

Desarrollo de software “Búsqueda Multibase”

Una nueva alternativa para consultar bases de datos bibliográficas de forma remota

Gonzalo Reséndiz Cansino*, Mauricio Fabián Sánchez,**
Isela García Bravo**

RESUMEN

El documento describe las características y funcionamiento de una aplicación *web* cuyo objetivo es realizar búsquedas de manera simultánea en diferentes bases de datos bibliográficas, las cuales se caracterizan por estar en diferentes plataformas de sistemas operativos y arquitecturas de *software* para automatización de bibliotecas.

PALABRAS CLAVE: Multibase, Protocolo HTTP, URL, OPAC, base de datos, interfaz, Perl, PostgreSQL, metabuscador, OpenURL, Z39.50, *software* libre.

Abstract

This document describes the features and operation of a web application, which the main purpose is to do searches simultaneously in different bibliographical databases, that are characterized by different operating systems platforms and software architectures for library automation.

KEYWORDS: Multibase, HTTP Protocol, URL, OPAC, databases, interface, Perl, PostgreSQL, metasearch, OpenURL, Z39.50, open source software.

Introducción

El avance tecnológico en el área de cómputo y las telecomunicaciones ha permitido el desarrollo constante del llamado “Sistema Integral de Gestión de Bibliotecas”. La década de los 90’s da inicio a la generalización y extensión de la automatización de éstos, siendo los OPAC (Online Public Access Catalog) un componente fundamental que a través de los años ha presentado cambios constantes; en un inicio las búsquedas realizadas dependían de un menú con distintas opciones soportadas por un lenguaje de comandos, siendo su principal desventaja la navegación entre los registros por no ser muy amigable. A finales de la década los OPAC evolucionan a un ambiente gráfico con mejores opciones de búsqueda.

Desde hace varios años atrás los catálogos en línea se consideran los mejores medios de difusión y acceso a las colecciones de libros de las bibliotecas universitarias. Algunos autores, como Charles R.

* Subdirector de Informática, Dirección General de Bibliotecas (DGB), Universidad Nacional Autónoma de México (UNAM), Circuito Interior s/n, Ciudad Universitaria, 04510, México, D.F., México. Correo electrónico: gonzalo@dgb.unam.mx

** Secretario Técnico y Técnica Académica de la Secretaría Técnica de Informática y Telecomunicaciones, Dirección General de Bibliotecas (DGB), Universidad Nacional Autónoma de México (UNAM), Circuito de la Investigación Científica, Ciudad Universitaria, 04510, México, D.F., México. Correo electrónico: mauricio@dgb.unam.mx; isela@dgb.unam.mx

Hildreth¹, mencionan que los catálogos en línea deben ser interactivos, expandibles y públicos; interactivos, porque deben permitir la comunicación con el usuario y responder a sus necesidades.

Según Rowley, el concepto de OPAC no debe limitarse a un catálogo electrónico local, sino también debe acceder a recursos tanto internos como externos, lo que se conoce como la tercera generación de OPAC's.

Esta transición constante de arquitecturas, plataformas, lenguajes de programación, bases de datos y la enorme gama de herramientas *web* que existen hoy en día, exigen desarrollar una interfaz de usuario para búsqueda de información, la cual se fundamenta en dos elementos o características principales: *accesibilidad* y *funcionalidad*.

Asimismo, y en consecuencia, es preciso contar con una interfaz que mediante una sola consulta permita acceder y recuperar información de los catálogos bibliográficos de diferentes bibliotecas o instituciones educativas.

Antecedentes

Los OPAC se definen como catálogos de acceso público en línea, que permiten la consulta de los registros bibliográficos de una biblioteca realizando la función de intermediario entre los datos y el usuario. Si bien la mayoría de los sistemas de información comerciales para automatización de bibliotecas (ALEPH, INNOPAC, UNICORNIO, JANUIM, SIABUC, etcétera) incorporan un OPAC amigable y funcional, no ofrecen la posibilidad de realizar una búsqueda simultánea en catálogos bibliográficos externos o remotos de otras bibliotecas o instituciones educativas; por mencionar un ejemplo, el sistema Aleph 500 solo permite configurar múltiples bases de datos lógicas en el mismo servidor.

