

**Reto 6 – Diseño e Implementación de una Biblioteca Digital Federada (Yet Another Federation of Digital Libraries – YAFDL) mediante Web Services y HTTP.**

**Fundamentación en Integración o Federación de bibliotecas digitales:**

Una Federación de Bibliotecas Digitales (FBiDi) es la integración de varias Bibliotecas Digitales (BiDi) que permite de manera unificada ofrecer servicios de búsqueda y recuperación de información. A nivel de arquitectura, existen dos (2) modelos de integración de BiDi: a) una mediante recolección y b) por búsquedas en línea. En ambos casos se cuenta con uno o varios servidores de integración, quien recibe las consultas de los usuarios. Este servidor realiza la búsqueda de acuerdo a uno o ambos modelos de integración.

- a) El modelo de recolección consiste en la recolección de los metadatos en un servidor, de tal manera que las consultas se realizan en un solo servidor, ya que todos los metadatos están ubicados de manera centralizada. Para lograr esto, el servidor de la federación se conecta de manera periodica a cada una de las bibliotecas digitales remotas para descargar de manera total o incremental los metadatos de los servidores de bibliotecas digitales. Se pueden utilizar varios protocolos de recolección como HTTP, OAI-PMH (Open Archive Initiative Protocol for Metadata Harvesting) o mediante Web Services. En este modelo se realizan las consultas federadas de manera centralizada, ya que se concentran todos los metadatos de todas las bibliotecas digitales en un solo servidor, sin embargo la descarga de datos (recursos digitales) se realiza de manera directa desde cada una de las bibliotecas digitales sin pasar por el servidor central. Este modelo tambien se conoce como búsquedas centralizadas.
- b) El modelo de búsquedas en línea implica que cada vez que un usuario envia una consulta al servidor principal de la federación, dicho servidor reemite la misma consulta a cada una de las bibliotecas digitales remotas y espera la respuesta de cada una de ella para consolidarlas y presentarlas de manera unificada al usuario de la federación. Existen varios protocolos para implementar esta funcionalidad, tales como: Z39.50, SRU o SRW. Al igual que en el anterior modelo, las descargas de datos (documentos digitales), se realizan directamente desde cada BiDi. Este modelo tambien se conoce como búsquedas distribuidas.

### Descripción del problema:

Realizar la integración de bibliotecas digitales basadas en YADL (Ver la definición en el “Reto-05-YADL-WebXml”) y realizar la implementación de uno de los 2 modelos de integración de tal forma que permita a un usuario de la federación realizar búsquedas en todas las bibliotecas digitales. Cada YADL seguirá funcionando de modo autónomo para la gestión de los objetos digitales así como el módulo de búsqueda dentro de YADL.

El servidor principal de la federación, será conocido como YAFDL Server, y este leerá un archivo xml en donde contiene la información necesaria para localizar cada biblioteca digital remota (YADL) y su respectivo protocolo. La estructura de dicho archivo es la siguiente:

```
“yafdl.xml”

<?xml version="1.0" encoding="UTF-8"?>
<yafdl>
<yadl>
    <name>EAFIT</name>
    <protocol>OAI</protocol>
    <url>http://servidor-oai/yadl_oai_wss</url>
</yadl>
<yadl>
    <name>UDEM</name>
    <protocol>SRW</protocol>
    <url>http://servidor-srw:8080/yadl_srw_wss</url>
</yadl>
.
<!-- otros proveedores -->
.
</yafdl>
```

Para la implementación del modelo de integración a), utilizar la siguiente definición de servicio web:

```
[WebMethod]
public string getRecords();
```

El string de respuesta viene el documento yadl.xml del servidor de la biblioteca digital. **A continuación se muestra una respuesta típica de esta llamada:**

```

<?xml version="1.0" encoding="UTF-8"?>
<yadl>
  <record>
    <title>Java</title>
    <creator>autor1</creator>
    <creator>autor2</creator>
    <subject>programacion</subject>
    <subject>sistemas</subject>
    <subject>lenguajes</subject>
    <identifier>doi:1</identifier>
    <url>http://servidor/yadl/data/doc1.pdf</url>
  </record>
  <record>
    <title>C#</title>
    <creator>autor2</creator>
    <creator>autor3</creator>
    <subject>programacion</subject>
    <subject>microsoft</subject>
    <identifier>doi:2</identifier>
    <url>http://servidor/yadl/data/doc2.pdf</url>
  </record>
  <record>
    <title>Windows 2003</title>
    <creator>autor4</creator>
    <subject>sistemas operativos</subject>
    <identifier>doi:3</identifier>
    <url>http://servidor/yadl/data/doc3.pdf</url>
  </record>
</yadl>

