

Capítulo 5

Resultados

En este capítulo se presentan los resultados de haber utilizado Voai para la generación de servidores que comparten colecciones digitales reales. Cada colección digital tiene sus propias características y particularidades, lo que hace sumamente interesante a Voai pues es posible que, independientemente de la estructura de la colección, esta herramienta es capaz de generar un servidor OAI para dicha colección. Los servidores OAI generados para esas colecciones están oficialmente validados como proveedores de datos, lo que garantiza el correcto funcionamiento de esta herramienta.

En la sección 5.1 se proporciona una breve descripción de las principales características de las colecciones digitales que fueron usadas para la generación de servidores OAI. En la sección 5.2 se muestran las diferencias que existen en la información suministrada para la generación de dichos servidores. Un pequeño resumen de las características más sobresalientes de Voai es proporcionado en la sección 5.3. Finalmente en la sección 5.4 se da una conclusión de lo expuesto a lo largo del capítulo.

5.1 Servidores OAI generados

Para validar la implementación realizada, se utilizó a Voai para generar servidores OAI para diferentes colecciones digitales. Vale la pena señalar que estas colecciones son reales, es decir, no fueron creadas a la medida para que funcionaran adecuadamente con Voai, sino que fueron colecciones que realmente necesitaban un servidor OAI para compartir sus metadatos y por lo tanto fueron usadas como colecciones prueba para generar servidores y validar a Voai. A continuación se describe la exitosa aplicación de Voai para la generación automática de servidores OAI para tres colecciones digitales reales: colección de Tesis Digitales, colección de libros antiguos digitalizados, y colección de incunables.

5.1.1 Servidor OAI de Tesis Digitales

La colección Tesis digitales [Fernández y Sánchez 2003] es una colección que contiene más de mil documentos de tesis en formato digital, legado de los estudiantes de la Universidad de las Américas, Puebla. Esta colección fue diseñada a través de un modelo entidad relación e implementada bajo un RDBMS (*Relational Database Management System*) MySQL.

Cada tesis puede ser accedida a través de una dirección Web que está almacenada como un metadato de la base de datos. Esto permitió que el servidor OAI generado para esta colección pudiera compartir ese metadato, y por ende, que los usuarios pudieran acceder de manera directa al documento físico de la tesis digital.

El servidor OAI generador por Voai para esta colección está validado como un proveedor de datos oficial en la Iniciativa de Archivos Abiertos [Lagoze y Van de Sompel 2005]. Esta validación que hace OAI garantiza el correcto funcionamiento y la robustez del servidor OAI o proveedor de datos validado. Es por esto que es de suma importancia la validez de los servidores generados por Voai, porque es otra forma de garantizar el correcto funcionamiento de Voai. El servidor OAI que fue generado para la colección Tesis Digitales de la UDLA está disponible en:

http://ict.udlap.mx:9090/Tales/Oai_tesis.

5.1.2 Servidor OAI de Libros antiguos

Esta colección esta conformada por más de ciento treinta libros antiguos que han sido digitalizados y pueden ser accedidos a través de un software de navegación y visualización especializado denominado CIText [García 2002]. Al igual que Tesis Digitales, esta colección fue diseñada a través de un modelo entidad relación e implementada bajo un RDBMS MySQL.

Una característica particular de esta colección es que los libros antiguos digitalizados son recuperados directamente de la base de datos cuando se despliegan en CIText, y por tal motivo no existe una dirección Web en la base de datos que pueda ser compartida como un metadato en el servidor OAI. Para resolver el problema que representa la ausencia de una dirección Web que permita acceder directamente a un documento físico, Voai fue extendido para que generara un servlet que tiene la capacidad de recuperar el documento directamente de la base de datos, y de desplegarlo en un navegador de Internet. Por lo tanto,

la dirección de ese servlet será la dirección Web que será compartida como metadato por el servidor OAI.

También, al igual que con la colección Tesis Digitales, el servidor OAI generado para compartir esta colección de libros antiguos está validado ante OAI como un proveedor de datos oficial. Este servidor esta disponible en:

http://catarina.pue.udlap.mx:9090/u_dl_a/citext/Oai_citext

5.1.3 Servidor OAI de Incunables

La universidad de Valladolid ha construido esta colección digital para promover la preservación y diseminación de documentos incunables (libros antiguos impresos antes de 1501) resguardados por la Biblioteca Histórica de Santa Cruz. Esta colección también fue diseñada a través de un modelo entidad relación e implementada bajo un RDBMS MySQL.

