

Z39.50 en el siglo XXI: ¿estándar real o virtual?

Arantza López de Sosoaga Torija
SABINI Automatización de Bibliotecas

1.- INTRODUCCIÓN

El estándar Z39.50 (ISO 23950) normaliza el procedimiento para que dos ordenadores se comuniquen cuando se desea recuperar información. Dado que puede implementarse en cualquier plataforma, permite a los distintos sistemas informáticos interoperar para que el usuario final emplee una única interfaz pudiendo acceder con él a múltiples sistemas, con los comandos, formatos y estilos de presentación que le sean familiares.

La primera versión del estándar Z39.50 se aprobó en 1988, y ha quedado obsoleta. En los años siguientes se ha seguido trabajando en el desarrollo del estándar y se cuenta con participación de otros países como Canadá, Australia y países europeos, garantizando así un estándar de aplicación realmente internacional, que incorpora numerosas mejoras y reconoce como medio de aplicación TCP/IP e Internet.

En la actualidad, Z39.50 es un estándar maduro, con una amplia presencia en la comunidad bibliotecaria, al menos de algunos países. Pero, el desarrollo e implantación de Z39.50 convive con la popularización de páginas Web, cada vez mejor diseñadas y potentes, que se han convertido en una manera barata y muy extendida de ofrecer un acceso amigable a la información bibliográfica y cuya funcionalidad crece cada día.

No obstante, no hay dos sistemas Web cuya interfaz tenga las mismas características, no suele ser posible transferir directamente los resultados obtenidos para emplearlos en otras aplicaciones y, cuando un usuario realiza una búsqueda a través de Web de informaciones ofrecidas por distintas organizaciones, encuentra muchas veces un océano de datos faltos de estructura. En estas circunstancias, Z39.50 se sigue presentando como un método de acceso bien estructurado, que permite buscar en distintos servidores con una única interfaz y resultados consistentes que puedan emplearse para otros fines, pero también es cierto que aún queda mucho camino por andar.

2.- ¿CÓMO FUNCIONA Z39.50?

El estándar Z39.50 especifica formatos y procedimientos para gestionar el intercambio de mensajes entre un cliente y un servidor, permitiendo así que el usuario busque en bases de datos remotas, identifique los registros que se ajusten a determinados criterios, y recupere algunos o todos los registros identificados, así como otras

informaciones asociadas en función de la base de datos.

La ventaja de cumplir dicho estándar es conseguir un acceso uniforme a un gran número de fuentes de información diversas y heterogéneas, incluso de manera simultánea, superando las diferencias entre los sistemas informáticos, los motores de búsqueda y las distintas bases de datos.

Para conseguir esta interoperabilidad entre distintos sistemas, Z39.50 facilita un lenguaje común para realizar las dos operaciones básicas que garantizan la recuperación de información: selección de información y obtención de la misma. Por ello, Z39.50 contempla la estandarización tanto de los mecanismos de codificación (cómo deben codificarse los datos para ser transferidos), como de la semántica del contenido (modelización de los datos con una semántica común para cada comunidad específica).

Desde el punto de vista de los mecanismos de codificación, es importante recordar que Z39.50 simplemente es un estándar de comunicación, es decir, sólo especifica qué debe transferirse, pero adicionalmente necesita un servicio de transporte fiable como TCP. Z39.50 especifica la estructura (Abstract Syntax Notation One, ASN.1) de los mensajes (Protocol Data Units, PDUs) que se intercambian entre cliente y servidor.

Pero, Z39.50 es un estándar muy amplio que ofrece una gran funcionalidad y atiende muy diversos entornos, no sólo el bibliotecario. Ningún desarrollo comercial, ni particular, soporta el estándar completo definido para todos los entornos, aunque el propio estándar describe los mínimos que deben cumplir todos los desarrollos para garantizar la interoperabilidad. Estas diferencias de un desarrollo a otro conllevan ciertos problemas.

