

NORMA Z39.50, ACTUALIDAD, POSIBILIDADES. ¿ES NECESARIO UN CAMBIO DE ACTITUD?

Antonio de la Rosa Piñero*, José A. Senso Ruiz** y Ricardo Éfio Brun***

Resumen: La tecnología de los sistemas basados en la norma Z39.50 ofrece numerosas posibilidades, a priori bastante prometedoras. Conectividad de sistemas, recuperación de información muy estructurada, redefinición de aplicaciones (robots de búsqueda, agentes...), integración con sistemas propietarios, edición online y CD-ROM, protocolos... Sin embargo, en la actualidad se está favoreciendo muy poco el desarrollo de sistemas basados en esta norma. En este artículo se analizan las diversas posibilidades que ofrece la misma.

Palabras clave: Z39.50, recuperación de información, arquitectura cliente-servidor, análisis de clientes Z39.50, recuperación a texto completo, SQL, objetos digitales, pasarelas Z39.50, modelo STARTS.

Abstract: Z39.50 standard based systems technology offers many quite promising possibilities for the future: systems connectivity, retrieval of high-structured information, redefinition of applications (as search engines, agents and so forth), integration within proprietary software, online or CD-ROM publishing systems, protocols... However, the situation nowadays is very little advantageous in order to develop Z39.50 based systems. This paper analyses the range of possibilities offered by the mentioned standard.

Key words: Z39.50, information retrieval, client-server, Z39.50 clients review, full text retrieval, SQL, digital objects, Z39.50 gateways, STARTS model.

1 Introducción: el contexto profesional

Pensamos que en nuestros círculos existe una gran desconfianza hacia los nuevos medios que podrían amenazar con la pérdida del sentido del bibliotecario tradicional en su papel como intermediario de la información. En este punto, otros intermediarios tradicionales como editoriales y distribuidores (e incluso autores y grandes empresas mediáticas) también sufren una gran crisis de identidad en la que cada cual trata de situarse lo mejor posible dentro de un mercado en el que las nuevas tecnologías se han hecho omnipresentes e imprescindibles.

Lo que diferencia a las bibliotecas del resto de instituciones es, en muchos casos, su vocación mayoritariamente pública y sin ánimo de lucro. Esto es algo que se mantiene a pesar de la falta, a veces, de apoyo institucional y de las nuevas tendencias de autofinanciación.

* Groningen Graduate School for Behavioral and Cognitive Neurosciences. Holanda. Correo electrónico: a.de.la.rosa.pinero@let.rug.nl.

** Facultad de Documentación. Universidad de Granada. Correo electrónico: jsenso@platon.ugr.es.

*** Meta4 Recursos Humanos. Correo electrónico: ricardoe@meta4.es.
Recibido: 12-6-98.

Pero centrándonos en el tema, todo apunta a que los teleservicios van a aumentar espectacularmente su importancia en los próximos años, de una forma paralela al incremento de altas de usuarios de la WWW. En la actualidad, debemos admitir que los teleservicios que prestan las bibliotecas españolas a través de la WWW son como mínimo pobres (casi exclusivamente reducidos a intercambio de información bibliográfica o consulta online de catálogos).

Lo deseable... es la posibilidad de desarrollar todos los servicios de una biblioteca o centro de documentación (más algunos servicios añadidos) a través de redes telemáticas y concretamente a través de la WWW; lamentablemente es algo que en la actualidad no está ni mucho menos a nuestro alcance (salvo honrosas excepciones: iniciativas públicas concretas y muchas otras privadas e inaccesibles).

A pesar de que no esté a nuestro alcance, deberíamos estar preparados para estas nuevas condiciones. Sin embargo, existe una gran inadaptación, una gran reticencia en el mundo bibliotecario a ocupar este nuevo espacio al que parece inevitablemente destinado.

Todos estamos de acuerdo en que la adaptación debe ser un proceso gradual y mínimamente traumático, pero probablemente coincidiremos también en que lo que debe cambiar es el concepto de biblioteca (y no tratar de perpetuarlo sobre las nuevas tecnologías). Y quizá no nos estemos planteando con suficiente valentía ese cambio.

Sin embargo la tecnología avanza, con o sin nosotros, hacia los teleservicios integrales y comprensivos, el mayor protagonismo del usuario final (un usuario final mucho mejor formado en cuanto a búsqueda y recuperación de información y por lo tanto más exigente) y la disminución de la importancia de los intermediarios.

En este trabajo se pretende analizar hasta qué punto el estándar Z39.50 puede ayudarnos a afrontar estas perspectivas y por qué es deseable y útil la implantación de clientes Z39.50 muy sencillos en los navegadores de los futuros usuarios.

2 Introducción a la norma

El sistema de búsqueda en una base de datos tradicional (catálogo de biblioteca, CD-ROM...) tiene tres componentes básicos:

- Interfaz de usuario.
- Motor de búsqueda.
- Datos.

Un sistema Z39.50 sustituye la interfaz de usuario por el servidor-Z y añade una interfaz de usuario, que dependerá del programa cliente que se esté utilizando. Es decir:

- Cliente Z39.50.
- Servidor Z39.50.
- Motor de búsqueda.
- Datos.

La ventaja de este sistema radica en la separación de la interfaz de usuario del motor de búsqueda, lo que significa que una misma interfaz puede ser utilizada con

diferentes motores de búsqueda y bases de datos. El servidor se encarga de traducir el lenguaje, de forma que pueda ser comprendido por ambas partes de la cadena.

Z39.50 es un estándar que especifica un conjunto de reglas para gestionar las formas y procedimientos de interconexión remota de ordenadores, con el propósito de la búsqueda y recuperación de información.

La primera versión de este estándar nació en 1988. Dos años más tarde se crearon dos grupos que garantizaban su desarrollo y evolución. Por un lado nació ZIG (Z39.50 Implementors Group) y por otro Z39.50 Maintenance Agency, encargado de dar soporte al estándar. En 1992 se aprobó la versión 2, capaz de compatibilizar su uso en entornos OSI (Open Systems Interconnection). Recordemos que, un año antes, la ISO aprobó la norma Search and Retrieve SR (ISO 10162 y 10163). En 1995 se aprobó la versión 3, que fue aceptada como estándar ISO (con el número 23950) en marzo de 1997.

Si retomamos el esquema anterior, y lo integramos dentro de la estructura de red propia de Internet, nos encontraremos la siguiente estructura:

- Cliente WWW (un navegador).
- Pasarela WWW/Z39.50.
- Servidor Z39.50.
- Motor de búsqueda.
- Datos.

Tal y como afirma Félix de Moya, «no es posible que el cliente WWW sea capaz de enviar información a un servidor utilizando el protocolo Z39.50. Lo que hace el cliente es servirse de un gateway (pasarela), que es el que se entiende realmente con el servidor». Dentro de este modelo podemos diferenciar tres elementos:

- El cliente WWW y la pasarela WWW/Z39.50 forman el *frontend* (interfaz): un formulario en HTML suministra todas las variables de búsqueda —generalmente a través de un CGI—. El programa que gestiona la información del formulario actúa como una pasarela WWW/Z39.50 extrayendo todas las variables de la cadena que forman la petición y lo traduce a un lenguaje reconocible por el servidor Z39.50.
- La pasarela WWW/Z39.50 y el servidor Z39.50 gestionan la interacción entre cliente y servidor.
 - El cliente: se trata de un programa utilizado por el usuario final. Se encarga de lanzar las peticiones a los sistemas remotos y de presentar las respuestas producto de dichas peticiones.
 - El servidor: programa que recibe las preguntas del protocolo Z39.50. A partir de estas preguntas elabora una respuesta y la devuelve al cliente. Generalmente, las preguntas se traducen al formato de consulta de una base de datos, y la respuesta contiene los datos recuperados de la base de datos. Durante esta sesión se produce un proceso en el que se realizan las secuencias de *init*, *search* y *present*.
- El servidor Z39.50 constituye el *backend*. El *backend* del sistema es donde el servidor Z39.50 interactúa con un determinado motor de búsqueda, y éste con un grupo de datos (generalmente indizados).

Las etapas de la ejecución de este modelo son:

1. El usuario del OPAC selecciona la/s biblioteca/s objeto de la búsqueda a través de un menú de OPACs accesible en el cliente.
2. El usuario introduce su estrategia de búsqueda.
3. El cliente envía esta estrategia a la/s biblioteca/s objeto de la petición utilizando un servidor-Z (Z-Server) —que trabaja como una parte del programa de gestión bibliotecario—.
4. El servidor-Z traduce los términos de búsqueda a un lenguaje común (Z-Speak) y establece contacto con la/s biblioteca/s.
5. Existe una negociación preliminar entre el cliente-Z y el servidor-Z mediante la cual se determinan las reglas de asociación entre los dos sistemas.
6. El servidor-Z traduce el lenguaje común (Z-Speak) en una petición de información para que pueda ser comprendida por la/s biblioteca/s objeto, y recibe la respuesta sobre el número de registros encontrados.
7. El cliente-Z recibe los registros.
8. El OPAC presenta los registros al usuario.

Durante la negociación (paso 5), el cliente-Z (origen) y el servidor-Z (destino) intercambian una serie de mensajes. Cada uno de esos mensajes es identificado como un «servicio técnico». Los servicios se pueden agrupar en lo que la norma denomina *facility*. Existen 11, tal y como podemos ver en la tabla I.

El estándar también especifica el manejo y ordenación de los juegos de resultados, presentación de índices, apertura y cierre de sesiones y extensiones a otros servicios no definidos en el estándar mismo.

A pesar de que el origen de la norma Z39.50 se centra en la búsqueda y recuperación de información en bases de datos bibliográficas, su uso no se limita a

Tabla I
Funcionalidades (facility) del protocolo

<i>Facility</i>	<i>Notas</i>	<i>Versión</i>
1. Initialisation	Configuración de las reglas de comunicación, negociación de los niveles del servicio	2 y 3
2. Search	Envío de la ecuación de búsqueda a la base de datos y recogida de los primeros resultados	2 y 3
3. Retrieval	Recuperación de registros tal y como se especificó en el cliente-Z	2 y 3
4. Result-set-delete	Eliminación del conjunto de resultados almacenados en el servidor-Z	2 y 3
5. Access control	Proceso de verificación por parte del servidor-Z	2 y 3
6. Accounting-resource control	Gestión de cuentas	2 y 3
7. Sort	Ordenación del conjunto de resultados	3
8. Browse	Escaneo e índice en el servidor-Z	3
9. Extended services	El cliente-Z puede enviar peticiones de paquetes (préstamo interbibliotecario)	3
10. Explain	El cliente-Z interroga la base de datos en busca de información sobre la implementación del servidor-Z	3
11. Termination	Finaliza la negociación y la conexión	3

este campo. Puede utilizarse para consulta a bases de datos de texto completo, de sonidos, imágenes, contenidos multimedia. Esto se debe, en mayor medida, a que se ha creado partiendo de una norma «marco», que puede utilizarse para realizar búsquedas en diferentes lenguajes y recuperación de la información en diversos formatos.

