

Directorio y recolector de colecciones y recursos digitales
<http://roai.mcu.es>

Cómo encontrar la aguja



en el pajar
de la web

Grupo de Trabajo de Colecciones Digitales del Ministerio de Cultura y las Comunidades Autónomas¹

El gran volumen de información accesible en el entorno web hace cada vez más difícil su acceso por parte de los ciudadanos. Por ello, el Grupo de Trabajo de Colecciones Digitales del MCU y las CC.AA. ha desarrollado una aplicación de gran alcance: el Directorio y recolector de colecciones y recursos digitales. Por un lado, facilita la búsqueda de los proyectos e iniciativas de digitalización referidas a colecciones bibliográficas; por otro lado, recoge documentos digitalizados a los que se accede utilizando técnicas de recolección de metadatos.

La accesibilidad de la información en la red es uno de los mayores desafíos que han de solucionar aquellos organismos o instituciones que tienen como objetivo la creación de recursos digitales. El volumen de información accesible en la red es tan grande y se incrementa de forma tan constante que con frecuencia resulta difícil, cuando no imposible, acceder a esta información.

Son varios los proyectos e iniciativas que se han ido creando en los últimos años para solucionar este problema y no cabe duda, a la vista de los resultados, de que uno de los más eficaces y exitosos ha sido la denominada Iniciativa de Archivos Abiertos (OAI, *Open Archives Initiative*) promovida por la Universidad de Michigan, así como el protocolo de recolección de metadatos (PMH, *Protocol for Metadata Harvesting*) con el que está estrechamente ligada².

El éxito de los metadatos basados en la *Dublin Core Metadata Initiative*³, que desde 2003 es la norma ISO 15836⁴, los ha convertido en uno de los procedimientos más sencillos y ágiles para codificar la información, entre otros, de los recursos digitales.

Puesto que el objetivo de las colecciones digitales que se creen es contribuir

a la Biblioteca Digital Europea, es importante señalar que toda la normativa se ajusta a la utilizada por la Biblioteca Europea⁵, muy en particular por lo que se refiere a la aplicación del protocolo OAI/PMH tanto para la recolección como para la creación de repositorios de metadatos conforme a la norma Dublin Core / ISO 15386.

A la vista de estas circunstancias, el Grupo de Trabajo de Colecciones Digitales⁶ del Ministerio de Cultura y las Comunidades Autónomas, creado en las VII Jornadas de Cooperación Bibliotecaria (Toledo 2003), se planteó desarrollar una aplicación que permitiera conocer los distintos proyectos e iniciativas que buscaban la creación de colecciones digitales y acceder a los mismos con un doble objetivo:

- Por una parte, facilitar un instrumento que permitiese en cierto grado la coordinación de los proyectos de digitalización que se llevan a cabo en España para evitar, en la medida de lo posible, la digitalización dos o más veces de las mismas obras.
- Por otro lado, con ello se prestaba un importante servicio a investigadores y ciudadanos al proporcionar acceso directo a los recursos digitales que se iban elaborando.

