

# ZAGUÁN: REPOSITORIO INSTITUCIONAL DE LA UZ

MIGUEL MARTÍN

# ZAGUAN

Repositorio Institucional de la UZ



# Miguel Martín

## Ingeniero en Informática

Proy. Informatización de Bibliotecas Universidad de Zaragoza

[miguelm@unizar.es](mailto:miguelm@unizar.es) | [@miguelm\\_uz](https://twitter.com/miguelm_uz)



# ¿Informatic@ o bibliotecari@?



# Índice

Origen del Repositorio

CDS Invenio: arquitectura y módulos

Formatos de archivos a texto completo

Flujos de incorporación de información

Colecciones del Repositorio

Feedback: Google Analytics + AWStats

Plan de Preservación: la realidad de las Instituciones



# Índice

## Origen del Repositorio

CDS Invenio: arquitectura y módulos

Formatos de archivos a texto completo

Flujos de incorporación de información

Colecciones del Repositorio

Feedback: Google Analytics + AWStats

Plan de Preservación: la realidad de las Instituciones



# Qué es

- Repositorio = del lat.: repositorium. Armario, alacena, “*lugar donde se guarda algo*” [RAE]
- Repositorio = datos + infraestructuras + personas
- Misiones
  - **Archivar** (depósito)
  - **Preservar** (largo plazo, obsolescencia tecn)
  - **Difundir** (web / OAI-PMH)

# El origen

## **Digitalización de Fondo Histórico:**

Gobierno de Aragón: Ayudas a la generación y publicación de contenidos, años 2005, 2006, 2007 y 2008.

Ministerio de Cultura: Convocatorias de ayudas para creación y transformación de recursos digitales y su difusión y preservación mediante repositorios OAI, años 2007, 2008 y 2009



# Elección del Software

(análisis hace varios años)



**Desarrollado por el CERN**

**1.000.000 registros**

**+500 colecciones**

**MARCXML (MARC21)**

**Python + MySQL**

**Apache + mod\_python (mod\_wsgi)**

**Modular y flexible** (y menos complejo)

Servicio de soporte (¡barato!)

Licencia GNU/GPL

Búsquedas tipo Google (sencillas, rápidas)

Búsquedas con operadores booleanos complejos,  
expresiones regulares

Resultados similares (si no hay match directo)

Web 2.0 (review+comentarios)

Colecciones personales (cestas)

Alertas por email (automatizadas)



MIT (y Hewlett-Packard)

Adoptado por muchas instituciones

DublinCore

JAVA (PostgreSQL)

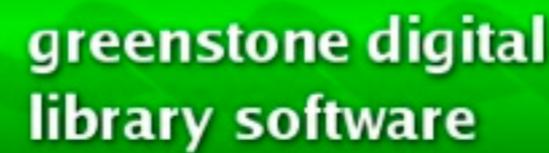
Tomcat

Customización (solo) del CSS

OpenURL

Identificadores CNRI Handle

**Relaciones (lógicas) entre objetos  
digitales**



Universidad de Waikato (New Zealand)

Soportado por UNESCO

DublinCore

PERL, GDBM, MG

Gestión de privilegios de usuarios más  
pobre

No admite subidas de contenido por  
usuarios que no sean "Content Managers".



Universidad de Cornell & Universidad de  
Virginia

Soportado por Melon Foundation y más  
tarde organización Fedora Commons

Funcionalidad de "core" (proporciona  
almacenamiento y mantiene objetos  
digitales, además de API's para  
desarrolladores)

¡Carece de interfaz propia! (requiere de  
esfuerzos iniciales importantes)

"Whenever you find yourself on the side of the majority, it is time to pause and reflect." - Mark Twain, Notebook, 1904

Más información: [http://nrgl.techlib.cz/images/Open\\_source.pdf](http://nrgl.techlib.cz/images/Open_source.pdf)

# Índice

Origen del Repositorio

CDS Invenio: arquitectura y módulos

Formatos de archivos a texto completo

Flujos de incorporación de información

Colecciones del Repositorio

Feedback: Google Analytics + AWStats

Plan de Preservación: la realidad de las Instituciones



# CDS Invenio



## Key features

- Configurable portal-like interfaces for hosting various kinds of collections
- Powerful search engine with Google-like syntax, including parallel searching of external collections
- Extensible metadata representation (MARC XML) to handle virtually any kind of document (articles, books, photos, videos and more)
- Flexible document type submission and approbation workflow
- User personalisation, including document baskets and email notification alerts
- User collaboration, reviews, comments, knowledge sharing
- Multilingual interface available in 20 languages, Unicode compliant (UTF-8)
- Compliant to Open Archive Initiative protocol for metadata harvesting
- Free software (GNU GPL)

Try the CERN server:  
Download Invenio:  
Contact us:

<http://cdsweb.cern.ch/>  
<http://invenio.cern.ch/>  
[cds.support@cern.ch](mailto:cds.support@cern.ch)

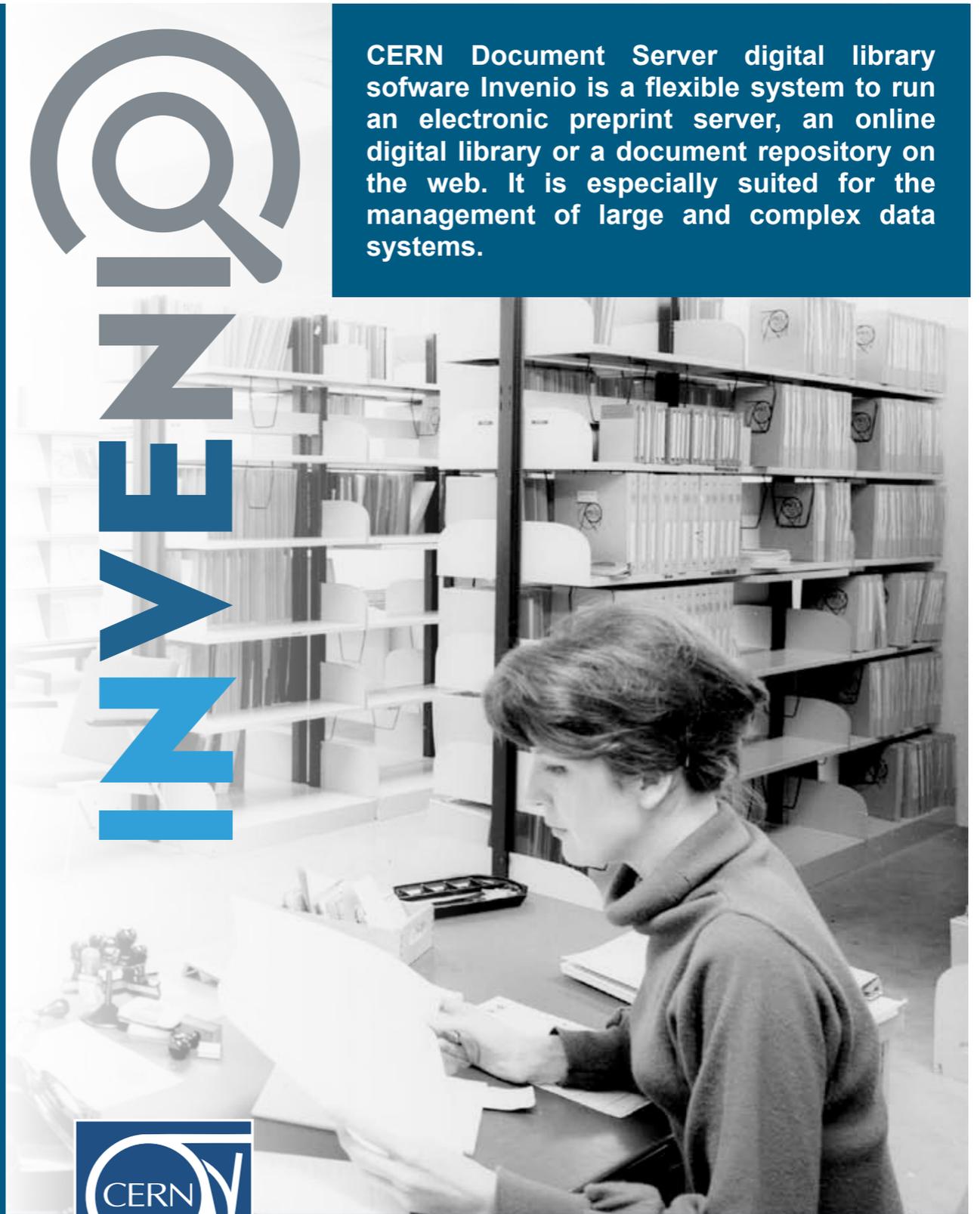


INVENIO

CERN Document Server digital library software Invenio is a flexible system to run an electronic preprint server, an online digital library or a document repository on the web. It is especially suited for the management of large and complex data systems.



<http://invenio.cern.ch/>



## Navigable collection tree

CERN Document Server

Over 800,000 bibliographic records, including 300,000 fulltext documents, of interest to people working in particle physics and related areas. Covers preprints, articles, books, journals, photographs, and much more.

Search 914,585 records for:  any field

Narrow by collection:

- Articles & Preprints (753,322)
  - Published Articles (283,895)
  - Preprints (364,367)
  - Theses (42,560)
  - Reports (5,937)
  - CERN internal notes (17,000)
  - Committee Documents (27,603)
- Books & Proceedings (90,694)
  - Books (37,174)
  - Proceedings (15,866)
  - Standards (3,754)
- Presentations & Talks (10,143)
  - Conferences, Announcements (14,119)
  - Academic Training Lectures (194)
  - Summer Student Lectures (199)
  - General Talks (32)
  - Workshops (29)
- Periodicals & Progress Reports (3,549)
  - Periodicals (2,897)
  - Progress Reports (652)
- Multimedia & Outreach (26,282)
  - Photos (9,742)
  - Videos (224)
  - Press (18,840)
  - Audio recordings (176)

Focus on:

- CERN Articles & Preprints (93,357)
  - CERN Published Articles (47,370)
  - CERN Preprints (14,642)
  - CERN Theses (4,899)
  - CERN Reports (1,199)
  - Committee Documents (97,600)
- CERN Series (2,000)
  - CERN Yellow Reports (1,119)
  - Academic Training Lectures (94)
  - Summer Student Lectures (99)
  - General Talks (32)
- CERN Departments (95,929)
  - Accelerator Technologies (AT) (4,891)
  - Accelerators & Beams (AB) (16,599)
  - Physics (P) (799)
  - Human Resources (HR) (3)
  - Information Technology (IT) (3,183)
  - Physics (PhD) (17,276)
  - Secondary/General (SG) (74,916)
  - Technical Support (TS) (1,195)
- CERN Experiments (15,123)
  - LHC Experiments (5,276)
  - LHC Experiments (8,479)
  - Neutrino Experiments (378)

- Documents organized in collections
- Regular and virtual collection trees
- Customizable portalboxes for each collection
- At CERN, about 1,000,000 documents in 500 collections

## Flexible metadata

CERN Document Server

Format: HTML | RSS | PDF | EPUB | MARC | MARCXML

Display: Life at CERN CERN-GE-808033

### Tim Berners-Lee : World-Wide Web inventor

Date: 26 Jul 1998

You can look at these photographs in the following format:

Conference "Internet Web, What's next?" on 26 June 1998 at CERN. Tim Berners-Lee, inventor of the World Wide Web and Director of the WWW, explains how the Web came to be and gives his views on the future.

Conference "Internet Web, What's next?" on 26 Jun 1998 at CERN. Tim Berners-Lee, inventor of the World Wide Web and Director of WWW, explains how the Web came to be and gives his views on the future.

© CERN. Copyright: the use of photos requires prior authorization from CERN. If you need pictures in high resolution, please send a request to the CERN PhotoLab with the reference of the picture: [photo/808033](#)

Appear in:
 

- Le Pays Genevois - 3 Jul 1998
- Le Temps - 27 Jun 1998
- La Tribune de Genève - 27 Jun 1998

Related links:
 

- CERN Press Release 2002-26 Jun 2002 (English)
- CERN Press Release 2002-26 Jun 2002 (French)
- Journal: Web, What's next? - 26 Jun 2002
- CERN Bulletin - 26 Jun 1998 (English)
- CERN Bulletin - 26 Jun 1998 (French)

Record created: 1998-07-02, last modified: 2008-01-12

Display: [HTML](#) [XML](#) [MARC](#) [EPUB](#) [RSS](#)

- Standard library metadata format (MARC)
- Handling articles, books, thesis, photos, videos, museum objects, and more.
- Customizable display and linking rules

Multiple output formats:



## Powerful search engine

- Specifically designed indexes to provide Google-like search speed for repositories of up to 1,500,000 records
- Parallel searching of external collections
- Customizable simple and advanced search interfaces
- Combined metadata, fulltext and citation search in one go
- Results clustering by collection
- Flexible ranking capabilities

Search:  any field

Search also:
 

- CERN Indico
- KISS Preprints
- SPIRES HEP

Search collections:

Sort by:  Display results:  Output format:

Results overview: Found 2,248 records in 0.43 seconds.

- Articles & Preprints: 1,366 records found
- Books & Proceedings: 131 records found
- Presentations & Talks: 12 records found
- Multimedia & Outreach: 111 records found
- Archives: 18 records found

Articles & Preprints 1,366 records found 1 - 10

1. **Causality and the speed of sound / Ellis, G; Maartens, R; MacCallum, M**  
A usual causal requirement on a viable theory of matter is that the speed of sound be at most the speed of light [...]  
gr-qc/0703121, 29 Mar 2007. - 6 p. [Fulltext](#)  
[Detailed record](#) - [Similar records](#)
2. **The effects of environment on morphological evolution between 0.4<math>z</math> and 1.2 in the COSMOS Survey / Capak, P L; Rosser, R G; Ems, H; Hoopes, S; Mowbray, B; Scoville, N Z; Spoto, C; Kistnerick, A**  
We explore the evolution of the morphology density relation using the COSMOS ACS survey and previous cluster studies. [...]  
astro-ph/0703088, 28 Mar 2007. - 12 p. [Fulltext](#)  
[Detailed record](#) - [Similar records](#)
3. **Polarization Effects in Weak Lensing / Hoopes, J; Massey, B; Ellis, R**  
Weak gravitational lensing is a promising probe of dark matter and dark energy requiring accurate measurement of the shapes of lens, distant galaxies. [...]  
astro-ph/0704171, 17 Mar 2007. - 17 p. [Fulltext](#)  
[Detailed record](#) - [Similar records](#)

## User personalization and collaborative tools

- User-defined baskets of documents
- User-defined automated email alerts
- Commenting and reviewing
- Knowledge sharing within user groups
- Search interface available in 20 languages

CERN Document Server

Home > Your Account > Group baskets > [Add a new basket](#) > [CERN baskets](#) > Display baskets

### Display baskets

Personal baskets Group baskets

it-dep-uds-ods-members@cern.ch [CERN (external)] (1)

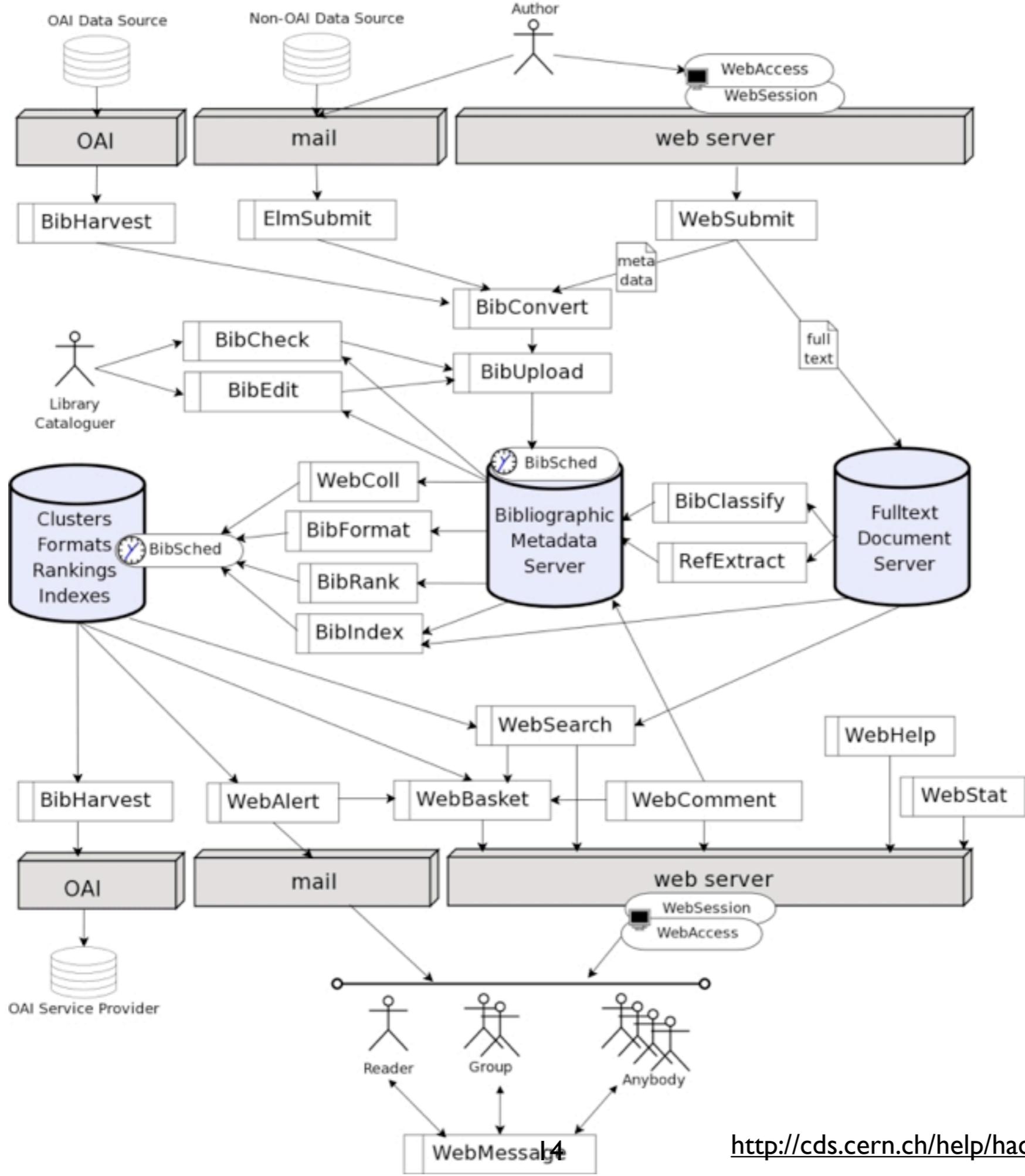
Cool stuff  
2 records - last update: 02 Apr 2007, 09:38

On the stability of four-legged tables / [Joshi, Anil](#)  
We prove that a perfect square table with four legs, placed on continuous irregular ground with a local slope of at most 1, can stand. [...]  
CERN-GE-808033, CERN-PH-TH-2005-197. Geneva : CERN, 17 Oct 2005. - 10 p. [Fulltext](#) - Submitted to: [Phys. Lett. B](#)  
[Details and comments](#)

Information management : a proposal / [Berners-Lee, Timothy J](#)  
CERN-DO-80-011-OC. Geneva : CERN, Mar 1989. - 15 p. [Fulltext](#) - CERN Series: [reports](#)  
[Details and comments](#)

# Cómo es

- Python + Apache + MySQL
- Pre-Sept-2009: Python 2.4 + mod\_python
- Post-Sept-2009: Python 2.6 + mod\_wsgi



<http://cds.cern.ch/help/hacking/modules-overview>

# Módulos de interés para el administrador

# webaccess: Autenticación (interna/LDAP/ otros) y autorizaciones (rol/acción/autorización)

- **ROL** = grupo (definición tipo firewall)

**ALLOW/DENY ANY/ALL**

**ALLOW/DENY FROM/UNTIL "YYYY-MM-DD"**

or

**ALLOW/DENY [NOT] *field* {one or more values}**

**uid:** an integer representing the user id  
**nickname:** the nickname of the user  
**email:** the email of the user  
**group/groups:** local or external group to which the user belong  
**guest:** 1 if the user is a guest (not logged), 0 otherwise  
**remote\_ip:** the remote ip address of the user who is browsing  
**remote\_host:** the remote hostname of the user who is browsing  
**referer:** the webpage from where the user is coming from  
**uri:** the uri the user is visiting  
**agent:** the agent string describing the user's browser

Ejemplo:

```
allow email /*@gmail.com/,/*@hotmail.com/  
deny group badguys  
allow remote_ip "155.0.0.0/24"  
allow uid 1  
deny all
```

# webaccess

- Acción = actuaciones que se pueden realizar en la plataforma. Centradas en quién puede acceder a menús de configuración (cfgXXX), quién puede ejecutar tareas (runYYY) desde línea de comandos y quién puede enviar (submit) y moderar (referee)

```
accdelegaterole
  cfgbibformat
  cfgbibharvest
  cfgbibindex
  cfgXXX...
  referee
runbibclassify
  runbibedit
  runbibformat
  runYYY
  submit
viewrestrcoll
viewrestrdoc
```

# webaccess

- Autorización = relación rol-acción  
(con/sin parámetros)
  - Rol - Action - Parameters
  - moderadorestesis - submit - act:DEMOTHE, doctype:APP

# websubmit

- **Módulo responsable de definir `doctype` y procesos de envío:**
- Varias “acciones” (workflows) predefinidas: `SBI`, `MBI`, `APP`, `SRV`
- Cada “acción” está formada por un formulario (`interface`) y funciones (`python`) que se ejecutarán una vez pulsado “enviar”.
- Hay “control de versiones” de ficheros a texto completo.