Dídac Margaix-Arnal² habla sobre la evolución de los OPAC y que los cambios más destacados son:

¹ HILDRETH, Charles R. "Beyond Boolean [en línea]... <http://www.ideals.uiuc.edu/dspace/bitstream/2142/7500/1/librarytrends35i4k_opt.pdf>

² MARGAIX-ARNAL, Dídac. *Acta del VI Workshop CALSI 2007* [en línea]... <http://www.calsi.org/2007/wp-content/uploads/2007/11/didac_margaix.pdf>

- El acceso remoto, cuando ya no fue necesario consultarlos en las propias bibliotecas.
- La aparición de interfaces *web*, que permitieron aprovechar las potencialidades del hipertexto.
- El enriquecimiento de los registros. Se añadió nueva información: cubiertas digitalizadas, resúmenes, etcétera.
- Integración con otras herramientas de información, especialmente a través de las tecnologías de enlaces como en OpenURL.
- Aportaciones desde disciplinas como la Recuperación de la Información, la Usabilidad, la Arquitectura de la Información, etcétera.
- Mejora de la experiencia de búsqueda con nuevas formas de visualización de la información y navegación.

Adicionalmente, las aplicaciones conocidas como Metabuscaadores permiten la conexión a catálogos bibliográficos mediante diferentes protocolos de comunicación; sin embargo, su adquisición e implementación representan costos elevados para una biblioteca.

Origen y objetivo del proyecto

La Dirección General de Bibliotecas de la Universidad Nacional Autónoma de México (DGB-UNAM) y la Red de Instituciones Mexicanas para la Cooperación Bibliotecaria "Grupo Amigos", acordaron realizar un proyecto académico institucional para el desarrollo de un aplicativo que permitiese proporcionar al usuario una interfaz común de consulta de información sobre catálogos bibliográficos de diferentes instituciones educativas y organismos de diversos sectores. Para tal efecto, la Subdirección de Informática de la DGB-UNAM procedió a ejecutar las tareas de análisis, programación e implementación de esta interfaz de *software*.

OPAC comerciales

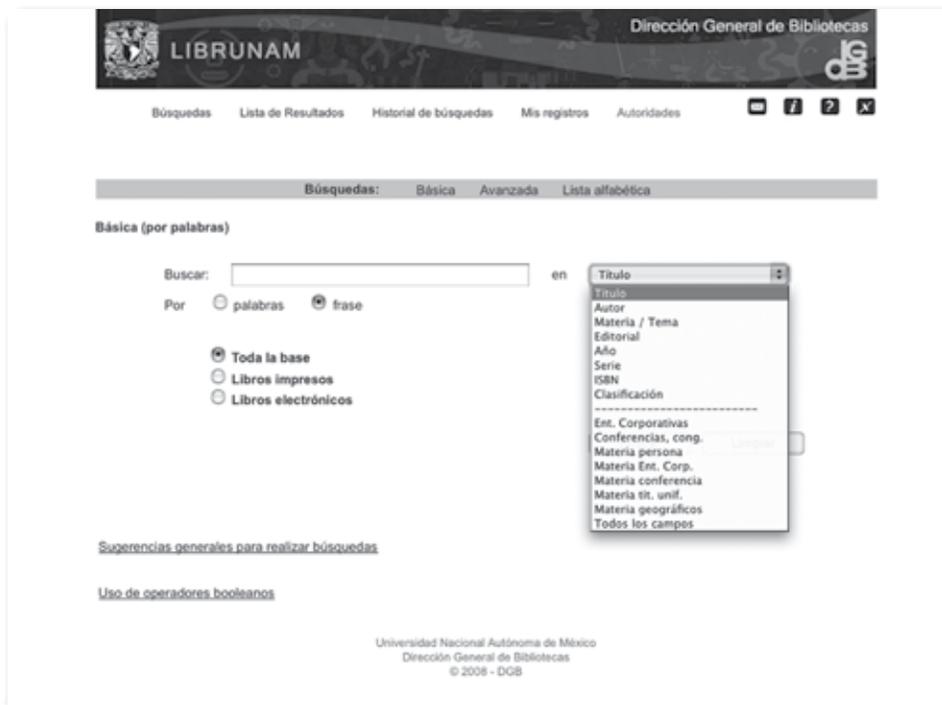
Hoy en día existen en el mercado infinidad de catálogos en línea que las bibliotecas pueden adquirir para brindar el servicio de consulta a sus usuarios. Por men-

