

# La nueva Biblioteca Digital de la Universitat de València

José Luis Canet Vallés

Director del S.I.B. (Universitat de València)  
C/ Nave, nº 2  
46003 VALENCIA  
[Jose.Canet@uv.es](mailto:Jose.Canet@uv.es)

**Resumen.** Las bibliotecas universitarias deben ofrecer cada vez más servicios en línea a sus usuarios. La Universitat de València, partiendo de su propio sistema de gestión bibliotecaria (DOBIS/LIBIS), ha buscado ofrecer nuevos servicios, en este caso dar la posibilidad al usuario de consultar vía web sus propios fondos (libros antiguos, periódicos, mapas, etc.) mediante imágenes digitalizadas de los mismos. Es el nuevo proyecto de Biblioteca Digital, y para su desarrollo se ha creado una nueva base de datos capaz de incorporar y gestionar dichos contenidos, enlazada con el catálogo general. Esta base de datos (*Digital Library*) será capaz, además, de gestionar contenidos museísticos y multimedia.

## 1 Antecedentes

La Universitat de València desde los años 87-90 inició una completa transformación de su sistema bibliotecario con un doble objetivo: la centralización de sus fondos y bibliotecas, por un lado, y la informatización de sus catálogos y recursos de información bibliográfica. Hoy en día, la estructura del Servei d'Informació Bibliogràfica está prácticamente configurada, mediante unos servicios centrales en donde se gestiona y mantiene el catálogo de toda la Universitat, el préstamo bibliotecario, la compra centralizada de revistas, libros para los estudiantes y bases de datos, etc., y bibliotecas de área (ciencias básicas, ciencias sociales, ciencias de la salud, humanidades). En este proceso de centralización podríamos incluir también la informatización, mediante la fusión progresiva en un mismo OPAC o IPAC de consulta de todos los recursos disponibles.

El año 1999, con motivo de la celebración del Quinientos aniversario de la Universitat de València, se procedió a crear la Fundació General de la Universitat - Cinc Segles, presidida por el Rector, y encargada de realizar una serie de eventos conmemorativos. Para ello, se pidió a la comunidad universitaria la presentación de proyectos, congresos, publicaciones y otras actividades para conmemorar el V Centenario. Uno de los proyectos seleccionados fue el de la *Biblioteca Digital*, presentado por Josep Lluís Canet, Director del Servei d'Informació Bibliogràfica de la Universitat de València.

## 2 Fase Inicial del Proyecto

El proyecto consistía en la creación de una biblioteca virtual o digital a texto completo, o mejor dicho a imagen completa de los libros, diarios, mapas, grabados, etc., capaz de ser visualizados mediante cualquier navegador y funcionando con los protocolos estándares TCP/IP, siguiendo las recomendaciones del Centro de Cálculo.

### 2.1 Objetivos

Las principales características de la nueva Biblioteca Digital podrían resumirse en los siguientes puntos:

