

<i>Estado (Aprobado)</i>	<i>UAH_eBuah_Proyecto</i>	<i>Vigencia desde 15-11-06</i>
--------------------------	---------------------------	--------------------------------

Proyecto "eBuah": Biblioteca digital de la Universidad de Alcalá

<i>Elaborado por:</i>	<i>Juana Frías Fernández</i>	<i>Firma:</i>
<i>Corregido por:</i>		<i>Firma:</i>

eBuah: Definición del proyecto

Introducción

La implementación de un repositorio OAI institucional en la Universidad de Alcalá, se enmarca dentro del proyecto “Madroño e-Ciencia” ¹del Consorcio Madroño.

Dicho proyecto pretende la creación de la red Madroño de repositorios abiertos, formada por los repositorios implementados en las distintas universidades participantes, asimismo pretende la creación de un Portal de acceso común a los contenidos almacenados en los distintos repositorios.

Independientemente del proyecto Madroño e-Ciencia, la creación de un repositorio de archivos abiertos en la UAH permite el acceso libre a la producción científica de la Universidad, así como garantiza la conservación de estos archivos digitales, aumentando la visibilidad de los contenidos generados por los miembros de la UAH.

El repositorio institucional de la UAH se denomina **Biblioteca Digital de la Universidad de Alcalá: “eBuah”** .

Modelo del proyecto

Un repositorio institucional es un sistema de gestión de contenidos, formado por un conjunto de:

- Objetos digitales que podrán ser almacenados y recuperados posteriormente.
- Metadatos que describen los objetos digitales para su recuperación posterior, conforme a los estándares vigentes.
- Procedimientos para el envío de objetos digitales, su conservación, licencia de uso y administración.

La Biblioteca Digital eBuah cubre tanto el nivel de datos (data provider) (archivo de objetos y estándares internacionales); como el nivel de servicio (service provider) (servicios al usuario, búsqueda, autoarchivo...)

El contenido debe ser²:

- Erudito: tanto de carácter científico como académico.
- Producido, enviado o patrocinado por un miembro de la comunidad científica o cualquier otro agente autorizado por la política de la Institución.
- No efímero; el trabajo debe estar completo y preparado para su distribución

¹ Proyecto “Madroño e-Ciencia”. Documento de trabajo

² Proyecto “Madroño e-Ciencia”. Documento de trabajo

- Licenciado a perpetuidad; el autor debe ser capaz y tener la voluntad de ceder, no en exclusiva, a la institución el derecho a preservar y distribuir su trabajo a través del repositorio.

Los objetos digitales incluidos en los repositorios institucionales pueden ser:

- Pre-prints de artículos enviados para su publicación
- Artículos aceptados para su publicación
- Ponencias, comunicaciones a congresos
- Monografías
- Informes técnicos
- Documentos de trabajo
- Tesis doctorales y tesinas
- Proyectos de fin de carrera
- Objetos de aprendizaje
- Colecciones de datos resultados de proyectos de investigación
- Software
- Contenidos audiovisuales y multimedia

Software

El software elegido para la implantación de la Biblioteca Digital eBuah en la Universidad de Alcalá es DSpace (<http://www.DSpace.org>)

DSpace es un sistema de biblioteca digital que captura, guarda, indexa, preserva y redistribuye en formatos digitales, la producción intelectual de las instituciones. Es un proyecto conjunto que nació de la colaboración entre MIT Libraries y Hewlett-Packard, y es en la actualidad, un producto libremente disponible para instituciones de investigación a nivel mundial como un sistema de fuente abierta que puede ser personalizado y extendido.

La instalación del software, así como la configuración del servidor: Java, Tomcat, Apache, etc. la realizan los Servicios Informáticos de la UAH.

Actualmente (noviembre 2006) se encuentra instalado en fase de pruebas y configuración en un servidor adquirido para ello. (DSpace versión 1.3.2)

Requisitos técnicos

- Identificador

Cada item del repositorio debe tener un identificador único. Este identificador permite localizar y enlazar de forma estable y unívoca ese item.

DSpace utiliza por defecto el CNRI Handle System (<http://www.handle.net/>). Para utilizarlo es necesario instalar el Handle Server (incluido en el software de DSpace) y solicitar el prefijo para la Universidad de Alcalá.

- Estándares

DSpace soporta el Open Archives Initiative Protocol for Metadata Harvesting³ v2.0 como proveedor de datos. Este un requisito indispensable.

Respecto a los metadatos que van a describir el contenido de los objetos del repositorio, DSpace utiliza la versión del Dublin Core Libraries Working Group Application Profile⁴.

La Biblioteca de la UAH tendrá la responsabilidad de administrar y definir los formatos de metadatos⁵ a utilizar, siempre de acuerdo a los términos de la Dublin Core Metadata Initiative⁶. De la aplicación correcta de estos estándares, deriva la posibilidad de búsqueda y recuperación del contenido repositorio.

