

LOS OPACS DE LAS BIBLIOTECAS DENTRO DEL ENTORNO MULTIMEDIA EN INTERNET

Félix de Moya, Pedro Hípola, Antonio Muñoz-Cañavate

Moya, F.; Hípola, P.; Muñoz-Cañavate, A. «Los opacs de la bibliotecas dentro del entorno multimedia en Internet", Barcelona, 1995. 5es Jornades Catalanes de Documentació.

Resumen:

Los usuarios de Internet se desaniman a menudo ante la escasa amigabilidad de los opacs de las bibliotecas, sobre todo cuando establecen comparaciones con la facilidad de uso de otros recursos informativos existentes en la Red.

Sin duda este problema ha de ser solucionado con urgencia. Los opacs al uso no sólo están desfasados, sino que además existe tanta diversidad de sistemas, que entre todos forman una tortuosa senda en la que los softwares de consulta a los catálogos bibliotecarios no aparecen como la ayuda deseada, sino que cada uno es un obstáculo más dentro de un laberinto.

El problema está quedando de manifiesto gracias al hecho de que hoy resulta posible conectarse en línea a más y más opacs.

En esta comunicación se defiende la posibilidad de utilizar de forma generalizada los clientes World Wide Web -un sistema único y normalizado- como interfaces para llegar, directa o indirectamente, a través de Internet, hasta los programas de gestión de las bases de datos bibliográficas de nuestros centros.

Esto puede llevarse a cabo de dos maneras:

- 1) La aplicación cliente WWW hace uso de gateways existentes en la Red que convierten las consultas de los usuarios en peticiones ajustadas al protocolo Z39.50
- 2) Uso de las prestaciones propias de los sistemas Web sin necesidad de recurrir a las citadas gateways, utilizando pequeños programas que actúan como interfaces entre un servidor HTTP y la base de datos catalográfica.

En el segundo caso, el más sencillo de los dos, son indispensables los siguientes elementos:

- a) un browser Web, que permite al usuario formular las consultas y gestionar la visualización y manipulación de los resultados;
- b) un servidor WWW, con todas las funcionalidades típicas de tal tipo de software;
- c) un desarrollo específico para cada sistema bibliotecario, que actúe como interface entre el servidor WWW y el gestor de la base de datos catalográfica.

Quizá lo más interesante de este esquema es resaltar el hecho de que los elementos b) y c) pueden estar implementados en cualquier punto de la Red, no necesariamente en el sistema informático que gestiona el catálogo bibliotecario, lo cual hace posible diseñar una arquitectura telemática muy flexible.

También es destacable que todo ello es posible gracias a la existencia de las funciones *form* existentes en el lenguaje HTML desde su segunda versión. El elemento b) envía los datos alimentados en a) hasta c). El ciclo se cierra con el flujo de datos en sentido contrario.

Introducción

La llegada de Internet y la generalización masiva del uso de ésta ha actuado como revulsivo en los sistemas bibliotecarios actuales. Sin duda alguna la biblioteca como proveedora de información tendrá que competir en el futuro inmediato con otros proveedores de información para atender el masivo uso de los recursos de información electrónicaⁱ, siendo en esencia las redes electrónicas de información el lugar más adecuado en el que se producirá dicha competencia.

Las nuevas aplicaciones en Internet han provocado un efecto inmediato, como el aumento de la comunicación entre los usuarios y las fuentes de información electrónica, que han tenido a su vez una rápida progresión. Esta nueva revolución de la información ha originado que surjan nuevas aplicaciones cliente/servidor y que en el centro de dicho avance se encuentren los sistemas multimedia y las redes electrónicas, como Internet, que ha generado una novedosa cultura de acceso a la información ligeramente diferente al tradicional mundo on-line o incluso al CD-ROM, sistema este último que seguirá dando que hablar.

Entendemos que en una posición muy destacada dentro de esta revolución se encuentra el hipertexto y el hipermedia, esto es, la gestión de información en la que los datos son almacenados en una red de nodos conectados por enlaces. Los nodos pueden contener texto, gráficos, audio, vídeo o cualquier otro tipo de datos.ⁱⁱ

En definitiva el paradigma hipertexto/hipermedia conlleva un sistema de acceso a los datos no secuencial formado por nodos de información y enlaces.