Además, Z39.50 define otros aspectos, como:

- Lenguaje de búsquedas basado en juegos de atributos.
- Sintaxis de resultados aceptables: MARC (Machine Readable Cataloging), GRS-1 (Generic Record Syntax)...
- Lenguaje de construcción de resultados para su transferencia.
- Facultad para que el servidor pueda explicar su estructura interna y sus capacidades.

2.1 Algunos detalles

2.1.1 Búsqueda

La versión 3 (1995) de la norma Z39.50 permite establecer un sistema de búsqueda muy potente, que puede incluir:

- Todos los operadores booleanos (que no implementan en la actualidad la mayoría de los clientes).
- Operadores de comparación de fechas (*greater than, equal to...*).
- Operadores de proximidad.
- Diversas opciones para realizar el truncamiento.
- Búsquedas completas (*part of field, complete field...*).

Además, existen otros rasgos adicionales que ofrecen múltiples posibilidades de búsqueda:

- Autenticación. Esto permite que el servidor-Z pueda controlar quién accede a las bases de datos.
- Control de los recursos y de los accesos (cuentas).
- Opción «explain», que permite obtener información sobre bases de datos remotas, servicios disponibles...
- Visualización del índice
- Definir el formato de los registros.

2.1.2 Servicios ampliados

También se define el uso de la norma para implementar lo que denomina «servicios extendidos». A pesar de no estar incluidos dentro de la norma, Z39.50 incluye los siguientes servicios para facilitar el control:

- Almacenar resultados.
- Guardar una petición de consulta.
- Definir un esquema de búsqueda.
- Solicitar un ejemplar.
- Actualización de la base de datos.
- Crear un fichero de exportación.

2.2 Posibilidades de la norma

A pesar de que la tecnología sobre la que se basa es bastante simple, las posibilidades que ofrece son muy grandes. Entre los principales beneficios que esta norma aporta al intercambio de información destacan:

- Los clientes-Z pueden solicitar una misma petición de información a diversas bibliotecas de forma simultánea.
- El formato básico con el que trabaja Z39.50 para el intercambio es MARC.
- Es posible ampliar los servicios que se ofrecen a través de esta norma, incluyendo la presentación de documentos, actualización de bases de datos y almacenamiento de ecuaciones de búsqueda. Todos los sistemas que incluyen esta norma —como ocurre con otros procesos, como el préstamo interbibliotecario— se convierten en abiertos.

Dentro del trabajo bibliotecario, las implicaciones de la norma se pueden observar en numerosos campos:

- OPACs.
- Catalogación (catalogación compartida y creación de catálogos generales).
- Préstamo interbibliotecario.
- Acceso a servicios de CD-ROM (por ejemplo, ERL v2.0 de SilverPlatter).
- Difusión selectiva de información (la última versión permite almacenar peticiones, y ejecutarlas de forma automática pasado un período de tiempo).
- Gestión de información a través de web (posibilidades de futuro para las bibliotecas, bibliotecas virtuales en WWW).
- Actualización de bases de datos. Un cliente-Z puede recuperar un registro, editarlo y volver a enviarlo a la base de datos. Por ejemplo, GeoCAT (de la compañía Geac Illustrates) permite realizar esta función.

La norma Z39.50 puede ser incorporada a numerosos productos, ya sea como módulo dentro de ellos, o como programa cliente con uso exclusivo. La independencia de plataforma la convierte en un sistema abierto, es decir, que un cliente-Z realizado para funcionar en el sistema operativo MS-DOS puede consultar servidores que residan en máquinas Unix o NT. Las diversas implementaciones realizadas en los productos actuales nos permiten dividir las aplicaciones Z en:

- OPAC: programas integrados dentro de sistemas de gestión bibliotecaria. Los OPACs basados en Z39.50 permiten a los usuarios realizar búsquedas en los

catálogos y definir un conjunto de bibliotecas para lanzar sus consultas. Se trata del uso más normal del cliente-Z, y se puede encontrar como programas hechos en Unix, Windows, clientes web y, cada vez más, clientes Java.

- Módulo de catalogación: generalmente este módulo se encuentra conectado con el resto del sistema integrado de gestión bibliotecaria (formando parte de él, o a través de un cliente dedicado exclusivamente a realizar esta función).
- Herramientas bibliográficas: Bookwhere o Znavigator.

Las posibilidades de la norma permiten aplicar diferentes grados de implementación durante determinados procesos de la conversación entre el cliente y el servidor. Así, por ejemplo, en la fase de inicialización, es posible establecer diversos «atajos» que permitan reducir tiempos de espera y facilitar una comunicación más fluida. Negociando sobre aspectos previamente especificados, es posible eliminar procesos superfluos.

Como se comentó anteriormente, la norma Z39.50 establece «sólo» las directrices que deben regir la comunicación entre un origen (cliente-Z) y un destino (servidor-Z), pero no especifica en absoluto la forma en que deben constituirse dichos sistemas, o cómo deben mostrar la información (desde la presentación de opciones hasta la lista de resultados, es decir, que no determina las características de la interfaz). Así nos podemos encontrar con multitud de arquitecturas:

- Cliente-Z basado en un programa para PC: Bookwhere, Znavigator...
- Terminal (ya sea conectado a un OPAC o a un navegador de Internet): SIRSI Unicorn.
- Navegador de Internet (por medio de *plug-in* o de módulos tipo *search://*, que se explicará a continuación).

Hasta ahora hemos observado la cantidad de opciones que ofrece el uso de la norma. Sin embargo, el amplio abanico de opciones puede ser tanto una ventaja como una desventaja. Uno de estos problemas lo podemos encontrar a la hora de decidir el nivel de servicio que deberá soportar el cliente. Este punto deberá evaluarse de forma diferente si estamos trabajando dentro de un sistema integrado de gestión bibliotecaria (SIGB), o si lo que se desea es realizar un cliente independiente. Durante el proceso de desarrollo de un cliente Z39.50 el diseñador debe decidir no sólo la versión de la norma con la que trabajará, sino también los servicios que piensa incluir.

3 Estado actual de la norma. Servidores de bases de datos

No sólo los fabricantes de sistemas integrados de gestión bibliotecaria han desarrollado módulos para garantizar la conectividad: los principales distribuidores de bases de datos también han estado trabajando en esta misma línea. En ocasiones, los fabricantes de sistemas integrados de gestión bibliotecaria ofrecen a sus clientes la posibilidad de participar en proyectos orientados hacia la creación de catálogos colectivos: eBase, de BCLMD, o la Red Libertas de SLS (Innopac) pueden ser ejemplos de estas iniciativas. Estos catálogos suelen hacerse accesibles, siempre con validación de acceso, a través de clientes Z39.50.

Dentro del mundo de la información online, las iniciativas más destacadas en la implementación de servicios Z39.50 son, sin duda alguna, las de SilverPlatter y Ovid.

SilverPlatter ofrece conectividad a sus servidores ERL, a partir de la versión 2, a través de Z39.50. De esta forma, cualquier usuario podrá interrogar las bases de datos de SilverPlatter utilizando un cliente Z39.50, versión 2 o versión 3, sin necesidad de trabajar con el software cliente propietario de SilverPlatter.

El servidor ERL soporta la versión 3. Esto permite que los clientes Z39.50 dispongan de los siguientes servicios cuando interrogan a un servidor ERL: Scan (visualización de índices), Explain (obtener información sobre la estructura de la base de datos), obtener registros en formato MARC, y SUTRS (Simple Unstructured Text Record Syntax), Termination (cierre de sesiones) y Segmentation (recibir resultados que exceden el tamaño de la transacción).

Para poder interrogar las bases de datos SilverPlatter, algunas de ellas se han estructurado de acuerdo con el formato MARC (AGRICOLA, ABI/INFORM, ABI/INFORM Full Text, INSPEC, CANCER-CD, Dissertation Abstract, MLA Bibliography, MEDLINE Express, PAIS International, etc.), y los servidores ERL reconocen una equivalencia entre los campos definidos en el grupo BIB-1 de la norma Z39.50 y los SilverPlatter Universal Fields, utilizados en la estructura de todas las bases de datos que ofrece este distribuidor.

El segundo ejemplo representativo de un distribuidor de información comprometido con la norma Z39.50 es el de OVID. Esta compañía implementa la versión 3 en su software cliente para Windows y vt-100 y en el software servidor. De esta forma, un cliente OVID podrá acceder a la información gestionada por otros servidores Z39.50, y desde cualquier cliente Z39.50 con posibilidad de autenticación se podrá acceder a los servidores de bases de datos OVID.

Como pasarela entre el WWW y los servidores Z39.50, Ovid Web Gateway incluye un componente que permite acceder a bases de datos Z39.50 a través de Ovid Server.

Otro distribuidor comprometido con la implementación de la norma es la OCLC, con sus bases de datos FirstSearch, WorldCat, etc.

La disponibilidad, no sólo de catálogos bibliográficos, sino de las grandes colecciones de información en línea distribuida por los mayores *hosts* online es uno de los factores críticos a la hora de considerar la conveniencia de implantar la conectividad a través de Z39.50, y las posibilidades abiertas para el futuro de la norma.

3.1 Clientes Z39.50

Uno de los mayores problemas que encontramos a la hora de evaluar la eficacia de los clientes Z39.50 está provocado, precisamente, por una de sus mayores ventajas: la gran facilidad con la que es posible realizar búsquedas y recuperar información. El poder consultar diversas bases de datos a la vez, utilizando una sesión y una misma ecuación de búsqueda, conlleva un problema obvio, y es que no todos los servidores implementan los mismos servicios en sus sistemas. Esto hace que, en muchas ocasiones, los registros que se recuperan no sean homogéneos, produciéndose numerosas incoherencias entre unos y otros.

Esto se puede deber a varias causas: error en la traducción de la petición, o du-

rante el proceso de negociación de inicio de sesión, incompatibilidad entre las versiones del cliente y el servidor, la configuración del servidor puede no ser la más idónea para que el cliente pueda desarrollar todas sus posibilidades de búsqueda o, simplemente, la base de datos del servidor no reconoce los servicios que solicita el cliente.

Si bien es cierto que la mayoría de estos problemas sólo aparecen cuando se pretende utilizar opciones más avanzadas, y fuera de las clásicas búsquedas simples con dos o tres términos unidos por operadores booleanos, la norma contiene las especificaciones básicas para realizar procesos de diagnóstico, que tienen como finalidad su eliminación.

