

LA BÚSQUEDA Y RECUPERACIÓN NORMALIZADA DE LA INFORMACIÓN BIBLIOGRÁFICA: Z39.50 Y Z39.50 INTERNATIONAL NEXT GENERATION.

I. INTRODUCCIÓN

El fin último y primordial de toda biblioteca es **hacer accesible el conocimiento** para así poder formar, entretener e informar. Para ello, resulta imprescindible **garantizar el éxito de las labores de búsqueda y recuperación de recursos**, especialmente ahora que las necesidades informativas de los usuarios son cada vez más exigentes, dado el enorme volumen documental producido en la Sociedad del Conocimiento. Es precisamente esta abundancia informativa la que hace que ninguna biblioteca actual, por grande que sea o amplios fondos que contenga, pueda considerarse autosuficiente a la hora de satisfacer las demandas de sus usuarios y busque la cooperación como solución prioritaria al problema. Hasta ahora, el mayor obstáculo para las iniciativas cooperativas ha sido, sin duda, la elección de diferentes sistemas de automatización por parte de las bibliotecas, lo que está intentando subsanarse a través de la elaboración de estándares y protocolos unificadores tan importantes como el Z39.50.

II. LA IMPORTANCIA DE LA NORMALIZACIÓN EN LOS PROCESOS DE BÚSQUEDA Y RECUPERACIÓN DE INFORMACIÓN

En principio, cada biblioteca o institución documental emplea sus propios sistemas de búsqueda y recuperación de información, caracterizados por unos determinados lenguajes de interrogación así como por unos procedimientos y métodos específicos para su manejo. El problema radica en que los usuarios no siempre encuentran lo que buscan en su centro habitual de información, de modo que necesitan ampliar sus consultas a otros centros complementarios. Teóricamente, esto obligaría a dichos usuarios a ir consultando diferentes portales, catálogos y bases de datos, con diferentes interfaces, mecanismos y modos de búsqueda hasta hallar los resultados deseados. Así, necesitarían conocer el funcionamiento de cada sistema individual para poder tener éxito en su tarea de rastreo de la información.

Para evitarlo, hace años que se realizan importantes esfuerzos por desarrollar normas y estándares capaces de posibilitar la interoperabilidad entre equipos heterogéneos y, con ello, la transparencia en el intercambio de información bibliográfica, tanto a nivel nacional e internacional. Uno de los más destacables es el protocolo Z39.50, destinado primordialmente a permitir el entendimiento entre aplicaciones diversas de búsqueda y localización de información.

Aunque los ordenadores presenten plataformas informáticas o sistemas operativos diferentes, características diversas y bases de datos de distinta naturaleza, el protocolo facilita la comunicación entre ellos en un lenguaje común.

Formalmente, facilita la **interconexión entre los usuarios y las bases de datos** necesarias, independientemente de su procedencia, a partir de una interfaz común de fácil manejo. Pero estudiémoslo más detenidamente:

III. EL PROTOCOLO Z39.50 – EVOLUCIÓN, ESTRUCTURA Y APLICACIONES

La abundante literatura sobre el Z39.50 sitúa sus orígenes a comienzos de los **años 80** en EE.UU., concretamente en el seno del **LSP** (*Linked Systems Project*), en el que participaban los grandes sistemas bibliotecarios y bibliográficos americanos, entre ellos los de la LC o el OCLC (*Online Computer Library Center*). El proyecto incluía entre sus objetivos el desarrollo de un protocolo experimental que buscara en diferentes bases de datos bibliográficas y permitiera transferir registros entre los diferentes sistemas participantes. Por lo tanto, uno de sus aspectos esenciales era la interoperabilidad, para lo que era necesario el uso de un lenguaje común caracterizado por la normalización sintáctica y semántica.