Algunos autores, como Cookⁱⁱⁱ o Raymond & Tompa^{iv}, han señalado las ventajas y desventajas de los sistemas hipermedia, y entre las ventajas indican que pueden almacenar gran cantidad de información y que permiten una navegación mucho más rápida a través de un gran número de entradas.

En esta nueva cultura electrónica se han originado términos y conceptos como "biblioteca virtual", que parece querer decir que podremos acceder a un universo de información en formato electrónico. Esta biblioteca virtual debe ofrecer a cualquier usuario el acceso a cualquier lugar del mundo (conectado), a cualquier hora del día o de la noche, a través de un ordenador usando las redes telemáticas^v. En este sentido el concepto de biblioteca virtual no se reduce a la biblioteca tradicional reciclada, sino que ésta forma parte de la citada globalidad de recursos de

información.

El poder del mundo World Wide Web, comparable a los efectos y capacidades de la radio y la televisión, radica en la facilidad de conectar recursos electrónicos en base a simples enlaces que permiten que un usuario pueda realizar múltiples conexiones desde su ordenador conectado a una red. Los documentos que soportan los servidores Web están escritos en un lenguaje especial. Si bien el texto es Ascii, éste se ve rodeado de una serie de etiquetas especiales que convertirán al archivo en un documento hipertexto una vez que sean reconocidos por una aplicación cliente o "browser".

Algunas de estas etiquetas hacen referencia a los mencionados enlaces o *links*, que conectan unos documentos con otros a lo largo y ancho de Internet. Estos enlaces requieren de un localizador de recursos. Para ello se utilizan los URLs (Uniform Resource Locators), con los que se especifica a qué recursos nos vamos a conectar (ftp, gopher, news, telnet, http, etc.), el dominio de Internet, el puerto, si existe, y el camino de directorios hasta llegar al archivo especificado.

Desde una simple página en formato HTML se puede acceder a recursos muy variados, de manera que el resto de las aplicaciones Internet quedan a disposición de los servidores http^{vi}. Entre otras funcionalidades se permite el acceso a bases de datos a través de formularios normalizados, lo que hoy por hoy es uno de los desarrollos más exitosos dentro del mundo WWW. Nos detendremos en ello en esta comunicación.

Durante los últimos meses ha habido interesantes aportaciones al mundo Web. Diferentes compañías han creado herramientas que hacen a los servidores http mucho más potentes. Entre estos desarrollos tenemos el Virtual Reality Modeling Language, un lenguaje para describir simulaciones interactivas con muchos participantes^{vii} utilizando el espacio tridimensional. La primera versión de VRML permitía la creación de mundos virtuales de una forma limitada. Estos mundos pueden contener una serie de objetos con enlaces a otros mundos, entre los que se encuentran los documentos HTML. Futuras versiones de VRML incluirán animaciones y mejores herramientas interactivas.

Los servidores World Wide Web han cubierto ya prácticamente todo el espectro temático. Las bibliotecas, al igual que los museos^{viii}, son fuentes básicas de conocimiento, y las versiones

electrónicas de estos recursos informativos terminarán por imponerse.

CGI (Common Gateway Interface)

Una de las más importantes funcionalidades de los servidores WWW proviene del uso de una herramienta de recuperación de información denominada Common Gateway Interface (CGI). Aunque está en período de transición, éste potente sistema permite recuperar documentos estructurados en una gran variedad de formatos, además del HTML.

Un *Common Gateway Interface* consiste, básicamente, en un interface que permite que funcionen programas externos dentro de un servidor de información http. Gracias al uso de esta herramienta, un usuario de las redes puede a través de un cliente Web realizar múltiples preguntas a otros servidores, ya que cada cgi puede llevar a cabo las búsquedas en una serie de bases de datos externas. Luego se devuelven los resultados de forma adecuada a los clientes para que los usuarios puedan visualizarlos.

