

Capítulo 2. Colecciones Digitales Federadas

El intercambio de información entre federaciones de bibliotecas digitales existentes se vuelve cada vez más necesario. De esta forma es importante contar con mejores servicios de búsqueda y recuperación de información en sistemas heterogéneos. También deben existir ambientes de usuario que faciliten el acceso abierto a las colecciones que cuentan con ciertos estándares como los establecidos en la Iniciativa de Archivos Abiertos (OAI).

Dentro de las colecciones digitales podemos encontrar diversos tipos de documentos en varios formatos. Cada biblioteca digital provee diferentes servicios y herramientas para realizar búsquedas, recuperación y almacenamiento en los acervos digitales. Para conocer los servicios proporcionados por las bibliotecas digitales participantes en este proyecto, y su funcionamiento, se proporcionará una breve descripción de sus herramientas, así como de su arquitectura.

2.1 U-DL-A

El objetivo principal es apoyar a la investigación y proporcionar servicios a la comunidad universitaria mediante el desarrollo de ambientes para facilitar la colaboración entre usuarios respondiendo a las preferencias personales de cada uno. La investigación realizada dentro de U-DL-A está basada en una arquitectura de biblioteca digital e interfaces de usuario, con el propósito de manejar grandes volúmenes de información. La Figura 2.1 ilustra los componentes básicos de dicha arquitectura.

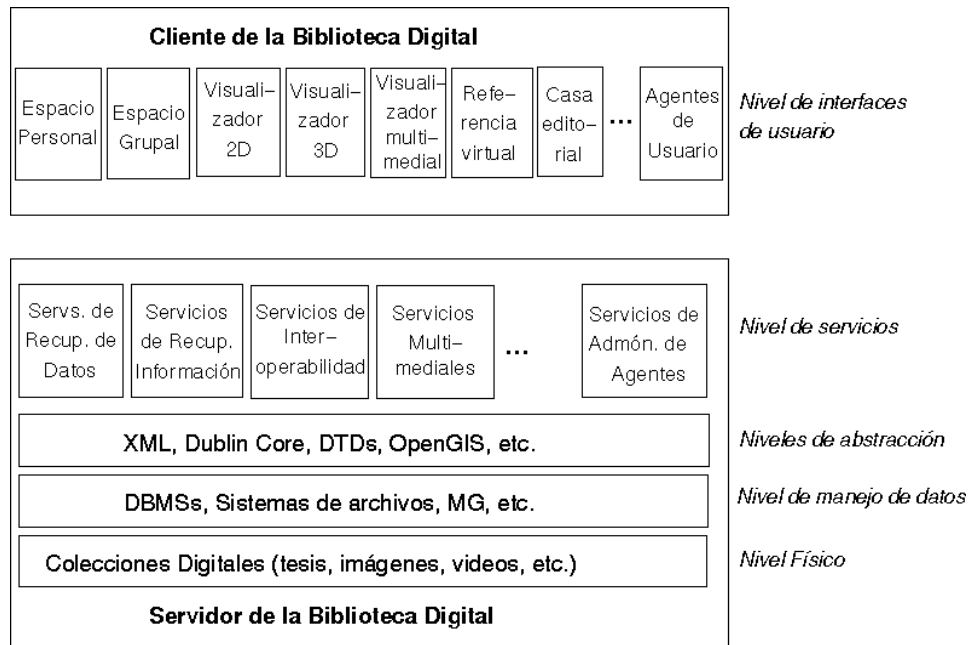


Figura 2.1 Arquitectura de U-DL-A [Sánchez 2003; Sánchez & Leggett 1997]

Entre los acervos digitales desarrollados como parte de la iniciativa U-DL-A se encuentra la colección de tesis digitales, el archivo presidencial de Porfirio Díaz, y la colección de publicaciones universitarias. El proyecto de tesis digitales define lineamientos y mecanismos para publicar las tesis de estudiantes en formato digital. En cuanto a la correspondencia presidencial de Porfirio Díaz se pretende digitalizar y decodificar los cerca de 70,000 telegramas y manuscritos que se encuentran en esta colección, para hacerlos públicos a una gran comunidad de usuarios. En lo referente a las publicaciones universitarias se pretende generar una colección que integre tanto publicaciones actuales como anteriores.

Como parte de los servicios alternos dentro de U-DL-A podemos mencionar:

Gracias al modelo cliente-servidor es posible tener acceso a los recursos desde cualquier lugar, los cuales se encuentran almacenados en el servidor de MiBiblio, sin necesidad de agregarlos cada vez que se entre a esta interfaz personalizable.

Espacios personales

Espacios grupales

Servicios generales de recuperación de información y de comunicación

UVA: Visualización de colecciones en 3D

Poseidón: Composición y anotaciones sobre tesis digitales

Syrex: Servicios de recomendación para usuarios de colecciones digitales

Viajerus y MAIDL: Agentes móviles para recuperación personalizada de información distribuida

Mecanismos de autofinanciamiento para bibliotecas digitales (comercio electrónico).

