

Bibliotecas digitales: concepto y principales proyectos

CRISTINA FABA PÉREZ

Universidad de Extremadura, Facultad de Biblioteconomía y Documentación
Badajoz, España Tel. (924) 25 99 10; Fax (924) 25 99 57
E-mail: cfaber@alcazaba.unex.es

FÉLIX DE MOYA ANEGÓN

Universidad de Granada, Facultad de Biblioteconomía y Documentación
Granada, España, Tel. (958) 24 39 43; Fax (958) 24 39 45
E-mail: felix@goliat.ugr.es

RESUMEN

Actualmente las bibliotecas digitales constituyen un campo de investigación muy activo en todo el mundo. Como parte de esta labor, se están desarrollando y evaluando numerosos proyectos nacionales e internacionales de bibliotecas digitales. Seis de los proyectos más avanzados y representativos conforman la iniciativa DLI (Digital Library Initiative), fundada conjuntamente por NSF (National Science Foundation), NASA (National Aeronautics and Space Administration) y DARPA (Department of Defense Advanced Research Projects Agency) de los Estados Unidos. Estos proyectos están trabajando en la próxima generación de herramientas para el análisis, recuperación, gestión y localización de información. Partiendo de esta realidad, el objetivo de este trabajo es doble: por una parte pretende clarificar e identificar el concepto de biblioteca digital y por otra, introducirse en el análisis de cada uno de los proyectos que integran la iniciativa DLI.

THE CONCEPT AND PRINCIPAL PROJECTS OF DIGITAL LIBRARIES
CRISTINA FABA-PÉREZ
FÉLIX DE MOYA-ANEGÓN

ABSTRACT

Research into the form and function of digital libraries is growing world-wide. As part of this trend, numerous international digital library projects are under development and evaluation. Among the most advanced and representative of these is the Digital Library Initiative (DLI) founded jointly by the National Science Foundation (NSF), the National Aeronautics and Space Administration (NASA) and the Department of Defense Advanced Research Projects Agency (DARPA). These projects are currently working on the development of the next generation of analysis, retrieval, management and addressing tools. The thrust of this paper is twofold: first to identify and clarify the concept of the digital library and second to analyze the DLI projects.

Artículo

Trabajo recibido
el 24 de marzo
de 1999

*

Trabajo aceptado
el 26 de abril de
1999

CONCEPTO Y TEMAS DE INVESTIGACIÓN EN BIBLIOTECAS DIGITALES

¿Qué es una biblioteca digital?

Durante los últimos años, el concepto de biblioteca digital ha llegado a ser tan ambiguo como la mayoría de los conceptos relacionados con las tecnologías de la información. Como prueba de ello observamos, por una parte, la variedad de calificativos que coexisten para acompañar al mismo término (biblioteca digital, virtual, electrónica, abierta, etc.); y por otra, los diferentes significados atribuidos a este concepto, que van desde la digitalización de colecciones e informatización de bibliotecas tradicionales hasta la publicación electrónica de cualquier texto.

Tanto la definición como las funciones que asumen las bibliotecas digitales varían entre los diferentes autores. La descripción más generalizada del concepto se estableció en el Workshop on Distributed Knowledge Work Environments, celebrado en Santa Fe, Nuevo México, en 1997 (Gross, 1998):

El concepto de biblioteca digital no es sólo equivalente a una colección digitalizada con herramientas que gestionan información. Más bien es el conjunto de colecciones, servicios y personas que participan en el ciclo vital completo para la creación, difusión, uso y conservación de datos, información y conocimiento.

Para B. Schatz y H. Chen (Schatz, 1996):

Las bibliotecas digitales almacenan principalmente materiales en formato electrónico y gestionan eficazmente grandes cantidades de dichos materiales. Así, investigar en bibliotecas digitales significa realmente investigar en sistemas de redes de información. Los aspectos clave de la tecnología se centran en cómo buscar y mostrar las selecciones solicitadas desde y a través de las grandes colecciones. Mientras que la práctica en bibliotecas digitales debe enfocarse hacia temas relacionados con los costes de acceso y la tecnología de digitalización, la investigación se centra en cómo desarrollar la infraestructura necesaria para la gestión masiva y eficaz de información en la Red.

