

La difusión del fondo histórico de la Universidad Complutense de Madrid: Proyecto Dioscórides Web

Pilar Moreno García, Víctor Herreros Villanueva.
Universidad Complutense de Madrid

Raquel Minteguía Apodaca.
Azertia

1 Introducción

La Biblioteca de la Universidad Complutense ha venido desarrollando en los últimos años un proyecto de digitalización de Fondo Antiguo en el área de Ciencias de la Salud¹ conocido como Biblioteca Digital Dioscórides. La riqueza de los fondos depositados en la Biblioteca Histórica *Marqués de Valdecilla*, donde se ha reunido el patrimonio bibliográfico de la Universidad repartido anteriormente por las diferentes Facultades, plantea que el acceso a la base de datos de imágenes sea posible desde cualquier punto del Campus universitario. Hasta este momento la visualización de estos libros sólo era posible desde una aplicación instalada en terminales situados en las bibliotecas de Farmacia, Medicina y Filología; por lo que surge la necesidad de publicar éstas imágenes vía WEB y hacerlas así accesibles desde Internet a los investigadores interesados en conocer cualquier aspecto de este fondo bibliográfico tan importante para el desarrollo y la historia de la Ciencia.

2 Situación de partida

Las imágenes que componen esta Biblioteca Digital se digitalizan de dos formas:

en blanco y negro a 400 dpis en 256 niveles de grises según se trate de textos impresos o manuscritos y grabados/ láminas respectivamente. Las imágenes en blanco y negro, se almacenan en formato TIFF Grupo IV y las imágenes en grises en formato JPEG.

Por otra parte cada imagen se indiza de forma independiente, y en este proceso se identifican como partes del libro, el lomo, cubierta, guardas, páginas, portada, láminas, preliminares, capítulos, colofón, índice general e índice de láminas, generándose así una estructura arborescente que permite la navegación a través del libro de una manera muy sencilla.

Las imágenes están almacenadas en INFORMIX 7.3.1 y hay actualmente 600.000 que ocupan aproximadamente 67 GB. La media de una imagen es de 114 KB.

3 Soluciones adoptadas para la publicación web

Teniendo en cuenta la situación de partida en la que las imágenes tienen muy buena calidad y por lo tanto son muy voluminosas y que los formatos de tamaño de los libros digitalizados son muy diversos, y pensando en su publicación WEB se han adoptado las siguientes soluciones:

¹ Este proyecto se ha realizado con la financiación de la Fundación de Ciencias de la Salud Glaxo SmithKline

- Los usuarios pueden consultar las imágenes sin necesidad de ningún software adicional (plugin, activeX), sólo el navegador.
- Las páginas son accesibles desde los navegadores mas utilizados:
 - Internet Explorer versión 5.0 o superiores
 - Netscape versiones 4.7 o superiores.
- Dado que las imágenes son objetos ‘voluminosos’, el usuario solo se descargará la página del libro que esté consultando en cada momento.
- Una imagen publicada es siempre legible, pero su tamaño, entre 50 y 80 KB, no ralentizará su acceso a través de Internet.
- Se mantiene la estructura arborescente actual, que permite de forma muy intuitiva acceder a las páginas del libro, sin tener que hacerlo de forma secuencial.
- Está integrado con el catálogo bibliográfico de la biblioteca actualmente publicado en WEB (INNOPAC). Aprovechando por lo tanto todas sus posibilidades de recuperación de información.
- Se puede acceder a los grabados de todos los libros digitalizados que conforman una base de datos independiente, utilizando sus índices actuales que hacen posible la búsqueda y por tanto la recuperación por materia, autor y periodo cronológico del grabado.
- Se presupone que la resolución de pantalla media de los usuarios que acceden a las páginas es de 800x600.

4 Descripción tecnológica

Se ha realizado una conversión masiva de las imágenes a formato JPEG, para que el navegador sea capaz de visualizarlas sin necesidad de otro software adicional.

Aunque la resolución será un parámetro que el administrador podrá modificar, se ha estimado que por defecto, la resolución que utiliza el programa será 630 pixels en el eje X. Para estimar este valor se ha supuesto que:

- la resolución media de la pantalla de un usuario de consulta es 800x600
- la imagen debe aparecer en la ventana de visualización ajustada a lo ancho, de forma que el usuario no tenga que hacer un desplazamiento lateral de la misma
- como además de la imagen va a aparecer otra información en pantalla (el árbol con la estructura del libro), de los 800 tendremos disponibles para ver la imagen 630.

Por otra parte, y durante ese proceso de conversión, se aplican filtros de tratamiento de imagen (Interpolación bicúbica) para mejorar su legibilidad, ya que en su mayoría se trata de textos manuscritos e impresos y así se podrá eliminar el deterioro causado por el tiempo en las imágenes.

Con todo ello el tamaño de la imagen resultante está situado entre 50 y 80 KB.

El servidor WEB es un Internet Information Server 5.0 de Microsoft y la tecnología empleada para el desarrollo de las páginas de acceso es la de Microsoft (páginas ASP). Desde ellas se accede a Informix vía ODBC (ADO) para recuperar las imágenes y los datos de indización.

Para ciertas validaciones se utiliza JavaScript en los clientes, por las mismas razones esgrimidas en el apartado anterior: para permitir el acceso mediante el mayor número de navegadores de los disponibles en el mercado.