# websubmit

Firefox File Edit View History Bookmarks Tools Window Help

Configure Document Type - Repositorio Digital de la Universidad de Zaragoza

zaguan.unizar.es/admin/websubmit/websubmitadmin.py/doctypeconfigure?doctype=TAZ

[Edit Details](#)

Submissions of Document Type TAZ:

Action	Creation Date	Modification Date	Displayed?	No. Pages	Button Order	Status Text	Level	Score	Stpage	End Text	View Submission Interface	View Submission Functions	Edit Submission Details	Delete Submission
MBI	2010-01-26	2011-11-28	Y	1	2			0	0		<a href="#">view interface</a>	<a href="#">view functions</a>	<a href="#">edit submission</a>	<a href="#">delete submission</a>
SBI	2010-01-15	2012-11-19	Y	1	1		1	0	1		<a href="#">view interface</a>	<a href="#">view functions</a>	<a href="#">edit submission</a>	<a href="#">delete submission</a>
SRV	2010-05-21	2011-09-08	Y	1	4		1	1	0		<a href="#">view interface</a>	<a href="#">view functions</a>	<a href="#">edit submission</a>	<a href="#">delete submission</a>

Add a new Submission:

[APP] Approve Record

Categories of Document Type TAZ:

Categ ID	Description				
PFC	Proyecto Fin de Carrera			<a href="#">edit</a>	<a href="#">delete</a>
TFM	Trabajo Fin de Máster			<a href="#">edit</a>	<a href="#">delete</a>
TFG	Trabajo Fin de Grado			<a href="#">edit</a>	<a href="#">delete</a>

Add a new Category:

ID:  Description:

Manage Referees for Document Type TAZ:

Referee

General Referee(s)

hperez@unizar.es  
hypatia@unizar.es  
miguelm@unizar.es  
zaguan@unizar.es

Find: 2415

20

# websubmit

<http://www.leccionespracticas.com/cds-invenio/cds-invenio-understanding-websubmit/>

# bibupload

- Otra forma de depositar contenidos (línea de comando) vía fichero que contiene los metadatos de los registros en formato MARCXML (<http://www.loc.gov/standards/marcxml/>)
- Desventaja para personal no bibliotecario: Hay que conocer el formato MARCXML (y seguirlo a rajatabla no es sencillo...)
- `bibupload -i /home/miguelm/my_record_list.xml`

# bibharvest/oaiarchive

- Bibharvest permite obtener registros desde otros repositorios OAI-PMH (<http://www.openarchives.org/pmh/>) y poner nuestros registros a disposición de otros (oaiarchive)
- Más detalles sobre preguntas a servidores OAI-PMH: <http://www.leccionespracticas.com/cds-invenio/oai-pmh-verb-parameters/>

# bibedit

- Permite editar los metadatos (MARCXML) de los registros depositados
- Versiones 0.99.x tienen pequeños bugs (duplican etiquetas, etc)
- Versiones 1.x han modificado este módulo y disponen de editor más completo.

**Record #1**

Your changes are TEMPORARY. To save this record, please click on submit.

Record #1  
Action: [Cancel](#) Record: [Add Field](#) | [Delete](#) Display: [Verbose](#) | [MARC](#)

037__	\$\$a	CERN-EX-0106015	 
100__	\$\$a	Photolab	 
245__	\$\$a	ALEPH experiment: Candidate of Higgs boson production	 
246_1	\$\$a	Expérience ALEPH: Candidat de la production d'un boson Higgs	 
260__	\$\$c	14 06 2000	 
340__	\$\$a	FILM	 
520__	\$\$a	Candidate for the associated production of the Higgs boson and Z boson. Both, the Higgs and Z boson decay into 2 jets each. The green and the yellow jets belong to the Higgs boson. They represent the fragmentation of a bottom and anti-bottom quark. The red and the blue jets stem from the decay of the Z boson into a quark anti-quark pair. Left: View of the event along the beam axis. Bottom right: Zoom around the interaction point at the centre showing details of the fragmentation of the bottom and anti-bottom quarks. As expected for b quarks, in each jet the decay of a long-lived B meson is visible. Top right: "World map" showing the spatial distribution of the jets in the event.	 
595__	\$\$a	Press	 
65017	\$\$2	SzGeCERN	 
	\$\$a	Experiments and Tracks	 
6531__	\$\$a	LEP	 
8560__	\$\$f	neil.calder@cern.ch	 
8564__	\$\$u	<a href="http://invenio-demo.cern.ch/record/1/files/0106015_01.jpg">http://invenio-demo.cern.ch/record/1/files/0106015_01.jpg</a>	 
8564__	\$\$q	<a href="http://invenio-demo.cern.ch/record/1/files/icon-0106015_01.gif">http://invenio-demo.cern.ch/record/1/files/icon-0106015_01.gif</a>	 
	\$\$x	icon	 
909C0	\$\$o	0003717PHOPHO	 

# Enriquecimiento de registros

- **BibClassify:** Extraer keywords basados en la recurrencia de términos específicos, elegidos de un vocabulario controlado. Vocabulario controlado es un tesoro de todos los términos que son relevantes en un contexto específico. BibClassify se basa en taxonomías expresadas en RDF/SKOS. Algunos tesauros: [http://www.fbi.fh-koeln.de/institut/labor/Bir/thesauri\\_new/thesen.htm](http://www.fbi.fh-koeln.de/institut/labor/Bir/thesauri_new/thesen.htm)
- **RefExtract:** Extraer listas de referencias de los PDF's [<http://invenio-software.org/wiki/Development/Modules/RefExtract>]

# bibformat

- Permite distintas visualizaciones de un mismo registro y niveles de caché sobre el registro formateado
- Formatos básicos: HTML breve (para listados de múltiples registros, cacheado por defecto), HTML detallado (para visualización de registro única, sin cache por defecto) y podemos construir formatos personalizados (p ejemplo RSS).

# bibformat

Gestionar los formatos de salida – Repositorio Digital de la Universidad de Zaragoza

zagan.unizar.es/admin/bibformat/bibformatadmin.py/output\_formats\_manage?ln=es

Código	Nombre	Descripción	Estado	Fecha de la última modificación	Acción [?]
EXCEL	Excel	Excel csv output	OK	Tue Aug 26 11:27:58 2008	Delete
HB	HTML brief	HTML brief output format, used for search results pages.	OK	Wed Sep 7 10:41:05 2011	Delete
HBGEOL	HTML Brief Geol	HTML Brief para Grupos	OK	Tue Oct 20 11:00:39 2009	Delete
HC	photo captions only	HTML caption-only output format for photos.	OK	Tue Aug 26 11:27:57 2008	Delete
HD	HTML detailed	HTML detailed output format, used for Detailed record pages.	OK	Wed Sep 7 11:57:50 2011	Delete
HDACT	Actions mini-panel	Used to display actions in mini-panel of detailed record pages.	OK	Fri Nov 12 08:03:23 2010	Delete
HDFILE	File mini-panel	Used to show fulltext files in mini-panel of detailed record pages.	OK	Thu Apr 28 09:04:32 2011	Delete
HDREF	References tab	Display record references in References tab.	OK	Tue Aug 26 11:27:59 2008	Delete
HLIGHT	HTML Light (invocaciones externas)	HTML brevísimo para devolver a las peticiones Ajax externas que se hagan.	OK	Mon Oct 19 11:07:45 2009	Delete
HP	portfolio	HTML portfolio-style output format for photos.	OK	Tue Aug 26 11:27:57 2008	Delete
HS	HTML similarity	Very short HTML output for similarity box ( <i>people also viewed...</i> ).	OK	Mon Apr 11 11:04:26 2011	Delete
HX	BibTeX	BibTeX.	OK	Tue Aug 26 11:27:58 2008	Delete
METS	METS	Formato METS.	OK	Mon Nov 24 13:55:03 2008	Delete
PREMIS	PREMIS	PREMIS	OK	Fri Dec 5 11:36:53 2008	Delete
UNTLD	Untitled		OK	Fri Sep 26 12:35:49 2008	Delete
UNTLD2	Untitled		OK	Tue Oct 20 10:58:27 2009	Delete
XD	Dublin Core	XML Dublin Core.	OK	Wed Nov 24 11:52:09 2010	Delete
XE	EndNote	XML EndNote.	OK	Tue Aug 26 11:27:58 2008	Delete
XM	MARCXML	XML MARC.	OK	Tue Aug 26 11:27:58 2008	Delete
XN	NLM	XML NLM.	OK	Tue Aug 26 11:27:58 2008	Delete
XOAIDC	OAI DC	OAI DC.	OK	Thu Aug 4 10:36:54 2011	Delete
XR	RSS	RSS.	OK	Fri Nov 5 14:33:35 2010	Delete
XW	RefWorks	RefWorks.	OK	Tue Aug 26 11:27:59 2008	Delete

Añadir un nuevo formato de salida

Find: 2415    Next    Previous    Highlight all    Match case

Instrucciones depósito

# bibformat

```
<span class="quicknote">
  Formatting record 4504 with output format XM.
</span>
<br/><span class="quicknote">
  Found preformatted output for record 4504 (cache updated on 2012-09-11 11:46:30).
</span><br/><record>
<controlfield tag="001">4504</controlfield>
<datafield tag="260" ind1=" " ind2=" ">
  <subfield code="a">Zaragoza</subfield>
  <subfield code="b">Universidad de Zaragoza</subfield>
  <subfield code="c">2010</subfield>
</datafield>
<datafield tag="300" ind1=" " ind2=" ">
  <subfield code="a">mult. p</subfield>
</datafield>
<datafield tag="856" ind1="0" ind2=" ">
  <subfield code="f">miguelm@unizar.es</subfield>
</datafield>
<datafield tag="910" ind1="2" ind2=" ">
  <subfield code="a">Lenguajes y sistemas informáticos</subfield>
  <subfield code="b"></subfield>
</datafield>
<datafield tag="830" ind1=" " ind2=" ">
  <subfield code="a">CPS</subfield>
</datafield>
<datafield tag="653" ind1="1" ind2=" ">
  <subfield code="a">Fotografía computacional</subfield>
</datafield>
<datafield tag="653" ind1="1" ind2=" ">
  <subfield code="a">Mapeado de tono</subfield>
</datafield>
<datafield tag="653" ind1="1" ind2=" ">
  <subfield code="a">Tone mapping</subfield>
</datafield>
<datafield tag="653" ind1="1" ind2=" ">
  <subfield code="a">Reverse tone mapping</subfield>
</datafield>
<datafield tag="653" ind1="1" ind2=" ">
  <subfield code="a">Subexposición</subfield>
</datafield>
<datafield tag="653" ind1="1" ind2=" ">
  <subfield code="a">Sobreexposición</subfield>
</datafield>
<datafield tag="950" ind1=" " ind2=" ">
  <subfield code="a"></subfield>
</datafield>
<datafield tag="700" ind1=" " ind2=" ">
  <subfield code="a">Gutiérrez Pérez, Diego</subfield>
  <subfield code="a">di</subfield>
</datafield>
```

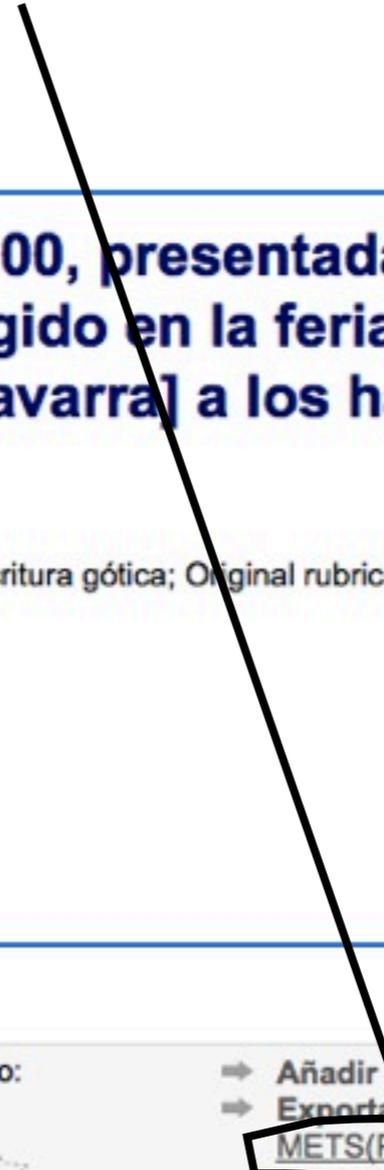
# bibformat

The screenshot shows a Firefox browser window with the following elements:

- Browser Title Bar:** Firefox File Edit View History Bookmarks Tools Window Help. System tray on the right shows 14, a speaker icon, a clock, a Spanish flag, and the date 'lun 13:00'.
- Page Title:** Fotografía Computacional para imágenes subexpuestas y sobreexpuestas | Trabajos académicos - Repositorio Digital de la Universidad de Zaragoza
- Address Bar:** zaguan.unizar.es/record/4504?of=HD&verbose=9
- Navigation Bar:** Buscar, Enviar, Ayuda, Sus opciones, Administracion. User name 'Miguel salir' is visible on the right.
- Breadcrumbs:** Página principal > Trabajos académicos > Fotografía Computacional para imágenes subexpuestas y sobreexpuestas | Trabajos académicos > HTML detailed
- Log/Status:** Formatting record 4504 with output format HD. No preformatted output found for record 4504. Formatting record 4504 on-the-fly. Using HTMLdetailedTAZ.bft template for record 4504.
- Record ID:** TAZ-PFC-2010-002
- Section Header:** **Fotografía Computacional para imágenes subexpuestas y sobreexpuestas**
- Authors:** Martín González, Miguel Gutiérrez Pérez, Diego (dir.)
- Institution:** Universidad de Zaragoza, Escuela de Ingeniería y Arquitectura, 2010  
Departamento de Informática e Ingeniería de Sistemas, Área de Lenguajes y Sistemas Informáticos
- Field:** Ingeniero en Informática
- Resumen:** Hoy en día, las imágenes de alto rango dinámico (High Dynamic Range Imaging, o HDRI) reciben una cuidadosa atención desde el punto de vista cinematográfico e industrial. El margen dinámico es el rango de iluminación (EV) en el que la cámara es capaz de ver detalle. Un margen dinámico elevado permite que, pese a que la instantánea tenga zonas con luces fuertes y sombras oscuras, se mantenga el detalle en ambas iluminaciones. Cuando se supera el margen dinámico, las zonas sobreexpuestas1 aparecen totalmente blancas y las sombras totalmente negras, de manera
- Link:** + Leer más...
- Palabra(s) clave (del autor):** Fotografía computacional ; Mapeado de tono ; Tone mapping ; Reverse tone mapping ; Subexposición ; Sobreexposición
- Área de conocimiento:** Lenguajes y sistemas informáticos
- Tipo de Trabajo Académico:** Proyecto Fin de Carrera
- Notas:** El trabajo está acompañado de un artículo presentado en el CEIG2008. La versión preprint de este artículo se encuentra en <http://miguelmartin.es/martin08.pdf>
- License:** CC BY-NC-SA. Registrado por la Universidad de Zaragoza bajo la licencia Creative Commons.
- Search Bar:** Find: 2415. Buttons for Next, Previous, Highlight all, Match case.
- Taskbar:** Contains icons for various applications including Firefox, Photoshop, and a calendar.

# bibformat

Permite la creación de nuevos formatos de salida...



Información | Discusión | Estadísticas de uso

## Testimonio de protesta de agravio, Estella, 1300, presentada por Sancho Pontz, ante el notario Thomas Xemenis, por el peaje exigido en la feria de Estella, alegando el privilegio otorgado por el rey Felipe [I de Navarra] a los habitantes de Sangüesa

**Editor:** Estella ; (Era 1338 [1300])  
**Notas:** Anotación manuscrita de época posterior en el verso; Escritura gótica; Original rubricado; Texto en lengua romance navarra  
**Idioma:** Español  
**Descripción física:** 1 h. (20 lin.) : 18x21 cm

[Ver en el catálogo](#)

Registro creado el 2011-10-28, última modificación el 2013-05-21 [Ver similares](#)

Texto completo  
 Transcripción  
Checksum (md5)

Registrado por la UZ bajo la licencia Creative Commons Reconocimiento-No comercial 3.0 España License.

Valore este documento:  
  
(Sin revisiones)

- ⇒ Añadir a la lista personal
- ⇒ Exportar como BibTeX, MARC, MARCXML, DC, **METS(PREMI)**, OAI-DC
- ⇒ Para ver el documento necesita instalar el plug-in gratuito DjVu.
- ⇒ Editar este registro

Me gusta <

# webstyle

Plantillas para formateo de sitio web.