La comunicación entre el cliente y el servidor se realiza por medio de la notación polaca inversa (Reverse Polish Notation, RPN). Este lenguaje (aplicado principalmente a cálculos matemáticos, no en vano es el utilizado por las calculadoras científicas HP) tiene un nivel de complejidad medio y presenta como característica más destacable la homogeneidad y simplicidad en su sintaxis. RPN une los términos de búsqueda entre sí, añadiéndoles operadores booleanos. Por ejemplo, la búsqueda: (INFORMATION AND RETRIEVAL) OR IRS se traduciría como: INFORMATION: RETRIEVAL: AND: IRS: OR. En general OPERANDO: OPERANDO: OPERADOR, ya que en RPN se lee la ecuación de búsqueda de derecha a izquierda (para ahorrarse los paréntesis). Algunas empresas han realizado modificaciones sobre esta notación (Znavigator utiliza Prefix RPN, que une los términos en pares).

A la hora de construir y presentar la ecuación de búsqueda, cada cliente presenta diferentes opciones, que hacen que el proceso sea más o menos fácil. Lo más común es encontrarse con árboles, generados a partir de una búsqueda en concreto (Bookwhere o Znavigator). De todas formas, las ecuaciones se pueden modificar en cualquier momento. También existe la posibilidad de realizar búsquedas complejas, en el lenguaje de consulta clásico, y que luego el cliente se encargue de traducir a RPN.

A continuación vamos a comentar las diversas opciones que ofrecen tres de los más conocidos clientes Z39.50.

3.1.1 BookWhere 2000

El núcleo de su sistema de búsqueda es YAZ (Yet Another Z39.50), conjunto de librerías de enlace dinámico creadas por la empresa danesa Index Data. La versión 2000, para Windows 95 y NT 4.0, contiene una gran cantidad de opciones, que nos permiten configurar el cliente a nuestro gusto. Es posible comenzar iniciando una nueva sesión o relanzar una previa.

Las búsquedas se pueden comenzar de dos formas: o bien escogiendo las bases de datos sobre las que lanzar las consultas, o bien sobre un grupo de bases de datos, que previamente habremos seleccionado. De esta forma, dependiendo del tipo de consulta que se quiera realizar, podemos acudir a un grupo o a otro (bases de datos de universidades, derecho, economía...). Los *hosts* se encuentran ordenados por nombre, tipo de biblioteca y área geográfica. Añadir nuevas bases de datos es muy fácil. Además existe una opción que permite hacer tests a tantos *hosts* como se desee, con el fin de conocer su disponibilidad antes de comenzar a buscar. Estos tests nos aportan datos como la versión de la norma implementada en el servidor o las opciones que soporta.

El cliente tiene dos opciones de búsqueda. La simple permite introducir términos en una serie de campos (que también pueden ser configurables) que se unen por defecto con el operador booleano *and*. La segunda opción, denominada «power», permite usar los operadores *and*, *or* y *not*, la introducción de múltiples términos de búsqueda, y modificar las opciones dentro de cada campo (truncar, posición de los términos, relación con otros términos...).

La búsqueda produce tres ventanas. En la primera de ellas, situada en la parte superior de la interfaz del cliente, se encuentra la lista de registros recuperados, ordenados alfabéticamente por nombre de autor (o título o fecha, según se decida). Al principio de cada fila, que representa un registro, aparece un icono, que especifica el tipo de soporte en el que se encuentra la obra.

La parte inferior del cliente se divide en dos zonas. En la de la izquierda se encuentran los *hosts* en los que se han realizado las búsquedas, y una lista alfabética de materias relacionadas con los términos introducidos en la consulta. Entre paréntesis se puede comprobar el número de registros que existen de cada caso. En la ventana de la derecha nos encontramos con el registro bibliográfico propiamente dicho, presentado como texto etiquetado o como formato MARC. El registro contiene una serie de términos resaltados. Si se pulsa con el ratón encima de alguno de ellos, se lanzará otra sesión, que buscará ese descriptor en las mismas bases de datos en las que se consultó en un principio.

Una vez seleccionado un grupo de registros, el programa ofrece la posibilidad de exportarlos en diferentes formatos: USMARC, Procite, Reference Manager, Ibid o texto plano.

A pesar de que BookWhere 2000 v2.2 está basado en la versión 3 de la norma Z39.50, no incluye todas las opciones que ésta ofrece. No obstante, destacan como novedades con respecto a la anterior versión, la posibilidad de filtrar los resultados, *scan* o la ordenación de grupos de resultados.

3.1.2 WinWillow

Software desarrollado por la Universidad de Washington. Una vez ejecutado el programa, aparecerá la lista de bases de datos a las que se podrá acceder. Lamentablemente, en la opción evaluada (la 1.42) no es posible añadir bases de datos a las que existen. Una vez seleccionadas, el programa nos presenta la pantalla de búsqueda, donde introduciremos los términos. En la opción «Search Criteria» es posible especificar los campos en los que se quiere realizar la consulta.

Realmente, la posibilidad de búsqueda más novedosa es la denominada «Basis Mode». Esta opción, que sólo se activa con determinadas bases de datos, permite buscar por *browsing*, en los campos de título, autor, materia y palabras clave como mínimo. Basta con introducir uno o varios términos, y el cliente nos mostrará su ubicación alfabética dentro del índice de la base de datos.

Una vez realizada la consulta, y seleccionado el término que más se adecúa a las necesidades de información, el programa abre la ventana denominada «Summary», donde aparece la lista de registros. Después de seleccionar el deseado, se activa la ventana «Record Retrieval» que muestra el registro completo (se puede configurar el número de campos a mostrar). En este momento el usuario puede salvar el resultado

en formato de texto, enviarlo a una dirección de correo electrónico, o copiarlo para luego pegarlo en cualquier aplicación Windows.

Excepto la opción «Basis Mode» y la posibilidad de configurar los campos de los registros, este cliente ofrece unas posibilidades muy pobres, mostrando grandes carencias e implementando pocas opciones de la última versión de la norma Z39.50.

3.1.3 ZNavigator

Forma parte del proyecto CaseLibrary (integrado dentro del programa Telematics for Libraries del DGXIII Libraries Programme). El método para comenzar a realizar búsquedas es similar al de los anteriores programas cliente comentados. Al principio ofrece una lista de servidores (es posible añadir nuevos), y tras la selección y conexión, se podrá comenzar la búsqueda. ZNavigator abre una ventana por cada uno de los *hosts* a los que intenta conectar. Esto quiere decir que si deseamos conectarnos a 8 bases de datos, y no se puede realizar esta conexión, el programa abrirá 8 ventanas en las que se nos indicará el error.

Las búsquedas se pueden introducir en la ventana destinada a la opción simple (introducir términos sin operadores, que se unen entre sí con el operador por defecto *and*) o la denominada «complex search» en la que los términos se unen por medio de una variante de la notación polaca inversa. Todas las consultas se quedan almacenadas en un fichero, por lo que pueden ser ejecutadas cuantas veces se quiera en diferentes sesiones.

Los resultados se obtienen en una tabla clasificada por orden de recuperación, aunque se pueden ordenar por cualquier otro parámetro. Cuando se realiza una consulta a varias bases de datos, el resultado de cada *host* se presenta en ventanas diferentes. Seleccionando con el ratón se accede al registro completo en formato MARC o por campos.

Una de las mayores desventajas de la versión actual de ZNavigator es que no permite almacenar los registros seleccionados en «full» MARC (esto viene dado porque en CaseLibrary no se proyectó realizar la catalogación utilizando el cliente), aunque se pueden copiar y se recuperan en MARC con los códigos de subcampo, pero sin indicadores.

En la tabla II se puede observar una comparación entre los diversos clientes Z39.50 analizados.

4 Z39.50 y otros sistemas

Desde su creación en 1988, Z39.50 sufrió una marcada dependencia con el mundo bibliotecario. El hecho de que el primer conjunto de atributos sobre los que se estableció la posibilidad de definir ecuaciones de búsqueda, BIB-1, utilice atributos usados tradicionalmente para la descripción de recursos bibliográficos es significativo de esta orientación inicial.

Todavía hoy, Z39.50 sigue estando asociada al mundo de las bibliotecas, y existe un prejuicio que hace que muchos proveedores de sistemas de información descarten el desarrollo de sistemas basados en la norma y no la consideren como una alternati-

Tabla II
Comparación entre los diversos clientes Z39.50 analizados

	<i>BookWhere 2000</i>	<i>WinWillow</i>	<i>Znavigator</i>
<i>Búsqueda</i>			
Permite múltiples sesiones	X	X	X
Múltiples hosts por sesión	X		X
Búsqueda hipertextual	X		
Truncamiento	X		X
Operadores de proximidad			
<i>Mostrar resultados</i>			
Ordenar registros recuperados en el cliente	X		X
Seleccionar el campo para ordenar	X		X
Opción de registro completo	X	X	X
Recuperar registro completo sin preguntar	X	X	
Ver registro en MARC	X		X
<i>Gestión de resultados</i>			
Almacenar sesión	X		X
Histórico de búsquedas			X
Filtrar duplicados			X
<i>Administración de hosts</i>			
Agrupar hosts	X		X
Hacer test a hosts	X		
Añadir Hosts	X		X
<i>Exportar registros</i>			
MARC	X		X
Procite	X		
Texto plano	X	X	X
Otros (Reference Manager, Inmagic, Citation 7...)	X	X	X

va a tener en cuenta para cubrir las necesidades de conectividad en sistemas de recuperación de información.

Si se descarta Z39.50, ¿hacia dónde podemos dirigirnos? En la evolución de Z39.50 no podemos obviar el impacto que la norma ha tenido en otros sistemas, y la situación contraria: es decir, la influencia que otros modelos para la recuperación de información han aportado sobre la definición, actual o futura, de la norma.

Dentro del primer apartado, es necesario considerar el desarrollo de WAIS (Wide Area Information System). Dentro del segundo, hemos de citar la propuesta STARTS y su posterior influencia en Z39.50.

4.1 Wais y Z39.50

El sistema de edición Wide Area Information Servers (WAIS) se diseñó para ayudar a los usuarios a encontrar información en una red de ordenadores. Los principios en los que se ha basado el desarrollo son:

- Un sistema de información de área amplia en red para buscar, mostrar y editar información.
- Basado en normas.
- Fácil de usar.
- Flexible y escalable.

Sobre estas bases un gran grupo de desarrolladores, editores, instituciones normativas, bibliotecas, agencias gubernamentales, instituciones de enseñanza y usuarios han ayudado al desarrollo del sistema WAIS.

La arquitectura software tiene cuatro componentes principales: *el cliente, el servidor, la base de datos y el protocolo*. El cliente es un programa interfaz de usuario que remite peticiones de información a servidores WAIS locales o remotos. Existen clientes para casi todos los entornos operativos. El servidor es un software que da servicio a los clientes WAIS y que está disponible sobre todo para plataformas UNIX. El servidor generalmente funciona en una máquina que a su vez contiene una o más bases de datos WAIS. El protocolo Z39.50-1988 se usa para conectar clientes y servidores y se basa en la versión de 1988 de la norma NISO Z39.50 Information Retrieval Service and Protocol Standard.