A partir del prototipo creado por el proyecto LSP, se fueron sucediendo las investigaciones hasta que, en **1988**, el Instituto Nacional Americano de Estándares (ANSI) acreditó el trabajo y aprobó la **norma Z39.50 versión 1**. No obstante, no fue hasta la publicación de su **versión 2**, en **1992**, cuando la comunidad bibliotecaria comenzó a interesarse por su aplicación. En los años siguientes se continuó trabajando en el desarrollo del estándar y se contó con la colaboración de otros países, como Canadá, Australia y naciones europeas. Esto dio lugar a la aprobación de su **versión 3** en **1995**, compatible con la versión 2, pero mejorada hasta tal punto que, en **1998**, la **Organización Internacional para la Normalización (ISO)** decidió adoptar la norma americana Z39.50 como estándar internacional bajo el nombre **ISO 23950-1998**. Además de adquirir una dimensión verdaderamente internacional, su aceptación como norma ISO consolidó el protocolo TCP/IP de Internet como su medio de aplicación.

Cabe señalar que su denominación como Z39.50 proviene de haber sido desarrollado por el **comité número 39 del mencionado Instituto ANSI** y de ser el **estándar número 50 de la NISO**, Organización Nacional para la Normalización en Información que representa a la ISO en EE.UU. Su **última versión como norma ANSI/NISO** data de **2003**.

* Se comprueba poniendo "Z39.50" en el sitio web de la ANSI:

<http://webstore.ansi.org/>

Su mantenimiento, como norma ISO y ANSI/NISO, corre a cargo de la **International Standard Maintenance Agency** de la LC. El **ZIG** (*The Z39.50 Implementors Group*), creado en 1990 bajo los auspicios de la LC, ha sido el **grupo de trabajo encargado del desarrollo y revisión** del estándar durante los últimos años. El estar compuesto en su mayoría por productores de Sistemas Automatizados de Bibliotecas, interesados en la interoperabilidad, ha redundado en una aplicación mucho más amplia y mejorada del Z39.50. En la actualidad, la práctica totalidad de los SIGB incorporan este protocolo en sus módulos.

ESTRUCTURA

El Z39.50 necesita cinco elementos fundamentales:

1. Un equipo cliente, denominado CLIENTE Z.
2. Un equipo servidor, denominado SERVIDOR Z.
3. Una estructura de datos.
4. Unas reglas de intercambio.
5. Un medio de transporte fiable para la información, como el protocolo TCP/IP.

Se basa pues en una estructura cliente-servidor, en la que un ordenador, que actúa como servidor, opera a modo de cliente enviando una consulta a otro ordenador, que actúa como servidor de la información. El software en el servidor ejecuta una búsqueda en una o más bases de datos y da como resultado un set de registros que se adecuan a la demanda informativa solicitada. Por lo tanto, el protocolo Z39.50 no hace sino definir un **medio normalizado** para comunicar ordenadores, a través de una estructura y una serie de reglas de intercambio, con el fin de recuperar información. Al quedar establecidas las especificaciones técnicas necesarias para que dicha comunicación pueda producirse, ésta resulta posible aunque los motores de búsqueda o bases de datos sean diferentes.

Pero además de solucionar el problema del “entendimiento” entre sistemas operativos, la clave del Z39.50 reside en la separación de la interfaz de usuario del motor de búsqueda, por lo que, como ya adelantaba, no es necesario conocer cientos de interfaces ni la estructura de decenas de bases de datos para encontrar la información deseada. El servidor se encarga de traducir el lenguaje, de forma que pueda ser comprendido por ambas partes de la cadena. Los resultados se presentan en el sistema local, en el formato y estilo con los que el usuario está familiarizado.

Lo ideal es que el cliente esté integrado en el sistema de gestión bibliotecaria, para que la información pueda ser recuperada sin abandonar el entorno habitual de trabajo. No obstante, si el sistema no dispone de tal funcionalidad, existen varios programas autónomos que pueden ser utilizados en un entorno Windows.