cionar los más comerciales, tenemos a ALEPH 500 del proveedor Exlibris, SIABUC desarrollado por la Universidad de Colima e INNOPAC-Millennium del proveedor Innovative Interfaces.

ALEPH 500 v.16

Ofrece una interfaz intuitiva para el usuario y permite la personalización de la base de datos, al igual que otras plataformas. Se caracteriza por utilizar tres métodos de búsqueda: Básica, Avanzada y Lista Alfabética. Se puede recuperar por diferentes campos o índices, como son: Título, Autor, Tema, Todos los campos, etcétera.

Aleph 500 puede configurarse para realizar búsquedas multi-base en bases de datos lógicas locales, es decir, que estén en el mismo servidor; sin embargo, incluir en este esquema bases de datos remotas involucra un costo para poder implementar una arquitectura que realice búsquedas a través del protocolo Z39.50.



SIABUC 8 y SIABUC Siglo XXI

Su mecanismo de consulta por *web* se basa en la utilización de un módulo CGI configurado en un servidor de *web*, que invoca el CGI a través de un formulario con el método “post” mediante una página *web*.³ El catálogo en línea permite realizar búsquedas avanzadas por Título, Autor, Descriptor y búsquedas libres.

³ SIABUC, Universidad de Colima [en línea]. <<http://siabuc.ucol.mx>>



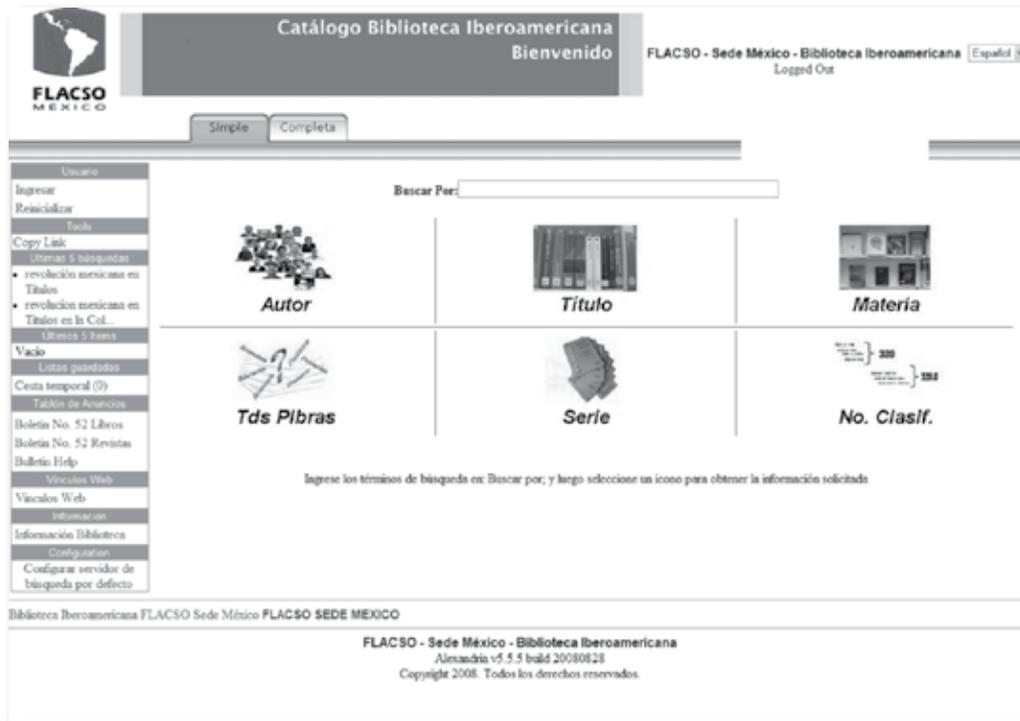
INNOPAC-Millennium

Ofrece acceso público a su base de datos, mediante navegadores *web*, a través de los cuales cualquier usuario puede buscar en el catálogo, así como gestionar sus datos personales, préstamos, renovaciones, etcétera. Se puede recuperar información a través del Título, Autor, Materia, Palabra clave, entre otros campos.



Alexandria