Los objetivos de diseño de WAIS fueron:

1. Proporcionar acceso para los usuarios a información bibliográfica y no bibliográfica incluyendo texto completo e imágenes.
Puesto que Z39.50-1988 surgió de la comunidad bibliográfica (fue diseñada para buscar y recuperar registros bibliográficos en formato USMARC) se han utilizado desarrollos adicionales para servir información no-bibliográfica. Esto fue necesario también para servir documentos en formatos diferentes, por ejemplo, RTF, Postscript, imágenes gif, etc.
2. Conservar una interface cliente/servidor simple e independiente de los cambios de funcionalidad del servidor. Originalmente la cadena de texto introducida por el usuario era transmitida al servidor sin dividirse en una consulta tipo-1 expresada en RPN, como en otras aplicaciones bibliográficas. En su lugar, WAIS definió un nuevo tipo de consulta (tipo-3) que contenía simplemente la cadena de texto. De esta forma, ni el cliente ni el usuario necesitaban conocer los atributos Z39.50 que soportaba el servidor. Además, el software cliente no tenía que modificarse para adaptarse al servidor cuando éste incorporaba algún cambio.
3. Proporcionar feedback de relevancia. Es decir, la posibilidad de encontrar un conjunto de documentos similar al seleccionado (o una porción del mismo) en primer término. WAIS incluía documentos usados en el feedback de relevancia como parte de la consulta de tipo 3.
4. Permitir al servidor operar de forma independiente. Los resultados de las búsquedas no son almacenados en el propio servidor. En términos Z39.50, el servidor ejercita su derecho de borrar unilateralmente el conjunto de resultados tan pronto como ha mandado la respuesta a la búsqueda.
5. Proporcionar al cliente la capacidad de recuperar documentos en partes. Podía realizarse de muchas formas con Z39.50-1988. Por eso se diseñaron supuestos específicos para implementar esta función. Se requería la posibilidad de

acceder a una porción de un documento tanto para la recuperación como para el feedback de relevancia.

6. Ejecutar sobre TCP. La norma Z39.50-1988 se diseñó para correr en la capa de aplicación, usando los servicios de presentación del modelo de referencia (OSI) Open Systems Interconnection. Debido a la mayor popularidad de los protocolos TCP/IP y en general de Internet, WAIS se diseñó para ser ejecutado sobre TCP.

Trabajando con el Grupo de Implementación de Z39.50 (ZIG), los desarrolladores de WAIS usaron una versión recomendada de Z39.50-1988 y supuestos específicos para satisfacer sus objetivos. Después, muchos de esos supuestos han ayudado en la definición de las siguientes versiones de Z39.50. WAIS soportaba los módulos *init* y *search* de Z39.50-1988. Las funciones de búsqueda y recuperación de información se implementaron usando el módulo *search*.

4.2 STARTS

STARTS es la propuesta para la definición de un protocolo para la recuperación de documentos distribuidos. Se desarrolló en el marco del proyecto de bibliotecas digitales de la Universidad de Stanford, y se hizo pública en agosto de 1996. Para su definición, el equipo encargado de la dirección del proyecto solicitó la colaboración de los principales fabricantes implicados en el desarrollo de tecnologías para la recuperación textual (Verity, Fulcrum, InfoSeek, PLS, Harvest, etc.), y para el desarrollo de Internet (Netscape Communication, Microsoft, Hewlett-Packard, etc.). En la lista de empresas que colaboraron en el desarrollo de la especificación también se incluyó WAIS, empresa dedicada a la comercialización del sistema de indización WAIS, descrito en el apartado anterior, que posteriormente sería adquirida por Fulcrum Technologies.

Dentro de los propósitos de STARTS se encontraba el desarrollo de una especificación que resolviese los problemas que surgen al tratar de recuperar documentos gestionados por sistemas heterogéneos. De hecho, si bien el modelo de recuperación propuesto por la norma Z39.50 ofrece una abstracción válida en un proceso de recuperación entre dos máquinas, cliente u *origin* y servidor o *target*, la generalización Internet nos obliga a enfocar el problema de la recuperación documental desde otra perspectiva: la necesidad de interactuar simultáneamente con un número variable de fuentes de información, y la conveniencia de obtener un conjunto unificado de registros. En este modelo, la procedencia de los registros incluidos en el conjunto resultado debería ser transparente para el usuario, y no tendría que tener ningún impacto a efectos de ordenación de resultados por relevancia o por cualquier otro criterio.

Es decir, en un sistema de búsqueda basado en este protocolo, el usuario lanzaría una ecuación de búsqueda a una serie de servidores de información. Tras procesarse las búsquedas, el cliente recibiría una única lista con los registros que cumpliesen las condiciones indicadas en la búsqueda, a partir de las distintas fuentes de información gestionadas por los servidores a los que se dirigió la petición.

El modelo STARTS propone un uso intensivo de la ordenación de registros recuperados a partir de su relevancia. A pesar de la importancia que se concede a este factor, en ningún momento se propone la adopción de algoritmos de ordenación de

terminados. La implementación de los sistemas de ordenación de registros se deja en manos de los proveedores de sistemas de recuperación textual que se ejecuten en los servidores. Para evitar las inconsistencias entre los sistemas de ranking, y minimizar las diferencias entre ellos, el protocolo propone:

1. Que sea el cliente el que indique en la ecuación de búsqueda el algoritmo de búsqueda que debe utilizar el servidor.
2. Que los servidores incluyan, junto al resultado de la búsqueda, información sobre el algoritmo de ordenación que haya utilizado.

Retomando la comparación entre el estado actual de Z39.50 y el modelo propuesto por STARTS, tenemos que señalar que, si bien las aplicaciones clientes evaluadas para la redacción de este artículo ofrecen la posibilidad de establecer múltiples sesiones simultáneas con dos o más servidores Z, ninguno de ellos es capaz de generar una lista unificada, en orden de relevancia, de todos los registros recuperados.

En la descripción del modelo de recuperación propuesto por la especificación STARTS, se distinguieron tres partes: arquitectura, estructura de la ecuación de búsqueda, y la información que deben ofrecer los servidores para poder crear una única visión unificada de todos los conjuntos-resultado.

El modelo de ecuación de búsqueda STARTS toma como base:

1. Las ecuaciones tipo-101 de la norma Z39.50 (ecuaciones booleanas).
2. Un subconjunto de la relación de atributos GILS, que a su vez incluye el conjunto bib-1 propuesto en la definición de la norma Z39.50. Estos atributos se utilizarán como propiedades descriptivas e identificativas de los documentos, y como un índice de puntos de acceso a los mismos.

Partiendo de esta base, una ecuación de búsqueda STARTS cuenta con dos partes diferenciadas: la ecuación de búsqueda propiamente dicha, es decir, la especificación de los datos que queremos buscar, y las instrucciones que debe utilizar el servidor para generar el ranking de los documentos recuperados.

La tabla III recoge todos los atributos disponibles en el conjunto de los atributos utilizados por STARTS. Como ya se ha indicado, se trata de un subconjunto del conjunto de atributos GILS. Para cada atributo se indica si es requerido o no. Por requerido entendemos que deben ser reconocidos y gestionados por los servidores de información que interactúan con la base de datos.

4.2.1 Ranking en la ecuación de búsqueda

La segunda parte de la ecuación de búsqueda recoge las instrucciones que debe seguir el servidor para ordenar los registros recuperados. En esta parte se indicará qué términos deben utilizarse para ponderar los documentos, y qué peso se quiere dar a cada uno de ellos. El peso de un término será un número comprendido entre 0 y 1.

Es posible indicar un número ilimitado de términos en esta parte de la ecuación. La sintaxis propuesta por STARTS propone la utilización del operador *list* para agrupar todos los criterios que deben tenerse en cuenta para generar el ranking.

Tabla III
Atributos utilizados por STARTS

<i>Atributo</i>	<i>Obligatorio</i>	<i>Descripción</i>
Title	Sí	Título del documento.
Author	Sí	Autor del documento.
Body-of-text	Sí	Cuerpo del documento, texto completo.
Document-text	Sí	Cuerpo del documento. Es igual que el anterior, pero Document-text se utiliza en la parte de la búsqueda que indica las condiciones de ranking.
Date/time-last-modified	Sí	Fecha y hora en la que se modificó el documento por última vez.
Any	Sí	Hace referencia a cualquier campo.
Linkage	Sí	URL del documento.
Linkage-type	Sí	Tipo MIME del documento.
Cross-reference-linkage	Sí	Lista de URLs que incluye el documento.
Language	Sí	Lenguaje en el que está escrito el documento.
Free-form-text	Sí	Ecuación de búsqueda en una sintaxis comprensible por el servidor, ajena a la especificación STARTS.

4.2.2 Información adicional en la ecuación de búsqueda

Además de los términos que deben incluir los registros para ser recuperados, y la información necesaria para generar el ranking, una ecuación de búsqueda STARTS podrá contener los siguientes elementos:

- Indicador booleano que permite al servidor saber si debe activar o desactivar una lista de palabras vacías.
- Conjunto de atributos por defecto que se utilizan para definir la ecuación.
- Lenguaje usado por defecto en la ecuación.
- Fuentes de datos adicionales a la fuente de datos seleccionada por defecto. La búsqueda se propagará a estas fuentes adicionales.
- Indicaciones sobre cómo debe ser la respuesta que envíe el servidor. Aquí se incluyen:
 - Atributos que debe devolver el servidor para los registros recuperados.
 - Ponderación mínima aceptable para que un documento se incluya en la lista de resultados.
 - Número máximo de documentos que se pueden incluir en la lista de resultados.

4.2.3 Generación de un listado único y eliminación de duplicados

Como se ha indicado antes, la ecuación de búsqueda se puede dirigir a un número variable de servidores. Cada servidor devolverá la referencia a los documentos que cumplan las condiciones indicadas en la ecuación de búsqueda. El siguiente paso en el proceso de recuperación será obtener una lista de resultados única, en la que se eliminen los duplicados, y en la que se ordenen los registros por un orden de relevancia obtenido mediante un criterio aplicable a todos los registros, independientemente

del servidor que los haya enviado como resultado, y de los criterios de ordenación que internamente utilice ese servidor.

Para garantizar la coherencia en los mecanismos de ranking, la especificación no propone la utilización de ningún algoritmo específico. En su lugar, se indica qué información debe acompañar a cada registro recuperado para poder obtener su ponderación.

Respecto a este último punto, el servidor deberá señalar, para cada documento recuperado y término incluido en la búsqueda:

1. Frecuencia del término en el documento.
2. Ponderación asignada al término, en ese documento, por el software servidor que lo ha recuperado de la base de datos.
3. Número de documentos en la base de datos en los que aparece el término propuesto como criterio de búsqueda.

A partir de estos datos, el sistema cliente será capaz de generar la ponderación de cada documento, independientemente del servidor del que se haya recuperado, y obtener una única lista ordenada con todos los resultados.