El proceso de generación del servidor OAI para esta colección fue muy parecido al seguido para Tesis Digitales, excepto por la construcción dinámica de direcciones Web que acceden directamente a un documento físico, y las cuales no están almacenadas en la base de datos. Estas direcciones fueron generadas fácilmente agregando el identificador de un recurso local a una dirección Web de una aplicación previamente desarrollada por la Universidad de Valladolid.

Finalmente es importante señalar que el servidor OAI generado para Incunables está validado ante OAI como un proveedor de datos oficial. El servidor esta disponible en:

http://gutenberg.dcs.fi.uva.es:8080/OAI_UVA/Oai_incunables.

5.2 Diferencias en la generación de servidores

Cada una de las colecciones usadas por Voai tiene sus propias características. De hecho, es sumamente importante señalar que fue posible generar un servidor OAI para cada colección independientemente de la estructura que tenía dicha colección. Debido a que cada colección tiene su propia estructura y diferente información de identificación, es obvio que la información proporcionada a Voai para la generación de un servidor fue distinta en cada ocasión. A continuación se muestran las diferencias que se dieron durante el proceso de especificación de información en cada colección utilizada.

5.2.1 Información especificada en el servlet Identify

Cada vez que se utilizó a Voai para crear un servidor para una colección, fue necesario proporcionar la información de identificación de dicha colección al servlet Identify. En la tabla 5.1 se pueden ver las diferencias de esta información de identificación de cada colección.

Tabla 5.1. Información de identificación del servidor de cada colección

Colección digital	Información proporcionada
	Nombre de la colección
Tesis Digitales	Universidad de las Américas, Puebla: Digital Theses
Libros antiguos	Universidad de las Américas, Puebla: CIText
Incunables	Universidad de Valladolid: Incunables
	Dirección Web
Tesis Digitales	http://ict.udlap.mx:9090/Tales/Oai_tesis
Libros antiguos	http://catarina.pue.udlap.mx:9090/u_dl_a/citext/Oai_citext
Incunables	http://gutenberg.dcs.fi.uva.es:8080/OAI_UVA/Oai_incunables
	Correo electrónico del administrador
Tesis Digitales	tesis@mail.udlap.mx
Libros antiguos	lulu@mail.udlap.mx
Incunables	jadiago@infor.uva.es

	Fecha más antigua de modificación de un registro
Tesis Digitales	1989-05-04
Libros antiguos	1660-01-01
Incunables	1467-01-01
	Soporte de historial de registros eliminados
Tesis Digitales	No
Libros antiguos	No
Incunables	No
	Granularidad de fecha soportada
Tesis Digitales	YYYY-MM-DD
Libros antiguos	YYYY-MM-DD
Incunables	YYYY-MM-DD
	Identificador de la colección
Tesis Digitales	thesisUDLAP
Libros antiguos	citext.UDLAP
Incunables	incunablesUVA
	Ejemplo de identificador de la colección
Tesis Digitales	oai:thesisUDLAP:24
Libros antiguos	oai:citext.UDLAP:abko
Incunables	oai:incunablesUVA:172

5.2.2 Información especificada en el servlet ListSets

De la misma manera en que existe diferente información de identificación para cada colección, también existen diferencias tanto en los parámetros de conexión a la base de datos como en las consultas en SQL para recuperar los conjuntos que soporta cada colección. Esas diferencias se pueden ver claramente en la tabla 5.2.

Tabla 5.2. Parámetros de conexión a la BD y consulta SQL para recuperación de conjuntos

Colección digital	Información proporcionada
	Nombre del servidor de la base de datos
Tesis Digitales	ict.udlap.mx
Libros antiguos	catarina.pue.udlap.mx
Incunables	gutenberg.dcs.fi.uva.es
	Puerto de conexión de la base de datos
Tesis Digitales	9097
Libros antiguos	9096
Incunables	3306

	Manejador del RDBMS
Tesis Digitales	org.gjt.mm.mysql.Driver
Libros antiguos	org.gjt.mm.mysql.Driver
Incunables	org.gjt.mm.mysql.Driver
	Dirección de la base de datos
Tesis Digitales	jdbc:mysql
Libros antiguos	jdbc:mysql
Incunables	jdbc:mysql
	Nombre de la base de datos
Tesis Digitales	iid
Libros antiguos	sangabriel
Incunables	IncunablesBHSC
	Nombre de usuario de la base de datos
Tesis Digitales	lectura
Libros antiguos	citext
Incunables	uvaoai
	Contraseña de acceso a la base de datos
Tesis Digitales	*****
Libros antiguos	*****
Incunables	*****
	Consulta SQL para recuperar los conjuntos que soporta la colección
Tesis Digitales	<i>Select id,name</i> <i>From department</i>
Libros antiguos	<i>Select id,nombre</i> <i>From coleccion</i>
Incunables	<i>Select idCiudad,nomIng</i> <i>From ciudad</i>