Los servicios mínimos que deben contemplarse permiten negociar el inicio de una sesión, realizar una búsqueda en una base de datos, crear un conjunto de resultados en dicha base que se ajusten a la búsqueda y obtener uno o varios de los registros del conjunto. Aunque la versión 3 del estándar completa considera 11 facilidades, alguna de ellas con varios servicios:

- Initialization: permite al cliente negociar una Z-asociación (funcionalidad soportada, juego de caracteres, idioma, etc.).
- Search: permite al cliente consultar las bases de datos de un servidor, crear el conjunto de resultados en el servidor y recibir información sobre dicho conjunto.
- Retrieval: incluye dos servicios distintos, Present, que permite al cliente solicitar uno o más registros de un conjunto de resultados, y Segmentation, que permite al servidor descomponer en varios segmentos la información solicitada en los casos necesarios.
- Result-set-delete: permite a un cliente solicitar que se borre un conjunto de resultados determinado, o todos ellos.
- Access Control: permite al servidor que evalúe al cliente, mediante palabras de

paso, etc.

- Accounting/resource Control: incluye tres servicios distintos, Resource-control, que permite al servidor controlar e informar al cliente de los recursos consumidos o estimados, Trigger-resource-control, que permite al cliente solicitar que se inicie el control de recursos o cancelar la operación, y Resource-report, que permite al cliente solicitar un informe de recursos tanto de una operación, como de una sesión completa.
- Sort: permite que el cliente solicite al servidor una ordenación del conjunto de resultados, o unir varios conjuntos y luego ordenar el resultado.
- Browse: permite recorrer una lista ordenada de términos (materias, títulos, etc.).
- Explain: ofrece detalles del servidor como bases de datos disponibles, índices, servicios disponibles, etc. con idea de que se puedan desarrollar clientes que se auto configuren en función de los servidores que encuentren.
- Extended Services: permite el acceso a servicios ajenos al protocolo que pueden perdurar una vez termine la Z-asociación como búsquedas periódicas, conservar conjuntos de resultados, etc.
- Termination: permite al cliente o al servidor interrumpir las operaciones activas e iniciar el cierre de la Z-asociación.

Para garantizar el otro aspecto de la interoperabilidad, conjuntamente a esta codificación estándar de mensajes, Z39.50 se apoya en el concepto de conocimiento de una semántica compartida. Los distintos ámbitos proveedores o consumidores de contenidos de información han ido acordando estructuras y atributos comunes en cada uno de ellos, lo que permite un acceso uniforme a información heterogénea dentro de cada ámbito o comunidad.

El modelo de arquitectura básico del estándar Z39.50 se apoya en este concepto de semántica en función del contenido. Es decir, cada servidor ofrece una visión de sus bases de datos en función del dominio o el ámbito en que nos encontremos, una representación virtual de los registros que contiene, donde la estructura lógica real de la base de datos permanece oculta, y sólo se refleja la que corresponde a la semántica de ese dominio.

Adicionalmente a la representación virtual de la base de datos, se definen y registran para cada ámbito los puntos de acceso que se pueden emplear en las consultas (Attribute Sets) y las maneras de estructurar los datos al facilitárselos al cliente (Schemas). En el caso de la comunidad bibliográfica, se ha registrado el Bib-1 Attribute set y distintos formatos para presentar las respuestas, como los distintos formatos MARC, el formato propio del OPAC a que accedemos, etc. Además, se sigue trabajando para ampliar las semánticas registradas en función de las carencias detectadas en la práctica.

En los diferentes ámbitos en que se ido adoptando Z39.50 como estándar para

acceder a distintos tipos de información, se han ido definiendo y registrando sus correspondientes Attribute Sets y Schemas. Aunque los puntos de acceso a la información suelen variar mucho de una semántica a otra, los formatos de presentación de los resultados son en muchas ocasiones válidos para todos ellos.

Entre otros, se han registrado ya semánticas propias de dominios como el bibliográfico, el científico-técnico, el geoespacial, las colecciones de museos, las bibliotecas digitales, etc.

3.- SITUACIÓN ACTUAL

En los últimos años se han implementado servidores Z39.50 en diferentes ámbitos geográficos y temáticos, pero si nos ceñimos al mundo bibliotecario, a 22 de septiembre de 1999 la empresa danesa Index Data ofrecía información detallada de 246 servidores y 377 bases de datos distintas, y Bookwhere llegaba a ofrecer 555 bases de datos.

