

Cadena documental en entornos multimedia

Pablo Aguayo de Hoyos
Universidad de Alcalá de Henares
coco@gtasn1.ciemat.es

Susana López Canora
Ministerio de Defensa C.G.A.
zyrya@nexo.es

RESUM

Aproximació teòrica i pràctica a la gestió de la informació multimèdia, com s'estructura i algunes pautes per a la seva documentació. Es revisen les fases de la cadena documental en entorns multimèdia.

RESUMEN

Aproximación teórica y práctica a la gestión de la información multimedia, cómo se estructura y algunas pautas para su documentación. Se revisan las fases de la cadena documental en entornos multimedia.

Introducción

La incorporación del ordenador en los centros de documentación y el tratamiento digital de la información multimedia están modificando el proceso documental en diversos aspectos relacionados con la adquisición, análisis, almacenamiento y difusión de la documentación, originándose nuevos conceptos como el denominado documento multimedia.

Estos cambios requieren un tratamiento documental específico de los fondos acorde con las características de la información y los nuevos medios disponibles. De ésta forma podrá atenderse a un número creciente de usuarios que esperan satisfacer sus expectativas informativas de manera cada vez más exigente.

Las nuevas tecnologías de información digital constituyen una herramienta útil que posibilita la recuperación de la información, siempre y cuando se organicen y documenten los contenidos asegurando su calidad y accesibilidad.

1. Multimedia y tecnologías de la información

1.1 Conceptos

Tanto el concepto de multimedia como el de documento multimedia pueden estar sujetos a definiciones diversas respondiendo a matices distintos. Las siguientes citas ponen de manifiesto esta situación.

Domènech (1996:186) hace hincapié en la interrelación de contenidos desde el punto de vista del comunicador: «mensaje compuesto por diversos estímulos visuales y auditivos, los cuales, si actúan en equilibrada conjunción, facilitan una mayor comprensión de lo que se pretende comunicar».

Por su parte, Tramullas (1997: 139) prefiere potenciar el aspecto dinámico de los documentos multimedia, a los que denomina *documentos electrónicos hipermedia*. Según su punto de vista las páginas WEB son un claro exponente de estos ya que «no limitan sus funciones a meramente informativas, sino que son capaces de interactuar con el usuario y con herramientas clásicas de gestión de bases de datos, sirviendo como intermediarios».

A efectos prácticos, en este artículo no se entra en disquisiciones terminológicas y se adopta la solución de hacer uso del término más difundido, *documento multimedia*. El enfoque de este artículo se ajusta más a una definición integradora de los dos aspectos citados anteriormente:

– Entidad en soporte digital capaz de aglutinar un conjunto de informaciones, disponibles en diferentes formatos, complementarias e interactivas.

El concepto de *documentación multimedia* también puede formularse de múltiples maneras, una de las cuales podría ser:

- *Generación de nuevo conocimiento a partir del análisis de información textual, sonidos e imágenes y su gestión mediante herramientas informáticas.*

De los conceptos anteriores es interesante destacar las siguientes ideas:

- El usuario no es un elemento pasivo en el proceso de comunicación, tiene la capacidad de seleccionar la información a través de mecanismos de interactividad en un entorno de fácil manejo.
- Los documentos multimedia recogen informaciones relevantes y complementarias en un mismo soporte y suponen dos ventajas:
 1. Incorporan más información a una entidad; un documento puede componerse de imágenes, texto y sonidos.
 2. Amplían el valor de un documento mediante la asociación con cualquier objeto en fuentes distintas.
- Es necesario un análisis documental (formal y contenidos) y la obtención de productos documentales para facilitar la creación y recuperación de los documentos multimedia.

1.2. Características de la información digital

Toda la información que alberga un ordenador está codificada como ristas de ceros y unos, los famosos bits. Para esta representación se emplea la base numérica binaria o digital, de ahí el origen del concepto *información digital*.

Algunas de las características de la información digital se resumen a continuación (Codina, 1996):

1. Reusabilidad. Cualquier información digital es susceptible de ser reutilizada para ser actualizada o generar nueva información, como por ejemplo la elaboración de estadísticas a partir de una serie de datos de entrada.
2. Interactividad. El usuario puede seleccionar uno u otro contenido informativo de los que se le ofrecen, definir criterios de búsqueda, etc.
3. Recuperabilidad. Facilidad de encontrar la información de una forma rápida y selectiva.
4. Facilidad de difusión. El número de copias que puede difundirse de un documento digital es prácticamente ilimitado. (Codina prefiere denominar a esta característica virtualidad).