Esta aplicación proporciona al usuario dos métodos de búsqueda: Simple y Completa, recuperando información por Autor, Título, Materia, Todas las palabras, Serie y por número de Clasificación.



Metodología para determinar el modelo y su implementación

En primer lugar, los catálogos se encuentran soportados por diferentes plataformas con características muy distintas, por lo que fue necesario analizar con detalle la organización e indización de cada uno de ellos. En este sentido, se revisaron pormenorizadamente los esquemas de búsqueda simple, avanzada y libre para poder determinar la viabilidad de cada una de ellas.

En esta primera fase se contemplaron las siguientes instituciones del Grupo Amigos:

Institución		Plataforma
CFN	Colegio de la Frontera Norte	Aleph ver. 500.16
COLMEX	Colegio de México	
INSTITUTO MORA	Instituto de Investigaciones "Dr. José María Luis Mora"	
ITAM	Instituto Tecnológico Autónomo de México	
UDLAP	Universidad de las Américas de Puebla	
UNAM	Universidad Nacional Autónoma de México	
UAS	Universidad Anáhuac México Sur	
UIA	Universidad Iberoamericana	
FLACSO	Facultad Latinoamericana de Ciencias Sociales	Alexandria ver. 5.5.3.4
CIDE	Centro de Investigación y Docencia Económicas A.C.	Horizonte ver. 5.3
BUAP	Benemérita Universidad Autónoma de Puebla	Millennium
TEC	Instituto Tecnológico y de Estudios Superiores de Monterrey	

La aplicación se desarrolló en cuatro fases:

1. Análisis detallado de la estructura del URL de búsqueda para cada plataforma, de los catálogos bibliográficos de instituciones que participan dentro del Grupo Amigos.
2. Diseño y creación del esquema de base de datos requerida para la aplicación.
3. Desarrollo del *software*, que construye el URL dinámicamente conforme a la petición del usuario.
4. Diseño de la interfaz.