VRef: Servicio de Referencia Virtual.

RDU: Reserva Digital Universitaria

2.2 MARIAN y ENVISION

Virginia-Tech desarrolló el sistema MARIAN, el cual puede definirse como una infraestructura de bibliotecas digitales, que soporta varias sesiones generadas en un sistema bibliotecario, es capaz de realizar secuencias breves de queries no relacionados. Podemos decir que MARIAN se describe como un sistema de multitareas distribuido. Cada módulo es funcionalmente definido por su relación con una tarea específica. Este sistema está constituido por cuatro módulos: Administración de sesión, interacción del usuario, búsquedas y administración de base de datos y archivos [France et al 1999].

Algunos puntos importantes de MARIAN son los siguientes: El uso de distintas técnicas de recolección de datos, arquitectura para construir un medio para unir colecciones digitales, medios para visualización de la información, compresión de información e indexamiento de documentos para realizar búsquedas flexibles [André et al 2000].

Cabe mencionar que ENVISION es un sistema que sirve de soporte para la visualización de resultados producidos por MARIAN, por medio de este se pueden consultar grandes volúmenes de información de una manera eficaz. [André et al 2000 citado en Rebollar 2002].

2.2.1 Ambiente

Este es un sistema multitareas. MARIAN está construido en un conjunto de módulos que se comunican y corren en una o más computadoras por medio de MACH UNIX. Cada módulo de MARIAN es definido funcionalmente por su relación a una cierta tarea o cuerpo de datos. Cada módulo presenta una interfaz bien definida para los métodos del resto del sistema. Estos métodos son implementados como llamadas a procedimientos remotos usando *MACH interface Generator* (MiG). [France et al 1999].

2.2.2 Arquitectura

MARIAN está compuesto por cuatro capas: capa de administración de sesiones, interacción con el usuario, capa de búsqueda, capa referente a la base de datos y a los archivos. La forma piramidal que presenta el sistema en la Figura 2.2 se refiere, a que las capas superiores son más pequeñas que las de abajo, esto en términos de módulos funcionales, y en tiempo de ejecución entre cada capa.

La capa de administración de las sesiones está compuesta de un sólo módulo encargado de administrar las sesiones de los usuarios, así como la conexión y cierre de las sesiones que han estado inactivas durante mucho tiempo. También da seguimiento a las peticiones de las preguntas. Las preferencias de los usuarios son almacenadas y modificadas por el módulo encargado de administrar las sesiones, y posteriormente se envían a otros módulos que las necesiten.

El sistema MARIAN, como otros sistemas de información, puede ser dividido en dos partes: cliente, y servidor. En un principio MARIAN fue diseñado para soportar solo sesiones uni-usuario, sin embargo, ahora muchos usuarios interactúan a través del Web. El Web actúa como un servidor para los usuarios, y como un cliente para MARIAN. Así pues MARIAN puede soportar varios clientes de manera simultánea. [France et al 1999].

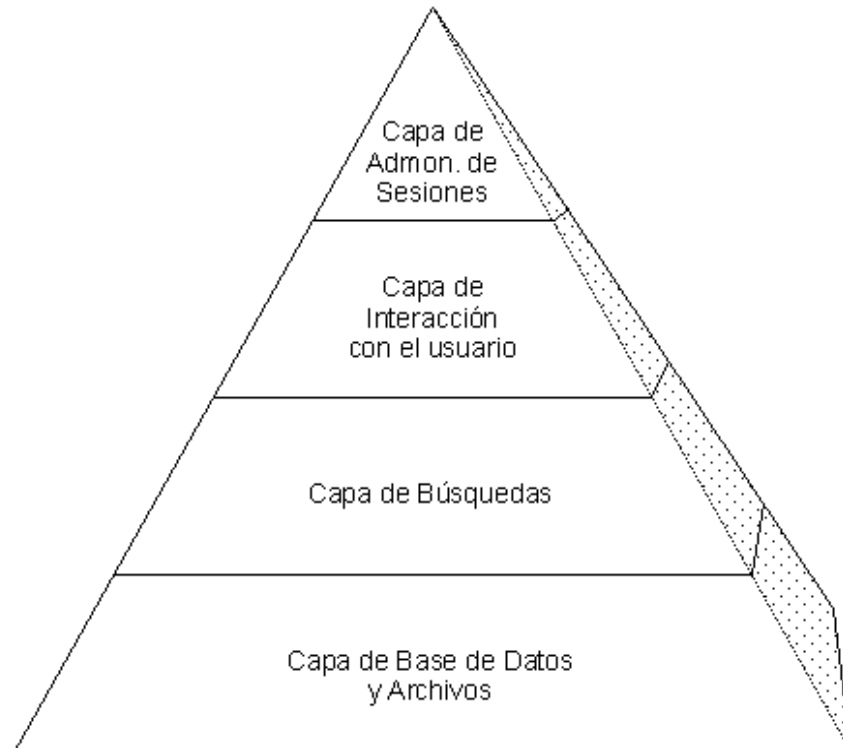


Figura 2.2 Arquitectura del Sistema MARIAN (Adaptada de Gonçalves et al [2001])

La interacción entre los clientes de MARIAN y el servidor consiste en la búsqueda de resultados que corresponden a un query hecho. La búsqueda en MARIAN comienza con un mensaje del cliente especificando el tipo de búsqueda. [France et al 1999].

