



EL Z39.50 EN EL AMBIENTE DE TRASFERENCIA Y RECUPERACIÓN DE INFORMACIÓN

Martha Helena Arango

EL Z39.50 EN EL AMBIENTE DE TRASFERENCIA Y RECUPERACIÓN DE INFORMACIÓN

MARTHA HELENA ARANGO DE V.*

earango@javeriana.edu.co

Aunque no tenga una denominación fácil, el Z39.50 es un protocolo desarrollado y mantenido por y para bibliotecas, incluso podría considerarse como la norma más relevante para el mundo de las bibliotecas y de la documentación desde la aparición del formato MARC.

Nació hace dos décadas como protocolo para la recuperación de información bibliográfica en el seno del Linked Systems Project y alcanza su mayoría de edad cuando es reconocido como la norma internacional ISO 23950.

El Z39.50 es un estándar que especifica un conjunto de reglas para gestionar las formas y procedimientos de interconexión remota de computadoras, con el propósito de buscar y recuperar información. Su aplicación es ahora mucho más amplia, incluye la consulta y el intercambio de datos bibliográficos, y también la intercomunicación de índices y resúmenes, de información geoespacial, de documentos oficiales, de objetos digitales o de metadatos que describen los documentos de las bibliotecas electrónicas o digitales.

Las especificaciones del Z39.50 permiten que un ordenador cliente (origen) sea capaz de buscar información en un ordenador servidor (destino) y recuperar el resultado de esa búsqueda. Una vez resuelto el problema fundamental del "entendimiento" entre dos sistemas con hardware y software diferentes, está al alcance de la mano la realización de la idea futurista de un "cliente universal" que permita la consulta de cualquier recurso local o remoto de información bibliográfica, textual o de otro tipo, sin necesidad de conocer cientos de interfaces de usuarios ni la estructura de decenas de bases de datos.

**Profesional en Información y Documentación de la Pontificia Universidad Javeriana. Actualmente es docente en esta Universidad e investigadora en el proyecto "Biblioteca Digital desde Colombia"*

La solución ideal es que el cliente esté integrado en el sistema de gestión que utiliza la biblioteca o el centro de documentación; de esta forma se pueden recuperar datos remotos sin abandonar el entorno habitual de trabajo.

En los últimos años, especialmente a partir de 1995, es bastante habitual encontrar menciones y referencias al protocolo Z39.50 en publicaciones del sector bibliotecario, y en catálogos de software, ya que los proveedores de programas de gestión bibliotecaria van paulatinamente incluyendo productos basados en el protocolo e integrados en los sistemas de gestión documental. Sin embargo, la información que se suele encontrar sobre el Z39.50 es excesivamente técnica, y prácticamente inexistente en español.

La norma ANSI/NISO Z39.50 es un estándar para la recuperación de la información basado en la estructura cliente/servidor, que facilita la interconexión de sistemas informáticos. Uno de los beneficios básicos del protocolo, en el ámbito de las bibliotecas y de los centros de documentación, es que hace posible la comunicación entre sistemas que utilizan diferente hardware y software. Permite la realización de búsquedas simultáneas a múltiples bases de datos, utilizando una única interfaz de usuario, recuperar la información, ordenarla, y exportar los registros bibliográficos.

La primera versión del estándar se liberó en 1988. Dos años después se formaron dos grupos de trabajo que garantizaron el desarrollo y evolución de la norma: el ZIG (Z39.50 Implementors Group), grupo de desarrolladores que con el apoyo de la Biblioteca del Congreso de Washington vienen trabajando en la normalización y ampliación de las funcionalidades del estándar, y el Z39.50 Maintenance Agency, encargados de dar soporte al estándar. En 1992 se aprobó la versión 2, con capacidad de hacer compatible su uso en entornos OSI (Open Systems Interconnection). En 1995 se aprobó la versión 3, que fue aceptada como estándar de la ISO con el número 23950 en marzo de 1997. Las nuevas facilidades que se han ido añadiendo tienen un carácter modular, y pueden irse implementando progresivamente de forma independiente. Los sistemas de gestión documental, evolucionan gradualmente hacia la versión 3, manteniendo además la compatibilidad con versiones anteriores del estándar.

Con el Z39.50 el proceso de consulta de la información es más sencillo y ágil para los profesionales, por eso también son más fluidas otras funciones y servicios habituales en las bibliotecas y centros de documentación, como los trabajos de referencia e información bibliográfica, puesto que desde nuestro puesto de trabajo y sin cambiar de interfaz podemos acceder a las grandes bibliotecas de todo el mundo (LC, BL, BN Canadá) y a muchas bibliotecas especializadas.