4.2.4 Intercambios de información adicionales

En los apartados anteriores se ha descrito el modelo de recuperación básico propuesto por STARTS. La especificación también incluye otros intercambios de información que sirven de soporte al sistema:

1. Metadatos sobre la implementación de STARTS en el servidor: partes de la ecuación de búsqueda que soporta, conjuntos de atributos disponibles, etc.
2. Información sobre las colecciones de documentos que gestiona cada servidor.
3. Algoritmos de búsqueda, listado de palabras vacías, idioma, etc.

4.2.5 STARTS y Z39.50

Todos los mensajes que intercambian el cliente y los servidores en el modelo STARTS se enviarán a través del nivel de aplicación http de TCP/IP. La información se codificará de acuerdo con la sintaxis SOIF (Summary Object Interchange Format). En esto, la propuesta es mucho más simple que el modelo descrito en la norma Z39.50.

Desde un punto de vista conceptual, STARTS no ofrece la flexibilidad de la norma Z39.50, ya que su definición es mucho menos abstracta. Sin embargo, STARTS presenta la solución a uno de los problemas más difíciles de afrontar por los implementadores de la norma Z39.50: obtener un único conjunto a partir de múltiples fuentes, con eliminación de duplicados y con la posibilidad de aprovechar las ventajas que ofrece la ordenación por relevancia.

Esto llevó, poco tiempo después de la redacción del cuarto borrador de la especificación STARTS, al desarrollo de un «perfil» de la norma Z39.50 que mostrase las po-

sibilidades de adaptarla para poder trabajar con el nuevo marco de trabajo STARTS. Así surgió ZDSR, para el que se desarrolló la adaptación de los siguientes servicios:

1. Inicialización-Init.
2. Búsqueda-Search:
 - Soporta ecuaciones tipo-1.
 - Incluye información sobre el ranking en el parámetro additionalSearchInformation.
 - Algoritmo de ranking solicitado por el cliente (figuraría también en Explain).
 - Ordenación de registros mediante el parámetro OtherInfo de la petición de búsqueda.
 - Recuperación de documentos en sintaxis GRS-1.
 - Relevance-feedback tipo RFDT, implementada como una nueva clase de ecuación.
 - Present.

La adaptación de la norma proponía un uso intensivo de las posibilidades de encapsulación, es decir, la posibilidad de enviar de una sola vez dos peticiones de servicios distintos desde el cliente hasta el servidor, o dos respuestas a servicios distintos desde el servidor hasta el cliente.

4.3 Conclusión

Si WAIS se desarrolló tomando como base Z39.50-1988, STARTS es un ejemplo de cómo nuevas definiciones y modelos de sistemas de recuperación pueden afectar al desarrollo de la norma. El caso de STARTS demuestra el carácter abierto del estándar y su adaptabilidad.

La adaptabilidad es, sin duda, el factor crítico en el futuro éxito de la norma. Pero no sólo debe orientarse hacia nuevos modelos de recuperación; también hay que tener en cuenta la capacidad de adaptarse a los cambios en las tendencias en el desarrollo de software y al modo de trabajo de todos los usuarios potenciales. Respecto a esto último, el impacto de la World Wide Web en los últimos años ha sido uno de los puntos críticos que parecían poner en entredicho la utilidad de la especificación. La rapidez con la que la comunidad de usuarios y desarrolladores ha adaptado la funcionalidad de la norma al nuevo marco de comunicación Internet, y las ventajas implícitas de la norma, nos obligan a seguir considerándola como un factor clave para el éxito de cualquier iniciativa que pretenda aproximarse al concepto de «biblioteca virtual».

En los siguientes apartados se describen las diferencias entre Z39.50 y el nivel de aplicación http, y cómo conviven y se complementan los dos sistemas.

5 Z39.50 y HTTP

Son dos niveles de aplicación preparados para funcionar sobre redes TCP/IP y para satisfacer demandas totalmente distintas: si Z39.50 define un modelo de recuperación

de información abstracto, en el que un cliente y un servidor intercambian ecuaciones de búsqueda y los registros resultados de la evaluación de esa ecuación en un back-end, http define un modelo de distribución de hiperdocumentos en el que un cliente solicita a un servidor la transferencia de un archivo en un formato especial con capacidad hipermedia.

Aparte de esta diferencia funcional, existen otras en el nivel de comunicaciones: la comunicación entre un cliente y un servidor en la norma Z39.50 se basa en el concepto de sesión. En este modelo se establece una conexión entre las dos máquinas, y esta conexión permanecerá hasta que se cierre de forma explícita. En el protocolo http, los intercambios entre cliente y servidor no se basan en el concepto de sesión, con lo que se obliga al cliente a establecer una nueva conexión al servidor http siempre que necesita hacer el mínimo intercambio o solicitud.

En este punto, Z39.50 ofrece una mayor flexibilidad: por ejemplo, con Z39.50 es posible interrumpir una acción determinada, sin que esto suponga la interrupción de la conexión. El hecho de que la conexión permanezca activa a lo largo de todas las interacciones entre el cliente y el servidor, permite mantener conjuntos de resultados a los que se puede volver a recurrir en cualquier momento.

Independientemente de estas diferencias, y de los problemas que pueda tener la definición actual del protocolo http, el WWW se ha convertido en el mejor puente para acceder a «bibliotecas virtuales», en las que resulta extremadamente fácil visualizar información manejando unas herramientas muy sencillas, los navegadores, con las que todos comenzamos a estar familiarizados.

En muy poco tiempo, los navegadores Web han dejado de ser meros visores de documentos HTML, para convertirse en las herramientas hacia las que se orienta el desarrollo de un número creciente de aplicaciones que aumentan sus funcionalidades hasta extremos insospechados.

Los fabricantes de soluciones informáticas para la gestión de información, y, dentro de este grupo, los fabricantes de sistemas integrados de gestión bibliotecaria, se percataron de la importancia que iban a llegar a alcanzar los navegadores Web con interfaz gráfica, y desarrollaron pasarelas entre el WWW y los subsistemas de recuperación de sus productos. Estas pasarelas consisten principalmente en programas script escritos para CGIs o en una integración más directa (caso INNOPAC Web), que recogen una petición de búsqueda formulada en un formulario HTML, la traducen para que sea comprendida por el módulo de recuperación del sistema integrado de gestión bibliotecaria, que envía los resultados de la búsqueda al navegador que solicitó la petición, en forma de página HTML.

Las pasarelas entre el WWW y los subsistemas de recuperación permiten acceder a la biblioteca desde cualquier ordenador, en cualquier parte del mundo, que esté conectado a la red Internet. Además, es posible implementar fácilmente funcionalidades adicionales: autenticación del usuario a través de un navegador, transferencia y recuperación de registros en formato MARC, solicitudes de préstamo inter-bibliotecario, etc.

La tabla IV muestra algunos de los fabricantes de sistemas integrados de gestión bibliotecaria que han desarrollado pasarelas entre sus sistemas propietarios y el WWW.

Muchos de estos fabricantes también desarrollaron módulos que facultan a sus sistemas trabajar con la especificación Z39.50. Distinguiremos dos tipos de implementaciones:

Tabla IV

Algunos de los fabricantes de sistemas integrados de gestión bibliotecaria que han desarrollado pasarelas entre sus sistemas propietarios y el WWW.

Un comentario aparte merece Talis (sistema que trabaja con un servidor basado en Unix y cualquier tipo de cliente), que incluye acceso a una base de datos, eBase con 16 millones de registros, gestionada desde un servidor Z39.50-compliant. Por otra parte, SLS Information Systems pasó a ser la división europea de Innovative Interfaces Inc. desde abril de 1997, comenzando un plan de migración de Libertas a INNOPAC Millenium

<i>Fabricante</i>	<i>Sistema</i>	<i>Pasarela HTTP</i>	<i>Incluye posibilidad Z-client</i>
Ameritech Library Services	Horizon - 1991	WebPac	Sí
Ameritech Library Services	NOTIS LMS	WebPac	Sí
Baratz	Absys	Absys Opac Web	Sí
BLCMP	Talis	TalisWeb	Sí
Aleph	Aleph500 - 1996	WWW Server	Sí
Information Dimensions	TechLib Opac	Sí. con cualquier servidor web y con Basis Plus	Cliente z
Elías	DOBIS-LIBIS	IPAC 2.0	
VTLS	Virtua	Virtua-Web-Gateway	Sí
VTLS	VTLS	VTLS-Web Gateway	
Innovative Interfaces	INNOPAC MILLENIUM	INNOPAC Web Server	Sí
SLS Information Systems			

— Pasarelas o módulos que permitiesen acceder a su servidor propietario a través de cualquier cliente Z39.50 conectados a través de Internet. Por ejemplo, SLS desarrolló el módulo SLS Gateway Server, EliAS el software servidor Z39.50 para DOBIS-LIBIS 3.0 VMS y AMICUS 2.0, soportando distintas versiones de la norma (DOBIS-LIBIS y AMICUS la versión 2, Virtua la 3), VTLS Z39.50 Server, INNOPAC Gateway (versión 2), etc.

Dentro de estos fabricantes, probablemente el caso más interesante, desde un punto de vista técnico, es el de DOBIS-LIBIS. Preparado para correr bajo mainframes IBM MVS y VSE, tuvo que afrontar un doble problema: el desarrollo del módulo servidor conforme a la norma Z39.50 que interactuase con el subsistema de recuperación de su sistema integrado, y el desarrollo de una herramienta que facilitase la conectividad entre TCP/IP (protocolo a través del cual se reciben las peticiones de clientes Z39.50), y APPC (redes SNA para conectar mainframes en un entorno «síncrono» IBM). Este segundo problema se resolvió con un tercer componente, TCPPGATE, que recibe peticiones Z39.50 a través de una conexión TCP/IP, y las convierte a mensajes APPC. El software TCPPGATE se ha implementado sobre PCs con OS/2.

— Clientes propietarios con capacidad de establecer sesiones Z39.50 con cualquier servidor capacitado para interpretar la norma. El fabricante del software que ejecute este servidor podrá coincidir o no con el fabricante que añadió la capacidad de trabajar con Z39.50 en su cliente.

Dentro de esta línea hay que destacar VTLS Z39.50 Client, para el que se produjo un módulo de integración con el navegador Mosaic para MS-Windows.

El desarrollo de pasarelas directas entre el subsistema de recuperación del SIGB y el WWW, y la implementación de la tecnología Z39.50 en la parte cliente o servidor de los sistemas propietarios, hizo que se encontrase en un mismo equipo servidor la capacidad de responder a peticiones de clientes Z39.50 y a peticiones http.