A este respecto, algunos especialistas en redes consideran que Internet es la biblioteca de la Era Digital. Para corroborar su afirmación, establecen una comparación entre los índices-WWW y los catálogos de una biblioteca tradicional argumentando que la finalidad de ambos sistemas es la localización de información. Esta teoría es ampliamente criticada por C. Lynch (Lynch, 1997) quien considera imprescindible la cooperación entre bibliotecarios e informáticos para construir bibliotecas digitales que asuman las funciones tradicionales de una biblioteca (desarrollo, organización, acceso y conservación de colecciones) y que sean capaces de organizar el “caos” llamado Internet:

A veces se describe Internet como la biblioteca mundial de la Era Digital. Esta descripción no debe prevalecer ni siquiera en un examen superficial. Internet –y particularmente su colección de recursos multimedia conocida como World Wide Web- no fue diseñada para llevar a cabo la publicación y recuperación organizada de información, como hacen las bibliotecas. Ha evolucionado hacia lo que podría ser la idea de un archivo documental caótico para la producción colectiva de la imprenta digital mundial. Este depósito de información contiene no sólo libros y artículos, sino datos científicos originales, menús, anuncios, grabaciones de audio/vídeo y conversaciones interactivas o transcritas. Material efímero mezclado por todo el mundo con trabajos de importancia permanente.

Aunque, como acabamos de ver, son varias las tendencias relacionadas con la conceptualización de las bibliotecas digitales, K. Drabenstott (Drabenstott, 1994) establece los elementos comunes que caracterizan a todas ellas:

- ❖ Son entidades complejas;
- ❖ Requieren el uso de una amplia tecnología para reunir los recursos de información distribuidos a través del mundo;
- ❖ Las conexiones que establecen con los servicios de información son transparentes para el usuario final;
- ❖ Su objetivo es el acceso universal a las bibliotecas y servicios de información;
- ❖ Sus colecciones incluyen productos digitales que no pueden representarse o distribuirse en formatos impresos.

¿Cuáles son los principales temas de investigación analizados en los proyectos de bibliotecas digitales?

En 1995, C. Lynch y H. García-Molina (Lynch, 1995) organizaron el primer Congreso de IITA¹ dedicado específicamente a las “Bibliotecas Digitales”. Los asistentes al mismo las analizaron desde perspectivas diferentes y coincidieron en los siguientes aspectos como los principales temas de interés en la investigación sobre bibliotecas digitales: interoperatividad, descripción de objetos y archivos documentales, gestión y organización de colecciones, interfaces de usuario e interacción hombre-máquina y aspectos legales/sociales/económicos.

Interoperatividad

Generalmente cuando hablamos de la infraestructura de tecnologías, protocolos y herramientas que mantienen y facilitan el crecimiento de las bibliotecas digitales en Internet, estamos hablando de interoperatividad.

Sin embargo, definir la interoperatividad de las bibliotecas digitales no es tarea fácil y puede dar origen a varias interpretaciones. Los participantes del Congreso IITA establecieron dos tendencias claramente divergentes:

1 U.S. Government's Information Infrastructure Technology and Applications.

- ❖ Interoperatividad centrada en el uso de herramientas e interfaces comunes que suministran uniformidad para la navegación y el acceso, pero basada en la inteligencia humana para conseguir cualquier coherencia de contenido;
- ❖ Interoperatividad semántica o inter-dominio que se ocupa de la capacidad del usuario para acceder, de forma consistente y coherente, a clases similares de objetos y servicios digitales distribuidos a través de depósitos documentales heterogéneos, mediante el uso de un software que actualiza las variaciones encontradas en cada una de las localizaciones analizadas.

Descripción de objetos y archivos documentales

Como señala la *Figura 1*, R. Kahn y R. Wilensky (Arms, 1995) definieron una serie de conceptos clave alternativos para describir la arquitectura de un sistema de biblioteca digital.

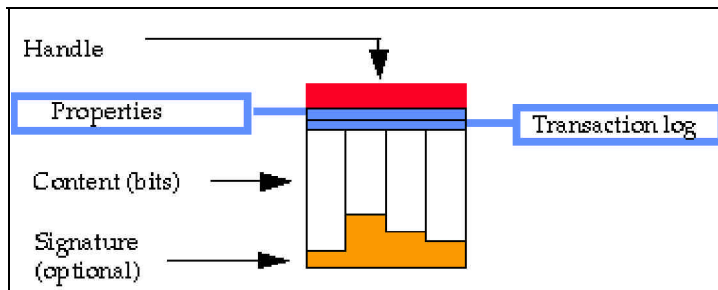
Figura 1
Terminología en la arquitectura de biblioteca digital

| TERMINOLOGÍA USADA | TERMINOLOGÍA KAHN/ WILENSKY |
|----------------------------------|--------------------------------|
| Fondos de la biblioteca digital | Objetos digitales |
| Información acerca de los fondos | Propiedades y metadatos |

Los **objetos digitales** representan una unidad fundamental a la hora de definir la arquitectura de la biblioteca digital. Sus principales características, tal y como se reflejan en la *Figura 2*, son las siguientes:

- a) Están constituidos por su propio contenido (conjunto de bits), las propiedades o metadatos asociados a él y un registro de las operaciones realizadas
- b) Se almacenan en archivos o depósitos documentales
- c) Se identifican mediante etiquetas o identificadores

Figura 2.
Identificadores y partes del objeto digital (Arms, 1995)



Los **identificadores** constituyen un bloque vital en la biblioteca digital. Se precisan, no sólo para identificar los objetos digitales sino también para recuperar la información y para establecer relaciones entre los objetos, y son independientes de la localización en la que se almacena el objeto digital.