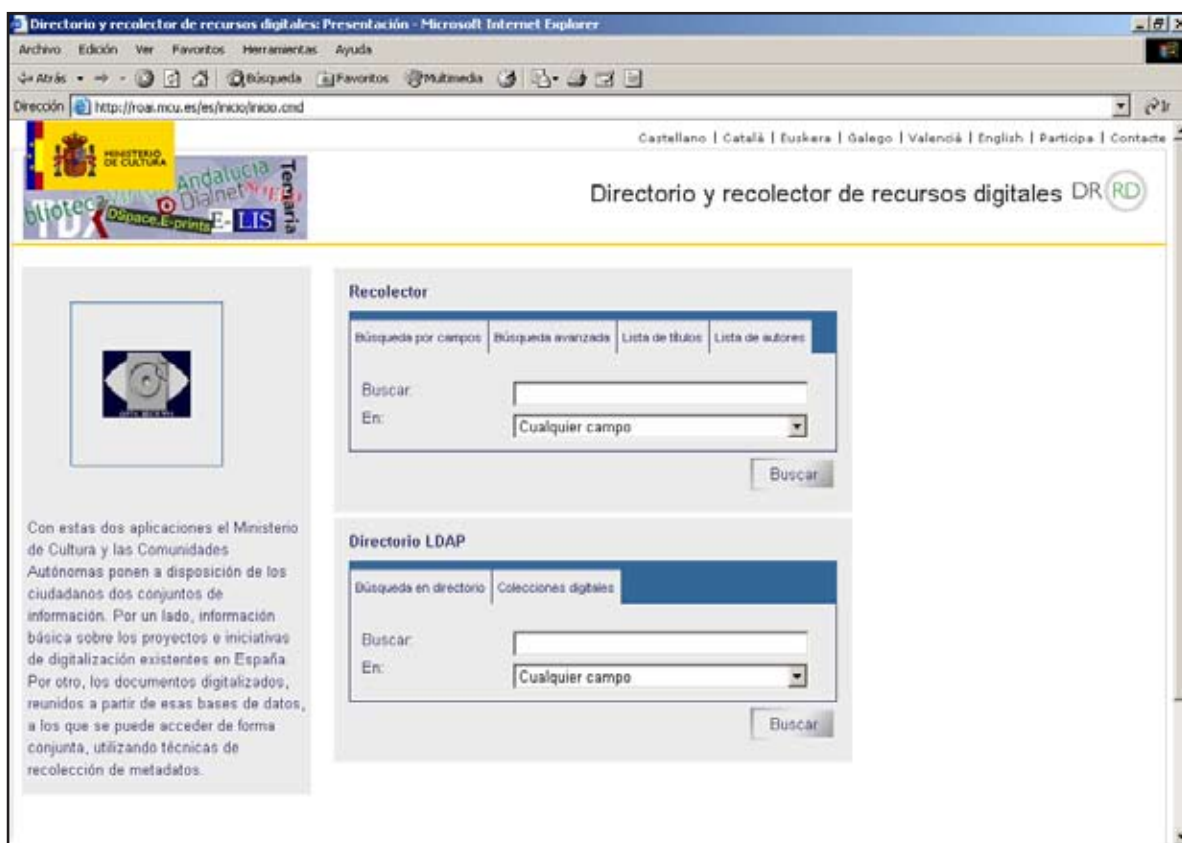
Directorio y recolector de recursos digitales: Directorio > Colecciones digitales - Microsoft Internet Explorer

Castellano | Català | Euskera | Galego | Valencià | English | Participa | Contacte

Directorio y recolector de recursos digitales DR RD

1 al 50 de 105 ▶▶ 1 2 3

1. [AAGRAFA \(Adquisición y Almacenamiento de material GRáfico de Arquitectura\)](#) - Universidad Politécnica de Madrid. Escuela Técnica Superior de Arquitectura
2. [ARCA \(Arxiu de Revistes Catalanes Antiques\)](#) - Biblioteca de Catalunya
3. [Archivo "Rodríguez Marín"](#) - Consejo Superior de Investigaciones Científicas
4. [Archivo Virtual de la Edad de Plata](#) - Fundación Residencia de Estudiantes
5. [Arxiu Gràfic - Biblioteca ETSAB - UPC](#) - Universitat Politècnica de Catalunya
6. [Biblioteca Abadia de Montserrat. biblioteca de ciències humanes i de teologia](#)
7. [Biblioteca de Catalunya : Fons Digitalitzats = Biblioteca de Catalunya : Fondos Digitalizados](#)
8. [Biblioteca de la Ciencia y de la Técnica del Renacimiento](#) - Universidad de Salamanca



Además, dado que se han establecido los correspondientes *site maps* con los principales buscadores (Google, Yahoo, etc.), se ha aumentado considerablemente la visibilidad de los recursos digitales.

Este proyecto se complementó con un directorio que describe las citadas colecciones digitales utilizando la estructura LDAP (*Lightweight Directory Access Protocol*), así como el *MARC21 Format for Community Information*⁷ a la hora de codificar la información descriptiva de las colecciones digitales o de la institución que promueve la citada digitalización.

Este directorio, cuyo objetivo es el disponer de una relación actualizada de los proyectos de digitalización elaborados en España —o fuera de España, pero referidos a sus colecciones bibliográficas—, permite, entre otras funcionalidades, conocer la lista de proyectos de digitalización en curso, finalizados o previstos; los organismos o instituciones patrocinadores de los proyectos de digitalización; las fases previstas para la realización de los mismos; los estándares utilizados; la cobertura temática y cronológica de los proyectos; los tipos de materiales digitalizados; las condi-

ciones de uso y política de protección de derechos de autor; y proporciona el enlace a los sitios web de los proyectos así como a las obras digitalizadas.