De esta manera el servidor puede mostrar información dispersa por diversas fuentes, y en este tipo de información incluimos cualquier tipo de base de datos. Estos cgi o gateways pueden ser utilizados en una gran variedad de servidores http^{ix}. En líneas generales se pueden resumir las funciones de los cgi en varias ideas básicas:

- permiten desde un formulario HTML *enviar las peticiones* a servidores WAIS (Z39.50) o a otras bases de datos, convirtiendo los resultados obtenidos en documentos HTML que se remiten al cliente Web.

- permiten *enviar información* desde un servidor por medio de formularios HTML.

Los correspondientes scripts pueden estar escritos en diferentes lenguajes de programación, como C, C++, PERL, TCL, C Shell, etc.

Cualquier propietario de bases de datos puede poner su información disponible al conjunto de usuarios de Internet. Para ello lo que tiene que hacer es crear un programa CGI. El "daemon" del Web transmitirá la información a la base de datos y reenviará los datos al cliente.

Dentro de las especificaciones HTML, las etiquetas relacionadas con los formularios, llamados "forms", están, evidentemente, asociadas a estos programas cgi, ya que actúan de forma directa entre el usuario (que usa un cliente Web en modo gráfico) solicitante de información y los servidores de información.

En la siguiente figura se puede observar cómo el usuario va rellenando los formularios que aparecen en pantalla y ejecuta la búsqueda de información.

[FIGURA1.RS]

Estas herramientas permiten dar mayor potencia a los servidores Web, que dejan de ser sólo fuentes de información (con todo lo que ello significa) para convertirse además en *buscadores* de información. Existen algunos programas muy potentes que buscan información en sus propias bases de datos sobre recursos en Internet. En el anexo 1 se ofrece una relación de algunos de estos sistemas que utilizan formularios preparados en HTML.

Un URL puede identificar un nodo que contenga un programa o script que se ejecuta en el momento en el que el usuario activa el enlace. Este programa, que puede estar en otro ordenador, se ejecutará y preparará un documento HTML que será enviado al cliente. Estos scripts suelen estar en directorios llamados cgi-bin.

Según resume Michael Grobe^x, cuando pretendemos implementar alguna de las múltiples bases de datos existentes en Internet podemos actuar de varias formas:

1. Utilizar un simple enlace telnet a un servidor donde no se requiera password, entrando directamente a la interface con la que se accede a las bases de datos o catálogos.
2. Desarrollar un cliente para cada servidor, lo que supondría desarrollar miles de clientes que tendrían que estar distribuidos por la Red.
3. Desarrollar un lenguaje de programación de clientes de bases de datos y cargar un pequeño programa para que cada base de datos pueda ser interpretada por cada cliente Web.

4. Desarrollar un lenguaje de display remoto.

5. Crear un lenguaje de formularios que permita que un cliente Web pueda cargar uno de estos formularios, el usuario lo rellene y envíe los datos para que sean procesados por el servidor.

Es decir, un cliente toma un formulario de un ordenador, el usuario rellena los datos y los envía a un servidor, que realiza la búsqueda y prepara los datos a través de un programa residente, reenviando los resultados al cliente que mostrará el producto de la búsqueda.

Las diferencias existentes entre el acceso vía telnet y el uso de formularios usando un cliente Web parecen claros, ya que con telnet el usuario no puede disponer de las funciones normales de un cliente Web, como salvar archivos o enviar los resultados vía e-mail. Usar un cliente Web tiene muchas ventajas ya que un cliente actúa como *pantalla final* de muchas bases de datos.

El hipertexto y los catálogos de bibliotecas.

Durante mucho tiempo se ha especulado acerca de la posibilidad de hacer opacs mucho más fáciles de utilizar. No es raro observar cómo los usuarios de catálogos automatizados muestran sus quejas sobre el funcionamiento de los opacs de las bibliotecas, y sobre la diversidad de sistemas con multitud de softwares diferentes que no son una ayuda sino un impedimento que retrae al usuario a la hora de realizar búsquedas.

Por el contrario estos nuevos sistemas que hemos analizado en páginas anteriores son una excelente y cómoda herramienta de mejorar las prestaciones de la recuperación de información en bibliotecas, y eliminar el problema del acceso telnet "*puro y duro*".