2.3 Phronesis

Por un lado, se encuentra el desarrollo hecho por el ITESM Campus-Monterrey, denominado Phronesis. Este proyecto está orientado a desarrollar e investigar la tecnología de bibliotecas digitales, con la finalidad concreta de obtener una herramienta que permita la fácil creación de colecciones, así como el fácil acceso a documentos digitales almacenados en ellas y, a su vez, es un medio para la investigación y el desarrollo. La colecciones digitales utilizando Phronesis brindan la oportunidad a cualquier persona de realizar búsquedas no sólo mediante el título o palabras claves del documento, sino también con base en el contenido del documento y sus metadatos (información acerca del documento como autor o tema) [Garza et al 1999].

Los objetivos del Sistema Phronesis son los siguientes:

Se presenta como un solo sistema para la integración de las funcionalidades dentro de una biblioteca digital, así como para la administración y la creación de colecciones digitales.

No se restringe a un dominio específico. Es posible crear colecciones de diversos temas.

Phronesis se ha construido utilizando componentes de software de dominio público y está disponible para cualquier usuario.

Usando Phronesis, el tamaño de una colección digital siempre es más pequeño que el tamaño original de los documentos que son parte de la colección. El tiempo requerido para construir una colección (la indexación, y la compresión de los documentos) es eficiente [Garza et al 1999].

Phronesis contempla los siguientes aspectos:

Indexamiento y búsquedas en texto completo o en los metadatos.

Control en el acceso y flujo de usuarios que pueden proporcionar material a la biblioteca.

La interfaz está basada en Web, y se presenta en español, y en inglés.

Se pueden realizar búsquedas simultáneas en las instancias de Phronesis.

Soporta búsquedas de documentos en varios formatos.

Se puede almacenar cualquier tipo de documento, así como hacer búsquedas en base a sus metadatos.

Existe una compresión eficiente de los documentos digitales

Phronesis utiliza el sistema MG (managing gigabytes), el cual se utiliza para manejar los documentos existentes dentro de la biblioteca digital [Garza et al 1999].

2.4 Iniciativa de Archivos Abiertos

Dentro del contexto de las colecciones digitales federadas, es importante mencionar la Iniciativa de Archivos Abiertos (OAI, por sus siglas en inglés), integrada por un grupo de universidades, archivos y otros participantes. Esta iniciativa comprende acervos digitales de distintas disciplinas, así como diferentes tipos de publicaciones. El objetivo principal de OAI, es abarcar todo el material digital generado por universidades y centros de investigación quienes promueven el acceso libre a la información. La meta del protocolo de la Iniciativa de Archivos Abiertos para la recopilación de metadatos, es promover la aplicación de una estructura robusta que tenga como principal característica la interoperabilidad. Esto contribuirá a que la información divulgada por un grupo de usuarios sea consultada de una manera más eficiente. [OAI 2001].

La principal meta de la Iniciativa de archivos abiertos para la recolección de metadatos, es promover una aplicación independiente, pero sobre todo que sea interoperable. Este protocolo contempla 2 participantes:

Proveedor de Datos: está encargado de administrar el sistema que da soporte a lo establecido en el protocolo de archivos abiertos, esto es, dar a conocer los metadatos que se refieren al contenido en sus sistemas.

Proveedor de Servicios: tiene como tarea enviar las peticiones que genera el protocolo de archivos abiertos a los sistemas encargados de proporcionar los datos [OAI 2001].

Finalmente podemos decir que la Iniciativa de Archivos Abiertos utiliza para realizar el intercambio de información en los diferentes tipos de archivos, el formato Dublin Core, el cual puede describirse como un conjunto de metadatos, los cuales permiten la interoperabilidad entre los acervos digitales. La existencia de estos metadatos proporciona una gran ventaja, pues es posible definir otras estructuras para crear nuevos metadatos, si algún proveedor de datos lo cree necesario.

2.5 Resumen del capítulo

La construcción del nuevo ambiente personalizable permite la integración de servicios ya existentes como: MAIDL, Sophia, Syrex, Poseidon, Reserva Digital. También es posible agregar información digital de las colecciones U-DL-A, MARIAN y Phronesis. De esta forma se cuenta con un espacio personal uniforme aparte de que pueden integrarse otras colecciones digitales que cumplan con los estándares establecidos en OAI. Así como cualquier tipo de aplicación la cual permita el acceso a la información digital siempre y cuando esté se haga a través de documentos XML.