Además hay que citar otra característica no menos importante de los documentos multimedia: su enorme tamaño. Por ejemplo, un documento que contenga una imagen con mucha resolución precisa un gran espacio de almacenamiento con los consiguientes problemas de ocupación en un soporte limitado, velocidad para su correcto tratamiento y transporte por las actuales redes de telecomunicaciones.

1.3. El documento multimedia

La información que se presenta al usuario está dispuesta en un contenedor virtual común, de forma que éste recibe los datos sin tener que ocuparse de dónde se encuentran físicamente ni cual es la estructura que los soporta (Tramullas, 1997).

El contenido informativo del documento es responsabilidad de su autor, quién decide qué contiene cada entidad y cómo se presenta al usuario. Para la creación de los documentos multimedia suelen emplearse herramientas apoyadas en lenguajes de programación procedurales o de marcación como HTML.

Las herramientas clásicas empleadas para la confección de material multimedia se apoyan en lenguajes de programación como Visual Basic o C ++. En general, disponen de mecanismos para sincronizar las distintas informaciones en el tiempo (eventos) y el enlazado de objetos de forma fácil y dinámica. Esta particularidad los hace muy útiles para la creación de vistosas presentaciones multimedia, a cambio de un cierto grado de complejidad de programación. Típicamente se emplean en el desarrollo de aplicaciones en soporte CD-ROM.

Por su parte, los lenguajes de marcación de documentos son menos efectistas, pues en su origen fueron concebidos para describir documentos textuales, qué contienen y cómo se estructura su contenido. Actualmente se emplean para crear documentos multimedia; en cierta medida el éxito de Internet se debe a la facilidad del HTML para describir cualquier tipo de información multimedia y la forma en que ésta debe presentarse al usuario. Una

ventaja del HTML frente a los lenguajes de programación procedurales es su carácter abierto y universal. La complejidad de su uso depende del grado de elaboración y dinamismo de la información presentada.

Un ejemplo de documento multimedia es una página Web: puede integrar diversas informaciones, enlazarse con fuentes distribuidas y admitir acceso selectivo a sus contenidos.

2. Cambios en el proceso documental

Un Sistema de Gestión documental (SGD) captura, procesa, almacena y recupera información relacionada con el conocimiento cuya entidad básica es el documento. Su finalidad es la explotación de la información almacenada.

Sobre los documentos que alberga un SGD se pueden realizar una serie de operaciones que en conjunto se denomina *cadena documental* (Tramullas, 1995: 6). En ella se establecen a grandes rasgos las siguientes fases:

1. Adquisición de documentos interesantes para los usuarios procedentes de diversos soportes y formatos.
2. Análisis documental de las nuevas entradas conforme a su tipo y contenido.
3. Gestión de los datos con herramientas apropiadas.
4. Difusión de los fondos documentales.

En este proceso confluyen diversas tecnologías, algunas de las cuales están en evolución y por tanto no hay soluciones definitivas a muchas de las necesidades que se plantean actualmente, tales como:

- Estandarización de formatos y portabilidad. Es deseable que los documentos puedan ser distribuidos entre distintas plataformas de manera independiente a cómo se hayan generado.
- Información autocontenida. Los documentos deben formar parte de un interfaz de usuario que los trate como documentos compuestos por enlaces y objetos que sepan algo de sí mismos y de sus características. Por ejemplo, un documento gráfico debe poder desplegarse eficientemente en diversos equipos tomando información del entorno en que va a ser visualizado. La manera en que están comprimidos los datos, la calidad de presentación y otras informaciones pueden figurar como parte del documento en forma de metainformación y facilitar así su transmisión y visualización.

2.1. Cadena documental en un entorno multimedia

En el presente artículo se pretenden revisar las fases citadas anteriormente teniendo en cuenta las particularidades de un entorno multimedia. En la figura 1 se muestran de forma esquemática los componentes habituales de una cadena documental en entornos multimedia (Aguayo, 1997).

En la fila superior de la figura aparecen distintas tareas o eslabones de la cadena. En la segunda fila se relacionan los productos documentarios que constituyen la entrada o salida de cada tarea. La tercera fila representa el flujo de operaciones de forma esquemática. La última fila resume algunas de las herramientas informáticas usadas en cada fase.