Definición del marcado HTML, CSS, JS

Reescritura de plantillas por defecto usando  
convención de nombre:

```
/soft/cds-invenio/etc/invenio-local.conf:CFG_WEBSTYLE_TEMPLATE_SKIN = buz
```

```
/soft/cds-invenio/lib/python/invenio/websession_templates_buz.py  
/soft/cds-invenio/lib/python/invenio/webstyle_templates_buz.py  
/soft/cds-invenio/lib/python/invenio/webalert_templates_buz.py  
/soft/cds-invenio/lib/python/invenio/websearch_templates_buz.py
```

# bibsched

- El *scheduler* de tareas
- Múltiples opciones (prioridades de tareas, tiempo de sleep, etc)
- Cambio de prioridades en MySQL (mejorable)

```
mysql> describe schTASK;
```

Field	Type	Null	Key	Default	Extra
id	int(15) unsigned	NO	PRI	NULL	auto_increment
proc	varchar(20)	NO		NULL	
host	varchar(255)	NO		NULL	
user	varchar(50)	NO		NULL	
runtime	datetime	NO	MUL	NULL	
sleeptime	varchar(20)	YES		NULL	
arguments	mediumblob	YES		NULL	
status	varchar(50)	YES	MUL	NULL	
progress	varchar(255)	YES		NULL	
priority	tinyint(4)	NO	MUL	0	

```
10 rows in set (0.00 sec)
```

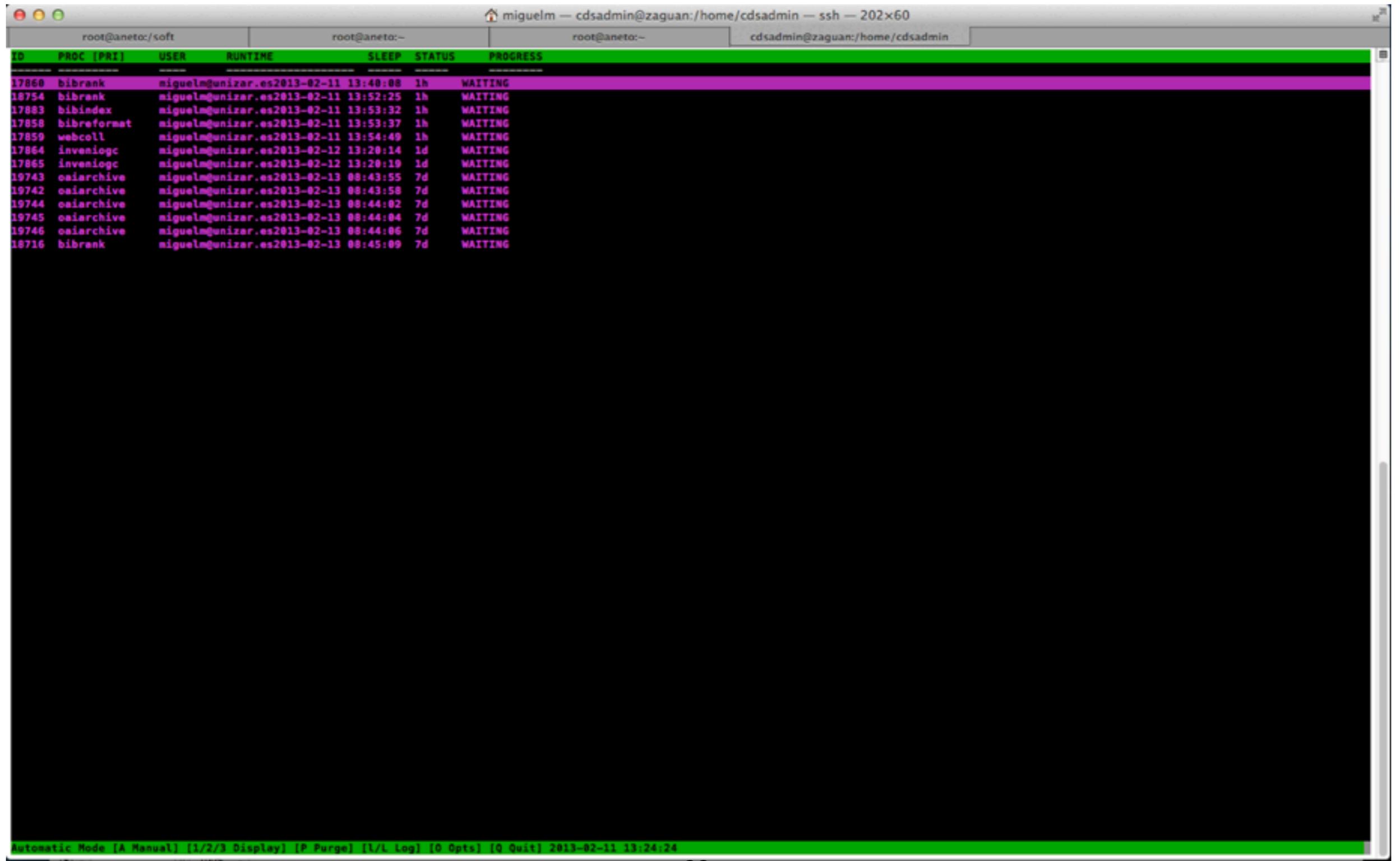
```
mysql> select distinct(status) from schTASK;
```

status
ACK_ERROR_DELETED
DONE
DONE WITH ERRORS_DELETED
ERROR
ERROR_DELETED
KILLED_DELETED
RUNNING
STOPPED_DELETED
STOPPING
STOPPING_DELETED
WAITING
WAITING_DELETED

```
12 rows in set (0.00 sec)
```

```
update schTASK SET priority=2 where priority=5;  
OK, 14882 rows affected (0.40 sec)  
atched: 14882 Changed: 14882 Warnings: 0
```

# bibsched



The image shows a terminal window with a title bar that reads "miguelm — cdsadmin@zaguan:/home/cdsadmin — ssh — 202x60". The terminal displays the output of a 'ps' command, showing a list of processes. The processes are all in a 'WAITING' state. The columns are: ID, PROC [PRI], USER, RUNTIME, SLEEP, STATUS, and PROGRESS. The processes listed are: bibrank, bibindex, bibreformat, webcoll, invenlogc, oaiarchive, and bibrank. The first row is highlighted in purple.

ID	PROC [PRI]	USER	RUNTIME	SLEEP	STATUS	PROGRESS
17860	bibrank	miguelm@unizar.es	2013-02-11 13:40:08	1h	WAITING	
18754	bibrank	miguelm@unizar.es	2013-02-11 13:52:25	1h	WAITING	
17883	bibindex	miguelm@unizar.es	2013-02-11 13:53:32	1h	WAITING	
17858	bibreformat	miguelm@unizar.es	2013-02-11 13:53:37	1h	WAITING	
17859	webcoll	miguelm@unizar.es	2013-02-11 13:54:49	1h	WAITING	
17864	invenlogc	miguelm@unizar.es	2013-02-12 13:20:14	1d	WAITING	
17865	invenlogc	miguelm@unizar.es	2013-02-12 13:20:19	1d	WAITING	
19743	oaiarchive	miguelm@unizar.es	2013-02-13 08:43:55	7d	WAITING	
19742	oaiarchive	miguelm@unizar.es	2013-02-13 08:43:58	7d	WAITING	
19744	oaiarchive	miguelm@unizar.es	2013-02-13 08:44:02	7d	WAITING	
19745	oaiarchive	miguelm@unizar.es	2013-02-13 08:44:04	7d	WAITING	
19746	oaiarchive	miguelm@unizar.es	2013-02-13 08:44:06	7d	WAITING	
18716	bibrank	miguelm@unizar.es	2013-02-13 08:45:09	7d	WAITING	

Automatic Mode [A Manual] [1/2/3 Display] [P Purge] [l/L Log] [O Opt] [Q Quit] 2013-02-11 13:24:24

# bibconvert

- Permite conversión entre distintos formatos XML (XSLT)
- Se proporcionan algunas plantillas de uso general (`$PATH_TO_cds-invenio/etc/bibformat/format_templates/*.xsl`)
- `bibconvert -ctemplate.xsl < input.xml`

# websearch/webcoll

- Permite organizar los registros en colecciones
- Cada colección es = Identificador + Nombre(s) + “Query” (al MARCXML) + Portalboxes (información estática en HTML q se muestra junto a las colecciones)
- Por ejemplo, si queremos que todos los documentos que contengan la keyword ‘bibliotecas’ estén en la colección ‘Documentos de bibliotecas’, haremos la query:
  - `653__a: 'bibliotecas'` (las keywords en MARCXML se almacenan en etiquetas 653, subíndice a)
  - Para abstraernos de las etiquetas y “aglutinar” conjuntos, se definen **índices** (`bibindex`)

# bibindex

- Permite definir índices de búsqueda
- Por ejemplo, en MARCXML los autores pueden ir en etiquetas `100__$` (autores principales) o `700__$` (secundarios). Podemos definir el índice “author” como:
  - `author: 700__%, 100__%`

# bibindex

4. Logical fields overview

Field	MARC Tags	Translations
any field	00%, 01%, 02%, 03%, 04%, 05%, 06%, 07%, 08%, 09%, 1001_a, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%, 19%, 20%, 21%, 22%, 23%, 24%, 25%, 26%, 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 850%, 851%, 852%, 853%, 854%, 855%, 857%, 858%, 859%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%	en, es
title	245%, 245__%, 246__%	en, es
author	720%, 720__%, 1001__%, 700%, 100%, 100__%, 700__%	en, es
abstract	520__%	en, es
keyword	65%, 6531_a	en, es
report number	037__a, 909C0r, 088__a	en, es
subject	650%, 65017a, 65027a	en, es
reference	037__a, 999C5%	en, es
fulltext	8564_u	en, es
collection	980__%	en, es
division	909C0p	en, es
year	260__c	en, es
experiment	909C0e	en, es
record ID	001	en, es
isbn	020__a	None
issn	022__a	None
coden	030__a	None
journal	909C4%	en, es
places	752__%	en, es

Find: 2415    Next    Previous    Highlight all    Match case

Instrucciones depósito

# bibindex

## Variables de configuración de índices

CFG\_BIBINDEX\_CHARS\_ALPHANUMERIC\_SEPARATORS  
CFG\_BIBINDEX\_CHARS\_PUNCTUATION

Definen cómo “romper” las palabras. Por ejemplo, “director-general”: ¿debe incluirse el término como tal? ¿deben incluirse “director”, “general” y “director-general”?

CFG\_BIBINDEX\_PATH\_TO\_STOPWORDS\_FILE

Define la ubicación en el servidor del fichero que lista las “palabras vacías”

CFG\_BIBINDEX\_MIN\_WORD\_LENGTH

Define la longitud mínima de la palabra para que ésta se indexe

CFG\_BIBINDEX\_FULLTEXT\_INDEX\_LOCAL\_FILES\_ONLY

Permite que el sistema indexe únicamente ficheros locales o también ficheros en URLs externas

## Se permite el stemming por idioma/índice

Permite reducir el tamaño del índice. Por ejemplo, la palabra “information” se acortará a “inform”. Del mismo modo, “looking”, “look”, “looked” se acortarán a “look”.

# bibindex

Cuanto más grande sea el índice y mayor sea el número de registros, mejor definidos deberán estar los criterios de relevancia.

ID	Name	Fwd.Idx Size	Rev.Idx Size	Fwd.Idx Words	Rev.Idx Records
1	global	4.02 MB	3.749 MB	89,119	3,958
2	collection	0.001 MB	0.231 MB	3	3,958
3	abstract	2.029 MB	2.082 MB	44,193	3,704
4	author	0.5 MB	0.559 MB	11,691	3,958
5	keyword	0.349 MB	0.404 MB	8,336	3,958
6	reference	0.302 MB	0.314 MB	7,922	3,958
7	reportnumber	0.302 MB	0.314 MB	7,922	3,958
8	title	0.53 MB	0.563 MB	12,375	3,958
9	fulltext	0.0 MB	0.181 MB	0	3,958
10	year	0.002 MB	0.226 MB	5	3,958
11	journal	0.0 MB	0.181 MB	0	3,958

1. Overview of indexes					
ID	Name	Fwd.Idx Size	Rev.Idx Size	Fwd.Idx Words	Rev.Idx Records
1	global	6.437 MB	10.162 MB	109,525	11,537
2	collection	0.01 MB	0.766 MB	68	11,534
3	abstract	4.053 MB	5.484 MB	63,068	6,828
4	author	0.561 MB	1.256 MB	10,713	10,334
5	keyword	0.606 MB	0.869 MB	11,501	6,298
6	reference	0.398 MB	0.632 MB	7,767	6,967
7	reportnumber	0.398 MB	0.632 MB	7,767	6,967
8	title	1.158 MB	1.828 MB	23,685	11,531
9	fulltext	193.521 MB	124.567 MB	4,038,902	11,531
10	year	0.038 MB	0.607 MB	897	10,495
11	journal	0.0 MB	0.438 MB	0	11,701

# bibrank

- Permite definir relevancias de los registros a la hora de mostrar resultados de búsquedas. Varios criterios (número de citas, similitud de palabras, etc)

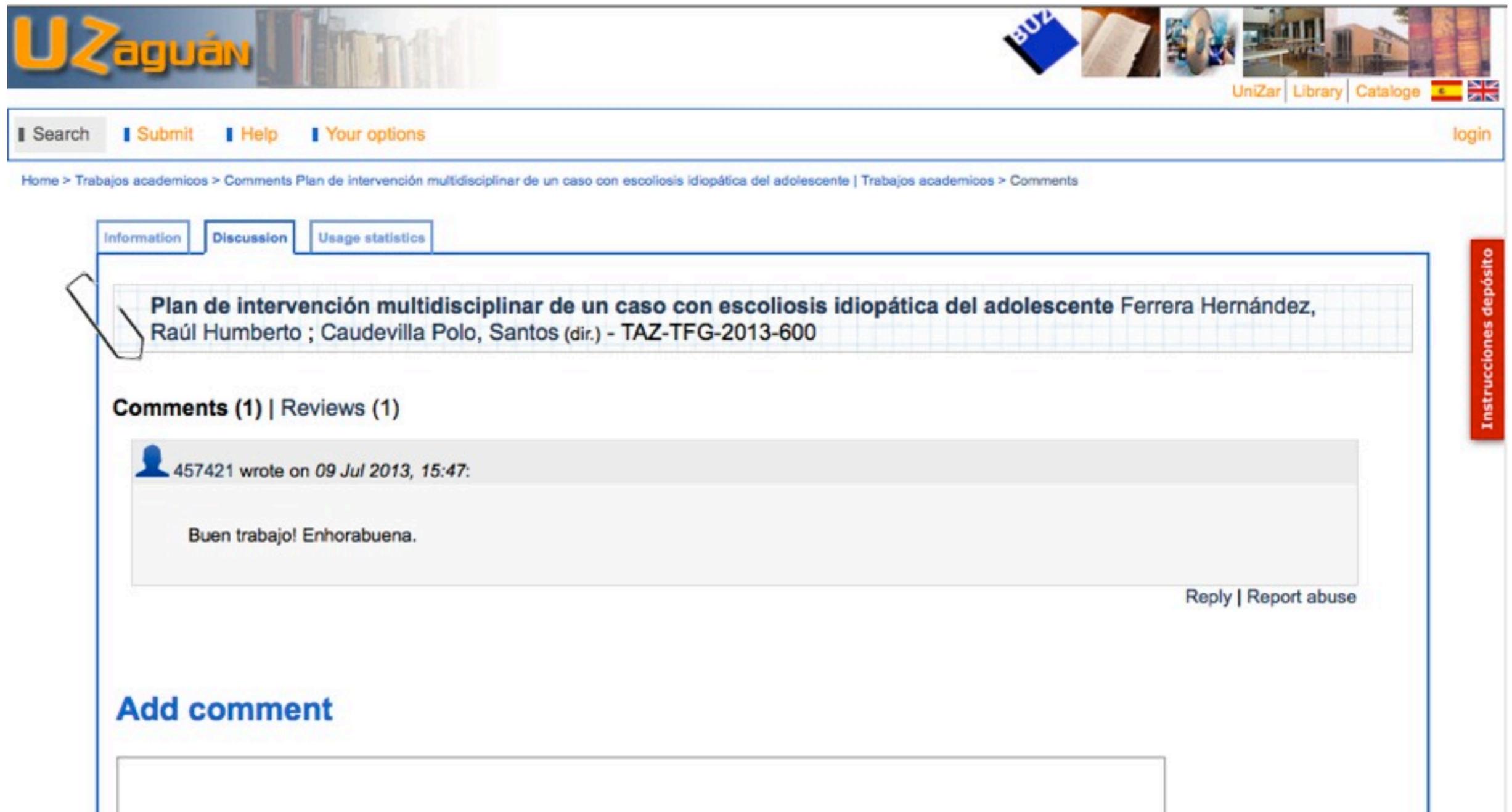
# Módulos de interés para el usuario

# webbasket y webalert

- El usuario puede definir sus “cestas/ listas” (personales o de grupo) como queries.
- Cuando hay registros nuevos que coinciden con la query definida, se muestran en las listas.
- Puede haber listas privadas o públicas
- Se permite generar alertas (que nos avise cuando entren registros que cumplan cierta query)

# webcomment

- Los usuarios pueden hacer comentarios sobre los registros, valorarlos, etc



The screenshot displays the UZaguán webcomment interface. At the top, the UZaguán logo is on the left, and navigation links for UniZar, Library, and Catalogue are on the right. Below the header, there is a search bar and navigation links: Search, Submit, Help, Your options, and a login button. The main content area shows a breadcrumb trail: Home > Trabajos académicos > Comments Plan de intervención multidisciplinar de un caso con escoliosis idiopática del adolescente | Trabajos académicos > Comments. There are three tabs: Information, Discussion, and Usage statistics. The document title is "Plan de intervención multidisciplinar de un caso con escoliosis idiopática del adolescente Ferrera Hernández, Raúl Humberto ; Caudevilla Polo, Santos (dir.) - TAZ-TFG-2013-600". Below the title, it shows "Comments (1) | Reviews (1)". A comment from user 457421, dated 09 Jul 2013, 15:47, reads "Buen trabajo! Enhorabuena." with "Reply | Report abuse" options. At the bottom, there is an "Add comment" section with a text input field.

# webstat

- Los usuarios pueden observar estadísticas sobre los registros (insuficiente...)



The screenshot shows a web interface with three tabs: 'Information', 'Discussion', and 'Usage statistics'. The 'Usage statistics' tab is active. Below the tabs, a record title is displayed: 'Plan de intervención multidisciplinar de un caso con escoliosis idiopática del adolescente Ferrera Hernández, Raúl Humberto ; Caudevilla Polo, Santos (dir.) - TAZ-TFG-2013-600'. Below the title, a section titled 'People who viewed this page also viewed:' lists several related records with their respective counts and titles. A 'Similar records' link is visible in the bottom right corner of the interface.

Information Discussion Usage statistics

**Plan de intervención multidisciplinar de un caso con escoliosis idiopática del adolescente Ferrera Hernández, Raúl Humberto ; Caudevilla Polo, Santos (dir.) - TAZ-TFG-2013-600**

People who viewed this page also viewed:

- (8) - Carmelitas
- (7) - Sáenz Gonzalo, Alberto - INPRO--2011-008
- (7) **Contribución al estudio de la prevención de riesgos laborales. Análisis de los ejercicios en pistas de entrenamiento con equipos respiratorios autónomos para profesionales del Servicio de Extinción de Incendios y Salvamento en Zaragoza.** Cester Martínez, Armando ; Gómez López, Luis Ignacio (dir.) ; Romero Alvira, David (dir.) - TESIS-2011-047
- (7) - García Landa, José Angel - ART--2011-024
- (7) - Calvo, M. *et al* - ART--2009-004

Similar records

# Mensajería interna

- Invenio cuenta con un sistema de mensajería interna. Permite envío de mensajes a usuarios o a grupos. Permite envío postpuesto.