Para obtener las máximas prestaciones de la integración de estas dos tecnologías, los principales sistemas incluyeron en sus pasarelas para http la capacidad de un cliente Z39.50. De esta forma, el sistema permitiría al usuario lanzar ecuaciones sobre el servidor que interactúa con el servidor web -que hace de pasarela-, y sobre otros servidores que, estando conectados a Internet, fuesen capaces de interpretar peticiones Z39.50. Este sistema de implantación es el más extendido en la actualidad. Un caso curioso dentro de este modelo de implementación es el de INNOPAC Web Server, de Innovative Interfaces. El software está preparado para recibir peticiones a través de http, y reenviarlas a servidores como mensajes APDU. La comunicación con otros servidores siempre se gestionará así, a menos que el software del servidor destino sea INNOPAC, caso éste en el que se omite totalmente la sesión Z39.50, y se trabajará directamente con el modelo de sesión propietaria del fabricante.

5.1 Z-Technology

Los grandes sistemas integrados ofrecen plena conectividad a través de la norma Z39.50: módulos en la parte servidor, en el software cliente, e incluso pasarelas entre servidores http-d y servidores Z39.50. La principal restricción de estas implementaciones de la norma era una marcada orientación hacia el tratamiento de registros bibliográficos. Como ya hemos indicado, no es ésta la única posibilidad de recuperación de información que ofrece el estándar.

Sin embargo, existe otra tendencia en la integración de la tecnología Z39.50 y el WWW: interfaces Z39.50 que incorporan utilidades para generar un índice o base de datos a partir de otras fuentes ya existentes, y/o permiten integrar el acceso a sistemas propietarios que soportan Z39.50 y a fuentes de información accesibles en Web a través de un navegador capaz de interpretar el lenguaje HTML estándar.

Estos sistemas actúan como pasarela entre el WWW y servidores Z39.50 ya existentes: es decir, implementan la funcionalidad de un cliente Z39.50, y las de un servidor http-d. Una característica importante de estos sistemas es que permiten interrogar simultáneamente a múltiples servidores. Dos sistemas representativos de este modelo de implementación son OCLC SiteSearch y LibroVisio, desarrollado por la empresa belga EliAS, propietaria de DOBIS-LIBIS y AMICUS. La versión 4.0 de OCLC SiteSearch se lanzó el pasado mes de abril.

OCLC SiteSearch es un software compuesto por tres elementos:

1. WebZ ofrece los siguientes servicios: gateway http: Z39.50. Que se encarga de convertir peticiones http a mensajes comprensibles según el estándar Z39.50. Cliente Z39.50: permite escalar peticiones a servidores Z39.50 locales o remotos. Servidor Z39.50: facilita que clientes Z39.50 accedan a los índices y bases de datos creados con SiteSearch Database Builder.

2. Database Builder permite crear bases de datos locales a partir de registros en formato MARC y documentos SGML (OCLC ofrece servicios de conversión adicionales). Los registros se convierten al formato ANS.1/BER, es decir, el que utiliza Z39.50 para codificar la información que se transmite entre clientes y servidores. El motor de indización y recuperación de Database Builder es Newton, también utilizado en el resto de bases de datos de la OCLC.
3. Image Support Package incluye herramientas software para escanear documentos impresos, indización mediante OCR, y conversión al formato ANS.1/BER utilizado en Database Builder. Image Support Package permite acceder mediante navegadores http y clientes Z39.50 a colecciones digitales de documentos a texto completo, con imágenes, etc. Para que un cliente Z39.50 pueda acceder directamente a una base de datos creada con Database Builder, Image Support Package permite asociar propiedades descriptivas a las imágenes, y modificar el valor de estas propiedades a través de cualquier cliente Z39.50.

El otro ejemplo de pasarela entre http y Z39.50 independiente de sistemas de gestión integrados propietarios lo encontramos en LibriVision, desarrollado por la empresa belga EliAS, a la que ya hemos hecho mención al referirnos a DOBIS-LIBIS y AMICUS. LibriVision ofrece funcionalidad de un cliente Z39.50 accesible a través de un navegador web. Este programa funciona en un equipo UNIX AIX (el UNIX de IBM) en el que esté instalado un servidor http-d de terceros. Los usuarios lanzarán peticiones a este servidor http-d desde sus navegadores, y LibriVision se encargará de convertir estas peticiones a peticiones Z39.50 que dirigirá a servidores capaces de trabajar con la especificación, presentando los resultados de forma unificada. LibriVision no ofrece las posibilidades de gestión de recursos informativos locales que encontramos en la última versión de SiteSearch.

Como conclusión, http y Z39.50 deben ser considerados como dos protocolos complementarios, y no excluyentes. Esta es una de las principales conclusiones a las que llegamos en vista del estado actual de la tecnología y de la aplicación de la norma: la necesidad de integrar las dos soluciones.

La integración de Z39.50 y http permite obtener un único punto de acceso a la riqueza documental disponible en sistemas de información con una orientación distinta (servicios bibliográficos y servicios de información en línea con colecciones a texto completo, documentos digitales, etc.), y, lo que es más importante, el acceso a sistemas heterogéneos desde la perspectiva hardware/software.

La importancia de Z39.50 se hará mayor a medida que se aumente la base instalada de servidores y clientes. El WWW, de momento, no ofrece un mecanismo uniforme para interrogar bases de datos múltiples, y pasarelas similares a LibriVision o SiteSearch están destinadas a convertirse en la interfaz del futuro para nuestros sistemas de información.

La integración de Z39.50 con la tecnología WWW ofrece un potencial enorme para aumentar la funcionalidad de la norma en aspectos claves como la visualización de objetos digitales asociados a descripciones bibliográficas (Image Support Package), procesar peticiones realizadas desde un navegador, etc.

La integración es, sin duda alguna, la línea que se debe seguir.

5.2 Módulo Search para acceder a servidores Z39.50 a través del navegador

Un ejemplo claro de integración de un módulo–cliente Z39.50 en un navegador lo podemos encontrar en el programa que desarrolló el CIIR (Center for Intelligent Information Retrieval) en colaboración con la Universidad de Massachusetts en 1995. Dicho software que soporta la versión 3 de la norma, se diseñó para trabajar con los navegadores de las empresas Netscape y Microsoft.

El funcionamiento es muy simple: cualquier usuario de Internet puede obtener gratuitamente este programa en la dirección: *ftp://www.usgs.gov/pub/ciir/dtic_a02*.

Una vez descomprimido se instalan las correspondientes DLL (Dynamic Link Libraries) en un directorio de trabajo y en el directorio *c:\windows\system*.

Para ejecutar el programa basta con abrir el navegador e introducir el URL de la forma: *search://nombrehost:puerto/basededatos*. Por ejemplo: *search://ibm2.loc.gov:210/books*. Esto activará la interfaz (una página HTML local, al igual que las ayudas de este módulo; todo localizado en el directorio de trabajo correspondiente), en la que introduciremos los términos de búsqueda en mayúsculas y en el lenguaje CCL (Common Command Language), para el que hay una ayuda online.

Entonces el cliente (y por lo tanto el navegador) se conecta al servidor Z39.50 especificado en el URL y ejecuta la búsqueda. Posteriormente el servidor devuelve una serie de resultados en el formato previamente seleccionado (USMARC, GRS-1, OPAC, SUTRS...) y tantos resultados como se hayan negociado a través del cliente. Si la conexión o la búsqueda han fallado, el cliente devuelve el correspondiente mensaje de error. También es posible realizar la búsqueda por campos.

A pesar de que el programa solamente sea una versión alfa, el objetivo del proyecto nos parece de sumo interés. La integración de clientes Z39.50 (por básicos que sean) en los navegadores es fundamental para difundir las posibilidades de la norma (entre usuarios especializados en Recuperación de Información y esa nueva clase de usuarios muy preparados que está surgiendo con el crecimiento del WWW).

Sería interesante observar también la posición que toman las dos empresas líderes dentro del mercado de los navegadores (Netscape y Microsoft) para aprovecharse de las posibilidades de las nuevas versiones de la norma. El interés (o la falta de interés) de estas compañías podría servir de barómetro al mundo bibliotecario, para medir con ciertas garantías su impacto fuera de estos círculos. También para anticipar el futuro y ver hacia donde está evolucionando nuestro campo en lo que se refiere a nuevas tecnologías.

6 Futuro de la norma

Es muy amplio, y se basa en el mismo prerrequisito que impulsó su desarrollo inicial: la necesidad de un protocolo común que permita aunar esfuerzos en el desarrollo de soluciones para la recuperación de información.

Este problema no se ha resuelto con la realización de nuevos sistemas de información: World Wide Web, Wais, Gopher, etc., sino que se ha acentuado. Continuamente recibimos noticias sobre el stress que motiva la sobrecarga de información (el llamado Infoglut), y después de cuatro años usando Internet, se han hecho manifiestas sus limitaciones:

- La dificultad para recuperar información que carece de estructuración y ordenación lógica.
- La dificultad de dar un grado de fiabilidad y confianza a la información que recibimos a través de la World Wide Web.

Ante esta perspectiva, el futuro de la norma sólo puede ser prometedor. De hecho, el proyecto IBM Digital Library manifestó en junio del pasado año la necesidad de adoptar un compromiso con Z39.50, y desarrollar herramientas para la conectividad entre sistemas Z39.50 y su Digital Library (respecto a este punto, hay que destacar la colaboración que estableció con la empresa belga EliAS). IBM considera Z39.50 como una garantía única para asegurar el acceso de todos los profesionales de la información a la IBM Digital Library.

El único punto oscuro que nos obligaría a poner en entredicho las posibilidades de la norma es el escaso, por no decir nulo, impacto que ha tenido entre los fabricantes de software para la recuperación de información: Verity, Excalibur, etc., que no han encontrado una respuesta favorable a la adopción del estándar. Únicamente en el caso de la empresa canadiense Fulcrum Technologies se pudo apreciar un interés mínimo.

Otro factor realmente preocupante es la escasa atención que las grandes editoriales que han puesto en marcha proyectos de biblioteca digital están prestando a la norma (Springer LINK, Elsevier, etc.)

Retomando las posibilidades futuras de la norma: ¿qué aspectos habría que desarrollar, o hacia dónde tendría que evolucionar para que aumentase su nivel de implantación?

- Gestión de metadatos.
- Agentes de nueva generación: ésta es una idea que está teniendo mucho impacto últimamente en los circuitos profesionales. Frente a los agentes que se han generalizado en el entorno http, se abrirían nuevas posibilidades si a los clientes Z39.50 se les dotase de capacidad de automatismo para interrogar simultáneamente información procedente de fuentes «fiables».
- Desarrollo de perfiles que simplifiquen la norma y garanticen la compatibilidad.
- Tratamiento de documentos compuestos (es decir, un mismo registro Z39.50 compuesto por un número variable de documentos o registros).
- Interpretar la riqueza semántica de registros SGML y XML.
- Capacitar al cliente para recuperar secciones específicas dentro de un documento a texto completo. Trabajar con los distintos formatos en los que puede encontrarse un documento (HTML, PDF, etc.), ofreciendo esta información al cliente, e información adicional sobre costes de transferencia, etc.
- Integración con servidores SQL y con software de indización propietario.

Vamos a ver, detenidamente, alguno de estos aspectos.