Las columnas están asociadas a cada una de las operaciones. La primera etapa corresponde a la *digitalización* de los datos de entrada. Si es necesario realizarla se emplean dispositivos *hardware* especializados, como escáner, tarjetas digitalizadoras de audio o vídeo, etc. Una vez introducido el documento en el ordenador es objeto de tareas de maquetación y corrección, en su caso, de la información introducida.

Durante la segunda fase o *documentación* se van a crear documentos secundarios con los productos obtenidos del análisis de la información: descriptores, índices, resúmenes y etiquetas HTML/SGML. Algunos de los productos citados pueden obtenerse de forma automática mediante herramientas incluidas en el Sistema de Gestión de la documentación.

En el *almacenamiento* los documentos multimedia ya elaborados, los objetos que forman parte de ellos y sus descriptores, se almacenan en la base de datos.

La *recuperación* pasa por la elaboración de una ecuación que recoja las necesidades informativas y el seguimiento de una estrategia de búsqueda.

La última fase es la *difusión*, en ella se hace llegar la información a su destinatario para que decida la pertinencia de los documentos extraídos. El uso de interfaces gráficas, de programas de fácil manejo y capacidad de navegación sobre los documentos dinamiza la presentación de la información al usuario, se encuentre o no en el centro de documentación.

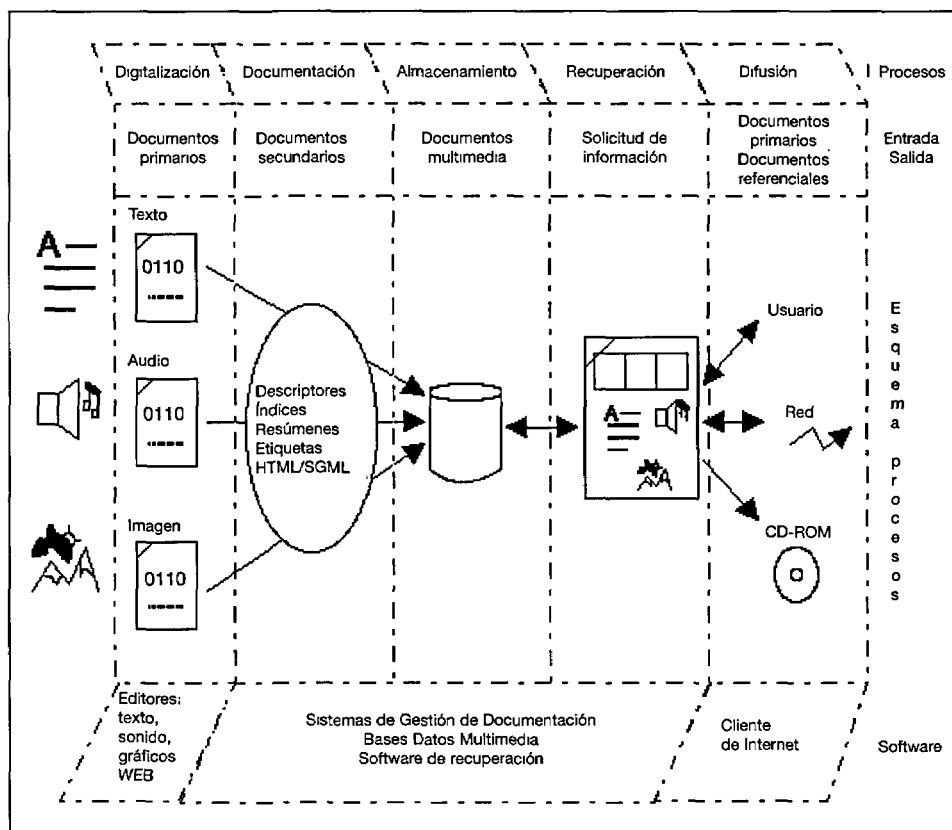


Figura 1. Cadena documental de entornos multimedia

A continuación se detallan más los elementos del proceso y se citan algunas repercusiones de ésta nueva panorámica sobre centros y documentalistas.

2.1.1. Entrada de datos

La adquisición de documentos se realiza a través de una serie de dispositivos *hardware* en combinación con *software*, que lo introduce en el disco duro. Los documentos que van a entrar a formar parte de la base de datos son en esencia de los tipos siguientes: textos, sonidos, imágenes y documentos electrónicos.