Todas las Comunidades Autónomas participantes son administradoras del sistema, gestionan en línea los datos de sus proyectos y efectúan las altas y modificaciones. El sistema incorpora, además, la posibilidad de introducir información sobre proyectos de otras instituciones, mediante un procedimiento de buzón abierto.

Los proyectos incluidos hasta la fecha se refieren en su mayoría a bibliotecas. Está previsto, de acuerdo con las conclusiones de las Jornadas de 2006, incluir los de archivos y museos.

El recolector de recursos digitales permite acceder, utilizando técnicas de recolección de metadatos, a los documentos digitalizados incluidos en las bases de datos del directorio. Es el primer —y hasta el momento único— servicio de este tipo desarrollado en España. Es uno de los 29 proveedores de servicios registrados de OAI⁸ y contiene más de 110.000 registros recolectados a partir de 25 repositorios, correspondientes la mayoría a iniciativas uni-

El volumen de información accesible en la red es tan grande y se incrementa de forma tan constante que con frecuencia resulta difícil, cuando no imposible, acceder a esta información.

versitarias. Asimismo está incluido en OpenDOAR⁹ (*Directory of Open Access Repository*) y en ROAR (*Registry of Open Access Repositories*)¹⁰.

Estas iniciativas están en línea con la Recomendación de la Comisión Europea sobre digitalización y accesibilidad en línea del material cultural y la conservación digital, de 24 de agosto pasado¹¹, cuyo objetivo es contribuir a la acción coordinada de los estados miembros en estos ámbitos. Entre las medidas recomendadas por la Comisión, que el grupo de trabajo está llevando a cabo, pueden destacarse las siguientes:

- Reunir información sobre la digitalización en curso y prevista para evitar la duplicación de esfuerzos como es el objetivo del Directorio de proyectos de digitalización.
- Recolección en la web (*web harvesting*). Este es el objetivo del reco-

lector de recursos digitales descrito anteriormente.

- Creación de un punto de acceso común multilingüe. Estas aplicaciones pueden consultarse en todas las lenguas oficiales de acuerdo con iniciativas como MICHAELPlus (*Multilingual Inventory of Digital Cultural Heritage in Europe*)¹², de la que está previsto que forme parte.

Cuando en marzo de 2006 se puso en producción el *Directorio y recolector de colecciones y recursos digitales*, coincidiendo con las Primeras Jornadas de Preservación del Patrimonio Digital¹³, se describían 55 colecciones digitales. Transcurrido un año, son ya 105 las que están descritas, es decir, prácticamente se han doblado.

El grupo de trabajo tiene previsto, contando con la financiación de la Subdirección General de Coordinación Bibliotecaria del Ministerio de Cultura,

The screenshot shows a web browser window displaying the 'Directorio y recolector de recursos digitales' website. The page title is 'Directorio y recolector de recursos digitales DRRD'. The main content area shows details for a digital collection:

- Recolector OAI: Ver datos | Registros recolectados
- Formato: Ficha | Aplicar
- Nombre: Biblioteca Virtual de Prensa Histórica
- Entidad responsable: Ministerio de Cultura. Dirección General del Libro, Archivos y Bibliotecas. Subdirección General de Coordinación Bibliotecaria
- Dirección: Plaza del Rey, 1
28071 Madrid
Tel.: + 34 91 701 71 37
E-mail: prensahistorica@dglab.mcu.es
- Descripción: La Biblioteca Virtual de Prensa Histórica es el resultado de un proceso de digitalización cooperativa del Ministerio de Cultura, las Bibliotecas Públicas del Estado y otras instituciones de la memoria. Iniciado en 2003, tiene como objetivos preservar y hacer accesibles unos materiales bibliográficos que se caracterizan por ser ejemplares únicos y, por lo tanto, de difícil acceso. Abarca el período comprendido entre finales del siglo XVII y los años cuarenta del s. XX. Utiliza los estándares XML MARC Schema, METS y DCMI. Incluye casi tres millones de páginas.
- Fechas: Creación: 2003
- Temas: Prensa-S. XVII-XX. BNE
- Número de control: DPO2005000001
- Dirección electrónica

Todas las Comunidades Autónomas participantes son administradoras del sistema, gestionan en línea los datos de sus proyectos y efectúan las altas y modificaciones.

umentar la capacidad de recolección de esta aplicación, ampliándola desde el mínimo que establecen las actuales *Directrices*¹⁴ (Dublin Core sin cualificar) a otros esquemas contemplados por las citadas directrices.