Si bien es cierto que es una cantidad importante de bases de datos a las que en teoría puede accederse mediante una única búsqueda como si fuesen recursos locales, también es cierto que casi el 40% ofrece sólo las funciones básicas del estándar, en sus diferentes versiones. Esta disparidad en los desarrollos, tanto por la manera de interpretar algunos aspectos del estándar, como por las opciones consideradas y las omitidas, conlleva que el resultado de una búsqueda en varios servidores puede ser poco fiable e inconsistente.

Aunque en algunas ocasiones este hecho ha llevado a plantearse la eficiencia del estándar Z39.50, en la actualidad existe un consenso mayoritario que achaca las deficiencias al uso que se ha hecho del mismo. Para paliarlo, las distintas comunidades interesadas han alcanzado acuerdos para hacer un uso común de las facilidades del estándar que les eran necesarias y la manera de desarrollarlas, definiendo así distintos perfiles (perfiles) que se registran a medida que se aprueban para que sean de utilidad a una comunidad más amplia.

Algunos de estos perfiles han sido definidos a escala nacional, como el de Finlandia o el de Francia. Pero buscando la interoperabilidad prometida por el estándar Z39.50, ya se está trabajando en un perfil internacional para catálogos virtuales que sea compatible con los definidos en otros proyectos, de manera que las búsquedas en distintos servidores simultáneamente optimicen la fiabilidad de los resultados obtenidos. La definición de este perfil internacional está coordinada por la Biblioteca Nacional de Canadá, y una vez se alcance el correspondiente consenso, el ZIG lo someterá al grupo de trabajo de ISO para su registro en el ámbito internacional.

Además, una vez superadas las funciones básicas del estándar, los desarrollos se enriquecen y contemplan servicios adicionales, como el Explain, que forma parte del estándar y permite al cliente preguntar al servidor las funciones que es capaz de realizar, una descripción del OPAC que sustenta, así como ayuda para gestionar los resultados esperados. Si bien en el momento actual sólo 11 de los 246 servidores mencionados han abordado esta fase de su desarrollo.

Pese al desarrollo de perfiles internacionales y al mejor comportamiento de clientes y servidores Z39.50, sigue siendo necesario estudiar y modificar las semánticas y estructuras en las que se basa el estándar, por ejemplo para su aplicación en bibliotecas musicales o en fondos de manuscritos modernos. Pero para cada una de estas situaciones se van creando grupos de trabajos específicos e internacionales.

También se ha intentado aplicar el estándar Z39.50 para la creación de uniones virtuales de catálogos, con idea de disminuir los costes técnicos y económicos del sistema tradicional, sin interferir en los métodos y procedimientos de trabajo de cada organización implicada, y persiguiendo una gestión simple del préstamo inter-bibliotecario con cualquier servidor. Esta aplicación del estándar ha sido desestimada en la práctica durante mucho tiempo fundamentalmente debido a la disparidad en la gestión de la información de holdings y circulación que presentaba cada sistema de automatización. Esta falta de consistencia entre sistemas ha desembocado, después de dos años de trabajo, en el OPAC/Holdings Schema recientemente aprobado que resuelve dichas diferencias basándose en el estándar Z39.50. Aunque quedan ciertos desarrollos por completar, por ejemplo para permitir búsquedas dentro de esa información, ya está anunciada su aplicación en distintos proyectos nacionales e internacionales.

Por otra parte, si el estándar Z39.50 aboga por la interoperabilidad, no pueden olvidarse otros muchos desarrollos ya existentes en la práctica. Se está trabajando en Z+SQL como una extensión del protocolo que una las ventajas de SQL a los servicios de Z39.50, en su compatibilidad con XML, en un conjunto de atributos consistente con el Dublin Core, con el acceso simultáneo a bases de datos en otros soportes como la aplicación de la tecnología ERL de SilverPlatter a los CD-ROMs, en la mejora de la negociación del juego de caracteres, etc. En algunos de esos campos los resultados empiezan a aplicarse, en otros queda mucho estudio por realizar.

La flexibilidad y la amplitud de miras del estándar Z39.50 ha permitido que se inicie su aplicación a campos mucho más extensos que la búsqueda puntual de información bibliográfica mediante una interfaz común. Las facilidades Extended Services de la versión 3 permiten definir criterios de búsqueda que se almacenan en los distintos servidores y se procesan periódicamente, de esta manera cada usuario puede localizar los servidores de su interés y definir su perfil para una difusión selectiva de información mediante una única interfaz, recibiendo la información solicitada en su correo electrónico, por ejemplo.