Textos

Escritos en soporte papel o digital. El objetivo es generar un documento electrónico que pueda ser utilizado por un procesador de textos y realizar sobre él las operaciones típicas de edición textual. Los medios empleados suelen ser un escáner en combinación con un programa para reconocimiento de caracteres, OCR, que realiza el reconocimiento uno a uno de los caracteres que componen el documento original y lo convierten a texto ASCII. Presentan ciertos problemas en función de la calidad del texto y de lo potente que sea el programa de reconocimiento.

Sonidos

Proceden de música grabada, voz, efectos sonoros o música sintética que son registrados en los típicos soportes de audio. El objetivo de su digitalización es disponer de ellos en formato electrónico. Algunos de estos documentos pueden manipularse como si se tratara de texto con editores de ondas sonoras, incorporando efectos, seleccionando parte de los mismos, etc. Los medios empleados suelen ser tarjetas digitalizadoras de audio. Estos dispositivos permiten captar una onda sonora con diferentes grados de calidad y generan ficheros de mayor o menor tamaño en función de los parámetros usados durante la digitalización.

Imágenes

Los documentos de imágenes están compuestos por fotos, animaciones, vídeo o gráficos generados con un ordenador. El tratamiento de las imágenes consiste en su digitalización para crear documentos gráficos o como

parte integrante de otros documentos. La información pasa a ser un fichero que representa la imagen mediante bits. Las operaciones más corrientes sobre estos ficheros consisten en visualizarla y editarla con programas gráficos que permiten la ampliación, selección de una parte, retocado, etc. Los medios típicos empleados son el escáner, tarjetas digitalizadoras de vídeo y programas gráficos. Normalmente dan lugar a ficheros de gran tamaño, por lo que a veces se requiere el uso de algoritmos de compresión/descompresión rápidos.

Documentos electrónicos

Los documentos digitales se obtienen a partir de:

- Consultas a bases de datos en línea,
- Servicios de Internet: mensajes electrónicos, páginas WEB, resultado de una búsqueda o ficheros de todo tipo,
- Soporte digital o magnético.

Es posible que el formato en el que están disponibles no sea el mismo que vamos a emplear en nuestro SGD, por lo que será necesario la traducción de formatos.

2.1.2. Documentación

La gestión de la documentación se realiza básicamente desde dos perspectivas bien diferenciadas: la técnica y la semántica. En la gestión técnica se consideran aspectos relativos a la representación física y manipulación de la información, incluyendo técnicas de almacenamiento, accesibilidad y representación. El componente semántico de la información es el responsable de reflejar la precisión con la cual la información almacenada expresa el contenido de los documentos y la eficacia con la cual los usuarios pueden satisfacer sus necesidades de información (López Canora, 1995).

Gestión semántica de la documentación

Para poder gestionar esta información hay que dotar al fichero de descriptores que identifiquen y representen el contenido del documento para la posterior recuperación del mismo.

Textos

En la documentación de un texto los descriptores se construyen con todo o parte del texto del documento.

Sonidos

Los documentos generados suelen documentarse en función de su tipo, así una pieza musical grabada queda en general perfectamente documentada por los créditos que figuran en el soporte original, además de una catalogación por estilo, etc. En el caso de ser una grabación de voz, la información que acompañe al fichero puede contener datos de la persona que habla, de qué habla, fecha de la grabación, etc.

Imágenes

La documentación de imágenes tiene el riesgo de dejar en manos del documentalista la introducción de información subjetiva (Valle, 1994). Se hace necesario por tanto el uso de lenguajes documentales específicos para la construcción de descriptores textuales que tengan en cuenta dos aspectos fundamentalmente:

- Las características técnicas de la imagen.
- El contenido de la imagen, su significado.

Además hay que añadir los datos de identificación del documento (autor, título, edición, etc.) que pueden ser catalogados mediante normas habituales.

Páginas Web

La documentación se realiza asociando descriptores de localización y del contenido. La notación empleada para la construcción de descriptores de localización es la utilizada para ubicar cualquier servicio o documento di-

gital en la red, su URL (Uniform Resource Locator). Para la descripción de una página Web pueden emplearse las etiquetas HTML que informan sobre su contenido, (meta etiquetas o etiquetas <META>). Éstas marcas se recogen en el estándar a partir de la versión 3.2 y proporcionan información útil para los robots o indizadores de páginas de la red. Algunas de ellas recogen información relacionada con el título, autor, resumen y palabras clave.

Existen iniciativas en marcha para enriquecer el conjunto de meta-etiquetas, de manera que puedan incluirse entre otras la descripción del tipo de documento que se describe (tesis, artículo, página WEB, gráfico, etc.), su fecha de creación, etc.