A: Usuarios   
Grupos

Asunto:

Mensaje:

¿Enviar más tarde?  Día  Mes  Año

Buscar usuarios o grupos:

# Organización directorios

- <http://cds.cern.ch/help/hacking/directory-organization>

# Más información

- CDS Invenio homepage: <http://cdsware.cern.ch/invenio/>
- CERN Document Server: <http://cdsweb.cern.ch/>
- CDS Invenio Demo installation: <http://invenio-demo.cern.ch/>
- Administrating CDS Invenio: <http://invenio-demo.cern.ch/help/admin/>
- Hacking CDS Invenio: <http://invenio-demo.cern.ch/help/hacking/>
- Mailing list: <http://cdsware.cern.ch/invenio/lists.html>

# Índice

Origen del Repositorio

CDS Invenio: arquitectura y módulos

Formatos de archivos a texto completo

Flujos de incorporación de información

Colecciones del Repositorio

Feedback: Google Analytics + AWStats

Plan de Preservación: la realidad de las Instituciones



# Formatos de archivos a texto completo

# FORMATO DJVU

<http://en.wikipedia.org/wiki/DjVu> [2013-06-18]

**DjVu** is a [computer file format](#) designed primarily to store [scanned documents](#), especially those containing a combination of text, line drawings, indexed color images, and photographs. It uses technologies such as **image layer separation of text and background/images**, [progressive loading](#), [arithmetic coding](#), and [lossy compression](#) for bitonal ([monochrome](#)) images. **This allows for high-quality, readable images to be stored in a minimum of space, so that they can be made available on the [web](#).**

DjVu has been promoted as an alternative to [PDF](#), promising smaller files than PDF for most scanned documents.<sup>[4]</sup> The DjVu developers report that color magazine pages compress to 40–70 kB, black and white technical papers compress to 15–40 kB, and ancient manuscripts compress to around 100 kB; a satisfactory [JPEG](#) image typically requires 500 kB.<sup>[5]</sup> Like PDF, DjVu can contain an [OCR](#) text layer, making it easy to perform [copy and paste](#) and text search operations.

**Free browser plug-ins and desktop viewers** from different developers are available from the [djvu.org](http://djvu.org) website. DjVu is supported by a number of multi-format document viewers and e-book reader software on Linux ([Okular](#), [Evince](#)), Android ([VuDroid](#)), Windows ([SumatraPDF](#)), iOS ([Stanza](#)), and BlackBerry OS ([DjVuBB](#)).

The DjVu Reference Library 3.5 was released by Lizardtech under the **GNU General Public License version 2**. DjVuLibre-3.5 was developed by Leon Bottou and others as a "Derived Work" of the DjVu Reference Library 3.5. As such, it is also subject to the GNU General Public License version 2.

DjVu divides a single image into many different images, then compresses them separately. To create a DjVu file, the initial image is first separated into three images: a background image, a foreground image, and a mask image. The background and foreground images are typically lower-resolution color images (e.g., 100 dpi); the mask image is a high-resolution bilevel image (e.g., 300 dpi) and is typically where the text is stored. The background and foreground images are then compressed using a [wavelet-based compression](#) algorithm named IW44.<sup>[5]</sup> The mask image is compressed using a method called JB2 (similar to [JBIG2](#)). The JB2 encoding method identifies nearly identical shapes on the page, such as multiple occurrences of a particular character in a given font, style, and size. It compresses the bitmap of each unique shape separately, and then encodes the locations where each shape appears on the page. Thus, instead of compressing a letter "e" in a given font multiple times, it compresses the letter "e" once (as a compressed bit image) and then records every place on the page it occurs.

Optionally, these shapes may be mapped to [UTF-8](#) codes (either by hand or potentially by a [text recognition system](#)), and stored in the DjVu file. If this mapping exists, it is possible to select and copy text.

# FORMATO PDF

<http://en.wikipedia.org/wiki/PDF> [2013-06-18]

**Portable Document Format (PDF)** is a [file format](#) used to represent [documents](#) in a manner independent of [application software](#), [hardware](#), and [operating systems](#).<sup>[1]</sup> Each PDF file encapsulates a complete description of a fixed-layout flat document, including the text, fonts, graphics, and other information needed to display it. In 1991, [Adobe Systems](#) co-founder [John Warnock](#) outlined a system called "Camelot"<sup>[2]</sup> that evolved into PDF.

While Adobe Systems made the PDF specification available free of charge in 1993, PDF remained a [proprietary format](#), controlled by Adobe, until it was officially **released as an [open standard](#) on **July 1, 2008****, and published by the [International Organization for Standardization](#) as ISO 32000-1:2008.<sup>[3][4]</sup> In 2008, Adobe published a Public Patent License to ISO 32000-1 granting [royalty-free](#) rights for all patents owned by Adobe that are necessary to make, use, sell and distribute PDF compliant implementations.<sup>[5]</sup>

# PDF vs DJVU



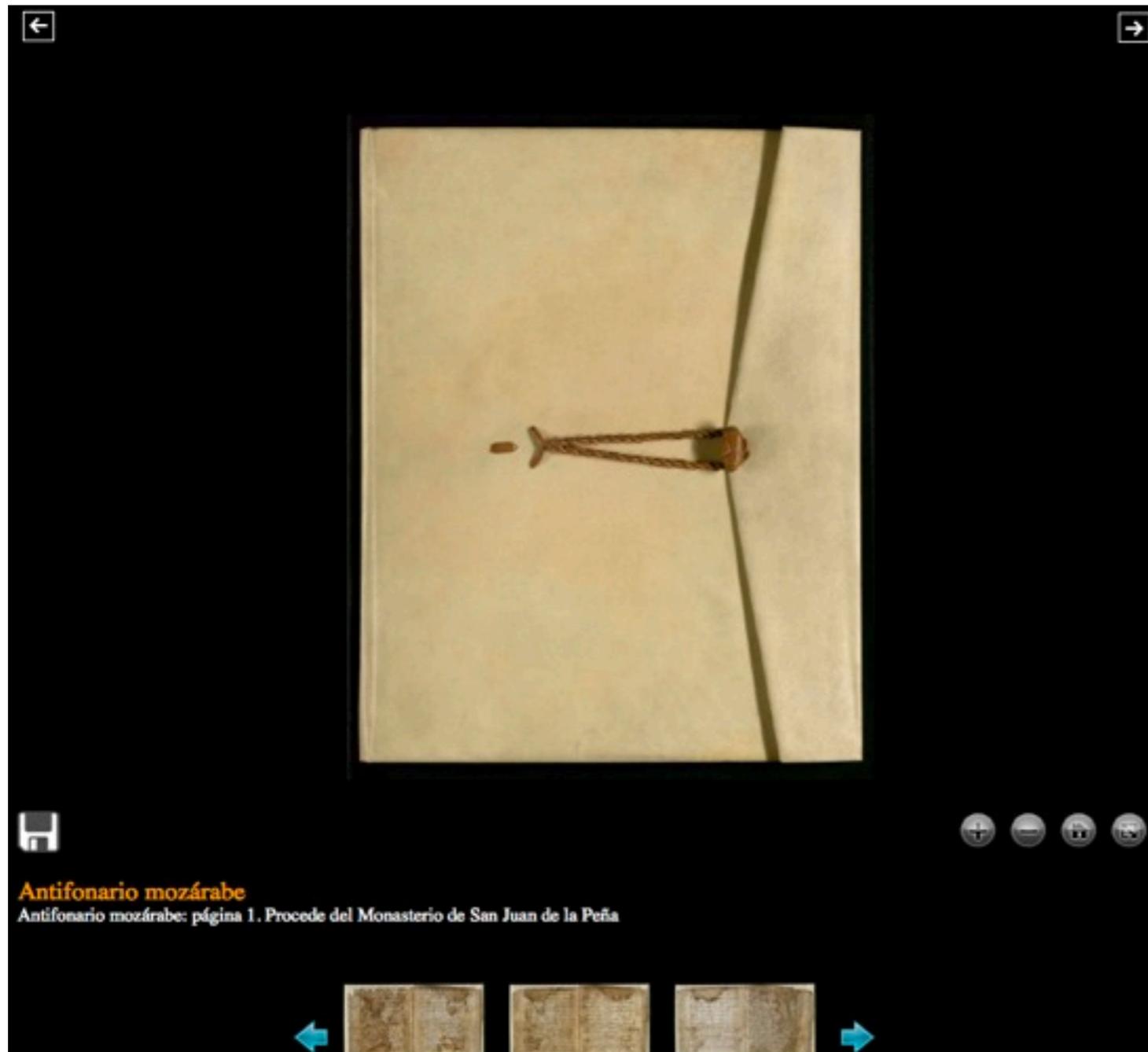
	<b>PDF</b>	<b>DJVU</b>
<b>Formato abierto</b>	Sí (desde 2008)	Sí
<b>Tamaño archivo [1]</b>	Mayor	Menor
<b>Velocidad de carga (lectura)</b>	Baja	Alta
<b>Almacenamiento de gráficos</b>	Vectorial	Rasterizado
<b>Plugin en navegador</b>	Por defecto	<b>Instalar</b>
<b>Compresión [2]</b>	Menos pérdida	Más pérdida
<b>Conocido por una mayoría</b>	Sí	<b>No</b>
<b>Resolución</b>	Mayor	Menor

[1][2] [http://djvu.org/resources/djvu\\_digital\\_vs\\_super\\_hero\\_pdf.php](http://djvu.org/resources/djvu_digital_vs_super_hero_pdf.php)

DJVU un 25% más pequeño que el PDF más pequeño...

Pérdidas por compresión asumibles. Cuando el material de origen es bueno, tanto PDF como DjVu obtienen ficheros legibles. Si el material origen es malo, ambos fracasan...

# FORMATOS INTERACTIVOS: DeepZoom (tiles) + SeaDragon



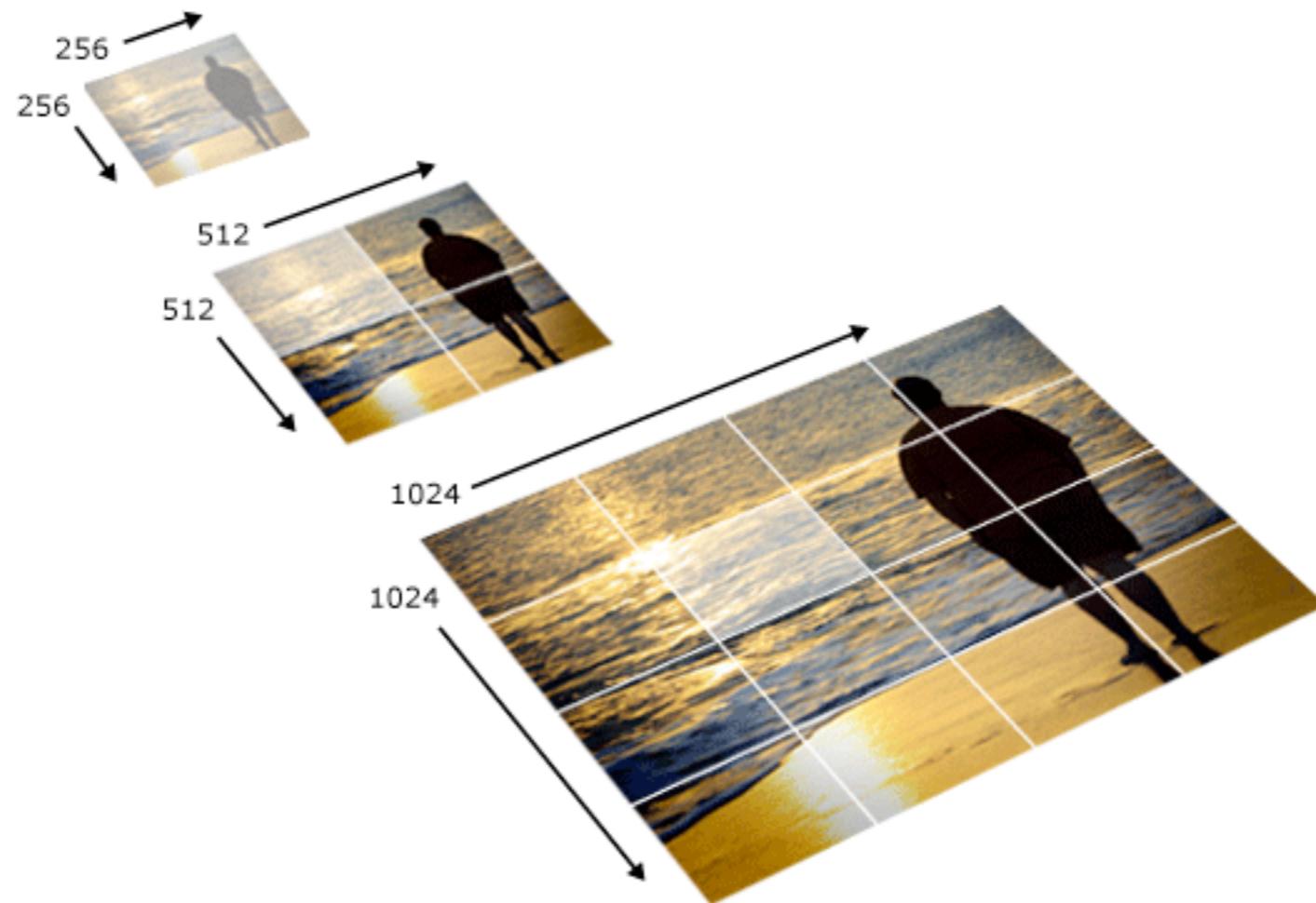
“Deep Zoom provides the ability to **interactively view high-resolution images**. You can **zoom in and out of images rapidly without affecting the performance of your application**. Deep Zoom enables smooth loading and panning by serving up multi-resolution images and using spring animations”

<http://msdn.microsoft.com/en-us/library/cc645050%28v=vs.95%29.aspx>

<http://zagan.unizar.es/deepZoom2/antifonario/index.html>

# Una imagen DeepZoom

Before you can create an application that uses Deep Zoom, you must create a Deep Zoom image. A **Deep Zoom image is composed of tiles of JPEG or PNG images at different resolutions that make up an image pyramid**. The tile size is typically 256x256, but this size can be modified. **Each tile is stored in a separate file, and each level of the pyramid is stored in a separate folder**. This enables Deep Zoom to fetch only those tiles required for the current size of the image on screen, instead of downloading the entire image. For example, if you zoom in to see only the highlighted middle part of an image, Deep Zoom loads only the highlighted tiles instead of the entire 1024x1024 image.



# Crear imagen DeepZoom



Deep Zoom Composer

<http://go.microsoft.com/fwlink/?LinkId=116569>

## DeepzoomIt

DeepZoomIt is a simple DeepZoom Composer application and library.

<http://gitorious.org/deepzoomit>

<http://gitorious.org/deepzoomit/mainline/archive-tarball/master>

<https://github.com/downloads/kennethreitz/osx-gcc-installer/GCC-10.7-v2.pkg>



# Ejemplo real: DeepZoom



[http://zaguan.unizar.es/deepZoom2/antifonario/mini/pagina1\\_rest.png](http://zaguan.unizar.es/deepZoom2/antifonario/mini/pagina1_rest.png)

[http://zaguan.unizar.es/deepZoom2/antifonario/mini/pagina1\\_selected.png](http://zaguan.unizar.es/deepZoom2/antifonario/mini/pagina1_selected.png)

Miniaturas de páginas

Un directorio por página

JPG para "guardar"

```
[root@zaguan antifonario]# pwd
/var/www/html/deepZoom2/antifonario
[root@zaguan antifonario]# ls -l
total 188
-rw-r--r-- 1 root root 5155 abr 12 2010 data.js
-rw-r--r-- 1 root root 33158 abr 12 2010 Default.xap
-rw-r--r-- 1 root root 3106 abr 12 2010 index.html
-rw-r--r-- 1 root root 33 abr 12 2010 index.html.md5
-rw-r--r-- 1 root root 33 abr 12 2010 index.htm.md5
drwxr-xr-x 2 root root 4096 abr 12 2010 mini
drwxr-xr-x 3 root root 4096 abr 12 2010 pagina1
drwxr-xr-x 3 root root 4096 abr 12 2010 pagina10
drwxr-xr-x 3 root root 4096 abr 12 2010 pagina11
drwxr-xr-x 3 root root 4096 abr 12 2010 pagina12
drwxr-xr-x 3 root root 4096 abr 12 2010 pagina13
drwxr-xr-x 3 root root 4096 abr 12 2010 pagina2
drwxr-xr-x 3 root root 4096 abr 12 2010 pagina3
drwxr-xr-x 3 root root 4096 abr 12 2010 pagina4
drwxr-xr-x 3 root root 4096 abr 12 2010 pagina5
drwxr-xr-x 3 root root 4096 abr 12 2010 pagina6
drwxr-xr-x 3 root root 4096 abr 12 2010 pagina7
drwxr-xr-x 3 root root 4096 abr 12 2010 pagina8
drwxr-xr-x 3 root root 4096 abr 12 2010 pagina9
```

```
[root@zaguan antifonario]# ls -l pagina1
total 572
drwxr-xr-x 15 root root 4096 abr 12 2010 dzc_output_files
-rw-r--r-- 1 root root 182 abr 12 2010 dzc_output.xml
-rw-r--r-- 1 root root 463696 abr 12 2010 pagina1.jpg
-rw-r--r-- 1 root root 74974 abr 12 2010 scene.xml
-rw-r--r-- 1 root root 389 abr 12 2010 SparseImageSceneGraph.xml
```

Define estructura del documento...

```
var thumbsPath = './mini/';
var URLPath = 'antifonario';
var data = [
  {
    title: 'Antifonario moz&aacute;rabe',
    thumb: 'pagina1',
    desc: 'Antifonario moz&aacute;rabe: p&aacute;gina 1. Procede del Monasterio de San Juan de la Pe&ntilde;a',
    dzi: 'pagina1/dzc_output.xml',
    xml: '<?xml version="1.0" encoding="utf-8"?><Image TileSize="256" Overlap="1" Format="jpg" xmlns="http://schemas.microsoft.com/deepzoom/2008"><Size Width="1572" Height="1048" /><ImageList /></Image>',
  }, {
    title: 'P&aacute;gina 2',
    thumb: 'pagina2',
    desc: 'Adquirido por compra por la Facultad de Derecho de la Universidad de Zaragoza en 1907.',
    dzi: 'pagina2/dzc_output.xml',
    xml: '<?xml version="1.0" encoding="utf-8"?><Image TileSize="256" Overlap="1" Format="jpg" xmlns="http://schemas.microsoft.com/deepzoom/2008"><Size Width="1572" Height="1048" /><ImageList /></Image>',
  }, {
    title: 'P&aacute;gina 3',
    thumb: 'pagina3',
    desc: 'Adquirido por compra por la Facultad de Derecho de la Universidad de Zaragoza en 1907.',
    dzi: 'pagina3/dzc_output.xml',
    xml: '<?xml version="1.0" encoding="utf-8"?><Image TileSize="256" Overlap="1" Format="jpg" xmlns="http://schemas.microsoft.com/deepzoom/2008"><Size Width="1572" Height="1048" /><ImageList /></Image>',
  }
];
```