La investigación en este área se basa fundamentalmente en tres aspectos:

- ❖ Definición y uso de metadatos, así como su captura desde los objetos digitales;
- ❖ Integración de archivos documentales heterogéneos con semánticas diferentes;
- ❖ Desarrollo de algoritmos para la clasificación y evaluación automática de ciertas propiedades de la información como son la calidad o el género.

Gestión y organización de colecciones

Algunos de los problemas principales de la gestión y organización de las bibliotecas digitales se centran en el diseño de políticas y métodos de control adecuados para incorporar los recursos informativos a la Red (Internet plantea serios problemas relacionados con la autoridad y calidad de los contenidos en las bibliotecas digitales).

Otros tienen que ver o bien con la captura, almacenamiento, indexación y recuperación de información multimedia (cuyo contenido es cada vez mejor gestionado en el entorno de las bibliotecas digitales), o con la necesidad de clarificar las funciones que los bibliotecarios deben desempeñar en la definición y gestión de colecciones digitales.

Finalmente, los problemas implicados en la conservación de los contenidos digitales durante largos períodos de tiempo (teniendo en cuenta las múltiples y variables generaciones de tecnologías hardware y protocolos software), a pesar de ser esenciales para el desarrollo eficaz de las bibliotecas objeto de estudio, no han constituido un tema de investigación principal en ninguno de los proyectos DLI.

Interfaces de usuario e interacción hombre-máquina

Podemos agrupar las principales tareas que se cuestionan los proyectos de bibliotecas digitales en relación con los interfaces de usuario e interacción hombre-máquina en las siguientes categorías:

- ❖ Comprender las necesidades, objetivos y comportamientos de los usuarios en el empleo de sistemas de bibliotecas digitales;
- ❖ Visualizar grandes colecciones de información y navegar por ellas utilizando herramientas de análisis adecuadas;
- ❖ Utilizar un interfaz de usuario apropiado;
- ❖ Adaptar los sistemas de bibliotecas digitales a los diferentes anchos de banda de las conexiones de red y estaciones de trabajo.

En este sentido, todos los proyectos DLI trabajan en el desarrollo de nuevas interfaces de usuario.

Aspectos legales/sociales/económicos

Las bibliotecas digitales no son simplemente tecnología. Están inmersas en un contexto legal, social y económico amplio y variado en el que tendrán éxito sólo en la medida en la que satisfagan las necesidades planteadas por este entorno.

Los principales problemas que se investigan en este área se relacionan con el diseño de modelos económicos para el uso de información electrónica y sus sistemas de facturación; los derechos de autor, la autenticidad e integridad de los datos; y otros aspectos íntimamente ligados con la gestión y el desarrollo de bibliotecas virtuales.

Además, a los anteriores problemas hay que sumar la diferencia legislativa que existe entre los países en materia de derechos de autor y depósito legal, diferencia que puede impedir el intercambio de los modelos de biblioteca digital entre los diferentes países.

PRINCIPALES PROYECTOS DE BIBLIOTECAS DIGITALES

Existe un número considerable de proyectos relacionados con las bibliotecas digitales que las analizan desde diferentes puntos de vista (véase, como ejemplo, la lista de proyectos americanos, europeos o asiáticos que la IFLA ofrece en la dirección electrónica <http://www.ifla.org/II/diglib.htm>). En muchos de ellos se nota una clara tendencia a identificar estas bibliotecas simplemente con la digitalización de las colecciones que constituyen la biblioteca tradicional, otros forman parte o están muy relacionados con el desarrollo de proyectos más amplios.

En primer lugar, el hecho de que consideremos estas bibliotecas como algo más que colecciones tradicionales digitalizadas y en segundo lugar, el resultado tan homogéneo obtenido tras realizar una búsqueda sobre proyectos de bibliotecas digitales en la base de datos LISA² donde un gran número de referencias apuntaban hacia los mismos programas, nos han conducido a seleccionar para nuestro estudio los seis proyectos de bibliotecas digitales que pertenecen a la Iniciativa DLI, ya que cumplen ambas condiciones.