El sistema está totalmente integrado con el catálogo bibliográfico de la Universidad (CISNE-INNOPAC):

- Por cada registro bibliográfico existe un enlace a una página ASP de DIOSCORIDES WEB en la que se pasa como parámetro la referencia bibliográfica a consultar, con lo que el usuario tiene accesibles todas las imágenes de un libro.
- Desde esta página se puede volver al registro bibliográfico de CISNE.
- También se puede acceder a la búsqueda en la base de datos grabados.
- Desde la página de grabados y una vez localizado por los índices de materia, autor y periodo cronológico, se puede acceder a la descripción bibliográfica de CISNE del libro que lo contiene, y a su vez desde esta última recuperar de nuevo todas las imágenes del libro solicitado.

5 Páginas de dioscórides web

Acceso a las imágenes desde la página de CISNE:

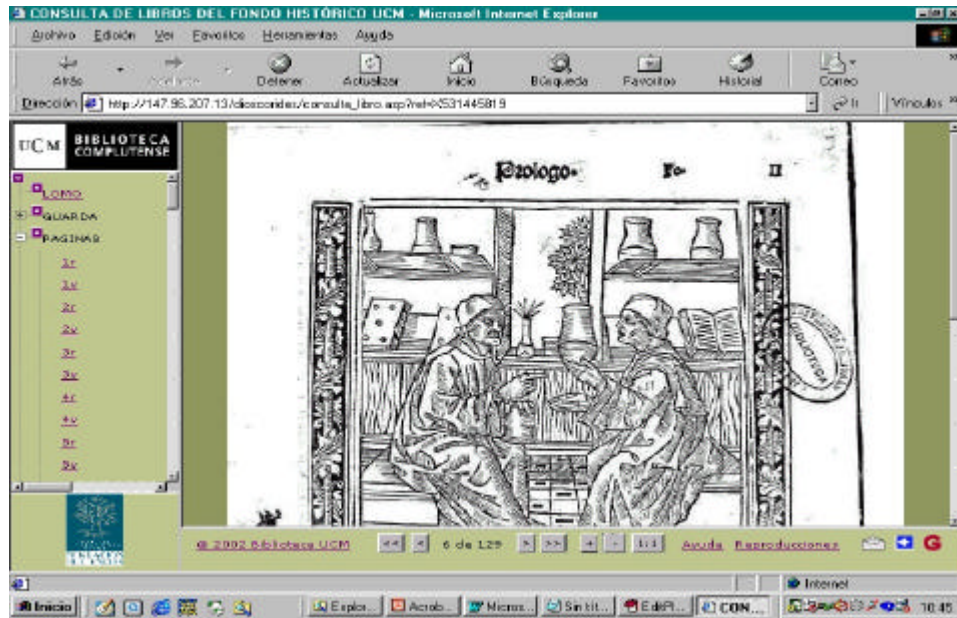
The screenshot shows a web browser window displaying the CISNE (Central Information System for Networked Electronic) interface. The page title is "CISNE". The main content area shows a record for a book with the following details:

- Autor:** [Kerham, Johannes de](#)
- Título:** [Fasciculus medicinarum = Epitome in medicina y cirugía o Compendio de la salud humana / Johannes de Kerham. De epidemia et peste \(en castellano\) = Tratado de la peste / Vasco de Toronta. Liber phytoprognosis \(en castellano\) / Michael Scotus](#)
- Publicac.:** Burgos : Mas de Burgos, 15 mayo, 1495
- Lugar impr.:** [España Burgos](#)
- Descripción:** LCVIII h. 4
- Colección:** [Proyecto Dioscórides. Serie Incunables](#)
- Nota:** [Códex](#)
Las ilustraciones son grabados litográficos
Las hojas XXV y XXVI son hojas dobles plegadas con grabados litográficos
También disponible la reproducción digital del ejemplar
C: 2301, P0145B5, Pol. 1402, G08E-1R, Hash. 246, Vndel VII.35, BBE 341D
Faltan las hojas XI, XIV, XV, XVI y XXV. El grabado de la hoja IV está cortado en su margen izquierdo. Anotaciones man. en los márgenes y en las guardas. Encuadernación de pergamino
- Clasif.:** 61
- Materia:** [DIOSCORIDES](#)
- Autor sec.:** [Medicina -- Obras anteriores a 1800](#)
[Toronta, Vasco de](#)
[Scotus, Michael](#)
[Toront, Juan de, imp. \(Burgos\)](#)

At the bottom of the page, there is a table with the following structure:

UBICACION	SIGNATURA	ESTADO	TIPO DE PRESTAMO	NOTAS

Vista general de las imágenes de un libro:



Zoom sobre el texto:



Consulta de Grabados:

The screenshot shows a web browser window displaying a search interface for engravings. The browser's address bar shows the URL: http://147.96.207.134/icos/oides/gravados/consulta_grabados.asp. The page header includes the UCM logo and the text "BIBLIOTECA COMPLUTENSE". Below this, there is a search section titled "Búsqueda de Grabados" with input fields for "Autor", "Materia", and "Periodo" (set to 1404 / 1495). A "Buscar" button is present. A list of search results is displayed, including entries for "Aristóteles, De caelo et mundo (latina)", "Cicero, Franciscus, Via Regii Machi", and "Johannes de Perdisio". To the right of the search results, a large anatomical engraving of a human figure is shown, with various parts of the body labeled with letters and lines pointing to them. The browser's taskbar at the bottom shows the Windows Start button and several open applications, with the system clock displaying 10:49.

Filtro bicúbico (Antes):



Filtro bicúbico (Después):