- Se trataba, ante todo, de una biblioteca cuyos materiales digitalizados fueran principalmente de investigación, y no únicamente de difusión. Al investigador le interesa el manejo de un ejemplar determinado, de un libro que es irreplicable en su configuración, de ahí que la mejor forma de presentar dicha información fuera la imagen del texto completo y no su transcripción en formato texto, tratando de imitar el funcionamiento tradicional de una biblioteca histórica, siendo el usuario quien maneja el libro sin otra ayuda que la de su relación personal y visual con el ejemplar solicitado. Lo único que se perdería con la Biblioteca Digital es el sentido del tacto, que quedaría para aquellos estudiosos del papel y de la encuadernación, etc., trabajos necesarios para la historia profundizada de la imprenta, y cuyos investigadores seguirán realizando sus consultas en las salas de lectura.
- Al estar encuadrado este proyecto dentro del V Centenario de la Universitat, que dicha Biblioteca Digital fuera capaz de mostrar los fondos valiosos depositados en su Biblioteca Histórica, enlazando así con otro de los proyectos aprobados, *Thesaurus*, encargado de la catalogación y difusión del patrimonio de la Universitat.
- La digitalización de los textos completos se haría mediante imágenes JPEG, por ser un estándar en Internet, sin necesidad de cargar applets o Plug-ins en los ordenadores de consulta. Además, porque era la mejor forma para poder reutilizar el material ya digitalizado en microfilm y microfichas a un bajo coste económico, así como la posibilidad de utilizar posteriormente escáneres y cámaras fotográficas para la digitalización de otros materiales más complejos, sobre todo de mapas, grabados, carteles, cuadros, material museístico, etc. Se prefirió el formato JPEG a otros formatos (PDF, por ejemplo, utilizado en algunas de las Bibliotecas digitales a texto completo -Proyecto Gallica de la B.N.F-), porque el PDF necesita de una elaboración más meticulosa, que implica un mayor coste de personal y también de un mayor ancho de banda en la red, pues se descarga todo el fichero completo repleto de imágenes, además de necesitar del Adobe Acrobat en el ordenador de consulta.
- La gestión de la captura de las imágenes y el posterior tratamiento debería ser lo más automatizada posible, de ahí que se insitiera en el desarrollo de diferentes módulos de captura, con indexaciones automáticas, si así fuera posible, y módulos de administración para detectar posibles errores, así como

que el módulo de captura contemplara la incorporación de marcas de agua como elemento identificativo de la Universitat de València.

- El Software de la Biblioteca Digital tendría que ser capaz de incorporar y gestionar cualquier formato multimedia, pensando siempre en ampliaciones futuras, caso de la incorporación de tesis (bien en formato TXT, RTF, PDF, etc.), vídeos (MPEG), música (MP3, MID), etc.
- El menor coste posible de mantenimiento y desarrollo posterior.
- Acceso vía Internet/Intranet
- La solución debería ser altamente escalable, con capacidad de crecimiento desde una única estación de trabajo hasta una red corporativa.
- Su funcionamiento mediante la estructura Cliente/Servidor, lo que permitiría integrar diferentes plataformas dentro de la propia solución.
- Capacidad de búsqueda avanzada.

## **2.2 Asesoramiento y Financiación**

Para el desarrollo del proyecto se hicieron diferentes consultas a grandes empresas especializadas en la gestión y desarrollo de bases de datos (IBM, ORACLE, BASIS, etc.), insistiendo en la necesidad de configurar una Biblioteca Digital capaz de gestionar miles y miles de imágenes JPEG, correspondiendo a los siguientes materiales:

- Prensa periódica antigua (S. XVIII y XIX)
- Libros del S. XV, XVI, XVII, XVIII y XIX
- Manuscritos
- Carteles / Grabados / Mapas
- La colección de monedas de la Universitat

A partir de estas premisas básicas y de la información recabada a las diferentes empresas, se pidió financiación a varios organismos y empresas privadas, siendo el Banco de Santander el primero en cofinanciar el proyecto, junto con la Fundación de la Universitat general-Cinc Segles. Posteriormente se pidió a las empresas con las que se había contactado en la fase previa que presentaran sus diferentes propuestas de hardware y software, indicándoles la posibilidad de que también fueran copatrocinadoras del Proyecto.

## **3 Licitación y Adjudicación del Proyecto**

Tres fueron los proyectos presentados, y para la adjudicación se constituyó una Comisión presidida por el Vicerrector de Investigación, el Director del Servei

d'Informació Bibliogràfica, el Director del Centro de Càlculo, el Director del Instituto de Robòtica y otro miembro más de dicho Instituto, decantándose la comisión por la Propuesta presentada por la empresa IBM, la cual se incorporaba como patrocinadora también del proyecto, aportando diferentes elementos del hardware y software.