```

En este modelo debe realizar las siguientes consideraciones:

- Cada cuanto se va a conectar a cada uno de las Bibliotecas Digitales remotas? Como se ejecuta este proceso? Recuerde que en este modelo hay 2 componentes en el servidor YAFDL: 1) el componente que realiza las búsquedas para los usuarios. 2) el componente que realiza la recolección, para este componente realizar la respectiva aplicación Web para ejecutar dicho proceso.
- Como realiza la consolidación de los datos xml de las BiDi.

Para la implementación del modelo de integración b), utilizar una versión simplificada de SRW (Search/Retrieve Web Services), consistente en la siguiente definición del servicio web:

```

[WebMethod]
public string query(string field, string term);

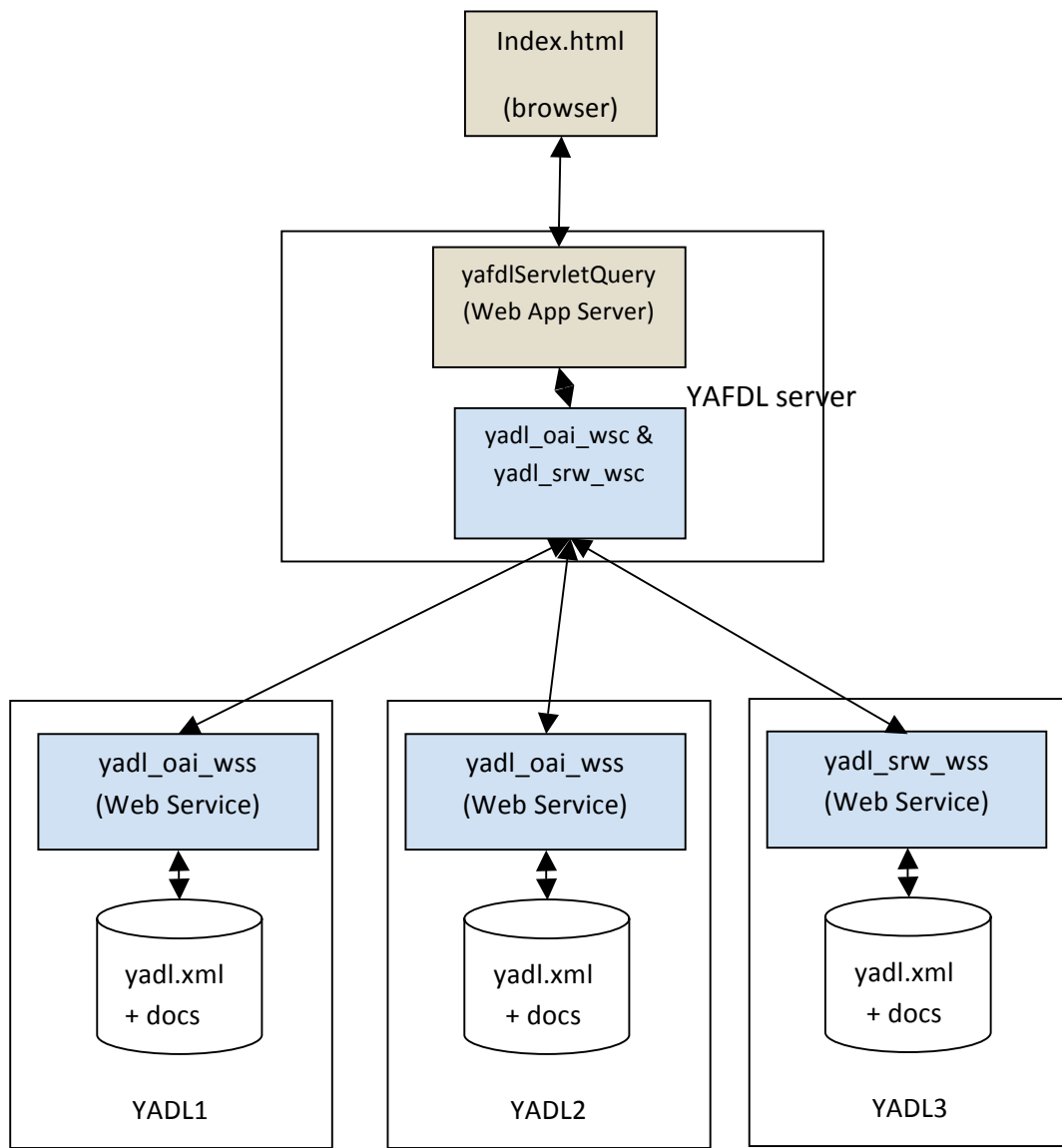
```

