

Proyecto: Agentes y Movilidad en Colecciones Multimediales Heterogéneas

**Reporte de Avance Técnico
Enero-Junio 2004.**

Instituciones participantes:

Universidad de las Américas, Puebla
Instituto Tecnológico y de Estudios Superiores de Monterrey, Campus Monterrey.

Investigadores responsables:

Dr. Alfredo Sánchez (UDLA)
Dr. David Garza (ITESM)

Resumen

De acuerdo a lo programado, el proyecto de colaboración UDLA-ITESM presenta avances en las siguientes áreas: (1) diseño de una arquitectura para integración de servicios, (2) diseño e implementación prototípica de servidores OAI 2.0, (3) Diseño de agentes ontológicos, (4) Generación de interfaces para dispositivos móviles (5) Difusión del proyecto. Se prevé la conclusión exitosa del proyecto en el tiempo programado originalmente.

Avances del proyecto durante el período Enero-Junio, 2004.

1. Comunicación de grupos remotos

Para facilitar la comunicación entre los grupos de trabajo en la UDLA e ITESM-Monterrey se decidió sostener por lo menos una reunión mensual por medio de videoconferencia en Internet 2. Asimismo, se estableció un foro compartido en línea para la discusión de ideas y almacenamiento de documentos de interés (<http://edison.mty.itesm.mx/cudi>).

2. Arquitectura general del proyecto

Cada una de las instituciones participantes ha realizado anteriormente desarrollos relevantes en las áreas de interés del proyecto. Para optimizar el aprovechamiento de esfuerzos previos se planteó la necesidad de integrarlos en el marco de una arquitectura compartida que permita el acceso transparente a los documentos de bibliotecas digitales personales y la interoperabilidad con colecciones Phronesis y con la biblioteca digital UDLA, desde diferentes plataformas incluyendo dispositivos móviles.

El diseño comenzó con el planteamiento de futuros escenarios de uso [1]. A partir de estos escenarios, se formularon requerimientos de software. Finalmente, tomando como base a estos requerimientos se creó el detalle de los subsistemas que conformarán la totalidad del sistema [2].

Arquitectura preliminar del sistema

Como se ilustra en la Figura 1, la arquitectura se compone de subsistemas clientes y servidores. Los subsistemas cliente son clasificados de acuerdo a su movilidad y a su arquitectura en:

1. Clientes móviles y ligeros. Estos clientes tienen bajo poder de procesamiento (por ejemplo, celulares), para que una aplicación funcione requiere en todo tiempo conexión a internet y protocolos de información alternos (WAP, WML) que se adapten a la interfaz del dispositivo.
2. Clientes móviles y pesados. Estos clientes tienen mayor poder de procesamiento, sin embargo, requieren soporte a la conexión y una eficiente utilización de recursos tales como energía y memoria.
3. Clientes fijos y ligeros. Estos clientes tienen una interfaz web y a ellos pueden tener acceso clientes fijos (por ejemplo, computadoras de escritorio).
4. Clientes fijos y pesados. Los clientes fijos y pesados son libres de limitaciones tecnológicas de los clientes fijos.

La presencia de clientes pesados en los cuales tanto código ejecutable como información reside en el dispositivo cliente es una característica que distingue a una aplicación móvil de una aplicación Web.