DLI (Digital Library Initiative)

El gobierno de los Estados Unidos ha hecho de las bibliotecas digitales el ejemplo más representativo de su NII (National Information Infrastructure) e intentado llevar la información a cada americano utilizando la Red como medio de comunicación. Para conseguir su objetivo, en 1994, tres entidades norteamericanas:

NSF (National Science Foundation), NASA (National Aeronautics and Space Administration) y DARPA (Department of Defense Advanced Research Projects Agency) patrocinaron la Iniciativa DLI (Digital Library Information) formada por seis grandes proyectos de bibliotecas digitales seleccionados entre un total de setenta y tres y con un presupuesto total de 24 millones de dólares.

Seguidamente mostramos la relación de estos seis proyectos, junto con la Universidad que lo está desarrollando, su principal área de investigación y “*homepage*” o página de presentación:

- 1. Universidad de Michigan:** “The University of Michigan Digital Libraries Research Project”. Área de investigación: Agentes inteligentes para localizar información.
<http://www.si.umich.edu/UMDL/>
- 2. Universidad de Illinois:** “Building the Interspace: Digital Library Infrastructure for a University Engineering Community”. Área de investigación: Archivos documentales relacionados con la literatura científica.
<http://dli.grainger.uiuc.edu/>
- 3. Universidad de California, Berkeley:** “The Environmental Electronic Library: a Prototype of a Scalable, Intelligent, Distributed Electronic Library”. Área de investigación: Planificación del medio ambiente y sistemas de información geográfica.
<http://elib.cs.berkeley.edu/>
- 4. Universidad Carnegie Mellon:** “Informedia: Integrated Speech, Image and Language Understanding for Creation and Exploration of Digital Video Libraries”. Área de investigación: Bibliotecas vídeo digitales
<http://informedia.cs/cmu.edu/>
- 5. Universidad de Stanford:** “The Stanford Integrated Digital Library Project”. Área de investigación: Desarrollo de instrumentos de interoperatividad entre servicios heterogéneos.
<http://www-diglib.stanford.edu/>
- 6. Universidad de California, Santa Barbara:** “The Alexandria Project: Toward a Distributed Digital Library with Comprehensive Service for Images and Spatially Referenced Information”. Área de investigación: Clasificación espacial de la información.
<http://alexandria.sdc.ucsb.edu/>

Cada uno de ellos consta de varios subproyectos y, aunque todos están siendo desarrollados por universidades, cuentan con la participación de otras organizaciones como IBM, Xerox, Microsoft u Oracle.

Con respecto a la tecnología que utilizan, los seis proyectos descansan en sistemas distribuidos y la mayoría de ellos están usando el protocolo “InfoBus” desarrollado en el proyecto Stanford, que favorece un acceso uniforme a la gran variedad de servicios y fuentes de información existentes en Internet, utilizando para ello intérpretes entre el protocolo InfoBus y el protocolo nativo (por ejemplo, Z39.50).

Análisis de los Proyectos DLI

Universidad de Michigan: *The University of Michigan Digital Libraries Research Project*

“A study of history shows that civilizations that abandon the quest for knowledge are doomed to disintegration” (Bernard Lovell).

Con esta acertada frase comienza la página de presentación del proyecto conocido como UMDL (University of Michigan Digital Library Project). Este proyecto es fruto de una actividad conjunta de carácter multidisciplinar entre los miembros de la Universidad de Michigan y algunos colaboradores externos tales como IBM, Elsevier Science o Apple Computer.

El corazón de UMDL ha sido la arquitectura basada en agentes, que fue definida durante la primera fase del proyecto y que consiste en la agrupación de los mismos a fin de suministrar servicios complejos combinando capacidades individuales limitadas. El diseño de esta arquitectura permite alcanzar el objetivo principal del proyecto, que consiste en la creación de un entorno en el que los usuarios dispongan de una biblioteca personalizada que recoja todas las colecciones disponibles en las fuentes de información distribuidas a través del mundo, situación que plantea la necesidad de utilizar interfaces de usuario avanzadas.

El contenido que configura la biblioteca digital del Proyecto UMDL, está formado básicamente por una colección diversa de materiales relacionados con las Ciencias del Espacio y de la Tierra capaz de satisfacer las necesidades informativas de un gran número de usuarios diferentes.

Con respecto al diseño, construcción y evaluación de este Proyecto, los responsables del mismo han elaborado un banco de pruebas (“*testbed*”) constituido por un conjunto de usuarios pertenecientes a escalas y campos de actividad diferentes (investigadores, estudiantes y público en general).