5.2.3 Información especificada en el servlet ListMetadataFormats

La tabla 5.3 es una de las más importantes pues muestra las consultas en SQL que se especificaron para cada colección. Cada una de esas consultas en SQL permite recuperar metadatos locales que serán mapeados a los metadatos del formato Dublín Core. Para este mapeo se utiliza el campo “correspondencia”, el cual me permite saber cuál es el metadato que corresponde al metadato Dublín Core para el que se está haciendo la recuperación de información.

También podemos darnos cuenta que en las consultas SQL existen símbolos “\$”. Si este símbolo esta en la consulta SQL, significa que aquel atributo que está entre dos símbolos “\$” será cambiado valores de parámetros de las peticiones OAI-PMH. En caso de que este símbolo este en el campo de correspondencia, entonces servirá para darle un formato de texto adecuado a los metadatos. Detalles de esta sintaxis son proporcionados en el apéndice C de este documento.

Tabla 5.3. Consultas en SQL para la recuperación de metadatos

Colección digital	Información proporcionada
	Title
Tesis Digitales	<i>Select title</i> <i>From thesis</i> <i>Where id=\$identifier\$</i> <i>Correspondencia: title</i>
Libros antiguos	<i>Select titulo</i> <i>From libro</i> <i>Where call_cutter=\$identifier\$</i> <i>Correspondencia: titulo</i>
Incunables	<i>Select titulo</i> <i>From titulo</i> <i>Where idIncunable=\$identifier\$</i> <i>Correspondencia: titulo</i>
	Creator
Tesis Digitales	<i>Select first_name,middle_name,last_name</i> <i>From person,author_thesis</i> <i>Where person.id=author_thesis.id_author and author_thesis.id_thesis=\$identifier\$</i> <i>Correspondencia: first_name,\$ \$,middle_name,\$ \$,last_name</i>

Libros antiguos	<p><i>Select nombre,apellido</i></p> <p><i>From autor,libro_autor</i></p> <p><i>Where libro_autor.id_autor=autor.id and libro_autor.call_cutter=\$identifier\$</i></p> <p><i>Correspondencia: nombre,\$ \$,apellido</i></p>
Incunables	<p><i>Select nomFamiliar,nomPersonal</i></p> <p><i>From Autor,IncAutor</i></p> <p><i>Where Autor.idAutor=IncAutor.idAutor and IncAutor.idIncunable=\$identifier\$</i></p> <p><i>Correspondencia: nomFamiliar,\$ \$,nomPersonal</i></p>
	Subject
Tesis Digitales	<p><i>Select name</i></p> <p><i>From topic,thesis_topic</i></p> <p><i>Where thesis_topic.id_topic=topic.id and thesis_topic.id_thesis=\$identifier\$</i></p> <p><i>Correspondencia: name</i></p>
Libros antiguos	<p><i>Select tema</i></p> <p><i>From libro</i></p> <p><i>Where call_cutter=\$identifier\$</i></p> <p><i>Correspondencia: tema</i></p>
Incunables	No existe información que corresponda a este metadato
	Description
Tesis Digitales	<p><i>Select content</i></p> <p><i>From summary</i></p> <p><i>Where summary.id_thesis=\$identifier\$</i></p> <p><i>Correspondencia: content</i></p>
Libros antiguos	No existe información que corresponda a este metadato
Incunables	<p><i>Select texto</i></p> <p><i>From IncNota</i></p>