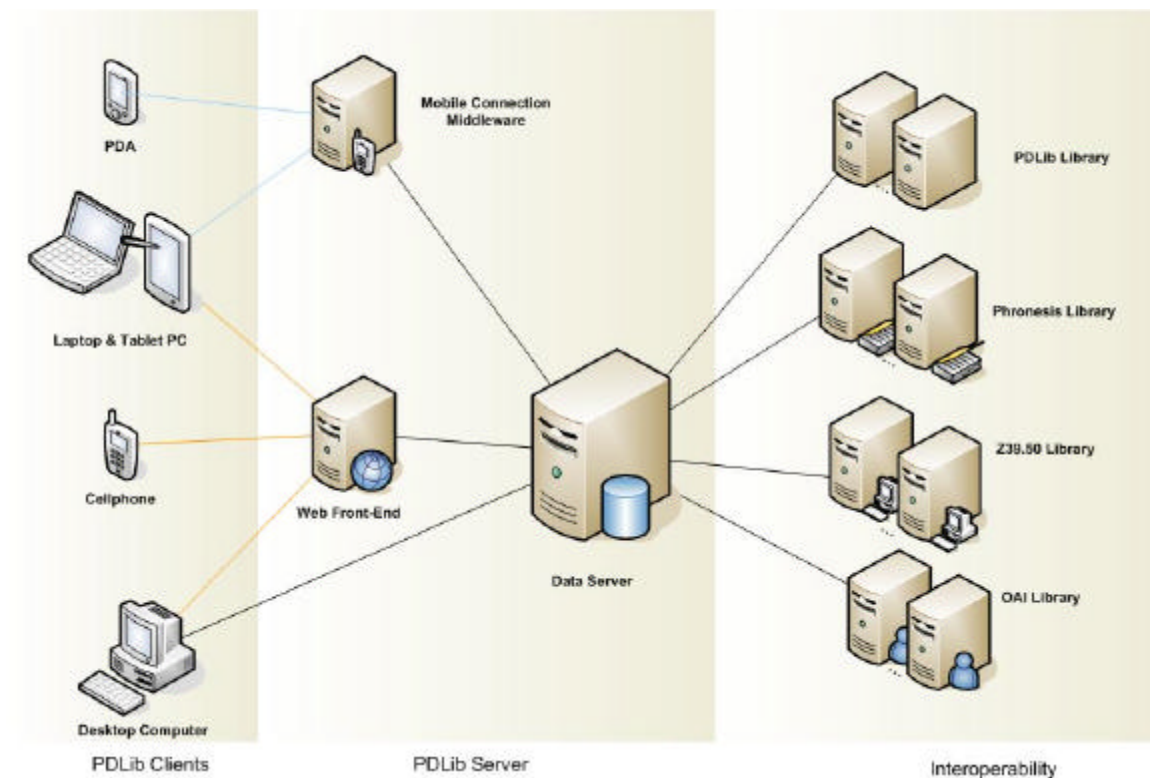


Figura 1. Arquitectura preliminar

Por otro lado, los subsistemas del servidor se clasifican en 3 grandes bloques funcionales:

1. El servidor de datos proporciona la mayor parte de la funcionalidad de un repositorio digital con los servicios de almacenamiento y recuperación de información así como la capacidad de comunicarse con otros servidores de datos. Además, esta entidad de software tendrá la posibilidad de comunicarse con repositorios Phronesis, OAI, y Z39.50.
2. El Mediador de Conexiones Móviles (MCM) es responsable de proveer soporte para los clientes pesados y ofrece servicios importantes como soporte a la conectividad, administración de los datos móviles y soporte para la interfaz con clientes pesados entre otros.
3. El servidor de acceso a Web (“Web Front-End”) provee acceso a las colecciones digitales multimediales.

Nueva arquitectura de Phronesis

Uno de los desarrollos previos que se aprovechará en el proyecto es Phronesis. Para ello ha sido necesario diseñar una nueva arquitectura de cuatro capas para este componente. Con estas capas (presentación, lógica de negocio, acceso a los servicios, servicios), cada servicio puede crecer, corregirse o mejorarse de forma independiente. Por medio de los servicios pueden almacenarse documentos, realizar búsquedas, administrar usuarios y recuperar documentos. Los servicios interactúan entre sí para crear bibliotecas digitales distribuidas [10].

Otra de las características de la arquitectura por capas es que se ha definido una capa especial llamada PROXY que permitirá la interoperabilidad de Phronesis con otras bibliotecas usando diferentes protocolos de comunicación como Z39.50 y OAI-PMH, necesarios para realizar la interoperabilidad de Phronesis con colecciones de U-DL-A y otras.

Dado que el ambiente móvil considerado en el proyecto incluye usuarios Web y el diseño del sistema contempla mecanismos para compartir información (incluso al nivel de una biblioteca completa) entre usuarios, se han encontrado áreas de oportunidad en las cuales los proyectos se beneficiarán mutuamente. En particular, se espera utilizar al servidor de datos de PDLib para implementar los servicios de búsqueda y recuperación del sistema Phronesis a partir de la versión 2.0.