SERVICIOS MÍNIMOS Y APLICACIONES BIBLIOTECARIAS

Los tres primeros y principales servicios que deben contemplarse son los relacionados con el proceso recién mencionado, es decir:

1. *Initialization* o **INICIALIZACIÓN**, precursora del trabajo real, en la que se establecen los parámetros básicos (reglas y negociación) de la sesión que se va a iniciar entre cliente y servidor. Permite asimismo la autenticación del usuario.
2. *Search* o **BÚSQUEDA**, funcionalidad más importante del estándar, que permite lanzar consultas simples o complejas con la misma herramienta a múltiples bases de datos que cuenten con un servidor Z. Las estrategias de búsqueda pueden utilizar operadores booleanos, de comparación, de proximidad, etc.
3. *Retrieval* o **RECUPERACIÓN**, una vez realizada la búsqueda, se produce la devolución de registros al equipo cliente y su visualización en el OPAC de usuario, como si de registros del propio OPAC se tratara.

El estándar ofrece otras muchas facilidades adicionales, aunque no es necesario que se implementen todas. Permite, por ejemplo, servicios extendidos como: guardar juegos de resultados para usos posteriores, guardar estrategias de búsquedas, actualizar bases de datos, ordenar la información recuperada, etc.

Durante los últimos años, el uso del Z39.50 se ha reservado casi exclusivamente al mundo bibliotecario. No obstante, factores como la ampliación de funcionalidades de la norma y la universalización en el acceso a Internet hacen que sus posibilidades de evolución futura sean enormes. Su implantación y desarrollo no van en detrimento de la Web, sino que deben converger.

Las aplicaciones más destacables a efectos de tareas bibliotecarias se han centrado fundamentalmente en los aspectos relativos a los OPAC, la catalogación compartida, la construcción de catálogos colectivos, el préstamo interbibliotecario o la difusión selectiva de la información. Todo ello ha sido posible gracias a que permite realizar peticiones a distintas bibliotecas, compartir registros bibliográficos, utilizar MARC como formato de intercambio

básico, especificar y grabar parámetros para ser ejecutados posteriormente, etc. Por lo tanto, sus ventajas, tanto para bibliotecarios como usuarios, son muchas.

Dado su origen, es en el mundo anglosajón donde existe una mayor aplicación de la norma. No obstante, el número de **servidores Z39.50** ha aumentado considerablemente en nuestro país a lo largo de los últimos años, por lo que encontramos su presencia en la BNE, REBIUN, REBECA, bibliotecas del CSIC, Insituto Cervantes, Congreso de los Diputados, diversas bibliotecas universitarias (UCM, UAM, UC3M, UPM, UCLM...), múltiples BPE, etc. El mapa actualizado de servidores en España puede consultarse en el sitio web de las **Bibliotecas de Castilla y León**, pioneras en este campo.

IV. EL Z39.50 INTERNATIONAL NEXT GENERATION (ZING)

Bajo este nombre se engloba el último proyecto promovido por el grupo de impulsores del protocolo Z39.50, en el seno de la mencionada *International Maintenance Agency*. Fue ideado con el propósito de conseguir una mayor accesibilidad para los contenidos intelectuales y semánticos del estándar así como para hacer de él una herramienta más atractiva de cara a los nuevos proveedores de información, desarrolladores, distribuidores y usuarios.

En la actualidad, el enlace ZING de la Agencia remite directamente al sitio web de una de sus iniciativas principales, **SRU**, ligada a otras dos, también previstas en el proyecto ZING: **CQL** y **ZeeRex**.

* Tecleando "http://www.loc.gov/z3950/agency/zing/" aparece la siguiente información: The ZING site has been removed and the information is now accessible from the SRU home page at <http://www.loc.gov/standards/sru/>

Tal y como se indica en dicho sitio web, **SRU** es el **protocolo**, **CQL** el **lenguaje de interrogación** y **ZeeRex** el **servicio de descripción**:

- **SRU** (*Search/Retrieval via URL*). Es un protocolo de búsqueda y recuperación de información en Internet, basado en la experiencia del Z39.50. De él ha tomado su funcionalidad y los aspectos más útiles, simplificando la complejidad mediante el uso de conexiones HTTP, navegadores Web y el lenguaje XML para las transacciones. Vio la luz en 2002 y su objetivo subyacente coincide con el del protocolo Z39.50, pues busca convertirse en un estándar para la búsqueda y recuperación de información, permitiendo que un único cliente acceda a varios servidores y que un servidor sea accesible por varios clientes de manera automática y fácil. Al hacer uso de XML facilita bastante el acceso a varias bases de datos y la legibilidad de los mensajes intercambiados. Las búsquedas deben ser articuladas usando la sintaxis del CQL.