6.1 Posibilidades de trabajo en texto completo e integración con SQL

Uno de los principales problemas de la norma Z39.50, y de su escaso nivel de implementación al margen del mundo bibliotecario, es la interpretación que se ha he-

cho de ella. Z39.50 no ha sido considerada como un modelo válido para cualquier proceso de recuperación de información automatizado, y su aplicación se ha visto muy restringida al tratamiento de registros bibliográficos.

Sin embargo, la norma define un proceso de recuperación de información abstracto, aplicable a cualquier tipo de información. El concepto de registro que gestiona el estándar es un concepto abstracto, general, y, por tanto, susceptible de aplicarse a procesos de recuperación de información en texto completo.

Por tanto, la vinculación del modelo propuesto por Z39.50 y los sistemas de recuperación a texto completo es directa e implícita en la norma.

Un buen ejemplo de integración de servicios Z39.50 y texto completo lo encontramos en un sistema al que ya hemos hecho referencia en un apartado anterior: OCLC SiteSearch. Este programa permite crear bases de datos a partir de documentos digitales (texto SGML o imagen tratada con OCR), y las almacena en el formato ANS.I/BER (Abstract Syntax Notation/Basic Encoding Rules), utilizado como sistema de codificación y transferencia por Z39.50, y basado en una norma ISO. Los registros de texto completo de SiteSearch se indizan con el motor de indización/recuperación Newton. Una vez creada la base de datos, un cliente Z39.50 podrá acceder a ella a través de los servicios de servidor-Z que hace un subsistema de SiteSearch (WebZ).

Si la posibilidad de trabajar con texto completo es una opción implícita en la norma Z39.50, no sucede lo mismo con la información almacenada en servidores de bases de datos relacionales. La información estructurada se suele gestionar con este tipo de herramientas, cuya principal característica es ofrecer al usuario, y a los programas, la percepción de los datos, y de las relaciones entre ellos, en forma de tablas.

Para trabajar con bases de datos relacionales existe un estándar, el lenguaje SQL (Structured Query Language), que permite acceder y actualizar la información gestionada por estos servidores a través de una sintaxis claramente definida. Tratándose de un estándar con un alto nivel de aceptación, no es extraño que se haya tratado de crear una extensión para la norma que permita acceder a la información gestionada por servidores SQL a través de una sesión Z39.50.

La versión actual de la norma no incluye esta posibilidad. Sin embargo, es más que probable que se incorpore en la 4, ya que en diciembre de 1997 se hizo pública una primera versión del módulo de integración, llamado Z + SQL, con la intención de que se abriesen proyectos piloto para su implementación.

Esta primera versión determina las extensiones necesarias que habría que establecer en la especificación de la norma. Entre ellas se encuentran:

- Aumentar los tipos de queries (peticiones de búsqueda) que acepta la norma, para incluir uno nuevo, la type-SQL, en el que se codificase una sentencia SQL completa.
- Ampliar la sintaxis del servicio explain, para añadir la posibilidad de obtener información sobre el conjunto de tablas, y su estructura, gestionadas por el servidor de bases de datos.

Servicios de la norma que, en un entorno SQL, dejarían de ser necesarios: por ejemplo, la opción *scan* para visualizar el contenido de un índice, ya que se utilizaría para ello una sentencia SQL. Algo similar sucede con *sort*, utilizado para indicar cómo se deben ordenar los registros, ya que esto se puede indicar en la definición de la sentencia *select*.

El principal problema que encontramos al referirnos a la integración de Z39.50 y servidores de bases de datos relacionales es que los fabricantes de estos sistemas quizá no lleguen a prestar atención a la posibilidad de implementar la norma, y al hecho de que el mundo de las bases de datos relacionales ya cuenta con sus estándares para garantizar la conectividad desde hace mucho tiempo: el propio SQL y otra interfaz propietaria que se ha convertido en un estándar de facto: la especificación ODBC de Microsoft.

6.2 Multimedia

La especificación Z39.50 para acceder a Objetos Digitales de Biblioteca (conocida como propuesta de especificación DL) es paralela a la especificación de Acceso a Colecciones Digitales, utilizada por la Biblioteca del Congreso (LOC).

Como propuesta para Z39.50, DL proporciona un subconjunto de características Z39.50 para soportar los requisitos funcionales y de usuario en la búsqueda y recuperación de información en colecciones digitales, específicamente las de la LOC. El uso de esta especificación puede ser uno entre los muchos mecanismos para acceder a dichos documentos digitales. Este mecanismo en particular se distingue por el uso de una estructura llamada Registro Descriptivo del Objeto (metainformación que lo describe) que debe encapsular al Objeto Digital en sí. Cuando el Registro Descriptivo no incluye el objeto del que informa, lo sustituye por un puntero al objeto en sí. En cualquier caso, dentro de esta especificación, el acceso al objeto se realiza a través de dicho Registro.

6.2.1 Modelo de Objeto Digital

Esta especificación aporta un esquema general y flexible para la estructura de un Objeto Digital. En este modelo, un Objeto Digital consta de partes constituyentes, que a su vez están formadas por otras subpartes y así sucesivamente. Todas estas partes se representan como elementos Z39.50.

Consideremos un Objeto Digital constituido por varias imágenes (por ejemplo, fotos o imágenes de texto). Aunque el conjunto de imágenes está contenido en un único Objeto Digital, cada una debe ser específicamente identificable, y el Objeto debe comunicar el hecho de que está compuesto por diversas imágenes, cuántas, y sus características individuales. Por eso, cada una se representa como un elemento separado de un registro Z39.50.

Éste se representa en forma de árbol, donde cada nodo no-hoja posee un número arbitrario de subárboles y/o hojas; mientras que los nodos hojas representan datos. Cada nodo (tanto hoja como no-hoja) tiene una etiqueta de caracteres cuyo propósito es comunicar al usuario que representa ese nodo. Si esa etiqueta no es lo suficientemente descriptiva, puede reforzarse con una descripción (descripción con meta-elementos) de cada nodo.

El modelo podría, por ejemplo, representar un Objeto Digital constituido por 10 cajas, cada una con 20 carpetas, con treinta fotos. Términos-etiqueta como «caja», «carpeta» o «foto» sirven para definir los tipos de elementos (el tipo de elemento se

transmite al usuario y no al cliente; la especificación no trata de definir tipos de contenido procesables por clientes, aunque esto sea posible con ciertos añadidos de programación).

Esta especificación soporta elementos repetidos y la designación de la ocurrencia ordinal de un elemento en un conjunto de elementos repetidos. Por ejemplo, consideremos un Objeto compuesto por muchos volúmenes: p. ej. «volumen 1», «volumen 2», etc. «Volumen 2» podría representarse mediante la segunda ocurrencia del elemento cuya etiqueta es “volumen”. Se podría hacer lo mismo con: «tercer capítulo», «quinta imagen» o «primeras cinco páginas». Esto es posible gracias a la especificación de elemento eSpec-1 y a la sintaxis de registros GRS-1, un formato jerárquico de propósito general usado en GILS (Government Information Locator Service).

6.2.2 Categorías de Objetos Digitales

La especificación define las siguientes categorías de Objetos Digitales:

- Basados en texto.
- Basados en imágenes.
- Basados en sonido.
- Basados en movimiento.

Categorías de descripción asociadas:

- Registro catalográfico (por ejemplo, registro MARC, Dublin core).
- Registro de archivo.
- Encabezamiento (encabezamiento de archivo).
- Página web.
- Verificación, derechos y permisos, control de acceso y control de recursos.
- Autenticación: capacidad de un cliente para comunicar la identidad de un usuario y su clave de acceso (password) al comienzo de una sesión Z39.50.
- Metadatos sobre derechos y permisos: capacidad del servidor para retornar «términos y condiciones» o «derechos (de autor) y permisos» asociados a la información digital.
- Control de acceso: capacidad del servidor de pedir una identificación al cliente antes de procesar sus peticiones.
- Control de recursos: capacidad del servidor de notificar al cliente los recursos (potenciales o actuales) necesarios para una determinada petición de búsqueda/recuperación de información. También de advertir y pedir permiso al cliente para continuar la transacción.

6.2.3 Representaciones de un Objeto Digital

Un Objeto Digital puede tener muchos tipos de representación, por ejemplo: «miniatura», «muy comprimido», «resolución alta», «original» o «imagen de referencia».

Cuando se aplican estas caracterizaciones a un Objeto Digital como un todo, se representan con variantes (Z39.50) y se aplican en la raíz del árbol del objeto (es decir, en la raíz de elemento del tipo de datos del Objeto). Las representaciones pueden pertenecer también a los nodos subordinados a la raíz, y se aplicarán las reglas de aplicabilidad y herencia que se señalaban en el punto anterior.

6.2.4 Acceso Z39.50 a Objetos Digitales

En la especificación de colecciones se puede acceder a un Objeto Digital vía Z39.50 o vía cualquier otro protocolo. Sin embargo, para la especificación DL, cuando a un Objeto Digital se accede a través del protocolo Z39.50, sólo puede hacerse mediante el Registro Descriptivo de Objeto. El uso de Z39.50 para buscar o recuperar directamente un Objeto Digital (no a través de su Registro Descriptivo de Objeto) no está soportado por esta especificación.

7 Metodología de desarrollo de un cliente Z39.50

Hasta ahora hemos hablado de clientes Z39.50, públicos o comerciales, propietarios o en Web..., pero además existe la posibilidad de desarrollar un cliente con la ayuda de herramientas como el API de la OCLC o YAZ de Indexdata.

Este cliente podría implementar en general las funciones más básicas que ofrece el protocolo Z39.50. No es el objetivo de este apartado explicar el desarrollo completo de la implementación de dicho cliente, sino exponer estas opciones como posibilidades alternativas. Vamos a comentar por encima lo que nos ofrecen dichas herramientas.

7.1 API de la OCLC

Se basa en un API (*Application Program Interface*), es decir, una interfaz entre el sistema operativo y programas de aplicación que especifica la forma en que los programas de aplicación se comunican con el sistema operativo y los servicios que éste proporciona. Por ejemplo, un API puede hacer posible para programas que corren sobre entorno Windows abrir ventanas, cuadros de diálogo etc.

Las funciones básicas de un cliente Z39.50 son: conexión, inicio de sesión, búsqueda, presentación y desconexión. El API de un cliente Z39.50 debe proporcionar estas funciones e incrustarlas en el programa que el usuario use como aplicación, en nuestro caso un cliente WWW.

Para ello es necesario poder construir mensajes Z39.50 y mandarlos y recibirlos usando los protocolos estándar de socket TCP/IP (en UNIX y algunos otros sistemas operativos, socket es un software que conecta una aplicación con un protocolo de red, por ejemplo TCP/IP). Un programa puede mandar y recibir mensajes TCP/IP abriendo un socket y leyendo o escribiendo los datos mediante ese socket, como ocurre en UNIX. Esto simplifica la tarea del programador.