2.1.3. Almacenamiento: gestión técnica de la documentación

La gestión técnica de la documentación va a apoyarse en la utilización de bases de datos y en la tecnología de hipermedios (Codina, 1996).

Bases de datos

En esencia estas deben permitir la creación de índices y la rápida recuperación de la información que contienen, además del control de acceso y seguridad. Las bases de datos usualmente empleadas en documentación multimedia disponen funciones documentales junto a la capacidad de gestionar objetos de cualquier tipología.

Hipermedios

Los hipertextos son documentos en los que aparecen enlaces que permiten el acceso de una forma no lineal a la información. Un documento no tiene por qué leerse de principio a fin, sino que puede saltarse de una parte a otra a través de los enlaces. Esta manera de relacionar unas informaciones con otras es comparable a la forma que tenemos los humanos de asociar ideas en el cerebro. De esta forma el documento puede considerarse interactivo, al ser el usuario capaz de tomar decisiones acerca de la forma en que va a consultar la información. Con ello se sugiere cierta idea de navegación por los documentos.

Los hipermedios son una extensión de los hipertextos; tienen la capacidad de establecer enlaces dinámicos entre cualquier objeto digital: imágenes, sonido, programas, etc.

2.1.4. Recuperación de la información

La recuperación de la información en un entorno multimedia no es diferente de la recuperación en un entorno textual tradicional, si se asignan de forma correcta los descriptores a cada tipología de documentos. El problema surge por la ausencia de estándares para la formalización de dichos descriptores: los tesauros publicados son muy restringidos y no existe un lenguaje normalizado para la realización de un análisis semántico de los documentos (Azorín, 1994).

Por otra parte, existen nuevas demandas informativas que los sistemas de gestión documental basados en descriptores textuales son incapaces de satisfacer. Por ejemplo, los sistemas que almacenan información gráfica y que admiten consultas a través de mecanismos de interrogación por imagen. Existen ya en el mercado algunos de estos sistemas que están sustentados por técnicas de reconocimiento de imágenes o patrones gráficos entre los documentos. De ésta manera pueden solicitarse de una base de datos imágenes mediante sus atributos de textura, forma o color.

Los sistemas de recuperación de la información presentan ciertos problemas. En el caso de mal uso de la tecnología de hipertextos puede darse el denominado «desbordamiento cognitivo» o fatiga de información. Este se produce ante un exceso de información ya que la mente humana es capaz de asimilar sólo una cierta cantidad. Otro problema es la desorientación a la que puede conducir la navegación entre los documentos enlazados. La solución a estos problemas se relaciona, con una modificación de los hábitos del usuario (más cantidad no es siempre sinónimo de mayor información) y con la utilización de herramientas que informen acerca de qué información se está visualizando y cómo se ha llegado hasta ella.

Si la verdadera medida de la multimedia reside en la interactividad entre usuarios e información, es evidente que cada vez deberán estar más a su alcance las fuentes informativas sobre las que realizar consultas. En Internet esto se traduce en la recuperación directa, permitiendo el acceso remoto a bases de datos de cualquier centro conectado.

2.1.5. Difusión

Los cambios se orientan hacia una difusión más rápida y barata de los documentos mediante soportes como el CD-ROM y la red Internet. Por otra parte, las características de estos soportes especializan el tipo de documentos que se difunden de una u otra forma (Martín, 1997).

El almacenamiento en CD-ROM permite la escritura de documentos en un espacio muy reducido, sin pérdida de datos y con una relación precio/rendimiento muy alto. Por ello, se emplean habitualmente para la difusión de documentos que no precisan de una actualización inmediata: obras de referencia, catálogos, bases de datos, etc.

Por su parte, las redes de telecomunicaciones son idóneas para atender a las necesidades de información de documentos perecederos, por la posibilidad de una actualización inmediata. Pero las redes aún presentan ciertos problemas relacionados con la capacidad de la red telefónica actual, los costes de utilización de la misma, etcétera.

3. Un caso práctico de la cadena documental en multimedia

El objetivo de este apartado es presentar un SGD como ejemplo de cadena documental en un entorno multimedia, según se ha descrito en el apartado anterior.

El sistema de gestión documental tomado como ejemplo es un sistema informático concebido para poner en contacto a suministradores de productos, bienes y equipos con entidades consumidoras del sector de la construcción. La información incluida en el sistema está compuesta por documentación técnica relativa a productos e información gráfica de gran calidad, imprescindible en un sector como el de la edificación donde la estética es un objetivo primordial.