- **CQL** (*Common Query Language*). Es el componente clave para la operación *búsqueda y recuperación de información* del protocolo SRU. Es un lenguaje formal para representar dudas y consultas, escritas por las personas, pero legibles por las máquinas. CQL pretende combinar la sencillez de las búsquedas que estamos acostumbrados a realizar en Google, Yahoo, Altavista, etc, con la potencia del Z39.50. Persigue, en definitiva, ser un lenguaje simple e intuitivo (con operadores booleanos, de comparación y limitación, etc.), pero manteniendo la expresividad de otros lenguajes más complejos.
- **ZeeRex** (*Z39.50 Explain, explained and Re-Engineered in XML*). Esta iniciativa surge como un intento de resolver dos problemas muy relacionados: la localización de recursos y, una vez ubicados, el conocer exactamente que podemos hacer con ellos. Tras varios intentos, unas veces demasiado complejos (Explain Classic) y otras demasiado centrados en proyectos particulares (Explain Lite) contamos ahora con ZeeRex, basado en el intercambio de registros XML. Estos registros serán los que nos ofrezcan la información que buscamos, bien de forma abreviada (puerto, nombre de bases de datos) o completa, incluyendo hasta la persona de contacto encargada del mantenimiento del recurso al que queremos acceder.

V. CONCLUSIÓN

En resumen, el crecimiento de la demanda informativa unido a la proliferación de bases de datos, cada una con sus propias características formales, de búsqueda y recuperación llevó al desarrollo de una norma que permitiera consultarlas bajo una misma interfaz y lenguaje de recuperación, el protocolo Z39.50.

Hoy es un protocolo maduro, que se erige como medio universal para acceder a distintos catálogos interconectando sistemas informáticos. Además se trata de uno de los pocos protocolos desarrollados y mantenidos por bibliotecarios. Por todo ello, para algunos constituye el estándar documental más importante desde la aparición del Formato MARC, especialmente por su capacidad para realizar búsquedas en catálogos remotos a través de Internet.

PUNTOS TRATADOS EN EL TEMA

LA BÚSQUEDA Y RECUPERACIÓN NORMALIZADA DE LA INFORMACIÓN BIBLIOGRÁFICA: Z39.50 Y Z39.50 INTERNATIONAL NEXT GENERATION.

I. INTRODUCCIÓN

II. LA IMPORTANCIA DE LA NORMALIZACIÓN EN LOS PROCESOS DE BÚSQUEDA Y RECUPERACIÓN DE INFORMACIÓN

III. EL PROTOCOLO Z39.50 – EVOLUCIÓN, ESTRUCTURA Y APLICACIONES

ESTRUCTURA

- 1 Un equipo cliente, denominado CLIENTE Z.
- 2 Un equipo servidor, denominado SERVIDOR Z.
- 3 Una estructura de datos.
- 4 Unas reglas de intercambio.
- 5 Un medio de transporte fiable para la información, como el protocolo TCP/IP.

SERVICIOS MÍNIMOS Y APLICACIONES BIBLIOTECARIAS

1. *Initialization* o **INICIALIZACIÓN**
2. *Search* o **BÚSQUEDA**
3. *Retrieval* o **RECUPERACIÓN**

IV. EL Z39.50 INTERNATIONAL NEXT GENERATION (ZING)

- **SRU** (*Search/Retrieval via URL*).
- **CQL**
- **ZeeRex** (*Z39.50 Explain, explained and Re-Engineered in XML*).

V. CONCLUSIÓN