

Utilización de estándares W3C y “Open Source” en la Biblioteca de la UNED: aplicación a la gestión de tesis doctorales

V workshop Rebiun

18 octubre 2005

Luis Zorita

Índice

- Contexto
- Representando nuestros datos en XML
- Transformación de los datos
- Gestión de los datos y la información con MVC
- W3C . Unas recomendaciones que vamos a seguir:
 - Xpath, Xquery, Xforms
- Tomcat . Orbeon o Cocoon . Xforms decide.
- La aplicación : cliente-servidor
- Reutilización para OAI y Fedora como repositorio

Agradecimientos

CONTEXTO

- Necesidades concretas:
 - Gestor de contenidos basado en XML
 - Proyecto OAI
 - Repositorio Institucional
- Información dispersa: tesis, sumarios, bases de datos, directorio de personal, recursos web...
 - Diferentes sistemas de gestión
 - Diferentes formatos

Buscando un sistema único de representación de nuestros datos que nos permitiera
reutilización

Representando nuestros datos en XML

- Nuestros registros de tesis, bases de datos, sumarios, etc estaban en formatos diferentes: access , html , word...
- Tras ver tendencias : marcxml , tei , etc . xml es una sintaxis adecuada para esquemas de metadatos.
- Hay que generar los datos y la información en XML
- Hay que aprender una nueva tecnología y su jerga: XML, XSL, X...

Transformación de los datos

- Se crea un esquema de representación que incluye Dublin-Core y una serie de elementos propios. [Esquema de datos](#)
- Se transforman los datos correspondientes a las tesis doctorales y a los sumarios de revistas que se encuentran en la base de datos Access a este esquema.
- Se establece una correspondencia entre los campos Access y ese esquema.
- Se realiza la transformación.
- Resultado: + de 3000 tesis y + de 11000 sumarios. [Ejemplo de tesis](#) [Ejemplo sumario](#)

Gestión de datos e información con MVC

- Decisión de trabajar en entorno MVC
 - Separación entre : los datos , la presentación de la información y la lógica.
- . Los datos se almacenan en una B.D xml nativa
- . La presentación se hace mediante transformaciones XSLT.
- . La lógica ha de trabajar con datos en XML y no puede incluir nada relacionado con la presentación.

W3C . Unas recomendaciones que vamos a seguir:

Xpath, Xquery, Xforms,...

- Elegir es fácil ... equivocarse también
- Obligados a trabajar en entornos abiertos y “open source”, el único blindaje posible lo da ajustarse a estándares abiertos.
- W3C y sus recomendaciones (Xpath , Xsl , SOAP , Xforms , Xquery , X...) han sido una guía a la hora de seleccionar tecnología.

La lógica : Xforms decide

- Inicialmente Cocoon como servlet
- El problema de Xforms nos lleva a elegir **Orbeon** que se ajusta al W3c
- Se puede trabajar con SQL y Xquery (Tamino, Exist, Oracle, etc)

Características de la plataforma tecnológica.

- XML como sintaxis
- XSL como lenguaje de presentación en internet y transformación
- Java como plataforma de desarrollo
- Bases de datos nativas. Exist Dbxml
- Xquery como lenguaje de interrogación
- Xforms como intercambio entre páginas
- Apache como servidor WEB
- Tomcat como servidor de servlet
- Orbeon como servlet

Se puede trabajar con SQL y Xquery (Tamino Exist Oracle etc

- Independencia de la plataforma UNIX-Windows

La aplicación : cliente-servidor

- Dispone de :
 - autenticación → Tomcat
 - búsqueda → Xquery , Xpath
 - Cliente web para formularios , edición , borrado y actualización
 - Visualización en múltiples formatos
 - Grabar en base de datos nativa.
 - <http://localhost/orbeon/autenticacion>
- . Se pueden utilizar las herramientas propias de un sistema de gestión de la Base de datos para creación de perfiles, indexar, etc. Ver cliente administrador Exist
- . Dispone además de un sistema de búsqueda especializado.
<http://localhost/exist/xquery/tesisdoctorales.xml>

Reutilización para OAI y Fedora como repositorio

- Salida OAI_DC.

[Xsl que transforma ud en oai](#)

[Resultado de la transformación](#)

- Salida FOXML para repositorio Fedora

[Xsl que transforma ud en foxml](#)

[Resultado de la transformación](#)

Agradecimientos

- Gratis no significa que no cueste
- Para poner en producción este sistema ha sido necesario cambiar metodología de trabajo:
 - Hay que entrar en los foros de Internet
 - Hay que trabajar en entornos abiertos
 - Hay que aprender tecnologías en continua evolución
 - Hay que cooperar en los foros

Por todo ello creo que es justo agradecer a:

- . Todos los que aportan y trabajan en Open Source
- Organismos que elaboran estándares
- Andreas Hartmann, Alessandro Vernet , Erik Brouchez, Wolfgang Meier, ...
- Alicia López , Isabel Calzas, Isabel Belmonte.