Igualmente esas facilidades de Extended Services permiten desarrollos para filtrar las áreas no deseadas de los contenidos a los que se acceden a través de Web, pudiendo cada usuario definir sus propios parámetros de filtro en su cliente y para cada motor de búsqueda. Hasta ahora no ha sido considerado un aspecto prioritario, pero con su generalización se mitigarían en gran medida la frustración y el tiempo requerido en las búsquedas a través de Web.

Otra funcionalidad contemplada por el estándar que ahora inicia su aplicación práctica es la actualización remota de bases de datos, es decir, permite no sólo recuperar un registro, sino además editarlo y devolverlo para actualizar la base de datos. Cuando se

depuren los desarrollos podrían obtenerse clientes de catalogación que trabajen contra distintas bases de datos de diferentes sistemas, actualizando varios servidores a un mismo tiempo y ofreciendo al usuario una interfaz única para catalogar en todos ellos.

También se está trabajando en el desarrollo del perfil Zthes, que unido a las facilidades de Browse del estándar, dotarán a los clientes de consulta y catalogación de funcionalidades habituales de los sistemas de automatización que hasta el momento actual eran difíciles de implementar.

4.- UTOPIA VS SITUACIONES REALES

En los últimos años se han ido desarrollando numerosos servidores y aplicaciones Z39.50, según indican los datos facilitados anteriormente. En un principio se instalaban básicamente en bibliotecas universitarias, que siguen representando un 35% de los servidores disponibles, pero paulatinamente se ha ido incorporando a todo tipo de bibliotecas: especializadas, públicas, nacionales, etc. Pero al mismo tiempo encontramos servidores que han abandonado esta línea de trabajo después de algunos años de experiencia, como la base de datos Ten-year de la California Digital Library o las de la National Library of Medicine.

Los desarrollos iniciales basados en el estándar Z39.50 atienden las funcionalidades básicas de recuperación de información y obtención de registros, de hecho, sólo los servicios Search y Present son soportados por el 100% de los servidores, cualquier otro servicio no tiene una implantación superior al 67%. Pero si nos limitamos a esa funcionalidad básica, es mucho más espectacular la proliferación de desarrollos a través de Web, que resultan más baratos y rápidos de implementar y permiten controlar la manera de presentar la información al usuario, aspecto al que algunos proveedores de información, especialmente los comerciales, no quieren renunciar a favor de una interfaz única de libre elección por parte del propio usuario.

Limitándonos a esa funcionalidad básica del estándar Z39.50, y en contraposición al acceso a través de Web, deberíamos contar con la ventaja de poder acceder de manera simultánea a distintos servidores y bases de datos. Pero la experiencia dice que los resultados obtenidos en estas búsquedas simultáneas suelen ser engañosos, inesperados e inconsistentes.

La definición del estándar Z39.50 es tremendamente amplia y flexible, de hecho en el apartado anterior se han mencionado numerosos aspectos del mundo y la actividad bibliotecaria que podrían beneficiarse si se mejoran y completan los desarrollos que se basan en dicho estándar. Pero cuando se abordan esos desarrollos, la misma flexibilidad del estándar lleva a diferencias de interpretación y de niveles de desarrollo que ponen fin a la interoperabilidad prometida.

Se podrían apuntar más situaciones como las anteriores, que parecen reflejar una idea común: las capacidades del estándar Z39.50 y los beneficios que prometen hacen ser entusiasta, pero los problemas prácticos en los desarrollos desaniman al evaluador.

En cualquier caso, es una situación prometedora saber que la comunidad bibliotecaria cuenta con un estándar de estas características y, conociendo la realidad de los esfuerzos que se están haciendo en los desarrollos, se puede mantener la confianza en alcanzar los beneficios deseados.

Hay un interés creciente en el estándar Z39.50 dentro de la comunidad bibliotecaria, especialmente ahora que se manifiesta abiertamente su aplicación internacional. Por este mismo motivo, los usuarios demandan cada vez una mayor funcionalidad y calidad en los resultados y su aplicación. Cada vez hay más países que definen rigurosamente perfiles nacionales, con el apoyo de los correspondientes organismos responsables de las normas de catalogación y que trasladan los resultados a los centros de formación de bibliotecarios.