Los suministradores de productos, bienes o equipos son las entidades que generan información documental usando el sistema como medio de difusión para hacerla llegar a las entidades usuarias del sector. Para ello exponen sus catálogos de productos en forma de documentos multimedia. Los suministradores de servicios, tales como instaladores, servicios técnicos o profesionales aportan al sistema información referencial, es decir, datos identificativos de su actividad y opcionalmente un informe sobre su trayectoria profesional. El sistema dispone por tanto de dos motores de recuperación, uno para información documental y otro para datos referenciales.

Los objetivos del presente sistema de gestión documental son:

- Control de calidad de la información: características técnicas y su vigencia.
- Difusión de la información independientemente de la localización geográfica de los usuarios y servicios.
- Recuperabilidad de la información, de forma sencilla y eficaz, basada en la aplicación de estándares de catalogación vigentes en el sector de la construcción.
- Interactividad entre el sistema y los usuarios, en un entorno intuitivo en el que sea posible seleccionar y definir criterios de búsqueda o realizar consultas dirigidas por índices temáticos.

Las siguientes secciones detallan las fases de la cadena documental: entrada de datos, documentación, almacenamiento, recuperación y difusión.

3.1. Entrada de datos

En esta fase se realiza la transformación de la documentación facilitada por los productores de información en documentos multimedia, compuestos por textos e imágenes, en un formato apropiado para su posterior tratamiento automático y clasificación de acuerdo a índices temáticos.

Los productores de información entregan su documentación en soporte papel y/o magnético. En el primer caso, mediante la utilización de un escáner se digitaliza la información, convirtiendo los textos en ficheros formateados tipo RTF con ayuda de un programa de reconocimiento de caracteres (OCR) y los gráficos a ficheros tipo TIFF. Ambos formatos son estándares *de facto* para la mayoría de las aplicaciones de edición.

Con la información gráfica y textual en soporte digital se inicia la composición del documento multimedia, marco integrador de los distintos tipos de información. En este sistema el documento multimedia se compone de dos secciones: una cabecera y un cuerpo. La cabecera está formada por una colección de datos de referencia que identifican unívocamente al documento en el sistema y permiten clasificarlo a efectos de consultas referenciales. El cuerpo recoge el texto completo, las imágenes y los enlaces o vínculos con otros documentos.

La entrada de datos al sistema termina con la maquetación de textos e imágenes definiendo así el aspecto final del documento multimedia.

3.2. Documentación

La fase de documentación consiste en un proceso de análisis y clasificación de los documentos mediante el cual se le asignan descriptores, que básicamente identifican el contenido del documento para su posterior recuperación.

El sistema dispone de información relativa a tres tipos de entidades: productos, suministradores de productos y/o servicios y agentes de construcción (consumidores de información). Cada tipo de entidad desarrolla una función distinta en el sector, por tanto sus atributos identificativos también son diferentes.

La información recogida en el sistema es analizada conceptualmente y sometida a un proceso de clasificación. De ésta manera se puede determinar a qué tipo de entidad hace referencia y qué información contiene sobre ella. Un documento contiene información relativa a una sola entidad: producto, suministrador o consumidor. Los criterios de ordenación adoptados en cada caso son:

- **Suministradores.** Se clasifican en función de los productos con los que trabajan y su relación con ellos. Así se tienen fabricante, distribuidor, instalador, servicio técnico, oficios, etc.
- **Consumidores.** Se organizan atendiendo a la actividad profesional desarrollada en el sector de la construcción. Están asociadas a organismos públicos, colegios oficiales, constructoras, promotoras e inmobiliarias, consultoras e ingenierías y estudios de decoración.
- **Productos.** Los productos se organizan en el sistema por atributos identificativos y descriptivos. Los atributos identificativos determinan el tipo de producto y el suministrador. Para la clasificación de productos se emplea un diccionario.¹ Los términos utilizados en los descriptores son técnicos, normativos y legislativos, que permiten diferenciar calidades y cualidades de los materiales.

3.3. Almacenamiento

La gestión técnica de la documentación se realiza con el programa Lotus Notes, marca registrada de Lotus Development Corporation. Una base de datos en Lotus Notes está constituida por un fichero que contiene tres tipos de objetos: vistas o índices, documentos y formularios.