Funcionalidad

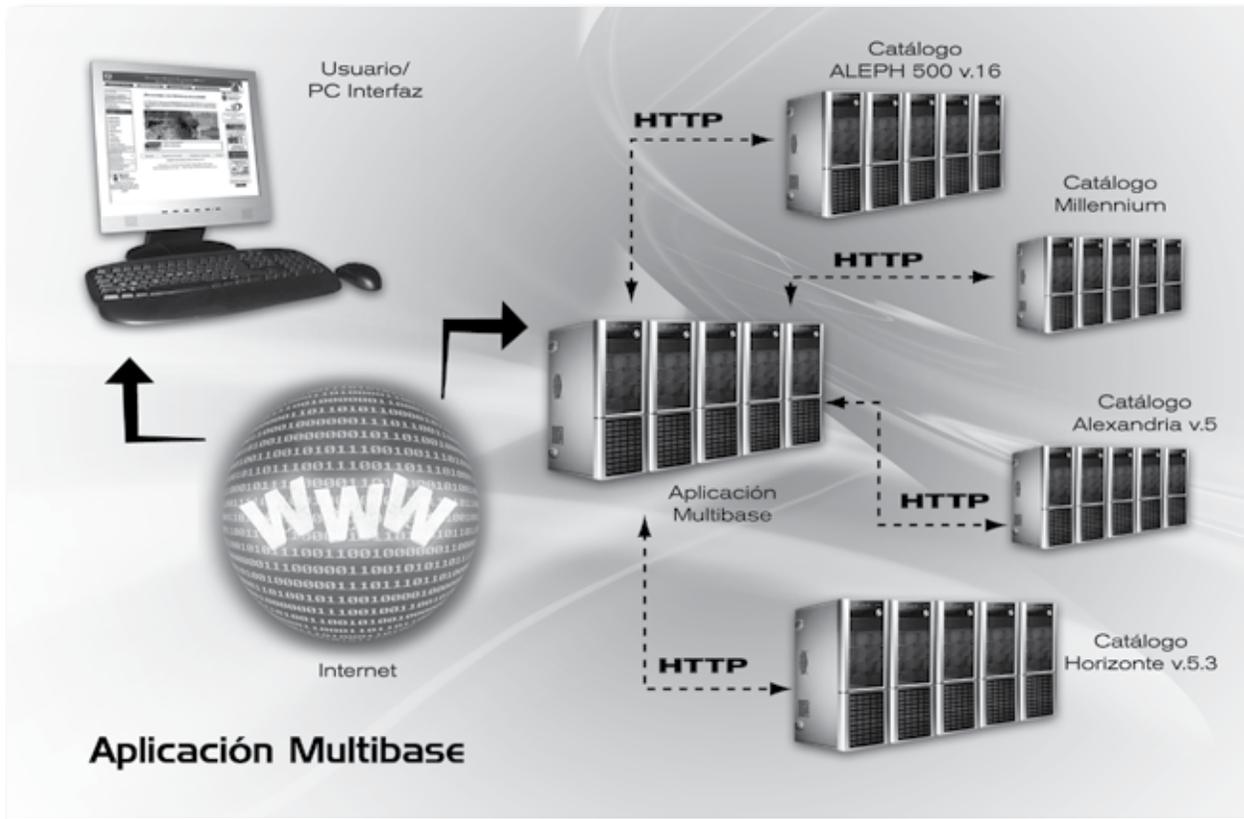
La conceptualización original consideró el desarrollo de una interfaz basada en una búsqueda simple, con el objetivo de facilitar la utilización a los usuarios, evitando la pérdida de tiempo que representa repetir consultas y comparar resultados manualmente. Su diseño se adapta a necesidades de recuperación de información poco exigentes, emplea métodos de búsqueda por un campo determinado, por

ejemplo: palabra, autor, tema y título, de ahí que el diseño sea muy sencillo para que el usuario entienda claramente cómo debe realizar la búsqueda de información.

Para el funcionamiento de esta aplicación no se requiere exportar los registros bibliográficos de cada institución para almacenarlos en un catálogo colectivo o repositorio, de ahí que el usuario vaya directamente a consultar los registros al propio catálogo de la biblioteca.

Otro aspecto importante de esta aplicación es lo referente al tiempo de respuesta de la consulta, ya que depende en gran parte de la disponibilidad y conexión de los servidores; se ha definido en esta aplicación un tiempo máximo de conectividad para cada uno de los catálogos bibliográficos, a fin de mejorar el tiempo de respuesta, es decir, en caso de que alguno de los catálogos no este respondiendo adecuadamente por tráfico en la red, la aplicación continúa buscando en los demás recursos.

La aplicación permite incorporar más plataformas, siempre y cuando la estructura de éstas lo permita, dependiendo en gran parte de las políticas de seguridad que existan en cada una de ellas.