# Ejemplo real: DeepZoom

```
[root@zaguan antifonario]# pwd
/var/www/html/deepZoom2/antifonario
[root@zaguan antifonario]# ls -l
total 188
-rw-r--r-- 1 root root 5155 abr 12 2010 data.js
-rw-r--r-- 1 root root 33158 abr 12 2010 Default.xap
-rw-r--r-- 1 root root 3106 abr 12 2010 index.html
-rw-r--r-- 1 root root 33 abr 12 2010 index.html.md5
-rw-r--r-- 1 root root 33 abr 12 2010 index.htm.md5
drwxr-xr-x 2 root root 4096 abr 12 2010 mini
drwxr-xr-x 3 root root 4096 abr 12 2010 pagina1
drwxr-xr-x 3 root root 4096 abr 12 2010 pagina10
drwxr-xr-x 3 root root 4096 abr 12 2010 pagina11
drwxr-xr-x 3 root root 4096 abr 12 2010 pagina12
drwxr-xr-x 3 root root 4096 abr 12 2010 pagina13
drwxr-xr-x 3 root root 4096 abr 12 2010 pagina2
drwxr-xr-x 3 root root 4096 abr 12 2010 pagina3
drwxr-xr-x 3 root root 4096 abr 12 2010 pagina4
drwxr-xr-x 3 root root 4096 abr 12 2010 pagina5
drwxr-xr-x 3 root root 4096 abr 12 2010 pagina6
drwxr-xr-x 3 root root 4096 abr 12 2010 pagina7
drwxr-xr-x 3 root root 4096 abr 12 2010 pagina8
drwxr-xr-x 3 root root 4096 abr 12 2010 pagina9
```

Un directorio por página

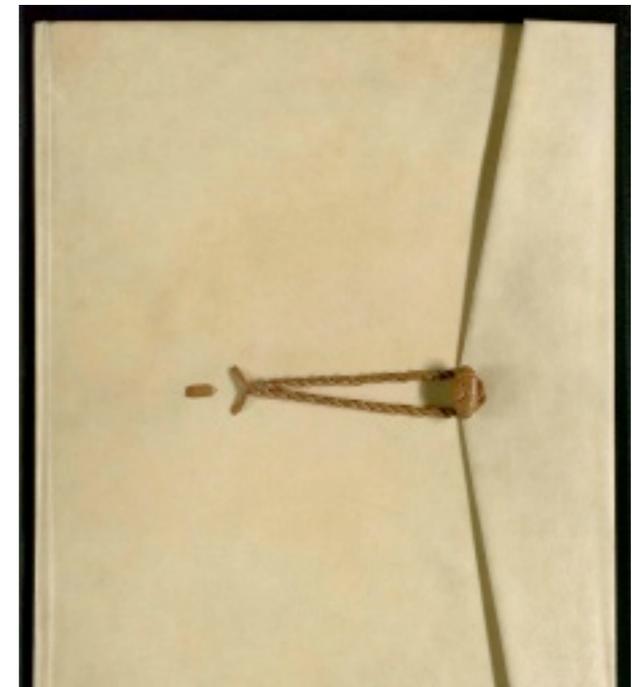
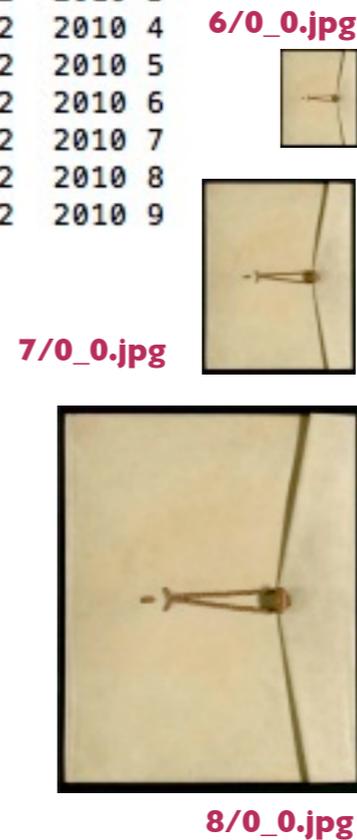
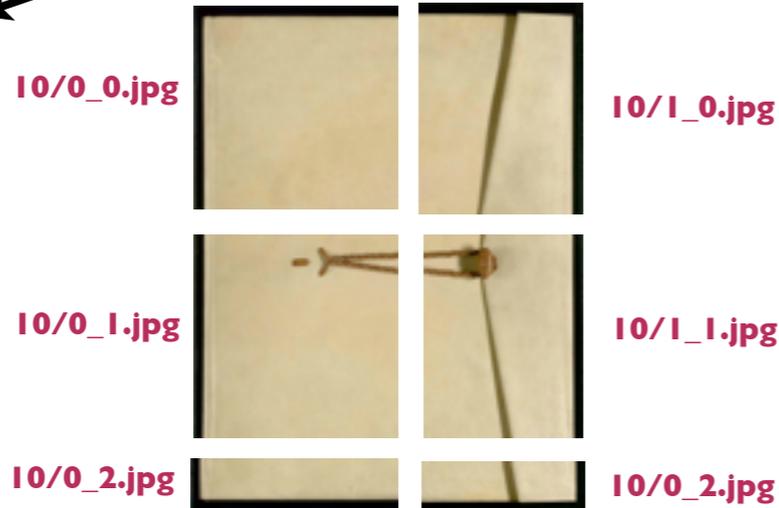
```
[root@zaguan antifonario]# ls -l pagina1
total 572
drwxr-xr-x 15 root root 4096 abr 12 2010 dzc_output_files
-rw-r--r-- 1 root root 182 abr 12 2010 dzc_output.xml
-rw-r--r-- 1 root root 463696 abr 12 2010 pagina1.jpg
-rw-r--r-- 1 root root 74974 abr 12 2010 scene.xml
-rw-r--r-- 1 root root 389 abr 12 2010 SparseImageSceneGraph.xml
```

```
[root@zaguan antifonario]# ls -l pagina1/dzc_output_files/
total 104
drwxr-xr-x 2 root root 4096 abr 12 2010 0
drwxr-xr-x 2 root root 4096 abr 12 2010 1
drwxr-xr-x 2 root root 4096 abr 12 2010 10
drwxr-xr-x 2 root root 4096 abr 12 2010 11
drwxr-xr-x 2 root root 4096 abr 12 2010 12
drwxr-xr-x 2 root root 4096 abr 12 2010 2
drwxr-xr-x 2 root root 4096 abr 12 2010 3
drwxr-xr-x 2 root root 4096 abr 12 2010 4
drwxr-xr-x 2 root root 4096 abr 12 2010 5
drwxr-xr-x 2 root root 4096 abr 12 2010 6
drwxr-xr-x 2 root root 4096 abr 12 2010 7
drwxr-xr-x 2 root root 4096 abr 12 2010 8
drwxr-xr-x 2 root root 4096 abr 12 2010 9
```

Distintos niveles de Zoom, cada directorio contiene las imágenes de menos a más zoom. En los directorios de menos zoom solo hay una imagen (0\_0.jpg) y en los directorios de más zoom va habiendo más imágenes.

El nombre del directorio indica el nivel de zoom.

```
pagina1/dzc_output_files/10/0_0.jpg
pagina1/dzc_output_files/10/0_1.jpg
pagina1/dzc_output_files/10/0_2.jpg
pagina1/dzc_output_files/10/1_0.jpg
pagina1/dzc_output_files/10/1_1.jpg
pagina1/dzc_output_files/10/1_2.jpg
```



# Ejemplo real: DeepZoom

```
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Transitional//EN" "http://www.w3.org/TR/xhtml1/DTD/xhtml1-transitional.dtd">
<html xmlns="http://www.w3.org/1999/xhtml">
<head>
  <meta name="description" content="zaguan - repositorio universidad de zaragoza" />
  <meta name="keywords" content="Repositorio Digital de la Universidad de Zaragoza, zaguan, Fondo Historico" />
  <script type="text/javascript">var _sf_startpt=(new Date()).getTime()</script>
  <title>Preview</title>
  <link rel="EditURI" type="application/rsd+xml" title="RSD" href="http://livelabs.com/api/rsd.ashx" />
  <link rel="wlwmanifest" type="application/wlwmanifest+xml" title="WLVManifest" href="http://livelabs.com/api/wlwmanifest.ashx" />
  <script type="text/javascript" src="..js/jquery-1.2.3.pack.js"></script>
  <script type="text/javascript" src="..js/jquery.jcarousel.pack.js"></script>
  <link rel="stylesheet" type="text/css" href="..css/jquery.jcarousel.mod.css" />
  <link rel="stylesheet" type="text/css" href="..css/skin_zaguan.css" />
</head>
<body>
  <div id="wrapper" style="margin-top: 0px; margin-left:auto; margin-right: auto; width: 810px;">
    <div id="viewer">&nbsp;</div>
    <div id="metadata">
      <div id="title">&nbsp;</div><div id="desc">&nbsp;</div>
    </div>
    <div id="nav" style="background: #000000; margin-left: 5px; width:800px; ">&nbsp;</div>
    <div id="thumbs">&nbsp;</div>
    <script type="text/javascript" src="..js/seadragon-branded-mod2.js"></script>
    <script type="text/javascript" src="data.js"></script>
    <script type="text/javascript" src="..js/zaguan_preview.js"></script>
  </div><!-- END wrapper -->
  <!-- SiteCatalyst code version: H.I. Copyright 1997-2005 Omniture, Inc. More info available at http://www.omniture.com -->
  <script language="JavaScript">var s_account="msnportallivelabs";</script>
  <script language="JavaScript" src="http://stj.msn.com/br/om/js/s_code.js"></script>
  <script language="JavaScript">
    <!--
      s.linkInternalFilters="javascript:.,live.,livelabs.";s.trackExternalLinks=true;
      s.server="livelabs.com";s.channel="livelabs.com";
      s.prop1="livelabs.com";s.prop2="en-us";
      /***** DO NOT ALTER ANYTHING BELOW THIS LINE ! *****/
      var s_code=s.t();if(s_code)document.write(s_code)//-->
  </script>
  <script language="JavaScript"> <!-- if(navigator.appVersion.indexOf('MSIE')>=0)document.write(unescape('%3C')+ '!'+ '-')//--> </script>
  <noscript></noscript>
  <!--/DO NOT REMOVE/-->
  <!-- End SiteCatalyst code version: H.I. -->
  </body>
</html>
```

# IMAGEN

Actualmente se está valorando la inclusión de una nueva colección de **Fotografías** (obras fotográficas) y **Carteles** en el Repositorio.

Formatos candidatos:

**PNG-8 / PNG-24** : abierto, libre, compresión sin pérdidas. Soportado por la mayoría de navegadores. Almacenamiento de másters de imagenes bitmap. <http://www.w3.org/Graphics/PNG/>

**JPEG**: Formato abierto, muy extendido. Compresión con pérdidas.

NO usar TIFF: Especificación abierta, pero bajo Copyright de Adobe... (éste y otros formatos a evitar en <http://www.openformats.org/esShowAll>)

**SVG**: abierto, desarrollado por el W3C, basado en XML+CSS+HTML. Para almacenamiento de imagenes vectoriales <http://www.w3.org/Graphics/SVG/>

# VIDEO: actualidad

En la actualidad el Repositorio Zaguan dispone de una colección de Videos.

Surge de una **necesidad apremiante** de dar cabida a material de este tipo. Un correcto almacenamiento de estos materiales, que permita alta disponibilidad, requiere una infraestructura costosa (almacenamiento del vídeo en RAW y varias calidades de visualización final, capacidad de red, etc).

En la actualidad, en espera de una infraestructura técnica adecuada, los vídeos están subidos a Vimeo y Zaguan almacena únicamente metadatos.

# VIDEO: futuro

Esperar a que se “estandarice” un formato.

WebM es un formato abierto  
<http://www.webmproject.org/>



# Índice

Origen del Repositorio

CDS Invenio: arquitectura y módulos

Formatos de archivos a texto completo

Flujos de incorporación de información

Colecciones del Repositorio

Feedback: Google Analytics + AWStats

Plan de Preservación: la realidad de las Instituciones



# Flujos de incorporación de información

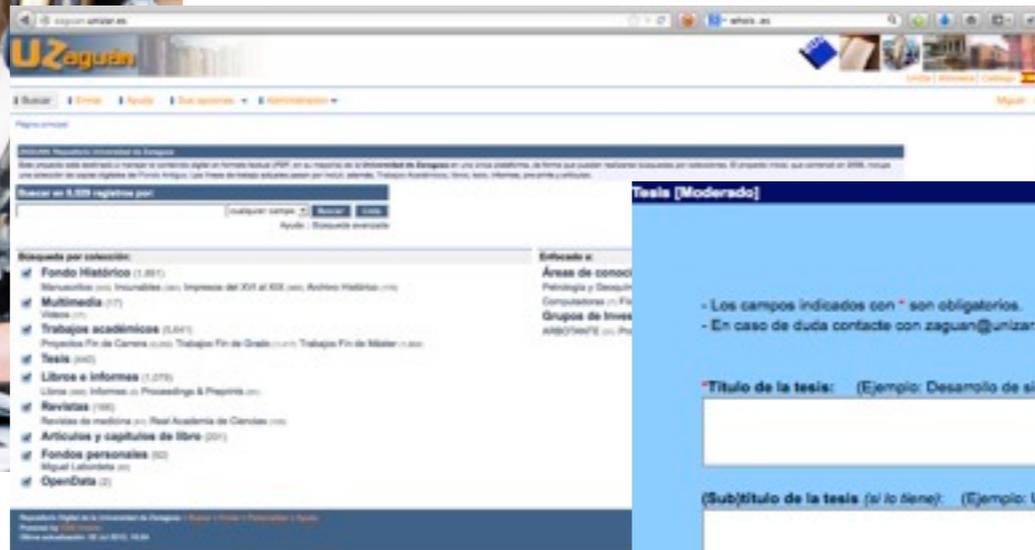
**Personal autorizado** (Tesis, Revistas)

**Autoarchivo** (Trabajos Académicos y libros)

**Flujos automatizados** (Fondo Histórico, Autoridades del Catálogo, Producción Científica)

# **Personal autorizado: Tesis**

# Personal autorizado: Tesis



Submit New Record page: 1

- Los campos indicados con \* son obligatorios.  
- En caso de duda contacte con zaguan@unizar.es

\*Título de la tesis: (Ejemplo: Desarrollo de sistemas y lentes holográficas)

(Sub)título de la tesis (si lo tiene): (Ejemplo: Una solución aproximada)

\*Apellidos, nombre del autor de la tesis: (Ejemplo: Alencia Carizo, Jesús)

Apellidos, nombre de los director(es) de la tesis: (uno por línea)  
(Ejemplo: Quintanilla Montón, Manuel)  
(Beta: Empezar a escribir y se autosugirán nombres de directores. Haga click sobre el que desea. Si no aparece entre las sugerencias, escríbalo nuevamente.)

\*Resumen: (Ejemplo: Desarrollo de una formulación completa del comportamiento como sistemas formadores...)

## Ventajas:

- Personal formado
- Evitan fallos, enriquecen metadatos
- En caso de fallo, transmiten incidencia de forma rápida y precisa

## Desventajas:

- Número limitado de personas autorizadas

# **Personal autorizado: Revistas producción UZ**

# Personal autorizado: Revistas UZ

Anteriormente se introducían directamente en Zaguan (poco flexible)

Actualmente la publicación de éste tipo de materiales se hace con la herramienta **OJS (Open Journal System)**: software libre, bajo licencia GNU General Public License.

Manual OJS (Universidad de Salamanca, OCW): <http://ocw.usal.es/enseanzas-tecnicas/taller-de-introduccion-a-ojs-open-journal-system>

Actualmente en <http://zaguan.unizar.es/ojs> pero en proceso de migración a <http://papiro.unizar.es>

Los registros se exportarán vía **OAI** a la herramienta Discovery de EBSCO [discovery.unizar.es](http://discovery.unizar.es) (actualmente en fase de implantación).

Ejemplo URL base para revista 'Tropelías': <http://zaguan.unizar.es/ojs/index.php/tropelias/oai?verb=>

OJS - zaguan.unizar.es  
Revistas electrónicas en la Universidad de Zaragoza

INICIO ACERCA DE... INGRESAR REGISTRO BUSCAR

Inicio > Open Journal Systems

## Open Journal Systems

### Journal of Sociocybernetics

The Journal of Sociocybernetics provides an international and interdisciplinary forum for the dissemination of research findings, reviews and theory in all areas of common interest to social scientists interested in the science of social systems.

The journal publishes material relevant to any aspect of systems science relevant to social science disciplines (eg. anthropology, economics, geography, psychology, social policy, sociology and ethics). It is particularly keen to publish findings or reviews which are of general interest to an international readership in social systems theory.

[VER REVISTA](#) | [NÚMERO EN CURSO](#) | [REGISTRAR](#)

### Acciones e Investigaciones Sociales

Acciones e Investigaciones Sociales es la revista científica de la Facultad de Ciencias Sociales y del Trabajo. Surgió con los siguientes objetivos: apoyar la actividad investigadora de los profesores, establecer un foro común de comunicación y debate entre los docentes e investigadores de los distintos Departamentos y Áreas de Conocimiento con representantes en el Centro, y ampliar esta comunicación, por último, con otras Universidades que estuvieran investigando en temas similares.

[VER REVISTA](#) | [NÚMERO EN CURSO](#) | [REGISTRAR](#)

### Tropelías: Revista de Teoría de la Literatura y Literatura Comparada

OPEN JOURNAL SYSTEMS

Ayuda de la revista

USUARIO/A

Nombre usuario/a

Contraseña

Recordar mis datos

Login

IDIOMA

Español (España)

CONTENIDO DE LA REVISTA

Buscar

Todos

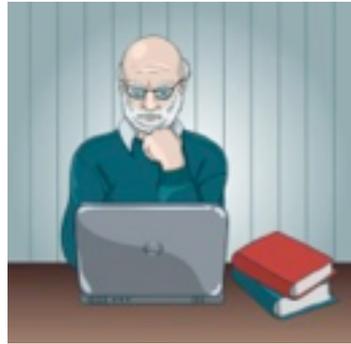
Buscar

TAMAÑO DE FUENTE

A A A

# **AUTOARCHIVO: Libros**

# AUTOARCHIVO: Libros



Enviar un libro:

\*Título: (Ejemplo: Agua bajo los puentes)

\*Apellidos, nombre del (de los) autor(es) del documento: (uno por línea)  
Ejemplo con tres autores:  
Muñoz Pasoual, Teresa  
Cantín Luna, Inmaculada  
Martín González, Miguel

Libros [Moderado] - unknown: Lista de los documentos revisados

Haga clic en un número de informe para más información.

Número de informe	Pendiente	Aprobado	Rechazado
DEMO-BOOK-2009-001			X
DEMO-BOOK-2008-001			X
BOOK-2009-003			X
BOOK-2009-004		✓	
BOOK-2009-002		✓	



Cualquier cuenta @unizar.es envía

El envío **pasa a estado “pendiente de validar”** y **ni metadatos ni texto completo son visibles**.

Personal autorizado recibe un correo indicando que hay “pendientes de validar”

Personal autorizado aprueba/rechaza el envío (verifican permisos, corrección de metadatos, etc). Se informa al original submitter de la decisión.

Si la decisión es “aprobar”, metadatos y texto completo pasan a ser visibles de forma simultánea.

# **AUTOARCHIVO: Trabajos Académicos (TAZ)**

Distinto workflow.

Los metadatos deben ser visibles desde el primer momento.