	<p><i>Where idIncunable=\$Identifier\$</i></p> <p><i>Correspondencia: texto</i></p>
	Contributor
Tesis Digitales	<p><i>Select first_name,middle_name,last_name</i></p> <p><i>From person p, evaluator e, evaluator_thesis et, user u</i></p> <p><i>Where p.id=e.id and e.id=et.id_evaluator and p.id=u.id and et.id_thesis=\$Identifier\$</i></p> <p><i>Correspondencia: first_name,\$ \$,middle_name,\$ \$,last_name</i></p>
Libros antiguos	<p><i>Select nombre</i></p> <p><i>From libro,acervo</i></p> <p><i>Where acervo.id=libro.acervo and libro.call_cutter=\$Identifier\$</i></p> <p><i>Correspondencia: nombre</i></p>
Incunables	Biblioteca Histórica de Santa Cruz
	Publisher
Tesis Digitales	<p><i>Select university.name</i></p> <p><i>From university, school, department, dep_major, author_major, author, author_thesis</i></p> <p><i>Where university.id=school.id and department.id_school=school.id and dep_major.id_dep=department.id and dep_major.id_major=author_major.id_major and author.id=author_major.id_author and author.id=author_thesis.id_author and author_thesis.id_thesis=\$Identifier\$</i></p> <p><i>Correspondencia: university.name</i></p>
Libros antiguos	Universidad de las Américas, Puebla
Incunables	Biblioteca Histórica de Santa Cruz, Universidad de Valladolid
	Date
Tesis Digitales	<p><i>Select day,month,year</i></p> <p><i>From thesis</i></p> <p><i>Where id=\$Identifier\$</i></p> <p><i>Correspondencia: year,\$-\$,month,\$-\$,day</i></p>
Libros antiguos	<i>Select anio</i>

	<p><i>From libro</i></p> <p><i>Where call_cutter=\$identifier\$</i></p> <p><i>Correspondencia: anio,\$01-01\$</i></p>
Incunables	<p><i>Select idAnyo</i></p> <p><i>From IncAnyo</i></p> <p><i>Where idIncunable=\$identifier\$</i></p> <p><i>Correspondencia: idAnyo,\$01-01\$</i></p>
	Type
Tesis Digitales	<p><i>Select major.degree</i></p> <p><i>From university, school, department, dep_major, author_major, author, author_thesis</i></p> <p><i>Where university.id=school.id and department.id_school=school.id and dep_major.id_dep=department.id and dep_major.id_major=author_major.id_major and author.id=author_major.id_author and author.id=author_thesis.id_author and author_thesis.id_thesis=\$identifier\$</i></p> <p><i>Correspondencia: major.degree</i></p>
Libros antiguos	Libros de fondo antiguo
Incunables	Incunable
	Format
Tesis Digitales	text/html, application/pdf
Libros antiguos	image/gif
Incunables	text/html
	Identifier
Tesis Digitales	<p><i>Select url_home, folder_name</i></p> <p><i>From url_thesis</i></p> <p><i>Where url_thesis.id_thesis=\$identifier\$</i></p> <p><i>Correspondencia: url_home, folder_name</i></p>
Libros antiguos	http://catarina.pue.udlap.mx:9090/u_dl_a/citext/VisualizarImg?call_cutter=\$identifier\$
Incunables	http://gutenberg.dcs.fi.uva.es/~bhscdes/modulos.php?op=ver_inc&inc=\$identifier\$
	Source
Tesis Digitales	No existe información que corresponda a este metadato

Libros antiguos	No existe información que corresponda a este metadato
Incunables	No existe información que corresponda a este metadato
	Language
Tesis Digitales	<i>Select language</i> <i>From thesis</i> <i>Where thesis.id=\$identifier\$</i> <i>Correspondencia: language</i>
Libros antiguos	No existe información que corresponda a este metadato
Incunables	<i>Select idLengua</i> <i>From IncLengua</i> <i>Where idIncunable=\$identifier\$</i> <i>Correspondencia: idLengua</i>
	Relation
Tesis Digitales	No existe información que corresponda a este metadato
Libros antiguos	No existe información que corresponda a este metadato
Incunables	No existe información que corresponda a este metadato
	Coverage
Tesis Digitales	No existe información que corresponda a este metadato
Libros antiguos	No existe información que corresponda a este metadato
Incunables	No existe información que corresponda a este metadato
	Rights
Tesis Digitales	http://www.udlap.mx/reglamte/capitulo2/t2c4.html
Libros antiguos	Derechos Reservados © 2005. Universidad de las Américas, Puebla
Incunables	Todos los derechos reservados a la Biblioteca Historica de Santa Cruz

5.2.4 Información especificada en el servlet GetRecord

En este servlet solo es necesario especificar una consulta que me permita recuperar un registro específico de la base de datos. En la tabla 5.4 podemos ver las diferentes consultas de cada colección.