También se facilitan la catalogación cooperativa, ya que el protocolo permite la descarga a menudo gratuita de registros MARC de distintas fuentes, y el préstamo interbibliotecario de un documento, solicitado a partir de los datos de ejemplares suministrados por un servidor Z.

Los usuarios también pueden obtener ventajas importantes de la aplicación de Z39.50 a los OPACs tradicionales. Las funcionalidades que añade el protocolo, convierten los

OPACs en puertas de acceso a cientos de catálogos de todo el mundo con la misma interfaz de consulta.

Los principales creadores de sistemas de gestión documental han tenido en cuenta las posibilidades del protocolo, de forma que no existe sistema de tercera generación que se precie que no incorpore la tecnología Z (servidor+cliente Z39.50) entre las herramientas disponibles, por ejemplo Horizonte de Ameritech, Unicornio de SIRSI, o ABSYS de Baratz. No son éstas, sin embargo, las únicas empresas interesadas en Z39.50; el registro de desarrolladores de aplicaciones Z que mantiene la LC incluye en la actualidad un centenar entidades, entre ellas Chemical Abstracts, Silver Platter, Ovid, LEXIS/NEXIS, OCLC, EBSCO y Dialog.

Una de las características que más ha contribuido a la difusión y popularización del Z39.50 es la posibilidad que tienen algunos clientes de realizar búsquedas en paralelo en varios servidores. Aplicando esta funcionalidad se abren amplias perspectivas de cooperación entre centros de información, ya que se facilita la creación de aplicaciones distribuidas y permite reagrupar funcionalmente recursos física y geográficamente dispersos.

Los proyectos más interesantes hacia el futuro, se basan precisamente en la capacidad que tiene el Z39.50 de crear agrupaciones lógicas de servidores de información. En este contexto ha surgido el concepto de "catálogo colectivo virtual" que permite que el cumpla de las funciones tradicionales asignadas a este tipo de catálogos, pero que sólo existe realmente, mientras se mantiene abierta la sesión entre un cliente y un conjunto de servidores Z. Esta posibilidad solo la proporcionan los clientes Z más evolucionados, que son los que pueden realizar búsquedas en paralelo en varios servidores.

Las características comunes que pueden motivar el agrupamiento de servidores pueden ser la localización geográfica (municipios, departamentos, regiones, estados); la materia en la que están especializados; el tipo de unidad (archivos, bibliotecas, servicios de resúmenes, museos, centros de información geoespacial), o el tipo de servicios concertados (préstamo interbibliotecario, acceso a documentos originales, catalogación cooperativa).

Con frecuencia los proyectos de catálogos virtuales van unidos a los de bibliotecas digitales, de forma que es habitual ver relacionados en la bibliografía conceptos como publicaciones electrónicas, objetos digitales, metadatos y Z39.50. En este nuevo ámbito de aplicación el protocolo puede desempeñar un importante papel en la recuperación de información digital, ya que puede aplicarse indistintamente a los documentos completos, partes o secciones de éstos, imágenes y sonido o a los metadatos que los describen.

El éxito de Z39.50 dependerá en gran medida de la cantidad de servidores disponibles. El número existente en EE.UU. y en Canadá es relativamente elevado y ha superado ya el punto crítico en donde comienza a ser productiva la utilización sistemática del protocolo para la búsqueda y recuperación de información. La Unión Europea contribuye decisivamente a la difusión del protocolo y ha subvencionado o está financiando 17 proyectos relacionados con Z39.50 dentro de su Programa de Bibliotecas; entre ellos,

cabe destacar por su alcance geográfico y por el gran interés de los objetivos que proponen, los proyectos ONE (OPAC Network in Europe) y UNiverse (Large scale demonstrator for global, open distributed library services). En los países hispanos es todavía muy escasa la implantación de Z39.50: los servidores operativos o en fase avanzada de desarrollo pueden ser menos de 10. Lo cierto es que sólo si aumenta el número de servidores activos en bibliotecas de todo tipo y crecen notablemente los recursos de información disponibles a través de ellos se podrá comprobar el potencial de este protocolo para establecer nuevos cauces de cooperación. Por esta razón, sería conveniente promover las iniciativas y proyectos que permitan dar a conocer la norma y fomenten su uso en los centros de información y documentación.

El protocolo constituye un gran avance en la interconexión entre sistemas de gestión documental al permitir superar las enormes barreras que conlleva operar con diferentes sistemas informáticos.

En otros ámbitos, como Museos y colecciones digitales, se está trabajando también en la adaptación y ampliación del protocolo.