- Las vistas son índices dinámicos de los documentos del sistema, permitiendo presentar los contenidos con diferentes niveles de confidencialidad en función del nivel de acceso de los usuarios.
- Los formularios controlan el modo en que la información debe introducirse, cómo será visualizada por el usuario y el formato de impresión.
- La captura de información a través de un formulario proporciona un documento, que puede contener datos textuales (numéricos o fechas), imágenes digitalizadas, mensajes sonoros o vínculos con otros documentos.

3.4. Recuperación

La recuperación de la información en el SGD se basa en un modelo que combina la recuperación por atributos identificativos y por atributos semánticos. El proceso consta de dos fases: selección del dominio de consulta y búsquedas de texto complejas en el dominio seleccionado.

3.4.1. Selección del dominio de consulta

La primera fase del proceso de recuperación determina el tipo de entidad sobre la que se desea obtener información. Para ello el sistema dispone de un motor de recuperación basado en índices que visualizan los documentos agrupados y ordenados por sus atributos identificativos. Por ejemplo, los documentos relacionados con «Iluminación de interiores», aparecen bajo la categoría «Edificación» y en la subcategoría de «Iluminación».

1. Elaborado por el Instituto Valenciano de Edificación.

3.4.2. Recuperación de documentos

El modelo de recuperación empleado en las búsquedas complejas se basa en la creación de un fichero invertido extendido. Este fichero es el resultado de un proceso de indexación automática, que incluye todos los términos que aparecen en cualquier campo de la base de datos con información de la frecuencia relativa del término y además el párrafo, la frase y la posición relativa en la que aparece.

Una vez creado el índice de texto es posible realizar búsquedas complejas. Una búsqueda compleja define un conjunto de criterios de selección que permiten localizar documentos que contengan palabras, frases, números, fechas, etc. Pueden estar formadas por una única palabra, frase o incluir funciones avanzadas para precisar la búsqueda. Las funciones avanzadas disponibles en Lotus Notes son: uso de caracteres comodín, palabras con guiones, frases entre comillas, operadores lógicos, operadores de proximidad, operadores de campo, asignación de pesos a los términos de búsqueda, etc.

Las ecuaciones de búsqueda se introducen en un cuadro de diálogo con funciones de ayuda, por lo que se libera al usuario de la necesidad de aprender los operadores y la sintaxis de uso de los mismos. El resultado de una búsqueda se presenta como una lista de documentos ordenados por la frecuencia relativa (ranking) o cronológicamente. Si se desea precisar más la búsqueda es posible formular otras consultas sobre el conjunto de documentos previamente seleccionados.

El acceso al documento desde las vistas es directo, basta con seleccionarlo para que el sistema visualice el documento íntegro con los términos de la búsqueda resaltados sobre el texto.

3.5. Difusión

Cualquier usuario del sistema puede disponer de una copia local de la base de datos, imprimir documentos, establecer contactos mediante correo electrónico, etc. También es posible realizar conexiones remotas al servidor donde reside el SGD.²

4. Tendencias

4.1. Agentes de conocimiento

La explosión informativa a la que estamos asistiendo nos obliga a emplear mecanismos de filtrado y organización de los documentos procedentes de fuentes distribuidas. Una de las herramientas que se perfila como indispensable a corto plazo es el empleo de un *agente personal de información*. Este tipo de programas conoce los intereses informativos del usuario y actúa de forma autónoma en su nombre. Inicialmente se alimenta con la estrategia de búsqueda definida por el usuario y posteriormente emplea el razonamiento con el que ha sido diseñado para modificarla en función de la pertinencia de los documentos entregados.

Con ésta tecnología un usuario puede personalizar sus preferencias de búsqueda en diferentes categorías mediante el uso de distintos perfiles. El agente se encargará de filtrar la información y proporcionársela con la periodicidad y forma de presentación que haya acordado con el usuario. Así éste puede visualizar cada cierto tiempo el resultado de la búsqueda a través de una página Web que contiene enlaces a los documentos o abrir una carpeta en su ordenador con los documentos clasificados temáticamente. (Chorafas, 1994; Jara, 1997).

4.2. Bases de datos Inteligentes

La efectiva explotación de los SGD en entornos multimedia está relacionada con su capacidad para solventar preguntas complejas que se caracterizan porque:

- Son polivalentes, pueden emplearse los mismos criterios de búsqueda en entornos diferentes.
- Su respuesta es probabilista, no responde con certeza a nuestros requerimientos ya que los documentos obtenidos pueden no satisfacer nuestra demanda.
- La respuesta es multimedia, proporcionan información acerca de lo que se busca en formatos diferentes.