Dentro de las propiedades del protocolo HTTP destacan las siguientes:

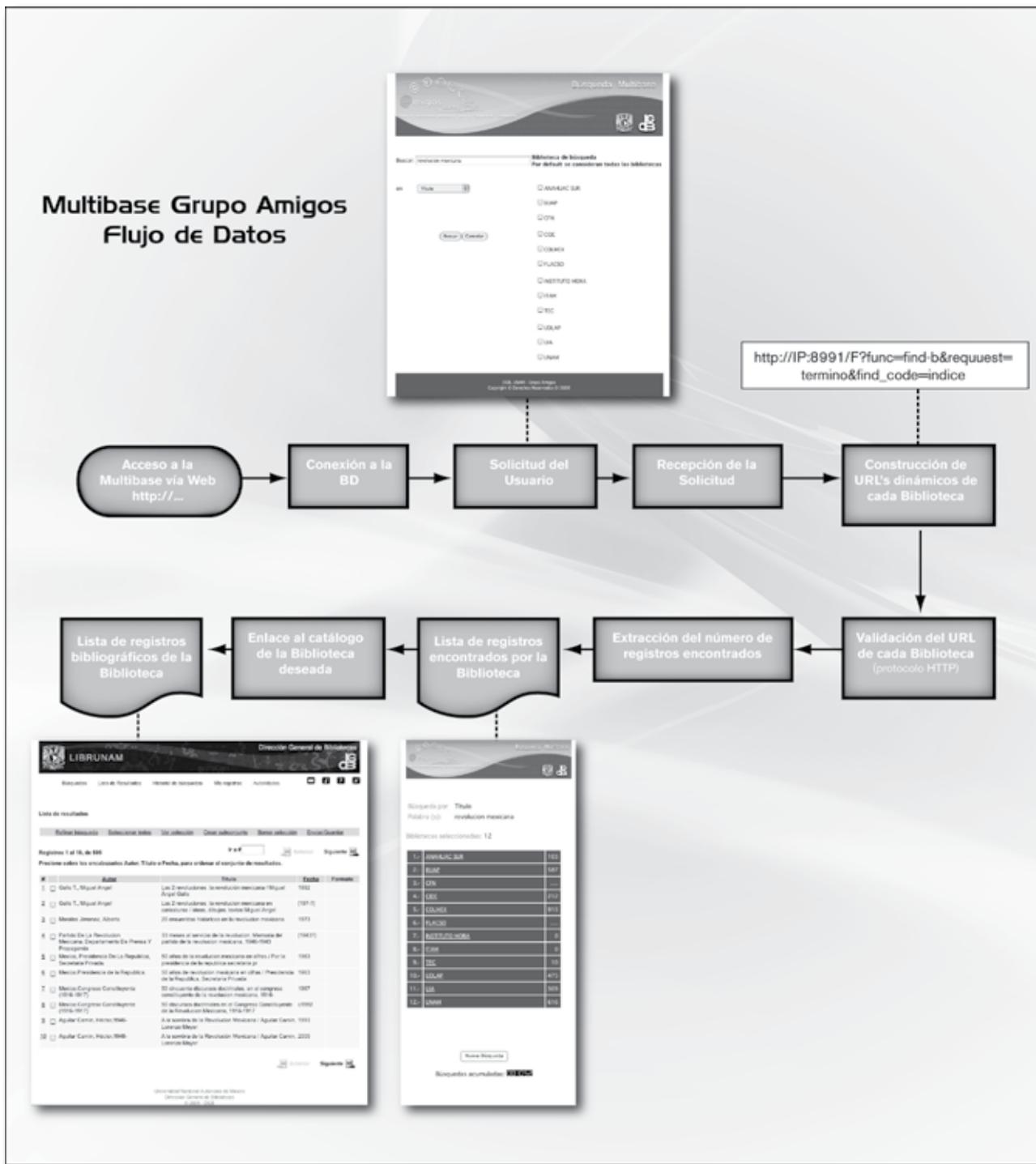
- **Direccionamiento:**
Utiliza el *Uniform Resource Locator* (URL), el cual es un localizador uniforme de recursos que se refiere a la dirección única que identifica una página *web* para acceder a ella.
Estructura: `servicio://servidor.dominio:puerto/ruta/archivo`
- **Arquitectura Cliente-Servidor:**
HTTP funciona bajo el esquema solicitud-respuesta, en donde el cliente abre una conexión TCP/IP⁴ con el servidor y hace su petición, el servidor responde con un mensaje y cierra la conexión.
- **Sin conexión y sin estado:**
Una vez que el servidor responde a la petición del cliente, se rompe la conexión entre ambos.