El “*testbed*” en el que descansa la investigación se basa en el conjunto de tres tipos de agentes software:

- ❖ **Agentes de interfaz de usuario:** realizan entrevistas con los usuarios para comprobar sus necesidades informativas con respecto al área de conocimiento que precisan y a la exhaustividad y cobertura de la información que solicitan;
- ❖ **Agentes intermediarios:** coordinan las búsquedas de las colecciones distribuidas en la Red según el orden de prioridad establecido por el agente de interfaz, lo que le permite al usuario buscar simultáneamente en muchas bibliotecas;
- ❖ **Agentes de colección:** están asociados con cada colección determinada y pueden realizar búsquedas de texto, imágenes, audio y vídeo dentro de colecciones específicas. La información requerida puede pertenecer a entidades u organismos diversos, que podrían demandar algún tipo de control a la

hora de diseminar sus contenidos o incluso una compensación económica para preservar los derechos de autor de su material.

Algunos de los proyectos relacionados con el Proyecto UMDL son:

- ❖ *JSTOR (Journal Storage Project)*: cuyo objetivo consiste en crear una réplica electrónica fiel de los archivos que contienen las principales revistas científicas en ciencias sociales y humanidades. La Universidad de Michigan es el servidor Web en el que descansa el proyecto JSTOR. Su participación en el Proyecto se relaciona con el uso de la tecnología empleada para la digitalización e indexación del contenido de las publicaciones periódicas. Entre 1995 y 1997 el número de títulos de revistas totalmente digitalizadas se incrementó de 12 a 25 y se prevé que para el año 2.000, el número de éstas se aproxime a la centena.
- ❖ *MoA (Making of America)*: esfuerzo cooperativo ente la Universidad de Michigan y la Universidad de Cornell para crear una biblioteca digital relacionada con la historia social norteamericana. En este proyecto se han digitalizado más de 1.600 libros y 50.000 artículos de revistas impresos en la segunda mitad del siglo XIX, con lo que se ha favorecido la preservación y el acceso electrónico a textos históricos relacionados con la educación, psicología, historia, religión, ciencia y tecnología de la sociedad norteamericana. Entre las ideas en las que trabaja actualmente MoA, se encuentra el propósito de que la Biblioteca de la Universidad de Michigan lleve a cabo la conversión digital a su programa "Libros Frágiles (*Brittle Books*)" del Departamento de Conservación, de tal forma que posteriormente estos nuevos materiales se incorporen a la localización MoA.
- ❖ *UARC (Upper Atmospheric Research Collaboratory)*: es una acción conjunta de investigadores en el campo de la física espacial, la informática y las ciencias del comportamiento. El Proyecto UARC ha desarrollado un laboratorio internacional de colaboración conectado en red, en el que la tecnología de las comunicaciones y la informática se combinan entre sí para permitir que los científicos, distribuidos geográficamente por todo el mundo, colaboren en la adquisición y análisis de datos.

Universidad de Illinois: Building the Interspace: Digital Library Infrastructure for a University Engineering Community

El Proyecto UIUC (University of Illinois, Urbana-Campaign) tiene la finalidad de desarrollar la arquitectura documental válida para hacer una búsqueda eficaz de contenidos técnicos en Internet. El esfuerzo del proyecto se centra en la elaboración de una biblioteca digital experimental con decenas de miles de artículos de revistas en texto completo de física, ingeniería e informática, que estarían disponibles en la Red incluso antes de ser impresos, y que utilizan el formato SGML³

3 Standard Generalized Markup Language.

Las áreas de investigación fundamentales del proyecto son dos: la investigación sociológica y la tecnológica. Mientras que la investigación sociológica incluye la evaluación de la efectividad del proyecto y realiza un seguimiento del uso que hacen de la colección cerca de un millar de facultades y estudiantes pertenecientes a la UIUC, la investigación tecnológica se centra en la recuperación semántica de la información y analiza la indexación de los contenidos documentales que posibilitan la búsqueda a través de fuentes múltiples.

Así pues, un componente fundamental de este proyecto es la investigación sociológica y la evaluación relacionada con el uso de la biblioteca digital, para lo cual, es imprescindible:

- ❖ Suministrar un “*feedback*” continuo con el usuario con el fin de desarrollar mecanismos de recuperación, sistemas de préstamo y otras funciones características de las bibliotecas digitales;
- ❖ Contribuir a la comprensión teórica de la infraestructura de la información que está transformando el trabajo y la comunicación en las bibliotecas;
- ❖ Desarrollar nuevos métodos basados en el usuario para dirigir la investigación en bibliotecas digitales;
- ❖ Comprender la concepción y el diseño de sistemas de información a gran escala.