7.1.1 Funcionamiento del API

Las rutinas descritas a continuación procesan mensajes IRP (Information Retrieval Protocol, el nombre del protocolo descrito por Z39.50), no los mandan ni los reciben sino que los procesan.

Cuando se llama una rutina de solicitud IRP esto produce un registro BER (Basic Encoding Rules, la sintaxis de transmisión elegida por ASN-1) que puede ser mandado a un objetivo Z39.50. A la inversa, un registro BER que contiene una respuesta IRP es procesado y una estructura es devuelta. Los registros BER son codificados y decodificados utilizando una serie de utilidades BER de la OCLC disponibles en ftp://ftp.rsch.oclc.org directorio pub/BER_utilities.

Hay dos capas para las rutinas de respuesta IRP. La que uno elige para usar depende del estilo de codificación y de la complejidad de la aplicación. La capa más baja (la aplicación más simple) se compone de las rutinas: *InitResponse* (), *SearchResponse* (), *PresentResponse* () y *ScanResponse* (). Estas rutinas son devueltas directamente con la respuesta Z39.50 cuando se aceptan solicitudes de inicio: (*InitRequest*), búsqueda: (*SearchRequest*), presentación o exploración. Estas rutinas rellenan las estructuras parametrizadas de la respuesta que proporciona el servidor.

Esta capa es apropiada cuando la aplicación es relativamente simple y puede saber qué clase de respuesta está recibiendo. En aplicaciones más complejas es difícil saber qué clase de respuesta ha llegado. Existe una segunda capa con dos rutinas algo más complejas: *WhatKindOfResponse* () y *Z3950_Response* (). La primera devuelve un valor que indica qué tipo de respuesta IRP se ha recibido. La segunda retorna una estructura que identifica el tipo de respuesta y además hace una llamada a la rutina de respuesta adecuada.

7.2 YAZ (Yet Another Z39.50)

El conjunto de utilidades desarrollado por la empresa sueca Indexdata ofrece diferentes niveles de acceso a los protocolos Z39.50 y SR. El nivel que cada usuario elige depende de sus necesidades y de si quiere ser servidor o cliente de esos protocolos. El nivel básico consiste en tres servicios:

- *ASN*, que proporciona una representación en C de los mensajes Z39.50/SR o (PDUs Protocol Data Units).
- *ODR*, que codifica y decodifica los mensajes de acuerdo con las reglas BER.
- *COMSTACK*, que intercambia los mensajes codificados (con BER) mediante una conexión de red.

El módulo representa la definición de la sintaxis ASN.1 del protocolo SR/Z39.50. Establece un conjunto de estructuras y tipos diseñados en C, asignándolos a cada una de las PDUs y a cada clase de sintaxis ASN.1.

El sistema ODR (Open Data Representation) es un mecanismo básico para representar tipos ASN.1 en C e implementar codificación/decodificación BER en cada uno de esos tipos.

Cuando se ha creado un buffer para los mensajes ya codificados en BER se puede

usar el sistema COMSTACK para transmitir o recibir datos a través de la Red. El módulo COMSTACK proporciona funciones simples para establecer una conexión (pasiva o activamente, dependiendo del papel de la aplicación cliente o servidor) y para intercambiar PDUs codificadas en BER a través de esa conexión. Cuando se crea un nodo de una conexión, debe especificarse qué protocolo se quiere usar, TCP/IP, SR, Z39.50...

La combinación de estos tres servicios ODR, ASN y COMSTACK constituye el API de Index Data para los protocolos Z39.50/SR

YAZ ofrece este conglomerado de servicios para que el usuario pueda decidir, por ejemplo, qué codificador/decodificador de BER quiere utilizar (podría preferir SNACC o BERUtils al que ofrece YAZ: ODR).

YAZ está implementado en algunos clientes Z39.50, por ejemplo BookWhere 2000 e IrTcl. Las ventajas que ofrece a estos productos son: flexibilidad, soporte de la notación polaca inversa (RPN) y reconocimiento de múltiples normas internacionales (por ejemplo, ISO-8777), lo que le confiere una gran versatilidad.

8 Conclusiones

Aunque Z39.50 se diseñó para buscar y recuperar registros bibliográficos con formato USMARC, actualmente existen organizaciones que usan este estándar para suministrar documentos en texto íntegro sobre consultas en lenguaje natural. Esto nos parece muy importante, porque nos indica que la norma se adapta bien a las nuevas tendencias en recuperación/presentación de información.

Otro punto importante es que, a diferencia de los sistemas actuales de comunicaciones, Z39.50 trabaja por sesiones, lo que facilita las tareas de verificación y protección de datos. Durante la sesión se establecen los mecanismos necesarios (utilizando metadatos, gestionando privilegios...) para favorecer la eficiencia y exactitud de todo el proceso de transferencia de información.

A través de este protocolo se accede a bases de datos cuya información está muy estructurada y que pueden servirla en una gran variedad de formatos (SGML, XML, HTML, PDF, etc.). Actualmente el WWW ofrece, en líneas generales, información muy poco estructurada (aunque se está avanzando en este sentido con el diseño y difusión de nuevos estándares) y en gran cantidad, lo que hace que pierda gran parte de su utilidad potencial.

Los mecanismos de control de la norma se basan en el hecho de que el contenido que gestionan se encuentra altamente estructurado. Es previsible que a medio plazo (debido en gran medida a las necesidades, cada vez más acuciantes, de información muy organizada y por lo tanto rentable), el WWW sufra importantes cambios.

Si esquematizamos mucho esta especie de cadena de la información, e identificamos el WWW con el depósito documental, el estándar Z39.50 como una posible clase de intermediario, y finalmente los usuarios, podemos asumir que tanto el depósito documental como los usuarios están en medio de un cambio cuantitativo, pero sobre todo cualitativo muy importante.

Pensamos que la norma Z39.50 puede adaptarse más fácilmente que otras (por su estructura y sus posibilidades) a este nuevo estatus. Pero esto sólo sucederá si se difunde lo más ampliamente posible, lo que depende de desarrolladores, empresas implementadoras, organismos normativos y consultivos y, por supuesto, usuarios.

Finalmente hay que decir que la relativa inadaptación del mundo bibliotecario a las nuevas tecnologías de las que hablamos en la introducción podría solucionarse en parte aprovechando las «herramientas» tecnológicas que fueron diseñadas para y desde nuestro campo, sobre todo aquéllas que ofrecen más posibilidades de futuro.

9 Bibliografía

- ADDYMAN, A. M. *Facilitating the creation of Z39.50 origins in the UK*. <http://ds.internic.net/z3950/z3950.html>.
- ANSI/NISO Z39.50-1995. *ANSI Z39.50: Information Retrieval Service and Protocol*, 1995. <http://lcweb.loc.gov/z3950/agency>.
- BIBLIO TECH INFORMATION TECHNOLOGY FOR LIBRARIES. *Z39.50: Part 1- an overview*, julio 1998. <http://www.biblio-tech.com/html/z39.50.html>.
- BIBLIO TECH INFORMATION TECHNOLOGY FOR LIBRARIES. *Z39.50: Part 2- Technical Details*, July 1998. http://www.biblio-tech.com/html/z39.50_part_2.html.
- BOOKWHERE. *Z39.50 Information Detail*, <http://www.bookwhere.com/z3950d.htm>.
- CARSON, S. M.; FREIVALDS, D. I. Z39.50 and LIAS: Penn state's experience. *Information technology and libraries*, 12 (2), junio 1993, p. 230-237.
- DEKKERS, M. *Implementing a Z39.50 in a multi-national and multilingual environment*. <ftp://ftp.loc.gov/pub/z3950/articles/makx.ps>.
- DEMPSEY, L.; RUSSELL, R.; KIRRIEMUIR, J. Towards distributed library systems: Z39.50 in a European context. *Program electronic library and information systems*, vol. 30, n.º 1, enero 1996. <http://www.aslib.co.uk/program/1996/jan/02.html>.
- HAMMER, S.; FAVARO, J. *Z39.50 and the World Wide Web*, Feb 1996. <http://www.oclc.org:5046/archive/webcat/0098.html>.
- HINNEBUSCH, M. Integrated library systems: the Z39.50 explain service. *Academic and library computing*, 9 (10), noviembre-diciembre 1992, p. 12-13.
- KUNZE, J. A. *Basic Z39.50 server concepts and creation*, 1995. <ftp://ftp.loc.gov/pub/z3950/articles/john.pdf>.
- KUNZE, J. A. Nonbibliographic applications of Z39.50. *The public access computer system review*, 3 (5), 1992, p. 4-30.
- LE VAN, R. *Building a Z39.50 client*. http://ftp.sunet.se/pub/z39.50/OCLC/z39.50_client_api/zclient.txt.
- LE VAN, R. *OCLC Ver Utilities*. <ftp://ftp.rsch.ubc.ca/pub/local/src/snacc>.
- LÉVÉJAC, A.-L. *Z39.50: L'Information bibliographique structurée sur le net*, Mar 1997. <http://www.perso.hol.fr/~alevejac/Z3950.htm>.
- LIBRARY OF CONGRESS. *Z39.50 Profile for Access to Digital Collections*, 1996, Final draft. <http://lcweb.loc.gov/z3950/agency/profiles/collections.html>.
- LYNCH, D. *Implementing Explain*. <ftp://ftp.loc.gov/pub/z3950/articles/denis.pdf>.
- LYON, J. El último eslabón entre Internet y los lectores. *Information World en Español* (Oxford) 1996, n.º 40, p. 12-13.
- MOEN, W. *The ANSI/NISO Z39.50 Protocol: Information Retrieval in the Information Infrastructure*. <http://www.cni.org/pub/NISO/docs/Z39.50-brochure/50.brochure.toc.html>.
- MOYA, F. de. Biblioteca virtual y redes: situación actual de las normas. *Information World en Español* (Oxford) 1994, n.º 25, p. 1-3.
- RFC-1729. *Using the Z39.50 Information Retrieval Protocol in the Internet Environment*. Diciembre, 1994. <http://ds.internic/rfc/rfc1729.txt>.
- STOVEL, L.; FUCHS, R.; CHANG, J. RLG's Z39.50 server: development and implementations issues. *Information technology and libraries*, 12 (2), June 1993, p. 227-230.

- TURNER, F.; ZEEMAN, J. *Z39.50 Keyword Searching of Bibliographic Systems: A Discussion Paper*, junio 1998. <http://www.nlc-bnc.ca/iso/z3950/keyword1.htm>.
- WAIS Profile of Z39.50 version 2, revision 1.4. Wais Inc. Abril 1994.
- WIBBERLEY, L. *Use of Z39.50 for Search and Retrieval of Scientific and Technical Information*. <ftp://ftp.loc.gov/pub/z3950/articles/peter.pdf>.
- Z39.50 versión 3. *Information World en Español* (Oxford) 1996, n.º 47, p. 19.
- ZIG Commentaries, julio 1998. <http://lcweb.loc.gov/z3950/agency/wisdom.html>.