- Formatos de archivos

Nivel de formatos soportados en eBuah:

Formato soportado (formatos de archivos abiertos, legibles con cualquier aplicación que soporte estándares)

Formato conocido (formatos de archivos abiertos, a incluir como soportados en el futuro)

Formato desconocido (formatos de archivos propietarios, pueden no ser reconocidos por los programas. Con * los pendientes de estudio posterior)

Texto	Presentaciones	Gráficos	Sonido	Vídeo
PDF	ODP	JPEG	OGG	OGG
TXT	PPT *	PNG	MP3 *	MPEG *
RTF		SVG		AVI *
PS		JPG		
DOC				
HTML				

El nivel de formato soportado supone el compromiso por parte de eBuah de preservar esa información y de establecer las medidas técnicas necesarias para que esos archivos sean visibles en el futuro.

- Definición de Sets

Los repertorios se organizan en "sets" recolectables a través del protocolo OAI. Es necesario definir los tipos de sets a utilizar, siempre de acuerdo con lo que se decida en el proyecto Madroño e-Ciencia.

³ <http://www.openarchives.org/OAI/openarchivesprotocol.html>

⁴ <http://www.dublincore.org/documents/library-application-profile/>

⁵ Véase los documentos: "UAH_Metadatos" y "UAH_MetadatosTES"

⁶ <http://dublincore.org/documents/dcmi-terms/>

El proyecto e-Ciencia utilizará sets para permitir recolectar datos por tipos de documentos y por grandes grupos temáticos.

Se utilizarán los siguientes Sets:

- **Sets por materias:**

Madroño acordó utilizar el sistema de clasificación CDU para definir las categorías temáticas generales.

Esto es independiente de la asignación de materias a los objetos, para lo cual eBuah se basará en lo establecido en las Categorías temáticas utilizadas en el SIGB Unicorn, y controladas por la Sección de Normalización de la Biblioteca.

- **Sets por tipo de documento:**

En función de los tipos que se vayan utilizando: tesis, artículos, etc.

El proyecto e-Ciencia plantea utilizar en todos los repositorios la misma tipología, basada en e-Prints Complutense:

- Artículo / Article
- Sección de libro / Book Section
- Ponencia de Conferencia, Seminario o Congreso / Conference, Workshop or Meeting Item
- Libro o Monografía / Book or Monograph
- Otro / Other
- Tesis / Thesis
- PeerReviewed
- NonPeerReviewed

Estructura de los contenidos: comunidades y colecciones

Dspace estructura sus contenidos en forma de Comunidades, que se corresponden en eBuah con los Departamentos, Servicios u otras entidades de la UAH. Dentro de cada comunidad pueden establecerse subcomunidades más pequeñas, por ejemplo grupos de investigación.

Dentro de las comunidades aparecerán las colecciones que sean necesarias para englobar su producción científica.

Los items pueden pertenecer por definición a una colección solamente, pero pueden aparecer en diferentes colecciones. De este modo, todas las tesis doctorales de la UAH estarán dentro de la comunidad "Biblioteca" en la colección "Tesis doctorales", pero también serán recuperables en sus correspondientes Departamentos, en la colección "*DepartamentoX – Tesis doctorales*".

Licencias

Los propietarios de eBuah debemos tener permiso del autor para poder poner en libre acceso sus documentos.

La UAH ha aprobado la licencia de tesis doctorales y está en trámite la

aprobación de una licencia para otro tipo de documento. Esta última licencia será la que los autores vean cuando suban un archivo a eBuah y deberán aceptarla.

Por otra parte, es posible establecer una licencia de cara al usuario que accede a eBuah. Se ha puesto en la página principal la **“Licencia Creative Commons”⁷ Reconocimiento-NoComercial-SinObraDerivada**.

Esta licencia establece que se permite copiar, distribuir y comunicar públicamente la obra, siempre que se citen el autor original y la institución que le ampara, y no se haga ningún uso con finalidades comerciales ni obra derivada. Por obra derivada se entiende aquel documento que ha sido editado, traducido, combinado con materiales de terceros, cambiado de formato, o modificado de cualquier otra forma.

Esta licencia limita el uso que los usuarios pueden hacer de la información que encuentran en el repositorio.

Política de preservación

Es necesario establecer las estrategias necesarias para garantizar la preservación de los contenidos de la Biblioteca Digital eBuah, a largo plazo.

Hay que tener en cuenta:

- Backup: respaldo de los datos del repositorio. Realizado por los Servicios Informáticos.
- Hardware y Software: cuando el medio físico (servidor) o el software estén obsoletos, hay que prever su sustitución.
- Migración de datos: habrá que prever la migración de los formatos de los datos según vayan quedando obsoletos. Es necesario establecer un *“Procedimiento para la conservación de archivos en eBuah”*.
- Estándares: hay que adaptar el repositorio a las normas vigentes en cada momento. Los metadatos deberán adaptarse al protocolo OAI en vigor.

⁷ <http://es.creativecommons.org/>