Cómo trabaja Z39.50

Al efectuar una búsqueda en una base de datos tradicional se encuentra una interfaz de usuario, un motor de búsqueda y los datos.

Un sistema Z39.50 sustituye la interfaz de usuario por un servidor Z, que accede al motor de búsqueda, y varios clientes Z, que dialogan con el servidor Z e interactúan directamente con el usuario.

El cliente Z: la aplicación cliente es un programa utilizado por el usuario final. Mediante este programa se pueden hacer peticiones a los sistemas remotos y se obtienen respuestas de dichos sistemas.

El servidor Z: la aplicación servidor es un programa que acepta las conexiones en redes y recibe las preguntas del protocolo Z39.50 en dichas conexiones. A partir de estas preguntas elabora una respuesta y la devuelve al cliente. Generalmente, las preguntas que recibe las traduce al formato de consulta de una base de datos y la respuesta contiene los datos recuperados de la base de 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.

Una de las funciones básicas de Z39.50 define cómo interroga el cliente al servidor. Este proceso se puede dividir en cinco fases:

-
1. El protocolo especifica formatos y procedimientos que deben regir el intercambio de mensajes entre el cliente y el servidor.
 2. El cliente puede iniciar peticiones; el protocolo dirige la comunicación entre las aplicaciones correspondientes de recuperación de información y entre cliente y servidor.
 3. El cliente puede proponer una búsqueda, especificando una o más bases de datos, e incluir una pregunta y los parámetros que determinan la forma de presentación del resultado de la búsqueda.
 4. El servidor responde con una cuenta de registros identificados.
 5. El cliente asume que los registros seleccionados por la búsqueda componen un conjunto ordenado (según el orden determinado por el servidor).

La búsqueda: Un ejemplo de búsqueda desde un OPAC con funciones de cliente Z, se puede dividir en ocho partes:

1. El usuario del OPAC selecciona la biblioteca objeto de la búsqueda a través de un menú de OPACs accesible en el cliente.
2. El usuario introduce los términos de búsqueda.
3. El cliente envía la estrategia a las bibliotecas objeto de la petición utilizando un servidor Z que trabaja como una parte del programa de gestión documental.
4. El servidor Z traduce los términos de búsqueda a un lenguaje común (Z Speak) y establece contacto con la biblioteca.
5. Existe una negociación preliminar entre el cliente Z y el servidor Z mediante la cual se establecen 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 las bibliotecas (base de datos) 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 llama “facility”. Existen 11 “facilities”:

1. Initialisation: Configuración de las reglas de comunicación, negociación de los niveles del servicio
2. Search: Envío de la ecuación de búsqueda a la base de datos y recogida de los primeros resultados
3. Retrieval: Recuperación de registros tal y como se especificó en el cliente Z.
4. Result set delete: Eliminación del conjunto de resultados almacenados en el servidor Z.
5. Access control: Proceso de verificación por parte del servidor Z.
6. Accounting resource control: Gestión de cuentas
7. Sort: Ordenación del conjunto de resultados
8. Browse: Escaneo e índice en el servidor Z

-
9. Extended services: El cliente Z puede enviar peticiones de paquetes (préstamo interbibliotecario, ...).
 10. Explain: El cliente Z interroga la base de datos en busca de información sobre la implementación del servidor Z.
 11. Termination: Finaliza la negociación y la conexión

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:

- 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...
- "Browsing" del índice.
- Definir el formato de los registros.

A pesar de no estar incluidos dentro de la norma ISO, el protocolo Z39.50 incluye los siguientes servicios para facilitar el control:

- Almacenar resultados.
- Almacenar peticiones de búsqueda.
- Definir un esquema de búsqueda.
- Solicitar un ejemplar.
- Actualizar la base de datos.
- Crear un fichero de exportación.

Implementación de los clientes y servidores Z39.50

Un problema que se encuentra al momento de evaluar los resultados del Z39.50, lo causa, a su vez, una de sus mayores ventajas: la gran facilidad con que permite realizar búsquedas y recuperar información en diversas bases de datos; el poder realizar esta búsqueda y recuperación en una sola sesión y con una sola ecuación de búsqueda, nos deja ver un problema obvio: no todos los servidores implementan los mismos servicios en sus sistemas, por lo tanto, en ocasiones los registros recuperados son incoherentes, ya bien sea por errores en la traducción de la petición, durante la definición de parámetros en el proceso de inicio de sesión, incompatibilidades entre las versiones del cliente y del servidor, la configuración del servidor, o la bases de datos del servidor no reconoce los servicios que solicita el cliente.