² El sistema de gestión documental descrito es un servicio privado en explotación ofrecido por el Colegio de Arquitectos de Madrid mediante suscripción.

Las bases de datos inspiradas en la tecnología de orientación a objeto pueden diseñarse con un cierto grado de conocimiento. Los modelos de datos (objetos) pueden enriquecerse con información semántica y dotar a los métodos que operan sobre los objetos de conocimiento. El objetivo puede ser, por ejemplo, presentar la información en contextos diversos o realizar operaciones distintas en función de determinados perfiles de usuario.

Los futuros sistemas de información tendrán que asumir el carácter impreciso de los procesos y datos que dominan el mundo real si quieren representarlo con mayor precisión. Una tecnología prometedora para solventar estos nuevos retos en el ámbito documental es la Lógica Borrosa o Difusa (*Fuzzy Logic*). La Lógica Difusa es una disciplina de la Inteligencia Artificial que puede englobarse dentro del área de «razonamiento con incertidumbre» y que se asemeja a la toma de decisiones del razonamiento humano, basándose más en el procesamiento de la información lingüística (aproximada e imprecisa) que en un procesado numérico.

En la confluencia de la tecnología de objetos y la ingeniería del conocimiento pueden estar las claves para la construcción de repositorios o sistemas de almacenamiento inteligentes (KBMS), capaces de explotar el acceso a la información distribuida y multimedia de una forma más eficiente (Chorafas, 1994).

Conclusiones

El objetivo en entornos multimedia no difiere del tradicional en todo proceso documentario: facilitar la identificación y localización de documentos de todo tipo. Un documento multimedia no es sinónimo de información de calidad por el mero hecho de contener textos, sonidos e imágenes en un mismo soporte. El valor añadido que supone disponer de mucha información, más rica por la interrelación de conceptos, rápidamente accesible y de mayor alcance, puede alcanzarse mediante el uso de herramientas informáticas y la correcta documentación y catalogación de la información.

Bibliografía

- AGUAYO, Pablo. (1997). Tema 1 de los apuntes de clase de la *asignatura Documentación Multimedia: Técnicas de Edición Contemporánea*. Curso 96-97. Alcalá de Henares: Universidad de Alcalá de Henares. <<http://www.cc.alcala.es/pablo/mmedia/tema1.html>>.
- AZORÍN, Virtudes; LÓPEZ BRUÑET, Trinidad. (1994). «Problemas que presenta la descripción de imágenes en la elaboración de bases de datos multimedia». En: *Actas de las IV Jornadas Españolas de documentación Automatizada. Documat 94*. Oviedo: Servicio de Publicaciones de la Universidad de Oviedo. P. 411-416.
- CODINA, Lluís. (1996). *El llibre digital. Una exploració sobre la informació electrònica i el futur de l'edició*. Barcelona: Centre d'Investigació de la Comunicació.
- CHORAFAS, D. N. (1994). *Intelligent multimedia databases*. Englewood Cliffs: Prentice Hall.
- DOMÈNECH RIERA, Eduald. (1996). «Información Multimedia». En: *Informe Fundesco sobre tendencias en Multimedia*. Madrid: Fundesco, p. 185-189.
- JARA, L. (1997). «Difusión selectiva de la información: clipping a través de Internet e Intranet». En: *Actas del II Congreso Nacional de Usuarios de Internet e Infovía*. Madrid. p. 73-76.
- LÓPEZ CANORA, Susana. (1995). *Modelos de recuperación documental. Aplicación para el sector de la construcción*. Madrid: Universidad Politécnica de Madrid. [trabajo fin de carrera].
- MARTÍN CARBAJAL, Francisco. (1996). «El sector de los servicios de información electrónica». En: *Informe Fundesco sobre tendencias en Multimedia*. Madrid: Fundesco, p. 79-83.
- TRAMULLAS SAZ, Jesús. (1995). «Una introducción a la informática documental». En: *Apuntes CCUZ*. Zaragoza: Universidad de Zaragoza. Nº 8, p. 6-10.
- TRAMULLAS SAZ, Jesús. (1997). «Recuperación de la información, planteamiento, herramientas y prespectivas». *Anuario SOCADI de Documentación e Información/1997*. Barcelona: Societat Catalana de Documentació i Informació, p. 137-145.
- VALLE GASTAMIZA, Félix del. (1994). «El análisis documental de la fotografía». En: *Cuadernos de documentación Multimedia*. Madrid: Universidad Complutense de Madrid. Nº 3, p. 43-56.