Además del Proyecto “Building Interspace”, la Universidad de Illinois ha desarrollado otros proyectos relacionados con las bibliotecas digitales, algunos de los cuales se muestran a continuación:

- ❖ *ADIL (Astronomy Digital Image Lab)*: base de datos en texto completo de imágenes radioastronómicas aportadas por la comunidad de usuarios, que está basada en la tecnología WWW.
- ❖ *Horizon Project*: proyecto cooperativo entre la NASA y la Universidad de Illinois que intenta mejorar la tecnología de la biblioteca digital (en concreto, la tecnología cliente/servidor del WWW) para ofrecer un acceso público de mayor calidad a los documentos relacionados con las Ciencias del Espacio y de la Tierra.
- ❖ *TM (The Daily Planet)*: archivo documental de acceso público basado en la tecnología WWW que proporciona acceso interactivo a una gran variedad de datos e imágenes climatológicas.

Universidad de California, Berkeley: The Environmental Electronic Library: a Prototype of a Scalable, Intelligent, Distributed Electronic Library

Este proyecto forma parte de la Iniciativa DLI y de CERES (California Environmental Resource Evaluation System). La investigación incluye al profesorado y a los estudiantes de la Computer Science Division de la School of Information Management & Systems, y del Research Program in Environmental Planning & Geographic Information Systems, así como la participación de diversas agencias gubernamentales y patrocinadores industriales.

Su objetivo se basa en el desarrollo de tecnologías para el acceso inteligente a grandes colecciones distribuidas de fotografías, imágenes vía satélite, mapas y documentos en texto completo, y la construcción de bases de datos de imágenes de varios terabytes de tamaño. En abril de 1998, el número de imágenes que formaban parte de la colección del Proyecto Berkeley ascendía a 58.000.

Debido a las características de estos documentos, el especialista del Proyecto Berkeley en bases de datos de imágenes centra su investigación en los métodos de acceso a las mismas, desarrollando algoritmos para localizar formas (por ejemplo, caballos, humanos, etcétera) en los documentos que circulan por Internet. Su meta consiste en construir nuevos algoritmos que puedan identificar formas originales en las diferentes colecciones de imágenes. Sin embargo, los resultados de estas investigaciones no siempre son favorables, lo que demuestra que la descripción de imágenes no es tarea fácil.

La búsqueda textual en las bases de datos se basa en Cheshire II que es un catálogo online de próxima-generación basado en un sistema de recuperación en texto completo que utiliza técnicas avanzadas de IR (Information Retrieval), que soporta Z39.50 e interopera con el InfoBus del Proyecto Stanford.

Como parte del Proyecto Berkeley también se encuentra el Programa TileBar que desarrolla interfaces de usuario para la toma de decisiones sobre qué documentos (o qué parte de los mismos) visualizar teniendo en cuenta el comportamiento distributivo de los términos de la consulta en los documentos.

Universidad Carnegie Mellon: *Informedia: Integrated Speech, Image and Language Understanding for Creation and Exploration of Digital Video Libraries*

“Informedia Digital Video Library” es un Proyecto de Carnegie Mellon University (CMU) que pertenece a la Iniciativa DLI y que estudia cómo crear y utilizar bibliotecas digitales de carácter multimedia.

El alcance y los objetivos del proyecto llevado a cabo por la CMU se pueden resumir en los siguientes:

- ❖ Establecer un sistema de biblioteca vídeo digital online interactiva;
- ❖ Capacitar a los usuarios para acceder, explorar y recuperar material relacionado básicamente con las materias de ciencias y matemáticas;
- ❖ Integrar las tecnologías capaces de realizar automáticamente los procesos de comprensión de imágenes, reconocimiento de la voz y procesamiento del lenguaje natural;
- ❖ Construir una biblioteca multimedia que contenga miles de horas de vídeo digital, audio, texto e imágenes y que se vaya incrementado automáticamente e indexando, segmentando y codificando datos;
- ❖ Investigar en la interacción hombre-máquina, en el precio y formas de pago para el uso del vídeo digital y en los mecanismos de seguridad y privacidad de los multimedia digitales.

Desde sus comienzos, el Proyecto "Informedia Digital Video Library" ha tenido un especial cuidado en el desarrollo de un interfaz de usuario que cumpla dos requisitos básicos: comodidad y eficacia.

La investigación centrada en el usuario, como centro neurálgico de análisis, descansa en tres aspectos fundamentales del proyecto:

- ❖ Cómo facilitarles a los usuarios la selección del segmento de vídeo más relevante cuando la búsqueda devuelve un gran número de ellos;
- ❖ Cómo permitirles ajustar la duración del segmento de vídeo para cubrir sus necesidades;
- ❖ Cómo facilitar la lectura rápida de segmentos de vídeo.