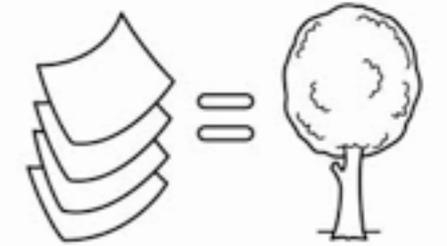
El (los) archivo(s) a texto completo únicamente estarán visibles cuando el autor de el visto bueno y el trabajo esté aprobado (>5.0) por el Tribunal.

Programación adhoc: <http://www.leccionespracticas.com/cds-invenio/cds-invenio-change-sbi-process-not-referred-records-restricted-fulltext-access/>

# Antes del autoarchivo de TAZ...

**Alumno:** Imprimir Trabajo (3 copias)

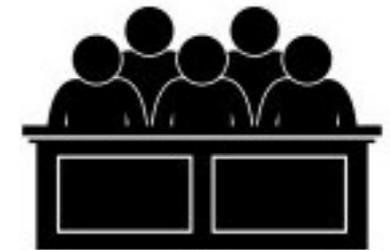
Decidir permisos: solo consulta (sala) o fotocopia



**Secretaría:** Valida que el alumno cumple los requisitos para Defensa, recibe las copias y las remite al Tribunal



**Tribunal:** evaluación (copias limitadas, no copias para todos los miembros...) remitir calificación y ejemplares a Secretaría



**Secretaría:** remitir ejemplar a Departamento, alumno y Biblioteca.



**Biblioteca:** custodiar ejemplar en papel para consulta en sala o fotocopia. catalogar (o al menos inventariar) dicho material



# Autoarchivo (de Trabajos Académicos)



Universidad  
Zaragoza

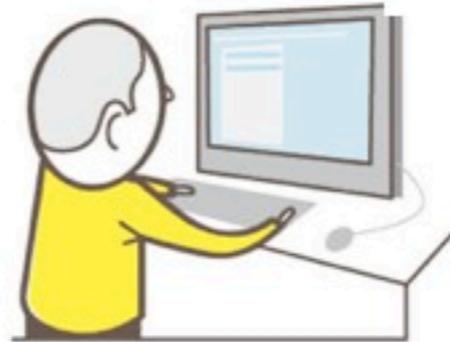
Autoarchivo

1

<http://zaguan.unizar.es>



2



3

Instrucciones de  
autoarchivo



4



5



autor



6



Validación  
(personal bibliotecario)

**Más detalles:** <http://prezi.com/mb78f-let8rg/zaguan/>

# Después del autoarchivo...

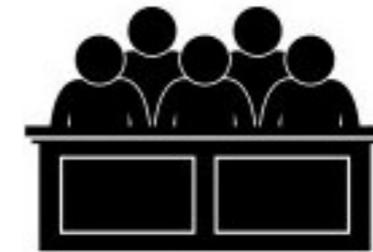
**Alumno:** Rellena la ficha en Zagan. Verifica que el PDF se ha cargado bien.  
Imprime el correo de confirmación (o no imprime nada)  
Decide licencia: restringido / creative commons



**Secretaría:** Verifica que el alumno ha realizado el Depósito



**Tribunal:** evaluación (copias ilimitadas, para todos los miembros...)  
Remite calificación a Secretaría



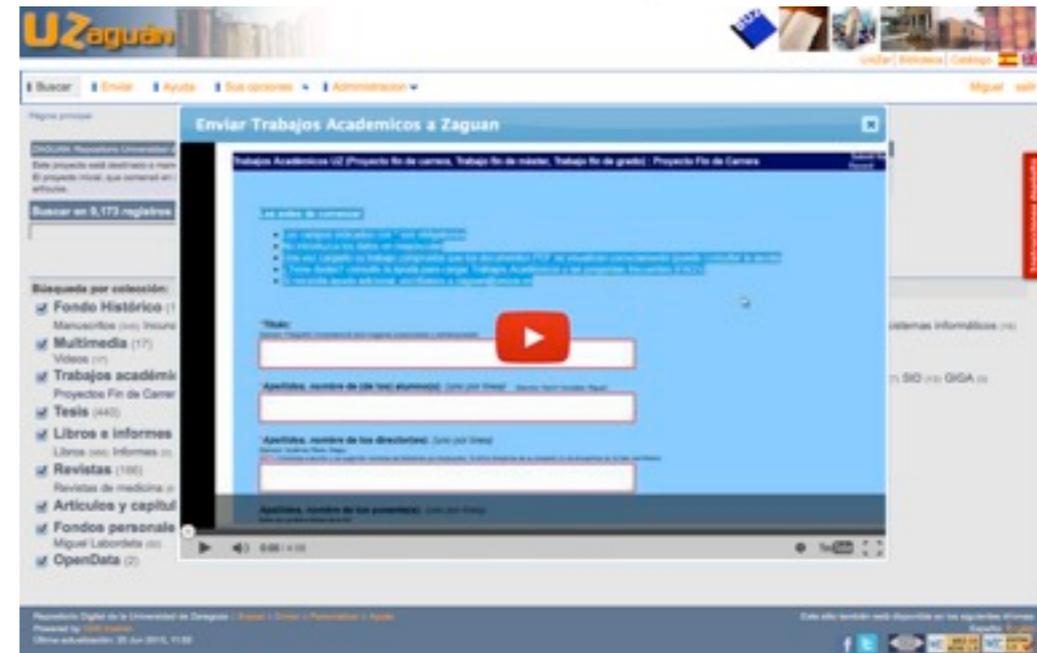
**Biblioteca:** Verifica que los metadatos sean correctos.  
Posibilidad de enriquecer metadatos  
“Valida” el registro



# Después del autoarchivo...

## No todo son ventajas...

- \* Más **dependencia del servidor** del Repositorio (en este caso, además, evidencia necesidad de alta disponibilidad...)
- \* Trabajo de respuesta a **dudas e incidencias** (Servicios Centrales)
  - \*\* Desarrollo de guías de Carga de Trabajos Académicos (no suelen leerlas) [http://zaguan.unizar.es/ayuda/deposito\\_zaguan\\_TAZ.pdf](http://zaguan.unizar.es/ayuda/deposito_zaguan_TAZ.pdf)
  - \*\* Desarrollo de vídeos explicativos con los pasos para realizar el Depósito (linkado en rojo y bien visible, se abre en contexto)



[http://www.youtube.com/watch?feature=player\\_embedded&v=PejtFTBbswU#at=14](http://www.youtube.com/watch?feature=player_embedded&v=PejtFTBbswU#at=14)

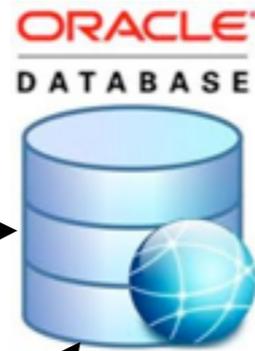
- \*\* Desarrollo de FAQ's (no suelen leerlas, pero sirven para responder a las dudas de forma consistente, con un link)

## FAQ's (PREGUNTAS FRECUENTES)

Seleccione la categoría de su duda/problema/incidencia

- Identificación en el sistema
- Trabajos Fin de Carrera/Grado/Master
- Envío de registros, derechos de autor y versiones de un mismo documento
- Moderación de contenidos

# Flujos automatizados: FH desde el Catálogo



crontab



Zaguan

```
[root@zaguan importaFHdesdeRoble]# sudo -u apache crontab -l
# Líneas anyadidas por Miguel -----
SHELL=/bin/bash
PATH=$PATH:$HOME/bin
ORACLE_HOME=/usr/lib/oracle/11.1/client64
LD_LIBRARY_PATH=/usr/lib/oracle/11.1/client64/lib
# -----
#Mins Horas Días Meses Día de la semana
# Para hacerlo los días 15 de cada mes a las 2:30 (cargar el Fondo Historico desde Roble)
30 02 15 * * /home/apache/importaFHdesdeRoble/importaFH.sh > /home/apache/importaFHdesdeRoble/salidacron.txt
# Para hacer cargas de registros de autoridad desde Roble (los días 1 de cada mes)
30 02 01 * * /home/apache/importaAUTORIDADESdesdeRoble/crea_uzPDI.sh

30 01 * * * /soft/cds-invenio/bin/alertengine
```

**importaFH.sh:** setea las variables de entorno necesarias, crea el fichero de “nuevos” y el fichero de “modificados” y valida su sintaxis marcxml...

**importaFH.py:**

1. Conexión a ORACLE para obtener los registros nuevos y modificados que siguen determinados criterios...
2. Verifica si dichos registros ya han sido cargados en el Repositorio Zaguan...
3. Si el registro ya estaba en Zaguan, crea el MARCXML para MODIFICARLO... (en fichero de “modificados”). En caso contrario, lo apunta en “nuevos”.
4. Realiza la carga de registros nuevos/modificados

# Flujos automatizados: Autoridades desde el Catálogo



## Problema

En el **Catálogo** se ha realizado un trabajo de **control de autoridades**.

**El Repositorio no dispone de control de autoridades** (al menos, de momento...)

El **autoarchivo** generaba inevitablemente muchas **variaciones de nombres al insertar “a mano”...**



## Solución (temporal)

Obtener autosugerencias al introducir nombres de autores en el flujo de AUTOARCHIVO

¿Cómo? JQuery + importación de autoridades de Roble + autoridades de Zaguan

<http://zaguan.unizar.es/preview/preview.js>

<http://zaguan.unizar.es/preview/autocompletar2.php>

# Flujos automatizados: Autoridades desde el Catálogo

\*Apellidos, nombre de los director(es)/tutor(es): (uno por línea)

(Ejemplo: Gutiérrez Pérez, Diego)

(Comience a escribir y se sugerirán nombres de directores/tutores ya introducidos. Si ellos directores de su proyecto no se encuentran en la lista, escribalos)

- Gabarre Raso, Sergio
- Gabas Trigo, Francisco
- Gabelas Barroso, José Antonio
- Gadea Rivas, María D
- Gaillard, François,
- Gairín Sallán, José M
- Galán Calvo, Francisc
- Galbán Bernal, Franc
- Galé Bornaio, Carlos
- Galé Pola, María del C
- Galé, José E.
- Galiano Martínez, Aid
- Galindo Ayuda, Ferna
- Galland, Pierre,
- Gallardo Gutiérrez, Ev

\*Apellidos, nombre de los director(es)/tutor(es): (uno por línea)

(Ejemplo: Gutiérrez Pérez, Diego)

(Comience a escribir y se sugerirán nombres de directores/tutores ya introducidos. Si ellos directores de su proyecto no se encuentran en la lista, escribalos)

- Guada Vallepuga, José Antonio
- Guajardo Cuervo, Tomás
- Guallar Leza, Chema
- Guallar Paracuellos, Jesús
- Guallar de Viala, Alfonso
- Gudmundsdottir, Unnur Stella
- Guedea Martín, Manuel
- Guedea Medrano, Isabel
- Guelbenzu Morte, Santiago
- Güemes Sánchez, Antonio
- Güemes Sánchez, Antonio Tomás
- Guerard, Brun
- Guerout, Fabr
- Guerra, Manu
- Guerrero Álva

\*Apellidos, nombre de los director(es)/tutor(es): (uno por línea)

(Ejemplo: Gutiérrez Pérez, Diego)

(Comience a escribir y se sugerirán nombres de directores/tutores ya introducidos. Si ellos directores de su proyecto no se encuentran en la lista, escribalos)

- Gutiérrez Palacio, Javier
- Gutiérrez Pérez, Diego

## DETALLES:

<http://www.leccionespracticas.com/cds-invenio/cds-invenio-implementing-autocomplete-for-authors-with-jquery-and-php-solved/>

# FLUJOS AUTOMATIZADOS: Producción científica



# **FLUJOS AUTOMATIZADOS: Producción científica**

## **EL PROBLEMA**

UZ paga al PDI por investigar (sueldo, infraestructuras, tiempo...)



El investigador desea publicar los resultados de investigación (difusión, méritos)



El copyright del **artículo terminado** pasa a ser de la revista/editorial (en ocasiones, también el preprint... si usa maquetación o membretes de la editorial...)



Investigadores de todo el mundo usan los artículos de la UZ





# **FLUJOS AUTOMATIZADOS: Producción científica**

## **SOLUCIONES**



PERO...

El investigador recibe incentivos en función a las publicaciones en revistas *peer-reviewed* (¡muchas de ellas de pago!), por lo que no siempre puede publicar en revistas *open-access*...

El prestigio y visibilidad de las revistas open access no siempre es el que el investigador desea...

# PrePrint

La propiedad INTELECTUAL siempre será del autor.  
La revista se *queda* los derechos de explotación

Difundir los preprints SIN MAQUETACIÓN de la editorial/  
revista :)



## PRECEDENTES...



# FASE I

Obtención de metadatos (SCOPUS)  
Limpieza y enriquecimiento de metadatos  
Importación a BD MySQL  
Exportación a MARCXML  
Importación al Repositorio

# FASE II

Obtención de preprints y flujo de carga automatizado

# WORK IN PROGRESS: FASE I - obtención de metadatos



Producción rama Ciencias

Búsqueda por Afiliación (Universidad de Zaragoza y variantes)

Exportar registros en varios formatos (nosotros elegimos bibtex)

“Extracción” de autores de los artículos (macro FileMaker cotejando información local de unizar). Matcheo de autoridades Scopus - “autoridades” unizar.

“Parsear” usando librería Python yapbib (biblio-py: <https://pypi.python.org/pypi/biblio-py/0.5.7> )

Exportación a BD MySQL

Generación de MARCXML ad-hoc con los campos necesarios para Zagan

# WORK IN PROGRESS: FASE I - obtención de metadatos

Búsqueda por Afiliación (Universidad de Zaragoza y variantes)

The screenshot displays the Scopus search interface. At the top, there are logos for SciVerse and Scopus, along with navigation links for Hub, ScienceDirect, Scopus, and Applications. Below this is a dark teal navigation bar with links for Search, Sources, Analytics, Alerts, My list, and Settings. The main search area has four tabs: Document search, Author search, Affiliation search, and Advanced search. The Advanced search tab is active. In the top right of the search area, there are links for Search tips and Field codes. The search query is entered in a large text box under the heading 'Outline query' and reads 'AFFIL(Universidad de Zaragoza)'. At the bottom of the search area, there are buttons for 'Clear form', 'Add Author name or Affiliation', and 'Search'.

# WORK IN PROGRESS: FASE I - limpieza de metadatos

Macro FileMaker (extracción y primera “limpieza” de autores: matcheo contra BD con información de publicaciones local - Algoritmo de desambiguación de autores)

The screenshot displays two FileMaker Pro windows. The background window shows a table with the following columns: nip, autor, autoridad\_buena, cod\_CDC, and contr\_area. The foreground window shows a search form with the following fields:

- cod\_CDC: 20110008
- nip: 388082
- titulo:
- autor: ADELL PASCUAL, JOSE ANTONIO
- autores: Adell, J. A.
- autores\_trampa: Adell, J. A.
- apellido1: ADELL
- autoridad1: Adell, J. A.
- buscafin1: 14
- autoridad2:
- buscafin2: 0
- autoridad3:
- buscafin3: 0
- cuanta: 1
- iniciales\_autor: JA
- iniciales\_autoridad\_1: JA
- iniciales\_autoridad\_2:
- iniciales\_autoridad\_3:
- autoridad\_buena: Adell, J. A.

The table in the background window contains the following data:

nip	autor	autoridad_buena	cod_CDC	contr_area
385672	ABAURREA LEBON, JESUS MARIA	Abaurrea, J.	20110001	HPM
144550	ASIN LAFUENTE, JESUS	Asín, J.	20110001	HPM
320225	CEBRIAN GUAJARDO, ANA CARMEN	Cebrián, A. C.	20110001	HPM
481722	LOPEZ PINAZO, JULIO	Lopez, J. L.	20110002	HNJ
134714	ABBCIA MARTINEZ, JOSE ALFONSO	Abecia, J. A.	20110003	HTN
386790	FORCADA MIRANDA, FERNANDO	Forcada, F.	20110003	HTN
134714	ABBCIA MARTINEZ, JOSE ALFONSO	Abecia, J. A.	20110004	HTN
386790	FORCADA MIRANDA, FERNANDO	Forcada, F.	20110004	HTN
427138	VALARES LOPEZ, JOSE ANTONIO	Valares, J. A.	20110004	HTN
429089	PALACIN ARIZON, INMACULADA	Palacin, I.	20110004	DEM
152610	BALCELLS TERES, JOAQUIM	Balcells, J.	20110005	HTN
117551	FONDEVILA CAMPS, MANUEL	Fondevila, M.	20110005	HTN
313464	BELENGUER FERRANDO, ALVARO	Belenguer, A.	20110005	HTN
439618	ABBCIA ALIENDE, LETICIA	Abecia, L.	20110005	HTN
386139	OVELLEIRO NARVION, JOSE LUIS	Ovelleiro, J. L.	20110006	HKL
149459	SARASA ALONSO, JUDITH	Sarasa, J.	20110006	HKL
183732	ACERETE GIL, JOSE BASILIO	Acerete, B.	20110007	DLG
128222	LLENA MACARULLA, FERNANDO	Llena, F.	20110007	DLG
122118	MONEVA ABADIA, JOSE MARIANO	Moneva, J. M.	20110007	DLG
388082	ADELL PASCUAL, JOSE ANTONIO	Adell, J. A.	20110008	HPM
388008	ASOREY CARBALLEIRA, MANUEL	Asorey, M.	20110009	HCJ
303121	AGUADO MARTINEZ DE CONTRASTA,	Aguado, M.	20110009	HCJ
388008	ASOREY CARBALLEIRA, MANUEL	Asorey, M.	20110010	HCJ
303121	AGUADO MARTINEZ DE CONTRASTA,	Aguado, M.	20110010	HCJ
605022	ROUBEAU, OLIVIER STEPHANE	Roubeau, O.	20110011	
107934	URRIOLABEITIA ARRONDO, ESTEBAN	Urriolabeitia, E. P.	20110012	IC1
437194	BIELSA SEBIO, RAQUEL	Bielsa, R.	20110012	HWG
472849	AGUILAR CAMAÑO, DAVID	Aguilar, D.	20110012	
114344	RABANAQUE HERNANDEZ, MARIA	Rabanaque, M. J.	20110013	HQK
187638	MARTOS JIMENEZ, MARIA DEL	Martos, M. C.	20110013	DZJ
386268	MORENO AZNAR, LUIS ALBERTO	Moreno, L. A.	20110014	DZJ
494117	BARBANA, CHOCKRY	Barbana, C.	20110015	HTV
121982	SANCHO SANZ, JAVIER	Sancho, J.	20110016	DEF
126725	MERINO FILELLA, PEDRO	Merino, P.	20110017	HXI
120697	TEJERO LOPEZ, TOMAS	Tejero, T.	20110017	HXI
387010	ALFARO GARCIA, MANUEL FELIX	Alfaro, M.	20110019	HNI