Ya no se plantea la aplicación del estándar compitiendo contra XML/SQL, sino como la única alternativa posible a la falta de estructura que conlleva cualquier desarrollo en HTML. Es decir, como un meta-lenguaje fiable y consistente, tanto desde el punto de vista de los datos, como de los requerimientos de los usuarios.

Se es consciente de que el estándar Z39.50 nunca dominará el mundo de los sistemas de información, pero seguirá creciendo y evolucionando en los entornos especializados.

Según refleja el estudio de William Moen sobre la evolución del estándar Z39.50, su gran logro ha sido permanecer siempre atento a los requerimientos prácticos. Aunque su desarrollo formal se inicia en 1979, realmente habría que describirlo como una evolución basada en individuos y organizaciones activas tanto en el entendimiento como en la resolución de los problemas a medida que se presentaban. Este estándar trata un problema muy común de su época, pero ofrece una solución genérica única y un modelo susceptible de evolucionar.

Esta característica del estándar, unida a una participación única de los propios desarrolladores en su evolución, ha permitido unir la estabilidad de los acuerdos al rango de estándar internacional. Es decir, pocas veces se ha contado con tal cantidad de grupos de trabajo dispuestos a validar, usar, proponer, definir requerimientos, más una Agencia de Mantenimiento que protege la integridad del estándar, más una marcada vocación educativa de la comunidad afectada. Ha sido básica la implicación del ZIG, donde figuran grupos interesados en la evolución del estándar, pero con una capacidad única tanto en aspectos técnicos, como en su decisión de compartir información y provenientes de las más distintas actividades.

Frente a los grupos que plantean aproximaciones alternativas a los estándares tradicionales, de implantación mucho más rápida como el W3C, Z39.50 se presenta como un híbrido entre la tradición y la innovación. Con el ZIG como grupo informal que potencia su evolución con una aplicación práctica, y la Agencia de Mantenimiento que garantiza su coherencia e integridad. Este hecho permite alcanzar un equilibrio entre anticipación y conservadurismo que lo convierte en un estándar internacional, pero entendido como “algo útil, no perfecto”.

5.- CONCLUSIONES

Por ello podemos concluir que Z39.50 como estándar internacional ha demostrado con creces cumplir sus objetivos iniciales, y aunque los desarrollos que se basan en dicho estándar siguen alcanzando numerosos problemas para su implementación, el esfuerzo de los numerosos grupos de trabajo que lo respaldan convierten estas dificultades en evolución del estándar, permitiendo que en la actualidad soporte funcionalidades no contempladas inicialmente y prometiendo un futuro de continua mejora.

En las puertas del siglo XXI, y considerando las bibliotecas y centros de documentación como piezas claves de la sociedad de información en que ya estamos inmersos, es importante que conozcan la situación de desarrollo e implantación del estándar Z39.50 en la comunidad internacional, en qué medida es un estándar “real” y en que aspectos debe considerarse “virtual”, pero en cualquier caso, debe contemplarse firmemente en las distintas actuaciones.

BIBLIOGRAFÍA

“ATTRIBUTE set for searching MARC structured records”, April 1999, <http://www.rlg.org/home/attrlib/marcattr.html>

BIBLIO TECH REVIEW Information Technology for Libraries: “Z39.50: Part 1- an overview”, May 1999, http://www.biblio-tech.com/html/z39_50.html

BIBLIO TECH REVIEW Information Technology for Libraries: “Z39.50: Part 2- Technical Details”, May 1999, http://www.biblio-tech.com/html/z39_50_part_2.html

“CIMI Standards Recommendations”, August 1999, <http://www.cimi.org/standards/index.html>

DISTRIBUTED SYSTEMS TECHNOLOGY CENTRE: “Z+SQL Distributed Interoperable Database Searching with Z39.50 and SQL”, <http://archive.dstc.edu.au/DDU/projects/Z3950/Z+SQL/>

FINNIGAN, Sonya; WARD, Nigel: “Z39.50 Made Simple”, <http://archive.dstc.edu.au/DDU/projects/Z3950/zsimple.htm>

GATENBY, Janifer: “Searching Serials and Multi-Part Items”, July 1999, <http://lcweb.loc.gov/z3950/agency/august/serial.html>