Estas dificultades son mas probables de encontrar al momento de potenciar las posibilidades del Z39.50 en sus opciones avanzadas, más que en las clásicas búsquedas sencillas con no más de tres términos unidos por operadores booleanos.

El Z39.50 y Http

El Z39.50 y el http, son dos niveles de aplicación preparados para funcionar sobre redes TCP/IP y para satisfacer demandas distintas: el Z39.50 define modelos de recuperación de información en abstracto, dónde un cliente y un servidor intercambian tanto ecuaciones de búsqueda como registros resultados en un backend; el http define una distribución de hiperdocumentos en el cual el cliente pide al servidor la transferencia de archivos en un formato especial con capacidad hipermedia.

En lo comunicacional, la comunicación entre un cliente y un servidor Z39.50 se basa en el concepto de sesión, en la cual se establece una conexión que permanecerá hasta que se cierre de forma explícita, lo cual permite interrumpir acciones determinadas sin que esto implique la interrupción de la conexión, permitiendo así mantener conjuntos de resultados a los cuales se puede volver a recurrir en cualquier momento. En el Protocolo http los intercambios no se basan en el concepto de sesión, el cliente debe establecer una nueva conexión al servidor http cada vez que necesite hacer un mínimo intercambio o solicitud.

Independientemente de estas diferencias, el WWW se ha convertido en el mejor puente para acceder a las "bibliotecas virtuales" en las que es fácil visualizar la información mediante los conocidos navegadores de la web, por lo cual estos navegadores se están convirtiendo en las herramientas hacia las que se orienta el desarrollo de un gran numero de aplicaciones que aumentan sus funcionalidades hasta extremos insospechados.

Los desarrolladores de sistemas de gestión documental, conscientes de la importancia y proyección de los navegadores web y su interfaz gráfica, desarrollaron pasarelas entre el WWW y los subsistemas de recuperación de productos, las cuales consisten en programas que recogen la petición de búsqueda en un formato html, la traducen para que sea comprendida por el módulo de recuperación del sistema de gestión documental, el cual envía los resultados al navegador en forma de página html. Estas pasarelas entre el WWW y los subsistemas de recuperación permiten acceder a la biblioteca desde cualquier computador en cualquier parte del mundo que esté conectado a internet. Además permiten implementar funcionalidades como la autenticación del usuario, transferencia y recuperación de registros en formato MARC, solicitudes de préstamo interbibliotecario, etc.

Muchos de estos desarrolladores también han creado los módulos que permiten a sus sistemas trabajar con la norma Z39.50. El desarrollo de pasarelas entre el subsistema de recuperación del sistema de gestión documental y el WWW, y la implementación del Z39.50 en la parte cliente y servidor de los sistemas propietarios, permitió a un mismo

servidor estar en capacidad de responder a peticiones de clientes Z39.50 y a peticiones http.

Para optimizar al máximo la integración de estas dos tecnologías, los sistemas de gestión documental incluyeron en sus pasarelas para http la capacidad de un cliente Z39.50, así el sistema permite al usuario generar ecuaciones en el servidor que interactúa con el servidor web que sirve de pasarela, y en otros servidores conectados a internet capaces de interpretar peticiones Z39.50; también se están generando interfaces Z39.50 que incorporan utilidades para generar índices o bases de datos a partir de fuentes ya existentes, y/o permiten integrar accesos a sistemas propietarios que soportan Z39.50 y a fuentes de información accesibles en web a través de un navegador. Por todo esto http y Z39.50 deben considerarse como protocolos complementarios y no excluyentes, pues su integración permite obtener un único punto de acceso a la riqueza documental disponible en sistemas de información con orientaciones distintas: servicios bibliográficos, servicios de información en línea con texto completo, documentos y objetos digitales, etc., y lo que es más importante, el acceso a sistemas heterogéneos desde la perspectiva hardware/software.

La importancia del Z39.50 crecerá en la medida que aumenten los servidores y clientes Z39.50 instalados. El WWW de momento no ofrece un mecanismo uniforme de interrogar múltiples bases de datos y las pasarelas integradas http Z39.50 serán la interfaz del futuro para los sistemas de información, pues será la que permitirá la visualización de objetos digitales asociados a descripciones bibliográficas, procesar peticiones realizadas desde un navegador, etc.