En el anexo 2 se puede ver una búsqueda de tales características. En el ejemplo se ha realizado una búsqueda en el catálogo de la Biblioteca de la Universidad de Utrecht. La búsqueda se puede realizar por diversos campos. Con un simple click del ratón (figura 2) se selecciona el campo autor, que queda resaltado, como se observa, colocando en el formulario a Cela como objeto de búsqueda.

En la figura 3 el sistema nos ha devuelto los resultados, entre los que aparece el autor que nos interesa, Camilo José Cela. Cada uno de los autores viene definido por un enlace hipertexto, de manera que presionando en ellos aparece la relación de sus obras en el catálogo de la biblioteca (figura 4). Las obras siguen estando definidas por enlaces hipertexto. La figura 5 reflejaría el registro de una de las obras.

Este simple juego de ratón y formularios simplifica el sistema de búsqueda, permite un acceso más rápido a la información deseada, evita el no siempre fácil manejo de un lenguaje de búsqueda -cuya dificultad (al usuario) se multiplica por cada lenguaje distinto de interrogación de cada software-, facilita la integración entre usuario y sistema, y en definitiva acerca el uso de sistemas automatizados de búsqueda a los usuarios al ser fáciles de aprender y usar.

Hemos realizado un análisis de aproximadamente un 75% de los servidores Web de bibliotecas, localizados en el catálogo que dispone y actualiza la Universidad de Washington en Seattle^{xi}. Se han separado cuatro áreas geográficas tal y como aparecen en dicho catálogo: Europa, Canadá, Estados Unidos y Oceanía (Australia y Nueva Zelanda).

El estudio se ha realizado sobre bibliotecas que han dispuesto sus sistemas de información a través de páginas HTML en un servidor Web, y centrados en el acceso a los catálogos automatizados. Se ha tratado de reflejar el porcentaje de bibliotecas que dentro de su página principal tienen o bien un enlace telnet o bien un acceso a través de formularios con scripts cgi que facilitan la búsqueda. Las gráficas apuntan claramente a que la mayoría de los catálogos aún están accesibles desde páginas HTML a través de enlace telnet (si bien ya es importante poder disponer de información general de las bibliotecas desde servidores WWW en Internet).

[GRAFICO TODOS.CHT]

Anexo 1

Infoseek Search, *<http://www.infoseek.com/>*

Está considerado como una de las mejores herramientas de búsqueda en Internet. Infoseek busca en servidores WWW, news de Usenet, en bases de datos a texto completo de revistas electrónicas, archivos de empresas, y en algunos servicios de noticias de agencias de prensa.

Lycos, *<http://query6.lycos.cs.cmu.edu/lycos-form.html/>*

Es un servicio mantenido desde la Universidad Carnegie Mellon. Realiza búsquedas en los títulos y contenidos de documentos en Internet. A 1 de mayo disponía de casi 4 millones de enlaces. El índice Lycos permite diariamente ir alimentando su base de datos con nuevos registros.

Yahoo, *<http://www.yahho.com/>*

Es una gran base de datos de enlaces, sobre un catálogo de materias organizado en forma de árbol. Contiene unas 44.000 entradas de servidores WWW.

WebCrawler, *<http://webcrawler.com/>*

Servicio de America Online, de reciente adquisición por esta empresa dentro de una estrategia de desarrollo y presencia en el mundo WWW.

Anexo 2

[FIGURA2.RS]

Figura 2. Selección del campo de búsqueda, definición del objeto de búsqueda y ejecución del proceso.

[FIGURA3.RS]

Figura 3. Resolución de la búsqueda por Cela.

[FIGURA4.RS]

Figura 4. Obras del autor.

[FIGURA5.RS]

Figura 5. Ejemplo de una de las obras de Camilo José Cela.

Anexo 3

Relación de documentos sobre cgi existentes en Internet.