GATENBY, Janifer; VAN LIEROP, Pieter: “Attribute set Holdings”, July 1999, <http://lcweb.loc.gov/z3950/agency/august/holdattr.html>

GILBY, John: "M25 Link Project", <http://www.M25lib.ac.uk/M25link/flyer.html>

IANNELLA, Renato; MILLER, Paul: "Proposed definitions for Dublin Core Version 1.1", June 1999, <http://archive.dstc.edu.au/RDU/DCAC/PR-DCV11.html>

JOINT Information Systems Committee: "Music libraries online", July 1999, <http://www.musiconline.ac.uk>

LÉVÉJAC, Anne-Lise: "Z39.50: L'Information bibliographique structurée sur le net", Mar 1997, <http://wwwperso.hol.fr/~alevejac/Z3950.htm>

LÓPEZ DE SOSOAGA Torija, Arantza: "OPAC-Web-Z39.50: ¿Redundantes o complementarios? La realidad es multilingüe", 1998, <http://lcweb.loc.gov/z3950/agency/papers/biblio.html>

LÓPEZ DE SOSOAGA Torija, Itziar: "El proyecto ARCA: La manera de recorrer las "autopistas de la información" sin necesidad de aprenderse los "códigos de circulación", 1996, <http://lcweb.loc.gov/z3950/agency/papers/biblio.html>

LUNAU, Carrol D.: "Availability of Summary Holdings Information from vCuc Targets", Decembre 1997, <http://www.nlc-bnc.ca/resource/vcuc/vchold1.htm>

LUNAU, Carrol D.: "The Need for an International Z39.50 Profile for Searching Virtual Catalogues", October 1998

LUNAU, Carrol D.; TURNER, Fay: "vCuc Pilot Project: Status Report and Preliminary Identification of Issues", June 1997, <http://www.nlc-bnc.ca/resource/vcuc/earticle.htm>

"MALVINE – Manuscripts and Letters via Integrated Networks in Europe", <http://www.malvine.org/index.html>

MOEN, William: "Handling Bibliographic Holdings Information in Z39.50: Implications for the Z Texas Profile", July 1999, <http://www.tsl.state.tx.us/LD/z3950/Holdings27Mar99.htm>

MOEN, William: "The development of ANSI/NISO Z39.50: a case study in standards evolution", August 1998, <http://www.unt.edu/wmoen/dissertation/DissertationIndex.htm>

"OPAC/HOLDINGS Profile 1: Profile for retrieving detailed library holdings in a bibliographic environment", January 1998, <http://www.nlc-bnc.ca/iso/z3950/holds6.htm>

"PROFIL Z3950 du Catalogue Collectif de France", Décembre 1998, http://www.ccf.fr/bnf.fr/rnbc_d_visu/apropos0.htm

"RIDING Z39.50 Gateway to Yorkshire Libraries", April 1999, <http://www.shef.ac.uk/~riding>

“SPECIFICATION of XML Encoding Rules (XER)”, April 1999,
<http://asf.gils.net/xer/standard.html>

ST-GELAIS, Rolande: “OPAC/Holdings Schema”, July 1999,
<http://www.dra.com/resources/z3950/Opac-schema-defs-4.pdf>

STOVEL, Lennie: “Bib-2 Attribute Set Second Draft”, July 1999,
<http://www.rlg.org/home/attrib/draft2.html>

TAYLOR, Mike: “Zthes: a Z39.50 Profile for Thesaurus Navigation”, July 1999,
<http://www.staff.tecc.co.uk/mike/zthes-03.html>

WELLS, Andrew; PEARCE, Judith; GROOM, Linda; LEE Bronwyn: “Connecting and Sharing: the Emerging Role of Z39.50 in Library Networks”, January 1998,
<http://www.nla.gov.au/nla/staffpaper/awells2.html>

“Z39.50 Cross-Domain Attribute Set”, August 1999,
<http://www.oclc.org/~levan/docs/crossdomainattibuteset.html>

“Z39.50 Target Statistics”, September 1999, <http://www.indexdata.dk/targettest/stat.shtml>

“Z39.50 Utility Attribute Set Draft 3”, July 1999,
<http://lcweb.loc.gov/z3950/agency/attrarch/util.html>

“ZIG Commentaries”, Septembre 1999, <http://lcweb.loc.gov/z3950/agency/wisdom.html>