Así pues, los estudios de usuario juegan un papel muy importante en el desarrollo y evaluación de mecanismos tales como la conservación de múltiples "imicom" (iconos móviles inteligentes) en ventanas diferentes; el ajuste de la duración de los segmentos de vídeo mediante la repetición o "playback"; y la realización de una lectura rápida que le permita al usuario seleccionar y resumir la información relevante del documento multimedia.

Para comprobar los resultados del proyecto, los investigadores han desarrollado un protocolo capaz de analizar las sesiones individuales de los usuarios para poder identificar a aquellos que han usado el prototipo y enfocar hacia ellos las preguntas que permitan evaluar el uso real de esta Biblioteca Digital Multimedia.

Universidad de Stanford: *The Stanford Integrated Digital Library Project*

Cada uno de los seis proyectos que conforman la Iniciativa DLI, analiza la biblioteca digital desde una perspectiva diferente. En el caso del Proyecto Stanford, este estudio se centra en el desarrollo de herramientas que permitan un acceso integrado al elevado número de servicios heterogéneos que existen en Internet; el acceso distribuido facilita la interacción de los procesos entre máquinas distintas, con arquitecturas y lenguajes diferentes. Es decir, su línea de investigación se relaciona con la interoperatividad de los sistemas.

El corazón del proyecto lo constituye el protocolo "InfoBus" que favorece un acceso uniforme a la gran variedad de servicios y fuentes de información existentes en Internet utilizando intérpretes entre el protocolo InfoBus y el protocolo nativo. InfoBus descansa en la arquitectura CORBA utilizando ILU de Xerox. CORBA (Common Object Request Broker Architecture)⁴ es una herramienta desarrollada por OMG (Object Management Group) que permite la comunicación entre procesos remotos y constituye un eslabón importante en el camino de la interoperatividad y normalización basada en objetos. ILU (Inter-Language Unification System) es un sistema de interfaz de objeto multilingüe que anula las divergencias de implantación entre lenguajes y sistemas operativos diferentes, por lo que puede usarse para construir bibliotecas orientadas a objetos multilingües. Utilizando estas herramientas, el

4 Para más información consultar la dirección electrónica: <http://www.omg.org/corba/>

Proyecto ha desarrollado DLIOP (Digital Library Interoperation Protocol) para la búsqueda y recuperación de información.

“Stanford Integrated Digital Library Project”, está formado por más de veinte subproyectos con objetivos y funciones muy concretas: por ejemplo, uno de ellos desarrolla el protocolo InfoBus, otro elabora el intérprete entre InfoBus y Z39.50, etc. Sin embargo, uno de los subproyectos más interesantes es el que se encarga de realizar el protocolo InterPay que posibilita la interacción entre servicios con sistemas de pago diferentes.

El Proyecto Stanford resalta el hecho de que la biblioteca digital, además de ser un depósito o archivo documental, no debe olvidar aspectos tan importantes como la comunicación entre los usuarios y el personal de la biblioteca, aspecto que incluye, por ejemplo, suministrar un servicio de publicación. Por ello, ha desarrollado el servicio InterBib que pretende: a) la conversión de bibliografías con formatos diferentes; b) el procesamiento de documentos que incluyen bibliografías; c) la agrupación de bibliografías que pueden ser consultadas.

Al igual que el resto de los proyectos DLI, Stanford también se ha preocupado de desarrollar una interfaz de usuario experimental llamada DLITE (Digital Library Integrated Task Environment) diseñada para facilitarle al usuario la interacción con servicios diferentes mientras realiza un trabajo determinado. Dichos servicios informativos son accesibles a través de InfoBus.

Los trabajos de interoperatividad llevados a cabo en este proyecto también le ofrecen la oportunidad de hacer avances similares a los sistemas bibliotecarios comerciales que no habían ido más allá del uso de Z39.50 para permitir el acceso a otras bases de datos bibliográficas.

Universidad de California, Santa Barbara: The Alexandria Project: Toward a Distributed Digital Library with Comprehensive Service for Images and Spatially Referenced Information

El Proyecto ADL supone la evolución natural de una iniciativa anterior llevada a cabo por el Map and Image Laboratory de la UCSB (University of California, Santa Barbara).

El objetivo de ADL es el desarrollo de una biblioteca digital amigable que suministre la información recogida en colecciones digitales de mapas, imágenes, texto y multimedia que se encuentra dispersa geográficamente alrededor del mundo. Para cumplir su objetivo, el proyecto diseña, desarrolla y evalúa una biblioteca distribuida de información indexada y espacialmente almacenada en formato digital. El sistema les permite a los usuarios el acceso electrónico a las colecciones de la biblioteca digital mediante interfaces textuales y visuales que hacen posible la búsqueda y recuperación de fondos específicos.