Tabla 5.4. Consulta en SQL para recuperar un registro específico

Colección digital	Información proporcionada
	Consulta SQL para la recuperación de un registro específico
Tesis Digitales	<i>Select *</i> <i>From thesis</i> <i>Where id=\$identifier\$</i> <i>Correspondencia: id</i>
Libros antiguos	<i>Select *</i> <i>From libro</i> <i>Where call_cutter=\$identifier\$</i> <i>Correspondencia: call_cutter</i>
Incunables	<i>Select *</i> <i>From Incunable</i> <i>Where idIncunable=\$identifier\$</i> <i>Correspondencia: idIncunable</i>

5.2.5 Información especificada en el servlet ListRecords y ListIdentifiers

Finalmente, en este servlet se especifican dos consultas. Una para recuperar los conjuntos a los que pertenece un registro, y la otra para recuperar los registros que pertenecen a un conjunto. Las consultas en SQL para las colecciones utilizadas se muestran en la tabla 5.5.

Tabla 5.5. Consultas en SQL para recuperar los conjuntos de un registro y viceversa

Colección digital	Información Proporcionada
	Consulta en SQL que recupera los conjuntos de un registro
Tesis digitales	<i>Select department.id</i> <i>From department, dep_majior, thesis_majior, author_thesis, thesis</i>

	<p><i>Where dep_major.id_dep=department.id and thesis_major.id_major=dep_major.id_major and thesis_major.id_thesis=thesis.id and author_thesis.id_thesis=thesis.id and and author_thesis.id_thesis=\$identifier\$</i></p> <p><i>Correspondencia: department.id</i></p>
Libros antiguos	<p><i>Select coleccion</i></p> <p><i>From libro</i></p> <p><i>Where call_cutter=\$identifier\$</i></p> <p><i>Correspondencia: coleccion</i></p>
Incunables	<p><i>Select idCiudad</i></p> <p><i>From IncEdicion</i></p> <p><i>Where idIncunable=\$identifier\$</i></p> <p><i>Correspondencia: idCiudad</i></p>
Consulta en SQL que recupera los registros de un conjunto	
Tesis digitales	<p><i>Select thesis.id</i></p> <p><i>From department, dep_major, thesis_major, author_thesis, thesis</i></p> <p><i>Where dep_major.id_dep=department.id and thesis_major.id_major=dep_major.id_major and thesis_major.id_thesis=thesis.id and author_thesis.id_thesis=thesis.id and and department.id=\$set\$</i></p> <p><i>Correspondencia: thesis.id</i></p>
Libros antiguos	<p><i>Select call_cutter</i></p> <p><i>From libro</i></p> <p><i>Where coleccion=\$set\$</i></p> <p><i>Correspondencia: call_cutter</i></p>
Incunables	<p><i>Select idIncunable</i></p> <p><i>From IncEdicion</i></p> <p><i>Where idCiudad=\$set\$</i></p> <p><i>Correspondencia: idIncunable</i></p>

5.3 Resumen de las características de Voai

Si recordamos, en la tabla 2.1 de este documento se presentaron las características más importantes de otras aplicaciones que también generan servidores OAI. Después de haber visto el diseño conceptual, la implementación y las pruebas de Voai, ahora es posible proveer un resumen de las características más sobresalientes de esta herramienta. Para esto, es indispensable retomar las características utilizadas para evaluar a las otras aplicaciones discutidas en el capítulo dos, y ahora se utilizan para evaluar a Voai.

Como se puede ver en la tabla 5.6, Voai soporta colecciones digitales implementadas en bases de datos multitaslas, y no es necesario que se modifique la estructura original de la base de datos para poder generar un servidor OAI que comparta a la colección. El usuario que construye un servidor OAI con la herramienta Voai no necesita realizar ningún tipo de implementación, ni tampoco es necesario que se involucre en el manejo de complejos archivos de configuración. El servidor OAI generado puede ser colocado inmediatamente en un servidor Web, y posteriormente ser accedido por aplicaciones o proveedores de servicios en el mundo: además, el servidor OAI resultante es capaz de ser validado ante la Iniciativa de Archivos Abiertos como un proveedor de datos oficial.

Tabla 5.6. Resumen de las características de Voai

	Soporta BD relacionales.	Soporta BD relacionales multitaslas.	Sin modificación de la BD.	No requiere implementación del usuario.
Voai	✓	✓	✓	✓

5.4 Conclusiones

Como conclusión para este capítulo se puede decir que la herramienta Voai fue utilizada exitosamente para generar servidores OAI para diversas colecciones digitales reales. Algo que es muy importante señalar es que Voai permite generar servidores OAI de manera automática independientemente de la estructura de las colecciones.

Las principales características de Voai son: soporta bases de datos relacionales multitablas, no es necesario modificar la estructura de la colección, ni tampoco es necesario lidiar con complejos archivos de configuración. Voai solo requiere que el usuario conozca perfectamente su colección, y que proporcione tanto información de identificación como consultas en SQL para generar el servidor OAI específico para esa colección.