceptuales, representaciones diagramáticas del conocimiento mediante diferentes lenguajes visuales (representaciones gráficas de los conceptos y de las relaciones entre ellos). Los *conceptos* son los *nodos* (cada uno con su tipo, nombre y contenido), y las *relaciones* son los *enlaces*. Son eficaces para organizar nueva información e integrarla en el conocimiento existente, ya que su construcción ayuda a reconocer nuevas relaciones entre conceptos y a refinar la comprensión de las relaciones existentes.

Finalmente, dentro de este grupo, debemos notar la aparición de la *Norma ISO/IEC 13250 Topic Maps* que ha venido a definir el modelo y la sintaxis de intercambio para formalizar Mapas conceptuales. Un mapa

<sup>37</sup> SALTON, G. *Automatic Text Processing: the Transformation, Analysis, and Retrieval of Information by Computer*. New York: Addison-Wesley, 1989.

<sup>38</sup> ZIPF, G. K. *Human Behaviour and the Principle of Least Effort: An Introduction to Human Ecology*. New York: Haffner, 1972.

<sup>39</sup> La obtención de relaciones se hace desde la integración de distintas técnicas, que trabajan en paralelo. Realizan conjuntamente un proceso de clusterización por el que se agrupa en clases aquellos descriptores que responden a una serie de características comunes: primero se identifica el centroide o concepto más significativo que actuará como raíz de una jerarquía. Luego se realiza la clusterización o agrupación en clases del resto de los descriptores mediante alguna técnica de clasificación.

<sup>40</sup> Max-min, K-vecinos, K-vecinos incremental, Isodata.

<sup>41</sup> Iniciada en la propuesta de HEBB, D. *Organisation of Behaviour*. New York: Wiley & Sons, 1949. En la actualidad los más aplicados son Kohonen, ART-1 y ART-2. Véase HILERA, J. R. y MARTÍNEZ, V. J. *Redes neuronales artificiales: Fundamentos, modelos y aplicaciones*. Madrid: Rama, 1995.

conceptual es un documento o grupo de documentos SGML o XML en los que se usan diferentes tipos de elementos para representar conceptos, los diferentes casos que ofrecen, y las asociaciones entre ellos. Esta estructuración semántica de los enlaces en la red ha permitido que se la denomine “el GPS del universo de la información”, por su capacidad para organizar y navegar por los gigantescos recursos de información, y para establecer un puente entre los campos del conocimiento y la gestión de la información.

Los Mapas Conceptuales proporcionan una notación normalizada para representar información intercambiable en torno a la estructura de los recursos de información utilizados para determinar los conceptos y sus relaciones. Un grupo de uno o más documentos interrelacionados que emplea la notación definida por esta Norma internacional es denominado un "mapa conceptual". Un mapa conceptual define un espacio conceptual multidimensional, donde las localizaciones son conceptos, y las distancias entre conceptos son apreciables a partir del número de conceptos intermedios que deben ser visitados si se quiere ir desde un concepto a otro, así como el tipo de relación que define la trayectoria de un concepto a otro, si la hay, a través de los conceptos intermediarios, si estos existen. Su estructura (nodo-enlace-nodo) es muy próxima a su equivalente en las redes hipertextuales y por consiguiente soportan la navegación de un modo muy natural. El uso de mapas conceptuales permite el desarrollo de mecanismos de representación y recuperación más robustos y eficaces, ya que las relaciones entre los conceptos se eligen teniendo en cuenta las necesidades y expectativas de cada usuario.

Cuando la representación se hace desde la posibilidad de crear una red de asociaciones de términos, basada en la co-ocurrencia con otros términos, hablamos entonces de los Mapas Lexicográficos. Una red gráfica que sitúa cada término en su interior por relación a los otros términos, precisando así la exactitud de su significado. Estos *mapas lexicográficos* son una prolongación de los métodos estadísticos.

#### **7.4. Tendencias en la gestión de imágenes digitales.**

Lo digital ha ampliado considerablemente la potencialidad del tratamiento y la difusión de otros soportes documentales más allá de la escritura. Muy especial, por la importancia social de su uso, ha sido el caso de las imágenes, para cuya recuperación es claramente perceptible una

orientación hacia sistemas visuales automáticos<sup>42</sup>: Se basan en la aplicación de funciones de similitud por atributos de color y estructura de objetos en la imagen<sup>43</sup>. Aún se ha investigado poco en el reconocimiento automático de objetos o de contenidos con un mayor nivel de abstracción, pues resulta muy difícil aplicar un proceso de información semántica de la imagen. Así las cosas, su diseño alcanza solo aún a una recuperación mediante caracteres *primitivos* (color, textura). Es muy difícil hacer recuperaciones por el nivel icónico, es decir, por clases de objetos como jarros, aviones, o personas, lo que hasta ahora se ha logrado tan sólo en aplicaciones restringidas al ámbito médico y militar. Los dos problemas pendientes de resolver en sistemas visuales algorítmicos son: generar un Vector de Características Gráficas para cada atributo de recuperación y generar un algoritmo de medida de similaridad entre Vectores de Características Gráficas. La interpretación es un proceso de atribución que relaciona lo percibido en la imagen con el estado de conocimiento del intérprete, por lo que depende de la percepción y el proceso cognitivo global, fenómenos ambos difícilmente automatizables. En el estado tecnológico actual no es posible el uso de técnicas de recuperación visual automática en sistemas e información en los que es preciso el empleo de atributos temáticos e icónicos durante la recuperación. Un sistema visual automático que pretenda la recuperación por atributos extrínsecos requiere una base de conocimientos que asocie la expresión de conceptos y nociones concretos (visualizables) con otros abstractos (no visualizables)<sup>44</sup>. Los sistemas de recuperación automática por conceptos se basan en asociaciones a través de la generación de una base de conocimiento que permite la sociación de conceptos a iconos o atributos plásticos derivados del contenido intrínseco.