Diagrama de flujo de datos

Una vez que el usuario se conecta a la aplicación a través del navegador, se hace una conexión a la base de datos para desplegar en pantalla las bibliotecas disponibles desde la *Multibase*, el usuario hace su solicitud o búsqueda seleccionando la(s) biblioteca(s) y la consulta de interés, la aplicación recibe la petición del usuario, construye el enlace en forma dinámica a la biblioteca enviando como parámetros de búsqueda el término a recuperar y el código del índice sobre el que va a buscar; éste varía dependiendo la plataforma de cada biblioteca, por ejemplo para catálogos en ALEPH 500 el código del índice para Título es "WTI", para Millennium es "t" y para Alexandria es "Titles". Con el URL construido la aplicación se conecta a través del protocolo "http" a cada uno de los catálogos (el tiempo máximo para esperar una respuesta es de 50 segundos, cabe señalar que este tiempo puede ser modificado; si en ese lapso de tiempo el catálogo no responde la aplicación trata de conectarse al siguiente catálogo), una vez que se valida el acceso al catálogo

⁴ TCP/IP (Protocolo de Control de Transmisión y Protocolo de Internet)

se extrae el número de registros recuperados y se despliega al usuario una lista alfabética de las bibliotecas y el número de registros encontrados; por último, el usuario selecciona la biblioteca de su preferencia para visualizar los registros bibliográficos recuperados.



Características de la aplicación

- Desarrollo en *software* libre:
 - Perl v5.8.5 (Lenguaje de programación)
 - PostgreSQL v8.0.1 (Manejador de base de datos)
 - Apache (Servidor *web*)
- Plataforma:
 - Workstation SUN BLADE 1500 (Procesador UltraSPARC-III, 1062MHz , 1 GB RAM)
 - Sistema operativo Solaris

Estas características permiten que la aplicación pueda ser fácilmente portable a otra plataforma con los mismos requerimientos de *software*.

Operatividad del prototipo desarrollado

<http://132.248.9.32/cgi-bin/proyecto/multibase.pl>

1. Solicitud del usuario

La interfaz ofrece una búsqueda básica por palabra en los catálogos bibliográficos, recuperando por el Título, Autor, Tema y en Todos los campos; además permite al usuario seleccionar una o más bibliotecas; de no hacerlo, por *default* buscará en todas las bibliotecas disponibles.

Búsqueda Multibase

@migos
Red de Instituciones Mexicanas para la Organización Bibliotecaria

Buscar: Biblioteca de búsqueda
Por default se consideran todas las bibliotecas

en

ANAHLIAC SUR
 BUAP
 CFN
 CIDE
 COLMEX
 FLACSO
 INSTITUTO MORA
 ITAM
 TEC
 UDLAP
 UIA
 UNAM

DGB, UNAM - Grupo Amigos
Copyright © Derechos Reservados © 2008

2. Resultados de la búsqueda

El usuario puede visualizar el número de registros encontrados en cada biblioteca y seleccionar el de su interés.



3. Consulta de registros bibliográficos

Se pueden consultar los registros bibliográficos recuperados y visualizarse desde diferentes plataformas, haciendo el enlace directamente al catálogo de la biblioteca, sin perder de vista los resultados de la búsqueda.

Es importante hacer hincapié que existe la posibilidad de que algunos de los catálogos consultados muestren registros "basura", como producto de una incorrecta indización de origen; este instrumento de búsqueda simultánea no corrige esos errores.



Conclusiones

La tendencia de los OPAC debe apoyarse en el desarrollo de aplicaciones de esta índole que permitan mejorar los servicios proporcionados por las bibliotecas a fin de garantizar la accesibilidad a las fuentes de información que proveen, además de reducir los costos que conlleva implementar una herramienta que recupere, en forma concurrente, información en múltiples bases de datos remotas.