El string de respuesta contiene todos los registros que cumplen con el criterio de búsqueda dado por *field* y *term*. Esta respuesta debe ser un documento XML bien formado, conteniendo los registros que cumplieron con la búsqueda, a continuación se presenta un ejemplo de la respuesta típica de esta llamada (field = "creator", term="autor2"):

```
<?xml version="1.0" encoding="UTF-8"?>
<yadl>
  <record>
    <title>Java</title>
    <creator>autor1</creator>
    <creator>autor2</creator>
    <subject>programacion</subject>
    <subject>sistemas</subject>
    <subject>lenguajes</subject>
    <identifier>doi:1</identifier>
    <url>http://servidor/yadl/data/doc1.pdf</url>
  </record>
  <record>
    <title>C#</title>
    <creator>autor2</creator>
    <creator>autor3</creator>
    <subject>programacion</subject>
    <subject>microsoft</subject>
    <identifier>doi:2</identifier>
    <url>http://servidor/yadl/data/doc2.pdf</url>
  </record>
</yadl>
```

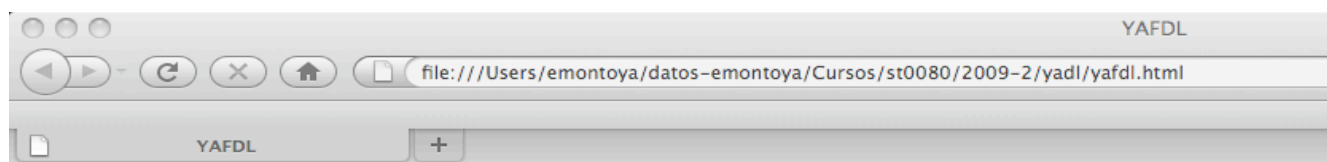
En field puede venir: "all", "title", "subject", "creator"

En term viene el termino de búsqueda. Ej: "java", "comp", "computador", etc. Puede contener un substring de cualquier campo.



### Especificación del servicio hacia los usuarios:

La página de acceso al usuario para la consulta, permitirá especificar los siguientes campos de entrada al momento de especificar una consulta:



# Yet Another Federation of Digital Libraries-YAFDL

## Query Documents

Field: ☒ All ☐ Title ☐ Creator ☐ Subject ☐ identifier

Query:

Digital Library:

☒ All

☐ EAFIT

☐ UDEA

☐ CES

☐ UDEM

[home](#)



# Yet Another Federation of Digital Libraries-YAFDL

## Query Results:

Repository (DL)	Title	Autor(s)	Subject(s)	Identifier	Document
EAFIT	Java	autor1;autor2	programacion;sistemas;lenguajes	doi:1	<a href="http://dl.eafit.edu.co/yadl/data/doc1.pdf">http://dl.eafit.edu.co/yadl/data/doc1.pdf</a>
UDEM	C#	autor3;autor2	programacion;microsoft	doi:2	<a href="http://dl.udem.edu.co/yadl/data/doc2.pdf">http://dl.udem.edu.co/yadl/data/doc2.pdf</a>
UDEM	Windows 2003	autor4	sistemas operativos	doi:3	<a href="http://dl.udem.edu.co/yadl/data/doc3.pdf">http://dl.udem.edu.co/yadl/data/doc3.pdf</a>

[home](#)

NOTA: PUEDEN TRABAJAR GRUPOS DE TRES (3) ESTUDIANTES REALIZANDO LA IMPLEMENTACIÓN DE LOS DOS (2) MÉTODOS DE INTEGRACIÓN DE MANERA SIMULTANEA. ESTO ES, EL SERVIDOR YAFDL AL MOMENTO DE EJECUTAR CONSULTAS, LAS REALIZARÁ LOCALMENTE POR RECOLECCIÓN Y DE MANERA DISTRIBUIDA POR SRW, PRESENTANDO AL USUARIO DATOS CONSOLIDADOS. RESTRICCIÓN: UNA YADL SOLO PUEDE ESTAR REGISTRADO UNA SOLA VEZ EN YAFDL.XML, YA QUE EN CASO CONTRARIO, LOS REGISTROS ESTARÍAN DUPLICADOS.