El Z39.50 tiende al desarrollo características como:

- Gestión de metadatos
- Agentes de nueva generación
- Desarrollo de perfiles que simplifiquen la norma y garanticen la compatibilidad
- Tratamiento de documentos compuestos
- Interpretación de la riqueza semántica de registros SGML y XML
- Capacitación del cliente para recuperar secciones específicas dentro de un documento a texto completo, trabajando con los distintos formatos (HTML, PDF, etc.), ofreciéndole esta información al cliente, e información adicional como costos de transferencia, etc.
- Integración con servidores SQL y con software de indización propietario.

REFERENCIAS

1. ANSI/NISO/ISO 23950 Information Retrieval Z39.50: Application Service Definition and Protocol Specification. National Information Standards, 1997. ISBN: 1880124378
2. Carrión Gútierez, Alejandro. De las virtudes del catálogo virtual. En: Dossier 2. Boletín de la SEDIC. p 2-3
3. De la Rosa Piñero, Antonio et al. Norma Z39.50, actualidad, posibilidades. : ¿Es necesario un cambio de actitud?. En: Rev. Española de Documentación Científica. Vol 21 No. 4 de 1998. p 416-447
4. Evans, Peter. Z39.50: Part 1- an overview [en línea] <http://www.bibliotech.com/html/z39_50.html> Biblio Tech Review Information Technology for Libraries Updated: 05 November 1999
5. Evans, Peter. Z39.50: Part 2 - Technical Details [en línea] <http://www.bibliotech.com/html/z39_50_part_2.html> Biblio Tech Review Information Technology for Libraries Updated: 05 November 1999
6. Finnigan, Sonya. Z39.50 Made Simple. [en línea] <<http://www.dstc.edu.au/cgi-bin/redirect/rd.cgi?>http://archive.dstc.edu.au/DDU/projects/ZINC/zsimple.htm>> Distributed Systems Technology Centre Pty Ltd University of Queensland, QLD 4072
7. Hakala, Juha. Z39.50-1995: Information retrieval protocol An introduction to the standard and it's usage. [en línea] <<http://renki.helsinki.fi/z3950/z3950pr.html>> Automation Unit of Finnish Research Libraries Helsinki University Library. 18 January 1996
8. Hammer, Sebastian y Favaro, John. Z39.50 and the World Wide Web. [en línea] <<http://mirrored.ukoln.ac.uk/lisjournals/dlib/dlib/dlib/march96/briefings/03indexdata.html>> D-Lib Magazine, March 1996 ISSN 1082-9873

-
9. Hinnebusch, Mark. Report to the CIC on the State of Z39.50 Within the Consortium [en línea] <<http://ntx2.cso.uiuc.edu/cic/cli/z39-50report.htm>> Prepared by May 19, 1998 Final Revision August 29, 1998
 10. Library of Congress Z39.50 Maintenance Agency Site. <<http://lcweb.loc.gov/z3950/agency/>>
 11. Library of Congress. The Z39.50 Document [en línea] <<http://lcweb.loc.gov/z3950/agency/document.html>>
 12. Lynch, Clifford A. The Z39.50 Information Retrieval Standard Part I: A Strategic View of Its Past, Present and Future [en línea] <<http://mirrored.ukoln.ac.uk/lis-journals/dlib/dlib/dlib/april97/04lynch.html>> D-Lib Magazine, April 1997 ISSN 1082-9873
 13. Martínez Gallo, J. Carlos. Cómo trabaja Z39.50. En: Dossier 2. Boletín de la SEDIC. p 4 carlos@baratz.es
 14. Moen, William. A Guide to Global Z39.50. National Information Standards, 1997. ISBN:1880124351
 15. Moen, William. The ANSI/NISO Z39.50 Protocol: Information Retrieval in the Information Infrastructure. [en línea] <<http://www.cni.org/pub/NISO/docs/Z39.50-brochure/50.brochure.toc.html>> [Consulta: 11 enero de 2000].
 16. Moreno, Alicia. El Z39.50. En: Dossier 2. Boletín de la SEDIC. p 1-2 amoreno@ifigenia.es
 17. National Information Standards. Information Retrieval (Z39.50): Application Service Definition and Protocol Specification (National Information Standards Series). National Information Standards, 1996. ISBN: 188012422X.
 18. NISO Z39.50 Resource Page <<http://www.niso.org/z39.50/z3950.html>>
 19. Randall, Sara (editor). Proceedings of the Iowa State Library Symposium on Ansi/Niso Z39.50. National Information Standards, 1997. ISBN:1880124335
 20. SILO Project State of Iowa Libraries Online. An Evaluation of Z39.50 <<http://www.silo.lib.ia.us/bluang.htm>> Blue Angel Technologies
 21. Z39.50 Implementors Group (ZIG). <<http://lcweb.loc.gov/z3950/agency/zig/zig.html>>