Ante estas dificultades, la tendencia actual en el desarrollo de sistemas de recuperación de imagen se circunscribe a los sistemas visuales mixtos. Son los sistemas más adecuados por su capacidad de representar y recuperar por contenido, en distintos grados de abstracción. Integran un

<sup>42</sup> RASMUSSEN, E. Indexing Images, en *Annual Review of Information Science and Technology* -ARIST-, (1997), 32: 169-195.

<sup>43</sup> TREAT ORT, H. *et al.* -«Searching Images Using Ultimeidia Manager», en *Information Services & us.*,(1996), 16: 15-24.

<sup>44</sup> CARRIDONI, M., BIMBO, A. Del, y VICARIO, E. «Image Retrieval by Color Semantics with Incomplete Knowledge», en *Journal of the American Society for Information Science*,(1998), 49, 3.



sistema visual puro con un sistema lingüístico, muy útiles en las demandas de imagen que deben corresponderse a un tema, pero con la presencia de unos atributos visuales determinados, difíciles de definir. Así, la tecnología QBIC de IBM analiza y recupera imágenes, o sus fragmentos, a partir del color, textura, forma y posición del objeto gráfico. Este sistema está basado en Redes Neuronales, O el modelo el ART MUSEUM, que integra un modelo de imagen y otro de usuario, en el que la máquina aprende la correlación entre las imágenes y los términos que el usuario ha relacionado con esas imágenes. Es un sistema sólo aplicable a una colección muy pequeña y con un número reducido de usuarios.

#### **7. 4. Otras herramientas de indización automática:**

Una cuestión clave es cómo encontrar lo que buscamos en un hipermercado con las estanterías repletas. La información se recupera mediante motores de búsqueda de información y robots automatizados. Aquellos indizan los documentos existentes en la red y aportan punteros (según un ranking de relevancia sobre la adecuación de la respuesta a la petición hecha) que señalan URLs donde se incluyen los contenidos solicitados. Su funcionamiento se realiza a través de tres módulos: a) Módulo de *recopilación automática de recursos existentes en páginas web* (origina bases de datos con las palabras indizadas); b) Módulo de indización; c) Módulo de *búsqueda de información* e interfaz con el usuario. No podemos olvidar que esta búsqueda textual sigue un criterio básicamente morfológico, por lo que suele recuperar recursos de información valiosos junto a otros de baja calidad o de ninguna relevancia. Como motores de búsqueda generales debemos citar a Altavista, AOL, Excite, Infoseek, LookSmart, Lycos, Web Crawler; junto a otros especializados como NewsBot (buscador de noticias de HotBot) o su paralelo NewsTracker de Excite, Hyperlink (textos en páginas web), WeebSeer o Altavista Image (para imágenes), ForumOne (para grupos de discusión), o StudyWEB (centros académicos y de investigación) y otros muchos. Cada día aparecen más buscadores, en adaptación al enciclopedismo temático de la web.

Se incluirían aquí también los Índices temáticos que organizan los recursos jerárquicamente en función del esquema de clases establecido por el administrador del sistema (Yahoo, Magellan). La búsqueda en los índices temáticos puede realizarse: *a)* navegando por el árbol de categorías, descendiendo por las subcategorías hasta encontrar lo que se busca; *b)* buscando por palabras clave en una categoría determinada o en todas.

## 8. CONCLUSIONES.

1. Las técnicas bibliotecológico–documentales se ofrecen como camino evolutivo hacia la solución de los problemas planteados por la información digital.
2. Los lectores necesitarán durante mucho tiempo información proveniente a la vez de recursos impresos y de los digitales.
3. Hay dudas sobre el tiempo que permanecerá almacenada la información electrónica y sobre los problemas legales que plantea su uso. El cambio continuo en los soportes, en los estándares para la edición, en los navegadores e incluso en los editores plantea problemas técnicos y de normalización a la hora de tratar la información electrónica.
4. Se hace difícil ofrecer a los usuarios el mejor servicio, pues las complicadas soluciones tecnológicas cuestionan el aprovechamiento adecuado de los fondos.
5. Reconociendo que los sistemas de información en red permiten la disponibilidad de los recursos para muchos más usuarios, no podemos olvidar que el tamaño y diversidad de estos recursos han dado paso a una seria falta de calidad.
6. La ausencia de controles de calidad exige la presencia de determinados filtros que permitan distinguir los buenos productos informativos de aquéllos inútiles o perversos.
7. A ello viene a contribuir el Tratamiento Documental, cuyo objetivo esencial es facilitar el acceso a determinados documentos que potencialmente son necesarios en un determinado colectivo de usuarios, suministrando información representativa que les permita discernir sobre su utilidad.