La próxima versión del *Recolector de recursos digitales* incorporará el esquema de datos XML MARC para recuperar registros bibliográficos constituyendo así catálogos colectivos bibliográficos virtuales, y aumentando su visibilidad en la red.

Además, se incluirán estructuras XML para los *e-prints* lo que hará posible una recuperación de recursos electrónicos vinculados al campo académico verdaderamente importante.

De la misma manera, se pretenden recolectar los metadatos de los archivos mediante la aplicación de la DTD (*Definición del Tipo de Documento*) de EAD (*Encoding Archival Description*) pues

con frecuencia iniciativas distintas caminan de forma paralela y, en último término, el usuario final no tiene por qué ver frustrados sus deseos de acceder a la información —máxime si esta es digital— por el tipo de institución en el que se encuentren depositados los originales.

Por tanto, y para concluir, el *Directorio y recolector de colecciones y recursos digitales* es una herramienta que permite potencialmente la coordinación del desarrollo de colecciones digitales, puesto que su consulta nos da a conocer qué obras y qué colecciones se han digitalizado. Por otra parte, suministra una fuente de información de primer orden mediante el acceso a los recursos digitales que recolecta. Y, por último, ensanchará y profundizará el acceso a otros muchos tipos de recursos digitales accesibles en la red como los catálogos bibliográficos automatizados, los *e-prints* y los documentos de archivos. ■

Notas

¹ Forman parte del grupo de trabajo Laura Cerezo (Junta de Andalucía), Mar Sarto (Gobierno de Aragón), Ana Isabel Cámara (Principado de Asturias), Alejandro Carrión (Junta de Castilla y León), Glòria Pérez-Salmerón (Generalitat de Catalunya), Belén Díaz (Junta de Extremadura), Ignacio Cabano (Xunta de Galicia), María Dolores Ramírez (Gobierno de La Rioja), María de la Villa (Comunidad de Madrid), Marta Natividad Martínez (Comunidad de Madrid), Ángel Peñalver (Región de Murcia), Jaime J. Chiner (Generalitat Valenciana), Dolores Rodríguez Fuentes (Ministerio de Cultura), María Isabel Cuadrado (Ministerio de Cultura) y María Luisa Martínez-Conde (Ministerio de Cultura), coordinadora.

² <http://www.openarchives.org/>

³ <http://es.dublincore.org/>

⁴ *ISO 15836:2003 Information and documentation : The Dublin Core metadata element set.* <http://www.iso.ch/iso/>

⁵ <http://www.theeuropeanlibrary.org/portal/index.html>

⁶ <http://www.mcu.es/bibliotecas/CE/ProgramasCooperativos/JornadasCooperacion.html>

⁷ <http://www.loc.gov/marc/community/eccihome.html>

⁸ <http://www.openarchives.org/service/listproviders.html> [Consultado el 6 de marzo de 2007].

⁹ <http://www.opendoar.org/>

¹⁰ <http://roar.eprints.org/>

¹¹ http://ec.europa.eu/information_society/activities/digital_libraries/doc/recommendation/recommendation/en.pdf

¹² <http://www.michael-culture.eu/>

¹³ *Correo bibliotecario: boletín de noticias de la Subdirección General de Coordinación Bibliotecaria*, n. 89 (marzo de 2006).

<http://www.correobibliotecario.com/>

¹⁴ *Implementation Guidelines for the Open Archives Initiative Protocol for Metadata Harvesting. Protocol Versión 2.0 of 2002-06-14.*

<http://www.openarchives.org/OAI/2.0/guidelines.htm>

Ficha Técnica

AUTOR: Grupo de Trabajo de Colecciones Digitales del Ministerio de Cultura y las Comunidades Autónomas.

ILUSTRACIONES: Ministerio de Cultura.

TÍTULO: *Cómo encontrar la aguja en el pajar de la web. Directorio y recolector de colecciones y recursos digitales* (<http://roai.mcu.es>).

RESUMEN: En este artículo se explica qué es y qué contiene el *Directorio y recolector de colecciones y recursos digitales* del MCU, una aplicación nacida hace un año que permite acceder a una gran información sobre las colecciones bibliográficas de España y fuera de España. Hoy por hoy es el primer servicio de este tipo desarrollado en nuestro país.

MATERIAS: Bibliotecas / Nuevas Tecnologías / Internet / Acceso a la Información / Digitalización / Bases de Datos / España.