Esta aplicación ya se ha implementado y está disponible a través de la página *web* de la Dirección General de Bibliotecas, realizando búsquedas simultáneas en las bases de datos centrales de Libros (LIBRUNAM), Tesis (TESIUNAM), Revistas (SERIUNAM), Mapas (MAPAMEX), Videos, Artículos de Ciencias Sociales y Humanidades (CLASE), Artículos de Ciencia y Tecnología (PERIODICA) y la Hemeroteca Latinoamericana (HELA).

Acceso: <http://132.248.67.133:8991/cgi-bin/multibasedgb/multibase.pl>

Bases de Datos Dirección General de Bibliotecas

Buscar:

En:

Por: Palabras Frase

Libros [LIBRUNAM]

Tesis [TESIUNAM]

Revistas Impresas [SERIUNAM]

Mapas [MAPAMEX]

Videos

Artículos de Ciencias Sociales y Humanidades [CLASE]

Artículos de Ciencia y Tecnología [PERIODICA]

Hemeroteca Latinoamericana [HELA]

Nota: Por default consulta todas las bases de datos, a menos que seleccione una o varias de ellas.

Universidad Nacional Autónoma de México
Dirección General de Bibliotecas
© 2008- DGB

A nivel del Sistema Bibliotecario de la UNAM también se ha iniciado la configuración de este esquema de búsqueda multibase simultánea, para los recursos locales que posee cada biblioteca, tal es el caso de:

- Biblioteca Campus Juriquilla
- Biblioteca Central
- Biblioteca Conjunta de Ciencias de la Tierra
- Dirección General de Divulgación de la Ciencia
- Facultad de Filosofía y Letras
- Facultad de Medicina Veterinaria y Zootecnia
- Instituto de Investigaciones Filosóficas
- Instituto de Investigaciones Históricas

Las cuales concentran distintas colecciones especializadas que están disponibles a través de sus páginas *web*. Este diseño de interfaz puede ser una alternativa para ser replicado a otros proyectos como ECOES (Espacio Común de Educación Superior), CUMEX (Consortio de Universidades Mexicanas), entre otros.

El desarrollo en *software* libre se ha convertido en una excelente alternativa para la implementación de herramientas complementarias de los Sistemas Integrales de Gestión de Bibliotecas comerciales, por lo que debe impulsarse su uso. ☞

Obras Consultadas

ARIE, Julie, WEAVER, Kent, TENNANT, Roy. *A Checklist of Considerations for Selecting Metasearch Software* [en línea]. <http://www.cdlib.org/inside/projects/metasearch/metasearch_checklist.pdf> [Consulta: 28 agosto 2008].

HILDRETH, Charles R. "Beyond Boolean [en línea]: designing the next generation of online catalogs". *Library Trends*, Spring 1987, 35, p. 647-667. <http://www.ideals.uiuc.edu/dspace/bitstream/2142/7500/1/librarytrendsv35i4k_opt.pdf> [Consulta: 6 octubre 2008].

Innovative Interfaces. [en línea]. <<http://www.iii.com/>> [Consulta: 6 octubre 2008].

MARGAIX-ARNAL, Dídac. *Acta del VI Workshop CALSI 2007* [en línea]: *el opac 2.0: las tecnologías de la Web 2.0 aplicadas a los catálogos bibliográficos*. <http://www.calsi.org/2007/wp-content/uploads/2007/11/didac_margaix.pdf> [Consulta: 10 octubre 2008].

NISO. *Metasearch Initiative October 2004* [en línea]. <http://www.niso.org/workrooms/mi/MI-Content_Provider_Survey.pdf> [Consulta: 22 agosto 2008].

SIABUC, Universidad de Colima [en línea]. <<http://siabuc.ucol.mx>> [Consulta: 10 octubre 2008]

TAMAR, Sadeh. "The Challenge of Metasearching [en línea]". *New Library World*, 2004, vol. 105, no. 1198/1199, p. 104-112. <http://www.exlibrisgroup.com/files/Publications/The_Challenge_of_Metasearching.pdf> [Consulta: 21 agosto 2008].