# WORK IN PROGRESS: FASE I - parseo e importación a MySQL

```
mysql> describe articulos_20130626;
```

Field	Type	Null	Key	Default	Extra
code	varchar(25)	YES		NULL	
tipo	varchar(25)	YES		NULL	
abbrev_source_title	text	YES		NULL	
abstract	text	YES		NULL	
affiliation	text	YES		NULL	
author_main	varchar(100)	YES		NULL	
author_keywords	text	YES		NULL	
coden	varchar(25)	YES		NULL	
correspondence_address	text	YES		NULL	
document_type	varchar(25)	YES		NULL	
doi	varchar(500)	NO	PRI	NULL	
firstpage	varchar(25)	YES		NULL	
issn	varchar(25)	YES		NULL	
journal	varchar(500)	YES		NULL	
journal_abbrev	varchar(25)	YES		NULL	
lenguaje	text	YES		NULL	
lastpage	varchar(25)	YES		NULL	
note	varchar(25)	YES		NULL	
issue	varchar(25)	YES		NULL	
citas	text	YES		NULL	
source	varchar(255)	YES		NULL	
title	text	YES		NULL	
url	varchar(500)	YES		NULL	
volume	varchar(25)	YES		NULL	
year	text	YES		NULL	
materia	varchar(25)	YES		NULL	
author_other	text	YES		NULL	
grupo_investigacion	varchar(250)	YES		NULL	
area_conocimiento	text	YES		NULL	
departamento	text	YES		NULL	
author_main_NIP	varchar(6)	YES		NULL	
date_insert_zaguan	datetime	YES		NULL	
date_lastmod_zaguan	datetime	YES		NULL	
recid_zaguan	varchar(25)	YES		NULL	
id_articulo	varchar(25)	YES		NULL	
keywords_index	text	YES		NULL	
author_other_NIP	text	YES		NULL	
factor_impacto	float	YES		NULL	
id	mediumint(9)	NO		0	

```
39 rows in set (0.00 sec)
```

```
90
```

Normalización de DOI  
Evitar duplicidades/inconsistencias DOI  
Truncado de abstract's  
Deduplicación de keywords

+ Trabajo colaborativo

# WORK IN PROGRESS: FASE I - exportación MySQL a XML

```
def valida_xml(file):
    # valida que el fichero 'file' sea marcxml valido
    path_to_xmlmarclint = '/usr/local/bin/xmlmarclint'
    command = path_to_xmlmarclint + ' ' + file + ''
    try:
        proc = subprocess.Popen(command, stdout=subprocess.PIPE, shell=True)
        (out, err) = proc.communicate()
    except:
        print "EXCEPCION al ejecutar '%s'" % command

    outwithoutreturn = out.rstrip('\n')
    if outwithoutreturn == "":
        return 1
    else:
        return 0

def db2marcxml(write_to_file=True, output_file_path='./registros/'):
    db = MySQLdb.connect(host, user, password, dbname)
    cursor = db.cursor(MySQLdb.cursors.DictCursor)
    #query = "SELECT * from " + tablename + " WHERE doi = '10.1016/j.crte.2010.11.003';"
    query = "SELECT * from " + tablename + " WHERE doi != '' AND doi NOT LIKE ' %' ORDER BY doi;"
    print "(db2marcxml) Ejecuta esta query: " + query
    #query = "SELECT * from articulos WHERE doi != '' AND doi NOT LIKE ' %' ORDER BY doi;"# LIMIT 1;"
    #query = "SELECT * from articulos WHERE doi = '10.2174/157341312802884544' ORDER BY doi;"
    cursor.execute(query)
    data = cursor.fetchall()
    db.close()
    output = '''<?xml version="1.0" encoding="UTF-8"?>
    <collection online="http://www.loc.gov/MARC21/xml">
    ...

    now = datetime.datetime.now()
    output_file_path += now.strftime("%Y%m%d_%H%M")
    output_file_path_new = output_file_path + "/nuevos"
    output_file_path_mod = output_file_path + "/modificados"
    output_file_path_inv = output_file_path

    # se crea el directorio y los subdirectorios...

    try:
        os.makedirs(output_file_path_inv)
        output_file_path_inv += '/invalidos.log'
    except Exception as e:
        print "EXCEPCION al crear el directorio de salida ('%s'): %s" % (output_file_path_inv, e)
        return

    try:
        os.makedirs(output_file_path_new)
    except Exception as e:
        print "EXCEPCION al crear el directorio de salida (%s): %s" % (output_file_path_new, e)
        return
```

# WORK IN PROGRESS: FASE I - exportación MySQL a XML

```

def record2xml(record, output_file_path_new, output_file_path_mod, output_file_path_inv, i, s_recid=True, s_doi=True, s_idioma=True, s_author_main=True, s_title=True, s_author_main_NIP=True, s_abstract=True, s_keywords_index=True, s_keywords_author=True, s_author_other=True, s_journal=True, s_collection=True, s_note=True, s_departamento=True, s_area_conocimiento=True, s_nip=True, s_year=True, s_grupo_investigacion=True, s_url=True, s_author_other_NIP=True, s_id_articulo=True, s_affiliation=True, s_correspondence_address=True, s_document_type=True, limpia_copyright_abstract=True):

    kw_separator = ';'
    author_separator = ';'
    author_other_NIP_separator = ';'
    departamento_separator = ';'
    area_conocimiento_separator = ';'
    grupo_investigacion_separator = ';'
    lenguaje_separator = ';'

    if i=="":
        i = 0;

    if (len(record) < 36):
        print "030! El vector del registro tiene %d elementos y trataré de exportar al rango 4-35!" % len(record)
        return ""

    output = ''

    if not record["doi"].startswith('10.'):
        # registro sin DOI...
        print "Registro sin DOI a ser DOI NO VALIDO -----> DONDRA"
        try:
            f = open(output_file_path_inv, "a") # 030! abrimos en modo Append pq todos los invalidos van en un fichero...
            f.write(record["doi"])
            f.write("\n")
            f.close()
        except Exception as e:
            print "No se pudo escribir en el fichero '%s': %s" % (output_file_path_inv, e)
            return -1

    elif record["recid_zaguan"] and record["doi"].startswith('10.'):
        # mostrar que el registro con recid=record["recid_zaguan"] ya esta en zaguan y se va a actualizar...
        print "Reg DOI='%s' con recid='%s' ----> ACTUALIZADO" % (record["doi"], record["recid_zaguan"])
        output_file_path = output_file_path_mod + '/' + record["doi"].replace('/', '_')

    elif record["doi"] and record["doi"].startswith('10.'):
        # mostrar que el registro cn doi X no esta en zaguan y se va a CARGAR...
        #print "Reg DOI='%s' sin recid -----> CARGAR" % record["doi"]
        output_file_path = output_file_path_new + '/' + record["doi"].replace('/', '_').replace(':', '_').replace('.', '_') + "_" + str(i)

    # si ya esta cargado en zaguan, tendrá RECID...
    if s_recid and record["recid_zaguan"]:
        output += '<record id="%s" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" >' % record["recid_zaguan"]

    # doi.....
    if s_doi and record["doi"] is not None:
        doi = ""
        if record["doi"]:
            doi = '<doi id="%s" xmlns="http://www.crossref.org/urn/urn:ietf:params:xml:ns:crossref-doi" >' % record["doi"]

```

# WORK IN PROGRESS: FASE I - importación al Repositorio (test)

ID	PROC [PRI]	USER	RUNTIME	SLEEP	STATUS	PROGRESS
35320	bibupload		2013-06-28 14:15:04		RUNNING	Done 388 out of 3958.

[Buscar](#)
[Enviar](#)
[Ayuda](#)
[Sus opciones](#)
[Administración](#)
[Miguel](#) [Salir](#)

Página principal > ARTICULOS\_20130627

Buscar en 3,958 registros por:

Ayuda :: Búsqueda avanzada

Últimas incorporaciones:

2013-06-28 14:16 Effect of magnet implant on iron biodistribution of fe@*c* nanoparticles in the mouse [Escribano, E.](#) ; [Fernández-Pacheco, R.](#) ; [Valdivia, J. G.](#) ; [Ibarra, M. R.](#) ; [Marquina, C.](#) ; [Queralt, J.](#)  
 Publicado en: *Archives of Pharmacal Research (APR)* v 35, issue 1 (2012) p. 93 - 100  
[Ver similares](#)

2013-06-28 14:16 Sewage sludge torrefaction in a fluidized bed reactor [Atienza-Martinez, M.](#) ; [Fonts, I.](#) ; [Ábrego, J.](#) ; [Ceamanos, J.](#) ; [Gea, G.](#)  
 Publicado en: *Chemical Engineering Journal (CEJ)* v 222, issue None (2013) p. 534 - 545  
[Ver similares](#)

2013-06-28 14:16 Single-trial classification of feedback potentials within neurofeedback training with an eeg brain-computer interface [López-Larraz, E.](#) ; [Iterate, I.](#) ; [Escolano, C.](#) ; [García, I.](#) ; [Montesano, L.](#) ; [Minguez, J.](#)  
 Publicado en: *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS (PAICIEMBSE)* v None, issue None (2011) p. 4596 - 4599  
[Ver similares](#)

2013-06-28 14:16 Endometrial expression of leptin receptor and members of the growth hormone-insulin-like growth factor system throughout the estrous cycle in heifers [Sosa, C.](#) ; [Carrquiry, M.](#) ; [Chalar, C.](#) ; [Crespi, D.](#) ; [Sanguinetti, C.](#) ; [Cavestany, D.](#) ; [Meikle, A.](#)  
 Publicado en: *Animal Reproduction Science (ARS)* v 122, issue 3-4 (2010) p. 208 - 214  
[Ver similares](#)

2013-06-28 14:16 Incidence, prevalence and mortality of kidney cancer in spain: estimates and projections for the 1998-2022 period [incidencia, prevalencia y mortalidad del cáncer renal en españa: estimaciones y proyecciones para el período 1998-2022] [Souza, D. L. B.](#) ; [Bernal, M. M.](#)  
 Publicado en: *Actas Urológicas Españolas (AUE)* v 36, issue 9 (2012) p. 521 - 526  
[Ver similares](#)

2013-06-28 14:16 Measuring spatial dynamics in metropolitan areas [Rey, S. J.](#) ; [Anselin, L.](#) ; [Folch, D. C.](#) ; [Arribas-Bel, D.](#) ; [Sastré Gutiérrez, M. L.](#) ; [Interlante, L.](#)  
 Publicado en: *Economic Development Quarterly (EDQ)* v 25, issue 1 (2011) p. 54 - 64

# WORK IN PROGRESS: FASE I - importación al Repositorio (test)

Buscar [Enter] Ayuda [Suas opciones] Administración [Miguel] [Salir]

Página principal > ARTICULOS\_20130827

Buscar en 3,958 registros por: cualquier campo [Buscar] [Lista] Ayuda : Búsqueda avanzada

### Últimas incorporaciones:

- 2013-06-28 14:16 Effect of magn Publicado en: f Ver similares
- 2013-06-28 14:16 Sewage sludge Publicado en: C Ver similares
- 2013-06-28 14:16 Single-trial clas Publicado en: f Ver similares
- 2013-06-28 14:16 Endometrial ex Publicado en: f Ver similares
- 2013-06-28 14:16 Incidence, prev Bernal, M. M. Publicado en: f Ver similares
- 2013-06-28 14:16 Measuring spi Publicado en: f

## Effect of magnet implant on iron biodistribution of $Fe_3O_4$ nanoparticles in the mouse

Escribano, E. ; Fernández-Pacheco, R. ; Valdivia, J. G. ; Ibarra, M. R. ; Marquina, C. ; Queralt, J.

Publicado en: [Archives of Pharmacal Research](#)  
APR v 35, issue 1 (2012) p. 93 - 100

Resumen: The in vivo biodistribution of  $Fe_3O_4$  nanoparticles (np) was tested in mice bearing an inflammatory focus induced by injecting carrageenan into an air pouch previously formed on their back. The animals were intravenously injected np with a high [60 mg/kg] or a low iron dose [6 mg/kg] and sacrificed 2 h later. blood and organ samples [liver, spleen, lung, and kidney] were obtained; washed exudates were also collected. iron concentration in plasma, blood cells, organs, and exudates was determined by flameless atomic-absorption-spectroscopy after digestion of organic material. pouch exudate volume increased in all groups of mice with experimental inflammation. after i.v. administration of the high and low dose of np, iron in exudate increased by 83.3% and 92.2%, respectively. a similar increase in hepatic iron appeared after the high dose [78%], but no increase appeared after the low dose. when the magnet was present, a 157% and 119% increase of iron in exudate appeared after both doses of nps, but only the high dose of np increased iron liver [60%]. the presence of a magnetic field in the pouch favored selective biodistribution of np in the inflammatory focus. these results indicate that mice with an inflammatory compartment are suitable for primary screening of different np types. they also show that selective biodistribution is greater when a low dose of np was used and that distribution in the target organ was increased by the magnetic field.

Idioma: English

Departamento: ICM,HBA

Palabra(s) clave: Magnet implant ; Magnetic nanoparticles ; Mouse ; Iron biodistribution ; Inflammatory focus

Notas: cited By (since 1996)1

DOI: 10.1007/s12272-012-0109-8  
Modificación el 2013-06-28

exto completo

Valore este documento:

- Añadir a la lista personal
- Exportar como [BibTeX](#), [MARC](#), [MARCXML](#)
- Editar este registro

(Sin opiniones)

Autores reconocidos (“autoridad” por NIP): subrayados

# WORK IN PROGRESS: FASE I - importación al Repositorio (test)

## Effect of magnet implant on iron biodistribution of fe@c nanoparticles in the mouse

Escribano, E. ; [Fernández-Pacheco, R.](#) ; Valdivia, J. G. ; [Ibarra, M. R.](#) ; [Marquina, C.](#) ; Queralt, J.

**Publicado en:**

[Archives of Pharmacal Research](#)  
APR v 35, issue 1 (2012) p. 93 - 100

**Resumen:** The in vivo biodistribution of fe@c nanoparticles [np] was tested in mice bearing an inflammatory focus induced by injecting carrageenan into an air pouch previously formed on their back. the animals were intravenously injected np with a high [60 mg/kg] or a low iron dose [6 mg/kg] and sacrificed 2 h later. blood and organ samples [liver, spleen, lung, and kidney] were obtained; washed exudates were also collected. iron concentration in plasma, blood cells, organs, and exudates was determined by flameless atomic-absorption-spectroscopy after digestion of organic material. pouch exudate volume increased in all groups of mice with experimental inflammation. after i.v. administration of the high and low dose of np, iron in exudate increased by 83.3% and 92.2%, respectively. a similar increase in hepatic iron appeared after the high dose [78%], but no increase appeared after the low dose. when the magnet was present, a 157% and 119% increase of iron in exudate appeared after both doses of nps, but only the high dose of np increased iron liver [60%]. the presence of a magnetic field in the pouch favored selective biodistribution of np in the inflammatory focus. these results indicate that mice with an inflammatory compartment are suitable for primary screening of different np types. they also show that selective biodistribution is greater when a low dose of np was used and that distribution in the target organ was increased by the magnetic field.

**Idioma:** English

**Departamento:** ICM,HBA

**Palabra(s) clave:** Magnet implant ; Magnetic nanoparticles ; Mouse ; Iron biodistribution ; Inflammatory focus

**Notas:** cited By (since 1996)1

**DOI:** 10.1007/s12272-012-0109-8

Modificación el 2013-06-28



texto completo

```
- <datafield tag="700" ind1=" " ind2=" " >
  <subfield code="a">Valdivia, J. G.</subfield>
</datafield>
- <datafield tag="700" ind1=" " ind2=" " >
  <subfield code="a">Ibarra, M. R.</subfield>
  <subfield code="g">387778</subfield>
</datafield>
- <datafield tag="700" ind1=" " ind2=" " >
  <subfield code="a">Marquina, C.</subfield>
  <subfield code="g">118994</subfield>
</datafield>
- <datafield tag="700" ind1=" " ind2=" " >
  <subfield code="a">Queralt, J.</subfield>
</datafield>
```

- Añadir a la lista personal
- Exportar como [BibTeX](#), [MARC](#), [MARCXML](#)
- Editar este registro

# WORK IN PROGRESS: FASE I - importación al Repositorio (test)

Al clickar sobre un autor reconocido, se lanza la búsqueda POR NIP

**Effect of magnet implant on iron biodistribution of fe@c nanoparticles in the mouse**  
Escribano, E. ; [Fernández-Pacheco, R.](#) ; Valdivia, J. G. ; [Ibarra, M. R.](#) ; [Marquina, C.](#) ; Queralt, J.

Publicado en:  
[Archives of Pharmaceutical Research](#)  
APR v 35, issue 1 (2012) p. 93 - 100

Resumen: The in vivo biodistribution of fe@c nanoparticles were intravenously injected with a high iron concentration in plasma, blood groups of mice with experimental inflammation at the high dose [78%], but no increase appeared in the increased iron liver [60%]. the presence of a suitable field for primary screening of different nanoparticles.

Idioma: English

Departamento: ICM,HBA

Palabra(s) clave: Magnet implant ; Magneti

Notas: cited By (since 1996)1

DOI: 10.1007/s12272-012-0109-8  
Modificación el 2013-06-28

155.210.47.102/search?sc=1&p=100\_g%3A419626+OR+700\_g%3A419626&ln=es

UZ - Zagon  
Repositorio Digital de Documentos de la Universidad de Zaragoza

Buscar Enviar Ayuda Sus opciones Administración

Página principal > Resultados de la búsqueda

Repositorio Digital de la Universidad de Zaragoza

Buscar:  
100\_g:419626 OR 700\_g:419626 cualquier campo Buscar Lista  
Ayuda : Búsqueda avanzada

Buscar colecciones:  
\*\*\* todas las colecciones \*\*\*

Ordenar por: el último primero desc. - o ordenar por - Mostrar resultados: 10 resultados Formato de visualización: reagrupar por colección HTML brief

Resultados globales: 4 registros encontrados en 0.01 segundos.

ARTICULOS\_20130627, Encontrados 4 registros  
Artículos y capítulos de libro, Encontrados 4 registros

ARTICULOS\_20130627 Encontrados 4 registros

1. Effect of magnet implant on iron biodistribution of fe@c nanoparticles in the mouse [Escribano, E.](#) ; [Fernández-Pacheco, R.](#) ; [Valdivia, J. G.](#) ; [Ibarra, M. R.](#) ; [Marquina, C.](#) ; [Queralt, J.](#)  
Publicado en: [Archives of Pharmaceutical Research \(APR v 35, issue 1 \(2012\) p. 93 - 100\)](#)  
Ver similares
2. Designing novel nano-immunoassays: antibody orientation versus sensitivity [Puertas, S.](#) ; [Moros, M.](#) ; [Fernandez-Pacheco, R.](#) ; [Ibarra, M. R.](#) ; [Grazu, V.](#) ; [Fuente, De L.](#)  
Publicado en: [Journal of Physics D: Applied Physics \(JPDAP v 43, issue 47 \(2010\) p. -\)](#)  
Ver similares
3. Nanoscale chemical and structural study of co-based febid structures by stem-eels and hrtm [Córdoba, R.](#) ; [Fernández-Pacheco, R.](#) ; [Fernández-Pacheco, A.](#) ; [Gloter, A.](#) ; [Magén, C.](#) ; [Stéphan, O.](#) ;  
Publicado en: [Nanoscale Research Letters \(NRL v 6, issue None \(2011\) p. 1 - 6\)](#)  
Ver similares
4. Low frequency noise in la 0.7ar 0.3mno 3 based magnetic tunnel junctions [Guerrero, R.](#) ; [Solignac, A.](#) ; [Feron, C.](#) ; [Pannetier-Lecoeur, M.](#) ; [Lecoeur, P.](#) ; [Fernández-Pacheco, R.](#)  
Publicado en: [Applied Physics Letters \(APL v 100, issue 14 \(2012\) p. -\)](#)  
Ver similares

**WORK IN PROGRESS: FASE II - recopilación de preprints y posterior incorporación a Zaguan**

## WORK IN PROGRESS: FASE III - flujo de carga automatizado desde sistemas locales de control de incentivos

- Investigador cobra incentivos en base a la cantidad/calidad de su producción científica.
- Detalla su CV en aplicativo web local (SIDERAL) para que se reconozcan sus méritos
- Modificación de SIDERAL para incorporación de datos bibliográficos necesarios para Zaguan

# Índice

Origen del Repositorio

CDS Invenio: arquitectura y módulos

Formatos de archivos a texto completo

Flujos de incorporación de información

Colecciones del Repositorio

Feedback: Google Analytics + AWStats

Plan de Preservación: la realidad de las Instituciones



# Colecciones del Repositorio

# Fondo Histórico

Manuscritos, Incunables, Impresos XVI-XIX, Archivo Histórico

# Fondo Histórico

zaguan.unizar.es/collection/Fondo Histórico?ln=es

**Uzaguan**

Buscar | Enviar | Ayuda | Sus opciones | Administración

[Página principal](#) > [Fondo Histórico](#)

### Descripción de la colección Fondo Histórico

La colección Fondo Histórico de la Universidad de Zaragoza pone a disposición pública y gratuita los documentos en formato djvu<sup>(?)</sup> de Fondo Antiguo (manuscritos, Incunables, Impresos del XVI al XIX y Archivo histórico) y permite su **descarga gratuita**.