3. Prototipo de la infraestructura

Se ha desarrollado un prototipo para la infraestructura del sistema, el cual comprende los siguientes elementos:

1. **Dos aplicaciones cliente:** El navegador y el intercambiador PDLib (PDLib Browser y PDLib Trader, respectivamente). El navegador es una aplicación que permite a un usuario consultar el contenido de su biblioteca digital personal, así como descargar documentos de la misma al dispositivo móvil en donde reside la aplicación cliente para su consulta en situaciones de desconexión. El intercambiador le permite al usuario tanto administrar su biblioteca personal (mover documentos, conceder acceso a otros usuarios, etc.) como consultar el contenido de bibliotecas personales de terceros, bajo una política de control de acceso. Se ha creado una implementación tanto del navegador como del intercambiador para clientes móviles y pesados sobre el perfil MIDP de la plataforma J2ME [3]. Las aplicaciones cliente se comunican con el MCM por medio de la implementación del protocolo XML-RPC del servidor de aplicaciones Java/XML Enhydra [4].
2. **El servidor de datos (PDLib DataServer).** El servidor de datos del PDLib proporciona servicios de almacenamiento y recuperación del contenido de las bibliotecas digitales personales. Se ha creado una implementación del servidor de datos en la plataforma J2SE [5]. Para el almacenamiento físico de los documentos y las colecciones de las bibliotecas personales, se emplea el manejador de base de datos MySQL [6]. La implementación actual utiliza como motor de búsqueda de texto completo a los índices del manejador de base de datos MySQL pero se espera reemplazar a este motor de búsqueda con el motor de búsqueda Lucene del proyecto Jakarta del grupo Apache [7].
3. **El MCM (Mobile Connection Middleware).** Se cuenta con una implementación del MCM realizada sobre la plataforma J2SE [5]. Actualmente, los servicios del MCM son accedidos por medio del protocolo XML-RPC implementado mediante las librerías XML-RPC del proyecto <Web Services /> del grupo Apache [8].

4. **El umbral Web (PDLib Web Front-end).** Se cuenta con una versión preliminar de la implementación del navegador y del intercambiador como aplicaciones Web. Esta implementación se ha realizado sobre el contenedor de Servlets Tomcat del grupo Apache [9].

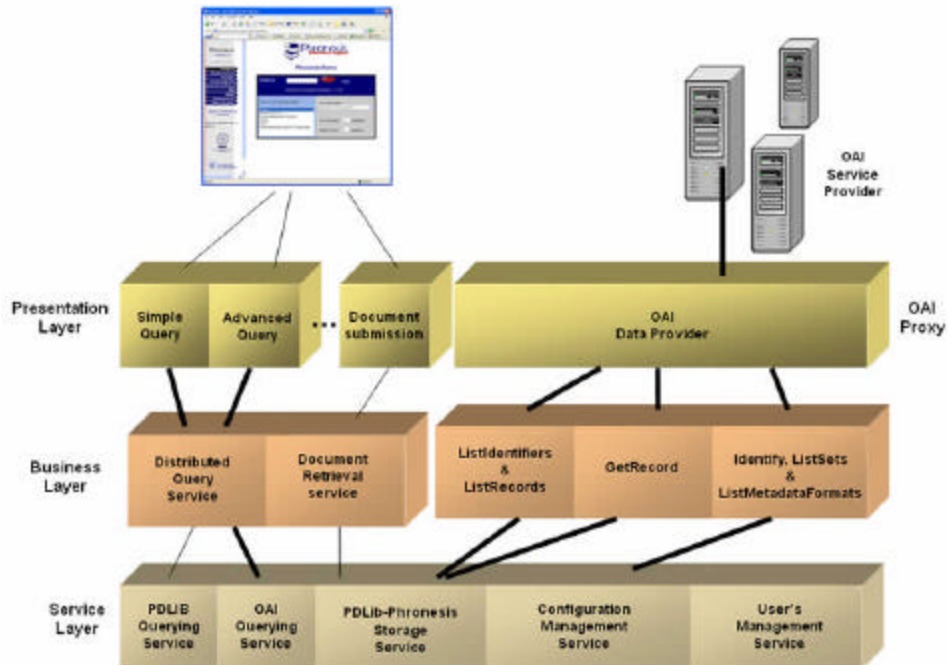


Figura 2. Servicios de interoperabilidad bajo OAI-PMH.

4. Interoperabilidad

Entre los componentes clave del proyecto se encuentran los servidores que permitirán el intercambio de datos a través del protocolo OAI-PMH [11]. Se han logrado avances en el diseño e implementación prototípica de un servidor para la versión 2.0 del protocolo que opera con el acervo de tesis digitales de la UDLA. En la fase siguiente se generalizará este desarrollo.