Sin entrar en excesivos detalles, podemos decir que la máquina propuesta y ya instalada es una RS/6000 Serv. Enterpris H70, y el software corresponde a Digital Library V2, junto con AIX 4.3, VisualAge for Java Enterprise, IBM C and C++ Compilers, y otros pequeños paquetes necesarios para la gestión de los RACs y otros sistemas de almacenamiento y de copias de seguridad, cuyas características específicas se pueden consultar en la web de la propia casa (<http://www.ibm.com>)

#### **4 Instalación del *Hardware* y Desarrollo del *Software***

Después de la adjudicación a la empresa IBM, se establecieron una serie de reuniones para poder perfilar las necesidades reales sobre la ubicación del *Hardware* (se buscó espacio en el propio Centro de Càlculo de la Universitat, y su instalación se realizó en menos de un mes), así como para el desarrollo del *Software*, sobre todo con lo referido a los módulos de catalogación, los de captura e indexación de las imágenes y el de consulta.

En un principio se pensó en la generación de un módulo de catalogación de los impresos antiguos y modernos, prensa periódica, grabados y mapas, etc., siguiendo la normativa internacional de catalogación USMARC o IBERMARC. Pero muy pronto nos apercebimos que este proceso no era ni más ni menos que reduplicar el trabajo, puesto que gran parte del material estaba ya catalogado dentro de nuestro sistema de gestión bibliotecaria Dobis/Libis.

También se vio la posibilidad de incorporar una pasarela Z39.50, capaz de recuperar las fichas catalográficas desde DOBIS y trasladarlas a la propia base de datos de la Biblioteca Digital. Pero ello suponía la configuración y mantenimiento de dos bases de datos simultáneas con fichas idénticas, o al menos muchas de ellas reduplicadas.

Finalmente se decidió convertir el sistema de la *Digital Library* en un módulo servidor de imágenes del propio sistema de gestión bibliotecaria Dobis/Libis. Ello conllevaba varias ventajas, desde una sola gestión del catálogo, hasta un único sistema de recuperación de la información.

Pero la solución no era fácil, ya que el programa Dobis/Libis no incorporaba el campo 856 capaz de gestionar direcciones URL, y por otra parte, al ser un sistema propietario y cerrado no podía comunicar directamente con la *Digital Library* para enlazar los títulos de las obras con sus imágenes digitalizadas.

Se convocó una nueva reunión entre los responsables de IBM, ELIAS (propietario actual de Dobis/Libis), nuestros informáticos del Centro de Càlculo y la empresa SCANBIT, para analizar las diferentes posibilidades de integración de ambos sistemas. En dicha sesión se pusieron sobre el tapete los pros y los contras de las diferentes posibilidades de integración, decidiendo finalmente la modificación de la

base de datos Dobis/Libis, incorporándole nuevos campos (capaces de gestionar direcciones URL), y un sistema de pasarelas que comunicaran directamente ambas bases de datos. Otro de los condicionantes exigidos a las empresas (en este caso por los miembros presentes del Servicio de Información Bibliográfica) fue la necesidad de modificar los módulos de exportación, para poder migrar en un futuro a cualquier otro programa de gestión bibliotecaria, manteniendo los registros enlazados con la *Digital Library*.

El objetivo perseguido consistía, pues, en establecer un mecanismo por el cual los usuarios pudieran realizar consultas sobre los textos digitalizados (almacenados en el sistema de archivo de imágenes DB2 *Digital Library*) juntos con los de información bibliográfica (gestionados por el sistema de Gestión de Biblioteca *Dobis/Libis*) a través de un único acceso ofrecido por el *Internet Public Access Catalogue* (IPAC), teniendo en cuenta, además, que en un futuro próximo puedan generarse otros enlaces entre la información de Dobis/Libis y su versión electrónica de los mismos a través de URL, pongamos por caso, de las revistas electrónicas.

Así, pues, se procedió inmediatamente a la incorporación del IPAC como sistema de consulta, pero manteniendo al mismo tiempo las otras formas de acceso al catálogo mediante el TN 3270 y el Z39.50. La razón para mantener los sistemas anteriores, sobre todo el Z39.50, fue porque en estos momentos se está desarrollando la integración de los diferentes recursos bibliotecarios, capaces de ser visualizados a través de un único formulario de búsqueda, siendo el usuario final capaz de elegir qué base o bases de datos quiere consultar, pudiendo hacerlo simultáneamente en todas ellas. En esta integración, el usuario podrá consultar incluso las bases de datos en formato CD-Rom (bibliográficas y a texto completo), las generadas por la propia universidad (sumarios de revistas), las gestionadas por diferentes editores a texto completo, todo ello junto con el propio catálogo de la universidad.