ADL le ha concedido una atención especial a los metadatos debido, fundamentalmente, a las propias características de la información que maneja. La principal carac-

terística de esta información es que sus datos –y metadatos- provienen de fuentes muy diversas, lo que significa que los formatos de origen pueden llegar a ser tan variados como sus fuentes, siendo precisa la conversión al formato utilizado por ADL (en la primavera de 1997 se convirtieron datos que provenían de siete formatos distintos).

El formato ADL es una mezcla entre las normas USMARC y FGDC⁵ para material cartográfico, sin embargo, ninguna de las dos normas cubre suficientemente la descripción de los materiales tan diversos que configuran la biblioteca digital de Alexandria, puesto que la biblioteca de la UCSB es una de las más grandes bibliotecas cartográficas de US (contiene más de 4 millones de fotografías aéreas y unos 500.000 mapas).

Con respecto a la cooperación entre ADL y otros proyectos pertenecientes a la Iniciativa DLI: a) utiliza el InfoBus desarrollado por Stanford; b) aunque trabaja en el reconocimiento de imágenes, al igual que el Proyecto Berkeley, no existe cooperación entre ambos, posiblemente porque su orientación técnica varía: mientras que Alexandria se centra en la identificación de las estructuras en fotografías aéreas en blanco y negro, Berkeley se basa mucho en los colores para el desarrollo de sus algoritmos.

CONCLUSIONES

Los proyectos que hemos presentado en este trabajo se pueden considerar una muestra parcial del “*state of art*” de las bibliotecas digitales hoy en día. Y decimos parcial porque el hecho de que la Iniciativa DLI sea una propuesta norteamericana, nos permite tener una visión algo sesgada de la realidad porque, no todos los proyectos de bibliotecas digitales cuentan con el respaldo de un millón de dólares por proyecto y año (como es el caso que nos ocupa) y, lo que es más importante, desgraciadamente no todos los países consideran las bibliotecas digitales como el ejemplo más representativo de su infraestructura informativa nacional.

A pesar de que los proyectos pertenecientes a la DLI tenían la pretensión de finalizar en 1998, los problemas afrontados por ellos son de tal envergadura que requerirán la puesta en marcha de la Iniciativa DLI 2.

BIBLIOGRAFÍA

- ARMS, William Y. Key concepts in the architecture of the digital library. *D-Lib Magazine*, Julio 1995. Accesible desde:
<http://www.cnri.reston.va.us/home/dlib/July95/07arms.html/>
- BARNES, Susan J. The electronic library and public services. *Library Hi Tech*, 1994, 47, p. 44-46
- DRABENSTOTT, Karen M. Analytical review of the library of the future. Washington, DC: Council Library Resource, 1994

5 Federal Geographic Data Committee

- GARRET, John. Digital libraries: the grand challenges. *EDUCOM Review*, Julio/Agosto 1993, v. 28, n. 4
- GROSS, Ben. Digital Library Information and Resources. Accesible desde:
<http://interspace.grainger.uiuc.edu/~bgross/digital-libraries.html>/ Última actualización 10/6/98
- HAKALA, Juha; Hormia-Poutanen, Kristiina. Digital Library Initiative Projects: a report of a study trip 21-25 April 1997. *Nordinfo-Nytt*, 1997, 1-2, p. 20-30
- LYNCH, Clifford. Searching the Internet. *Scientific American*, Marzo 1997. Accesible desde:
<http://www.sciam.com/0397issue/0397lynch.html>/
- LYNCH, Clifford; García-Molina, Héctor. Interoperability, Scalability and the Digital Libraries Research Agenda. *IITA Digital Libraries Workshop*, 1995. Accesible desde:
<http://www-diglib.stanford.edu/diglib/pub/reports/iita-dlw/>
- OGLE, Virginia; Wilwinsky, Robert. Testbed development for the Berkeley Digital Library Project. *D-Lib Magazine*, Julio-Agosto 1996. Accesible desde:
<http://www.dlib.org/dlib/july96/berkeley/07ogle.html>/
- PAEPCKE, Andreas. Summary of Stanford's Digital Library Testbed Design and Status. *D-Lib Magazine*, Julio-Agosto 1996. Accesible desde:
<http://www.dlib.org/dlib/july96/stanford/07paepcke.html>/
- Schatz, Bruce; Chen, Hsinchun. Building Large-Scale Digital Libraries. *Computer. Monográfico: US Digital Library Initiative*, Mayo 1996. Accesible desde:
<http://www.computer.org/computer/dli/>
- SOCIAL aspects of digital libraries. Los Angeles : Graduate School of Education & Information Studies, UCLA, 1996. Accesible desde: <http://www.gslis.ucla.edu/DL/>