Como ya se ha mencionado también se han desarrollado servicios que permitan al sistema Phronesis comunicarse a través de este OAI-PMH. La figura 2 muestra la arquitectura de capas del sistema. En ella se observan 3 niveles, y en los cuales se encuentran los servicios de interoperabilidad (OAI Data Provider, GetRecord, ListIdentifier, ListRecords, Identify, ListSets, ListMedatadaFormats)

El objetivo es que es que cualquier biblioteca de la federación, denotado en la figura como “OAI Service Provider” pueda consultar Phronesis por medio de los servicios antes mencionados. Actualmente se ha alcanzado el desarrollo de alrededor de un 70% de estos servicios.

5. Agentes

Se ha experimentado con el uso de agentes en un ambiente de referencia virtual y el uso de WordNet en el esquema de ontologías propuesto. Los resultados se aplicarán a la implementación de agentes ontológicos en una federación de bibliotecas digitales.

6. Interfaces para dispositivos móviles.

Se ha logrado, a partir de interfaces existentes en dispositivos convencionales, la generación de interfaces para PDAs y teléfonos celulares para el caso específico de espacios personales de U-DL-A. Se diseñaron las pruebas de usabilidad de estas interfaces, para lo que se crearon dos versiones: versión PDA y versión emulador. Estas pruebas son del tipo tutoradas/valoración retrospectiva y fueron diseñadas para aplicarse con el apoyo del Laboratorio de Usabilidad de la Universidad Tecnológica de la Mixteca. En la siguiente fase se planea generalizar la generación de interfaces.

7. Difusión

El proyecto en desarrollo ha despertado interés en diversos foros. Se han realizado las siguientes presentaciones:

- *Espacios personales genéricos en el contexto del Proyecto: “Agentes y movilidad en colecciones multimediales heterogéneas”*. Congreso Amigos 2004: Consorcios en la Sociedad del Conocimiento, Universidad de las Américas, Puebla, Feb. 2004.
- *Proyecto: Agentes y movilidad en colecciones multimediales heterogéneas*, Reunión de Primavera de CUDI, Manzanillo, Colima, Abril 2004.

Y se han generado las siguientes publicaciones:

- **Investigación:** Designing ontological agents: An alternative to improve information retrieval in federated digital libraries. *In Advances in Web Intelligence: Proceedings of the Second International Atlantic Web Intelligence Conference* (May 16-19). Lecture Notes in Computer Science 3034, Springer-Verlag, 155-163. ISBN 3-540-22009-7.
- **Difusión:** An Exploration of Network Technologies for Mobile Data Access in Digital Libraries, *WSEAS Transactions on Communications* (Issue1, Vol. 3, January 2004, -ISSN 1109-2742, pp. 104-109 (publicado simultáneamente en *Proceedings del 6th WSEAS International Conference on Telecommunications and Informatics* (TELE-INFO'04, ISBN 960-8052-98-X) pp. 485-213: 1-6.
- **Investigación:** Mobile Access to Digital Libraries via Generic Interfaces. Enviado al *Taller de Cómputo Móvil 2004* (TCM 2004) del Encuentro Nacional de Computación 2004 (ENC 2004), a celebrarse en Colima, Col., Sept. 2004.

Referencias

- [1] F. Alvarez, M. Guevara, A. Salinas. PDLib: Escenarios de ejemplo (reporte técnico).
- [2] L. Bass, P. Clements, R. Kazman. Software Architecture in Practice, 2nd Edition. Addison Wesley, 2003.
- [3] The Mobile Information Device Profile (MIDP). <http://java.sun.com/products/midp/>.
- [4] Enhydra kXML-RPC. <http://kxmlrpc.objectweb.org/index.html>.
- [5] Sun Microsystems Java 2 Platform, Standard Edition (J2SE™). <http://java.sun.com/>.
- [6] MySQL. <http://www.mysql.com/>.
- [7] Jakarta Lucene. <http://jakarta.apache.org/lucene/docs/index.html>.
- [8] Apache XML-RPC. <http://ws.apache.org/xmlrpc/>.
- [9] Tomcat. <http://jakarta.apache.org/tomcat/>.
- [10] A. Salinas, Arquitectura de multiples capas orientada a servicios para el sistema Phronesis, Tesis de Maestría, ITESM Campuns Monterrey, Mayo de 2004
- [11] Open Archives Protocol for Metadata Harvesting, <http://www.openarchives.org/>