El ideal perseguido ha sido desarrollar una tecnología adecuada para enlazar todos los recursos bibliotecarios de los que dispone la Universitat de València, de la manera más fácil y simplificada para un usuario no especializado. De ahí la necesidad de que a través del IPAC se pueda localizar cualquier texto, revista, plano, periódico, etc., que exista en la universidad, y que el resultado de la búsqueda incorpore enlaces activos, que permitan visualizar inmediatamente las imágenes de los fondos digitalizados, enlazar con las revistas electrónicas a texto completo, o incluso poder buscar al mismo tiempo en diferentes catálogos y múltiples bases de datos, mostrando un listado global en donde se indique mediante iconos gráficos cada uno de los items localizados de donde procede y que posibilidades tenemos para su consulta: inmediata a través de enlaces activos y a través del navegador o por el contrario en las salas de una biblioteca determinada.

#### **4.1 Módulo de Captura e Indexación de las Imágenes**

Posteriormente, se intentó definir el sistema de captura e indexación de las imágenes. Se escogió para éstas, como se ha dicho antes, el formato JPEG, por presentar una serie de ventajas:

- Ser un formato estandarizado, capaz de visualizarse en cualquier navegador.
- Gran capacidad de compresión.
- Tratar cada imagen de un libro o periódico, etc. de forma independiente.
- Digitalización de grandes formatos (desde el in-folio hasta el A3) y su posterior visualización mediante zoom.
- Coste aceptable por fotograma digitalizado.

Las características del módulo de captura e indexación de las imágenes se pensó para que tuviera el menor coste de mantenimiento posible. Las imágenes se podrán incorporar al sistema por FTP (desde las propias empresas que digitalizan o desde escáneres), a través de discos compactos, CD-Rom, DVD, etc. Cada grupo de imágenes, correspondientes a un libro, diario, revista, etc., se estructurarán en el interior de una carpeta, definida ésta mediante la signatura que lo individualiza. En el interior, y con formato texto, se incluye una pequeña descripción del ejemplar y un listado de todas las imágenes, estando éstas numeradas con una cadena de dígitos, que definen su especificidad (libro, mapa o plano, revista, prensa periódica), si es portada o índice, y su correlación o paginación.

El proceso seguido para la indexación y catalogación es el siguiente: cuando se recibe una carpeta perteneciente a un volumen, el operador escribe la signatura y el sistema busca en Dobis/Libis la ficha catalográfica, e inscribe en el campo URL la dirección donde se depositará en la Biblioteca Digital. Posteriormente, el operador utiliza un applet con el que valida la entrada de cada una de las imágenes, pudiendo así corregir cualquier error con posterioridad. El sistema, reconoce inmediatamente (por los dígitos de cada imagen) si corresponde a un libro, prensa periódica, grabado, etc., y la posición que ocupa cada imagen para su posterior visualización (portada, índice, paginación correlativa).

La indexación de las imágenes se realizará de forma muy automatizada, generando rutinas para que pueda el propio sistema indexarlas durante la noche, por ejemplo, para no sobrecargar el sistema.

## **4.2 Módulo de Visualización**

Después de realizada una búsqueda a través del IPAC en el catálogo de la Universitat, los ejemplares digitalizados incluirán un botón para su visualización en pantalla a texto completo. Pulsando dicho botón, accedemos a una nueva pantalla en la que aparece como cabecera el autor, título, impresor, año (en el caso de los libros), y en la parte inferior las imágenes de la portada y el índice en formato reducido (una resolución de unos 72 pixels para no sobrecargar las redes ni el propio servidor). En la parte superior de las imágenes se estructuran las herramientas para ir a una página concreta, avanzar o retroceder, o mostrar directamente la portada y el índice en una nueva ventana con resolución mayor. En el caso de la prensa periódica, se visualiza

una nueva pantalla con el título del diario y en la parte inferior una estructura arbórea realizada en HTML, indicativa del año, mes, día.