1. A Simple CGI Email Handler, <http://siva.cshl.org/email/index.html>
2. Amiga CGI documentation, <http://www.phone.net/amiga-docs/>
3. Applewebsearch-WWW gateway for use with MacHTTP 1.3.1b1 or later,
<ftp://ftp.uth.tmc.edu/public/mac/MacHTTP/applewebsearch.sit.hqx>
4. CGI for the non-programmer, <http://www.catt.ncsu.edu/users/bex/www/tutor/index.html>
5. CGI Form Handling in Perl, <http://www.bio.cam.ac.uk/web/form.html>
6. CGI Programmer's Reference, <http://www.halcyon.com/hedlund/cgi-faq/>
7. CGI-Wrap- Secure User Access to CGI's with httpd, <ftp://pluto.cc.umn.edu/pub/cgiwrap/>
8. CGI.pm- a Perl5 CGI Library, http://www-genome.wi.mit.edu/ftp/pub/software/WWW/cgi_docs.html
9. Cgiutils manual page, <http://www.w3.org/hypertext/WWW/Daemon/User/CGI/cgiutils.html>
10. EIT's CGI Library, <http://wsk.eit.com/wsk/dist/doc/libcgi/libcgi.html>
11. Felipe's Applescript CGI's Examples, <http://edb518ea.edb.utexas.edu/scripts/cgix/cgix.html>
12. Forms in Perl. some perl scripts relevant to the handling of CGI forms,
<http://www.seas.upenn.edu/~mengwong/forms>
13. Processing of CGI arguments under Tcl, <http://www.lbl.gov/~clarsen/projects/htcl/http-proc-args.html>
14. Script cgi para ser usado en servidores MacHTTP,
<http://www.lib.ncsu.edu/staff/morgan/email-cgi.html>
15. Shareware CGI, <http://128.172.69.103:8080/cgi-bin/cgis.html>
16. The Common Gateway Interface, <http://hoohoo.ncsa.uiuc.edu/cgi/overview.html>
17. Un-CGI- a frontend for processing queries and forms from the Web on UNIX systems,
<http://www.hyperion.com/~koreth/uncgi.html>
18. Win-httpd CGI-DOS, <http://www.achilles.net/~john/cgi-dos.html>
19. Windows CGI 1.1 Description, <http://www.city.net/win-httpd/httpddoc/wincgi.htm>
20. Writing Get Scripts, <http://hoohoo.ncsa.uiuc.edu/docs/setup/admin/script-data/GetScripts.html>

Notas:

- Shell, Eric. *Writing for the Web: A primer for Librarians*. <http://bones.med.ohio-state.edu/eric/papers/primer/webdocs.html>
- iiSmith, John & Weiss, Stephen. An overview of hipertext, *Communications of the ACM*, julio 1988
- iiiCook, Peter. *An Encyclopedia Publisher's Perspective, Interactive, Multimedia*. Microsoft Press, 1988
- ivRaymond, Darrell & Tompa, Frank. Hypertext and the Oxford English Dictionary. *Communications of the ACM*, julio 1988.
- vRooks, Dana. The Virtual Library: Pitfalls, Promises and Potential. *PACS-Review* 4, n°5 (1993), pp. 22-29
- viPrince-Wilkin, John. *A gateway between the World-Wide Web and PAT: Exploiting SGML through the Web*. The *PACS-Review*, vol. 5, n7 (1994), pp. 5-27.
<gopher://info.lib.uh.edu:70/00/articles/e-journals/uhlibrary/pacsreview/v5/n7/pricewil.5n7>.
 Recuperable también enviando un e-mail a listserv@uhupvm1.uh.edu, con GET PRICEWIL PRV5N7 F=MAIL en el cuerpo del mensaje.
- vii*Virtual Reality Modeling Language (VRML)*, <http://vrml.wired.com/vrml.tech/vrml10-3.html>
- viiiSe han publicado en Internet interesantes estudios acerca de la presencia de los museos en las redes:
 Argoski, Jason. *Virtual Museums: The Web Experience*, <http://mirror.wwa.com/mirror/rov-int/95/museums.html>.
 Bowen, Johatham. *Museums and the Internet*,
<http://www.comlab.ox.ca.uk/archive/other/museums>.
- ix*Introduction cgi*, <http://hoohoo.ncsa.uiuc.edu/cgi/intro.html>.
- xGrobe, Michael. Instantaneous introduction to CGI scripts and HTML forms,
<http://kuhttp.cc.ukans.edu/info/forms/forms-intro.html>
- xiCatálogo de bibliotecas con servidores WWW. Universidad de Washington (EE.UU),<http://www.lib.washington.edu/~tdowhing/libweb.html>