Las imágenes que aparecen en pantalla son activas, con lo que pulsando sobre ellas se activa una aplicación java que abre una nueva pantalla del navegador mostrando la imagen con una mejor resolución. Este desarrollo en java ha sido imprescindible para poder visualizar formatos in-folio hasta A3, con zoom incorporado para ver la imagen por partes, y dando la posibilidad de su impresión.

El sistema, además, incorpora a las imágenes mostradas en pantalla y para su posterior impresión un código de agua con el logo de la Universitat de València, que las identifica.

#### **4.3 Previsiones de Futuro**

Por otra parte, dicha Biblioteca Digital está también capacitada para la inclusión de cualquier material u objeto museístico. Buena prueba de ello es que desde sus inicios se ha desarrollado una base de datos paralela donde se encuentra la colección numismática de la Universitat, con más de 10.000 monedas, catalogadas y fotografiadas, pudiendo realizarse la búsqueda por años, periodos o siglos, lugares donde fueron acuñadas, peso, metal utilizado, etc., mostrando en pantalla la cara y el reverso de la moneda y sus características.

No descartamos en un futuro incorporar otros materiales museísticos, que forman parte del patrimonio de nuestra Universitat, como pueden ser los carteles, cuadros, objetos de muy diversa índole (caso del museo de medicina o de astrología, etc.).

Por supuesto, esta Biblioteca Digital podrá almacenar otros formatos, caso de los archivos PDF, documentos texto y/o Word, audios y videos, e incluso programas de ordenador, si así se considera oportuno.

## **5 Conclusiones**

Para concluir, podríamos señalar, que la nueva Biblioteca Digital de la Universitat de València estará compuesta básicamente por los siguientes elementos:

- Un sistema de gestión bibliotecaria (actualmente DOBIS/LIBIS)
- Un sistema de gestión de imágenes (*Digital Library* de IBM)
- Dos servidores para la gestión de bases de datos en CD-Rom (ERL y Metaframe)
- Un servidor para bases de datos de revistas electrónicas (sumarios, sobre todo)

- Tres servidores Z39.50 (correspondientes a DOBIS, Digital Library y bases de datos ERL) y dos clientes Z39.50.
- Un cliente IPAC (DOBIS), con dos interfaces diferenciados de consulta (uno general, y otro específico para los textos digitalizados)
- Y otros muchos servidores específicos para el préstamo interbibliotecario, para la gestión de las colecciones, para el web, etc.

Nuestra intención ha sido que a través del IPAC se tenga acceso a los fondos de la Universitat, bien a la ficha catalográfica de los ejemplares, bien a los propios textos digitalizados. Mediante el cliente Z39.50, se podrá acceder además al catálogo general de nuestra universidad y de otras muchas bibliotecas, a bases de datos bibliográficas, a servidores de revistas electrónicas a texto completo, etc., todo ello simultáneamente.

Se ha buscado, pues, la integración máxima de los recursos bibliotecarios de que dispone la Universitat de València, incluyendo el acceso a otras colecciones y fondos museísticos. Es un paso más ante el reto de proporcionar un acceso universal a los documentos y servicios de información, y para ello se ha intentado enlazar todos los sistemas informáticos de una forma transparente y con facilidad de manejo para el usuario final.

## **Bibliografía**

Direcciones en Internet:

- <http://www.ibm.com>
- <http://www.javasoft.com>
- <http://www.oracle.com>
- <http://www.sun.com>
- *Technical Recommendations for Digital Imaging Projects*, prepared by the Image Quality Working Group of ArchivesCom. New York, University of Columbia, Academic Information Systems, abril de 1997 (<http://www.columbia.edu/acis/dl/imagespec.html>)