### Información de licencia

Todos los registros de esta colección están licenciados bajo Licencia Creative Commons Attribution-NonCommercial 3.0 Spain.



Buscar en 1,887 registros por:

cualquier campo

[Ayuda](#) :: [Búsqueda avanzada](#)

### Búsqueda por colección:

- Manuscritos** (345)
- Incunables** (381)
- Impresos del XVI al XIX** (985)
- Archivo Histórico** (176)
  - [Libros de Aprobaciones](#) (14)
  - [Libros de Matrículas](#) (70)
  - [Libros de Gestis](#) (92)

Repositorio Digital de la Universidad de Zaragoza :: [Buscar](#) :: [Enviar](#) :: [Personalizar](#) :: [Ayuda](#)  
Powered by [CDS Invenio](#)  
Última actualización: 17 Jun 2013, 11:38

# Fondo Histórico

**Procedencia:** digitalización de documentos físicos

**Formato fulltext master:** TIFF

**Formato fulltext repositorio:** DJVU

**Licencia:** Creative Commons by-nc 3.0

**Envío:** Incorporado automáticamente desde Catálogo  
(registros pertenecientes al FH que contienen etiqueta  
856\_4)

# Tesis

# Tesis

## Tesis

Esta colección recoge las **Tesis de la Universidad de Zaragoza** desde 2008/2009 con archivo a texto completo descargable (formato PDF). La consulta de las **Tesis de la Universidad de Zaragoza** depositadas en el repositorio institucional Zaguán queda condicionada a la aceptación de las siguientes **condiciones de uso**.

- La difusión de esta tesis por medio del servicio TDR-UZa través del repositorio institucional Zaguán ha sido autorizada por los titulares de los derechos de propiedad intelectual únicamente para usos privados enmarcados en actividades de investigación y docencia.
- No se autoriza su reproducción con finalidades de lucro ni su difusión y puesta a disposición desde un sitio ajeno al repositorio institucional Zaguán.
- No se autoriza la presentación de su contenido en una ventana o marco ajeno a Zaguán (framing).
- Esta reserva de derechos afecta tanto al resumen de presentación de la tesis como a sus contenidos.
- En la utilización o cita de partes de la tesis es obligado indicar el nombre de la persona autora.

Si la Tesis que está buscando no se encuentra digitalizada, búsquela en el catálogo Roble

Buscar en 447 registros por:

[Ayuda](#) :: [Búsqueda avanzada](#)

## Últimas incorporaciones:

- 2013-07-02 14:26 **Auditoría de inteligencia: Un método para el diagnóstico de sistemas de inteligencia competitiva y organizacional** Vasconcelos Carvalho, Andréa ; Esteban Navarro, Miguel ; Universidad de Zaragoza, Prensas de la Universidad, Zaragoza, 2010  
Texto completo (spa):  
 [Ver el archivo PDF](#)  
[Ver similares](#)
- 2013-07-02 14:09 **La organización del territorio y las dinámicas sociales en Ribagorza durante la gran expansión medieval (1000-1300)** Tomás Faci, Guillermo ; Laliena Corbera, Carlos (dir.) ; Universidad de Zaragoza, Prensas de la Universidad, Zaragoza, 2013  
Texto completo (spa):  
 [Ver el archivo PDF](#)  
[Ver similares](#)
- 2013-07-02 13:26 **Sociedades cooperativas: Una aproximación desde la responsabilidad social corporativa y el desarrollo** Nachar Calderón, Pablo ; Marcuello Servós, Carmen (dir.) ; Universidad de Zaragoza, Prensas de la Universidad, Zaragoza, 2013  
Texto completo (spa):  
 [Ver el archivo PDF](#)  
[Ver similares](#)
- 2013-07-02 13:05 **Diseño, aplicación y evaluación de un programa de intervención escolar para incrementar los niveles de actividad física en los adolescentes** Murillo Pardo, Berta ; General ; Julián Clemente, José Antonio (dir.) ; Universidad de Zaragoza, Prensas de la Universidad, Zaragoza, 2013  
Texto completo (spa):  


# Tesis

**Procedencia:** nacido digital.

**Formato:** PDF (preferencia PDF/A ó PDF/X).

**Licencia:** Condiciones de uso específicas definidas en UZ y aceptadas por el autor.

- La difusión de esta tesis por medio del servicio TDR-UZa través del repositorio institucional Zaguán ha sido autorizada por los titulares de los derechos de propiedad intelectual únicamente para usos privados enmarcados en actividades de investigación y docencia.
- No se autoriza su reproducción con finalidades de lucro ni su difusión y puesta a disposición desde un sitio ajeno al repositorio institucional Zaguán.
- No se autoriza la presentación de su contenido en una ventana o marco ajeno a Zaguán (framing).
- Esta reserva de derechos afecta tanto al resumen de presentación de la tesis como a sus contenidos.
- En la utilización o cita de partes de la tesis es obligado indicar el nombre de la persona autora.

**Envío:** Personal autorizado recibe los materiales y los sube

# Trabajos académicos

Proyectos Fin de Carrera, Trabajos Fin de Grado y Fin de Máster

# Trabajos académicos

Buscar en 5,660 registros por:

cualquier campo ▾

Buscar

Lista

Ayuda :: Búsqueda avanzada

## Búsqueda por colección:

- Proyectos Fin de Carrera (2,255)
- Trabajos Fin de Grado (1,427)
- Trabajos Fin de Máster (1,978)

## Enfocado a:

### Centro (5,660)

Centro Politécnico Superior (605) Centro Universitario de la Defensa (0) Escuela de Ingeniería y Arquitectura (1,383) Facultad de Ciencias de la Salud (838) Escuela Universitaria de Enfermería de Huesca (89) Escuela Universitaria de Enfermería de Teruel (60) Escuela Universitaria Politécnica de La Almunia de D<sup>a</sup> Godina (0) Escuela Universitaria Politécnica de Teruel (0) Escuela de Ingeniería Técnica Industrial de Zaragoza (601) Escuela Politécnica Superior (Huesca) (97) Facultad de Ciencias (152) Facultad de Ciencias Humanas y de la Educación de Huesca (45) Facultad de Ciencias de la Salud y del Deporte de Huesca (18) Facultad de Ciencias Sociales y Humanas de Teruel (121) Facultad de Ciencias Sociales y del Trabajo (250) Facultad de Derecho (67) Facultad de Economía y Empresa (144) Facultad de Educación (628) Facultad de Filosofía y Letras (294) Facultad de Medicina (170) Facultad de Empresa y Gestión Pública (38) Facultad de Veterinaria (60)

### Departamento (3,343)

Anatomía, Embriología y Genética Animal (5) Anatomía e Histología Humanas (1) Anatomía Patológica, Medicina Legal y Forense y Toxicología (1) Ciencias Agrarias y del Medio Natural (43) Análisis Económico (7) Bioquímica y Biología Molecular y Celular (33) Ciencia y Tecnología de Materiales y Fluidos (146) Ciencias de la Antigüedad (12) Ciencias de la Documentación e Historia de la Ciencia (41) Ciencias de la Educación (40) Ciencias de la Tierra (18) Cirugía, Ginecología y Obstetricia (24) Contabilidad y Finanzas (36) Dirección de Marketing e Investigación de Mercados (7) Dirección y Organización de Empresas (193) Estructura e Historia Económicas y Economía (25) Física Aplicada (13) Ingeniería de Diseño y Fabricación (484) Informática e Ingeniería de Sistemas (306) Ingeniería Eléctrica (327) Ingeniería Electrónica y Comunicaciones (276) Ingeniería Mecánica (629) Ingeniería Química y Tecnologías del Medio Ambiente (274) Matemática Aplicada (22)

# Trabajos académicos

**Procedencia:** nacido digital.

**Formato:** PDF(s) (preferencia PDF/A ó PDF/X).

**Licencias:** A elegir por el autor..

Creative Commons by-nc-sa 3.0 (55% de los Trabajos)

Libre acceso y disseminación de metadatos. No acceso fulltext.

**Envío:** Autoarchivo (workflow adhoc)

# Revistas

(producción propia UZ)

# Trabajos académicos

Buscar en 166 registros por:

cualquier campo

[Ayuda](#) :: [Búsqueda avanzada](#)

## Búsqueda por colección:

**Revistas de medicina** (61)

Boletín médico escolar de Zaragoza (1) Revista 'La unión médica de Aragón' (7) Revista Clínica (Madrid) (2) Medicina: Publicación escolar (1) Anales de la Facultad de Medicina (4) Anales del Hospital Provincial de Zaragoza (1) Anuario de la Facultad de Medicina (1) Biblioteca médico-física (6) Revista 'El practicante' (17) Revista 'La medicina aragonesa' (18) Revista trimestral de histología normal y patológica (3)

**Real Academia de Ciencias** (105)

# Revistas

**Procedencia:** digitalización de fondos en papel y nacidas digitales

**Formato:** PDF(s) y JPG para (algunas) portadas.

**Licencias:** Creative Commons by-nc-sa

**Envío:** Personal autorizado (antes: importación directa vía XML procedente de diversos sistemas o por formulario web, ahora: utilización de la herramienta Open Journal System)

# Fondos Personales

(correspondencia Miguel Labordeta)

# Fondos Personales

## Fondos personales

La **Biblioteca de la Universidad de Zaragoza** ha recibido donaciones de material por parte de intelectuales y estudiosos de diversas áreas de conocimiento, pertenecientes o no a la comunidad universitaria. Éstos legados están constituidos por materiales heterogéneos (correspondencia personal, bibliotecas privadas, etc). Esta colección aglutina los contenidos de esta índole que estén digitalizados y cuyo acceso a texto completo sea libre.

## Buscar en 92 registros por:

[Ayuda](#) :: [Búsqueda avanzada](#)

## Búsqueda por colección:

- Miguel Labordeta (92)

# Fondos Personales

**Procedencia:** digitalización de fondos en papel (documentos originales) + transcripciones (nacidas digitales). Surge la necesidad apremiante de servir la colección a determinados investigadores.

**Formato:** PDF(s)

**Licencias:** Creative Commons by-nc-sa

**Estado:** Por completar

**Envío:** Alumnos de Biblioteconomía. Revisión a cargo de profesor de Biblioteconomía.

# Otras colecciones (en curso)

Multimedia (vídeos)

Libros e informes

Artículos y capítulos de libro

OpenData

# Índice

Origen del Repositorio

CDS Invenio: arquitectura y módulos

Formatos de archivos a texto completo

Flujos de incorporación de información

Colecciones del Repositorio

Feedback: [Google Analytics](#) + [AWStats](#)

Plan de Preservación: la realidad de las Instituciones



# Análisis de feedback

# Google Analytics

Útil para obtención de estadísticas de un sitio web:

cuándo se usa  
cuánto se usa  
cómo se usa  
quién lo usa

# Google Analytics

## Cuándo se usa



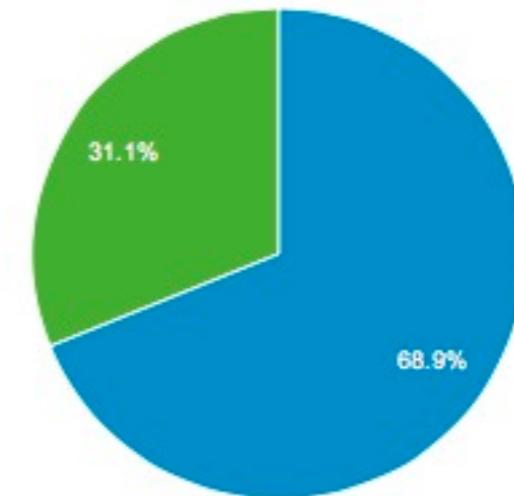
Página		Número de páginas vistas	% Número de páginas vistas
1. /		7.175	8,69%
2. /submit		3.715	4,50%
3. /submit?doctype=TAZ		2.176	2,64%
4. /collection/Trabajos Fin de Grado		1.253	1,52%
5. /collection/Trabajos Fin de Máster		1.200	1,45%
6. /collection/Trabajos academicos		878	1,06%
7. /youraccount/login		842	1,02%
8. /youraccount/login?referer=http://zaguan.unizar.es/		647	0,78%
9. /submit?doctype=TAZ&indir=running&act=SBI&startPg=1&mainmenu=/submit?doctype=TAZ&comboTAZ=TFM		545	0,66%
10. /collection/Trabajos Fin de Carrera		487	0,59%

# Google Analytics

## Cuánto se usa

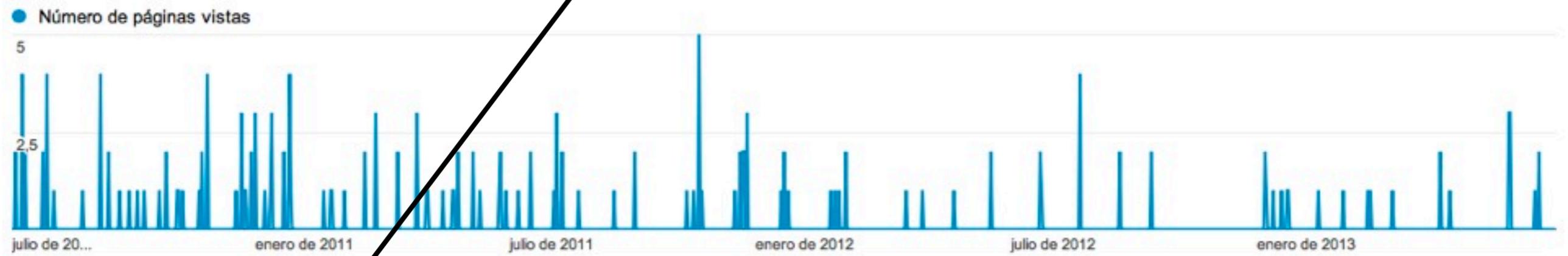
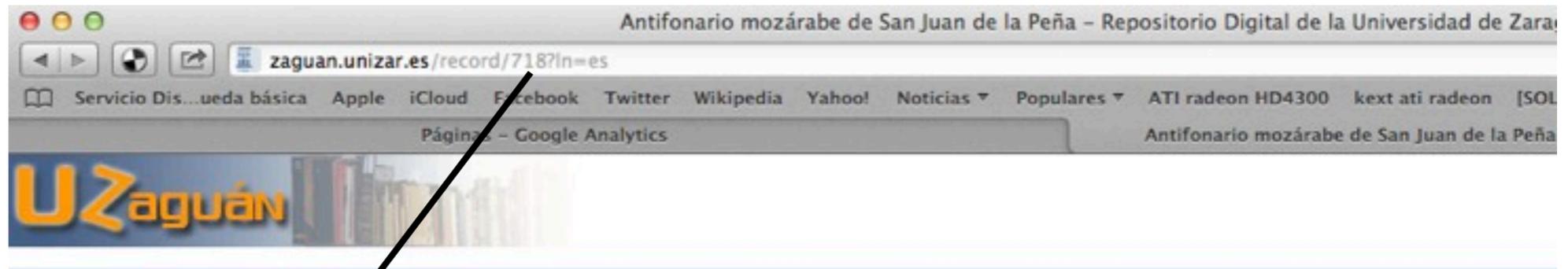


■ New Visitor ■ Returning Visitor



# Google Analytics

Cuánto se consulta (un registro en concreto, en un período determinado)



Dimensión primaria: **Página** Otros ▾

Incluir en gráfico  Dimensión secundaria ▾ Ordenar por tipo: Predeterminado ▾  avanzado

<input type="checkbox"/>	Página	Número de páginas vistas ?	Número de páginas vistas únicas ?	Promedio de tiempo en la página ?	Entradas ?	Porcentaje de rebote ?	Porcentaje de salidas ?	Valor de página ?
		<b>170</b> % del total: 0,01% (1.349.183)	<b>123</b> % del total: 0,01% (925.684)	<b>00:01:28</b> Promedio del sitio: 00:00:50 (76,29%)	<b>73</b> % del total: 0,03% (262.159)	<b>56,16%</b> Promedio del sitio: 53,97% (4,07%)	<b>44,71%</b> Promedio del sitio: 19,43% (130,08%)	<b>0,00 €</b> % del total: 0,00% (0,00 €)
<input type="checkbox"/>	1. /record/718	<b>170</b>	<b>123</b>	<b>00:01:28</b>	<b>73</b>	<b>56,16%</b>	<b>44,71%</b>	<b>0,00 €</b>

# Google Analytics

## Quién lo usa

Pais/territorio	Visitas	% Visitas
1. Spain	10.589	70,57%
2. Mexico	738	4,92%
3. Peru	607	4,05%
4. Argentina	518	3,45%
5. Colombia	463	3,09%
6. Chile	436	2,91%
7. Venezuela	346	2,31%
8. Ecuador	283	1,89%
9. Bolivia	121	0,81%
10. United States		

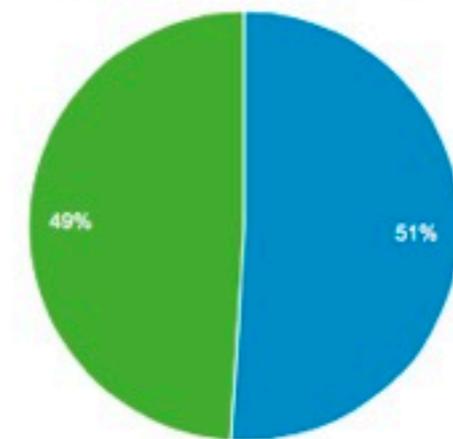
Navegador	Visitas	% Visitas
1. Chrome	6.366	42,42%
2. Firefox	3.640	24,26%
3. Internet Explorer	3.183	21,21%
4. Safari	1.198	7,98%
5. Android Browser	420	2,80%
6. Opera	48	0,32%
7. Safari (in-app)	40	0,27%
8. IE with Chrome Frame	37	0,25%
9. Opera Mini	27	0,18%
10. BlackBerry	22	0,15%

# Google Analytics

## Cómo se usa

48,99% de sus visitas han utilizado la búsqueda de sitios.

■ Visits Without Site Search ■ Visits With Site Search



# Google Analytics

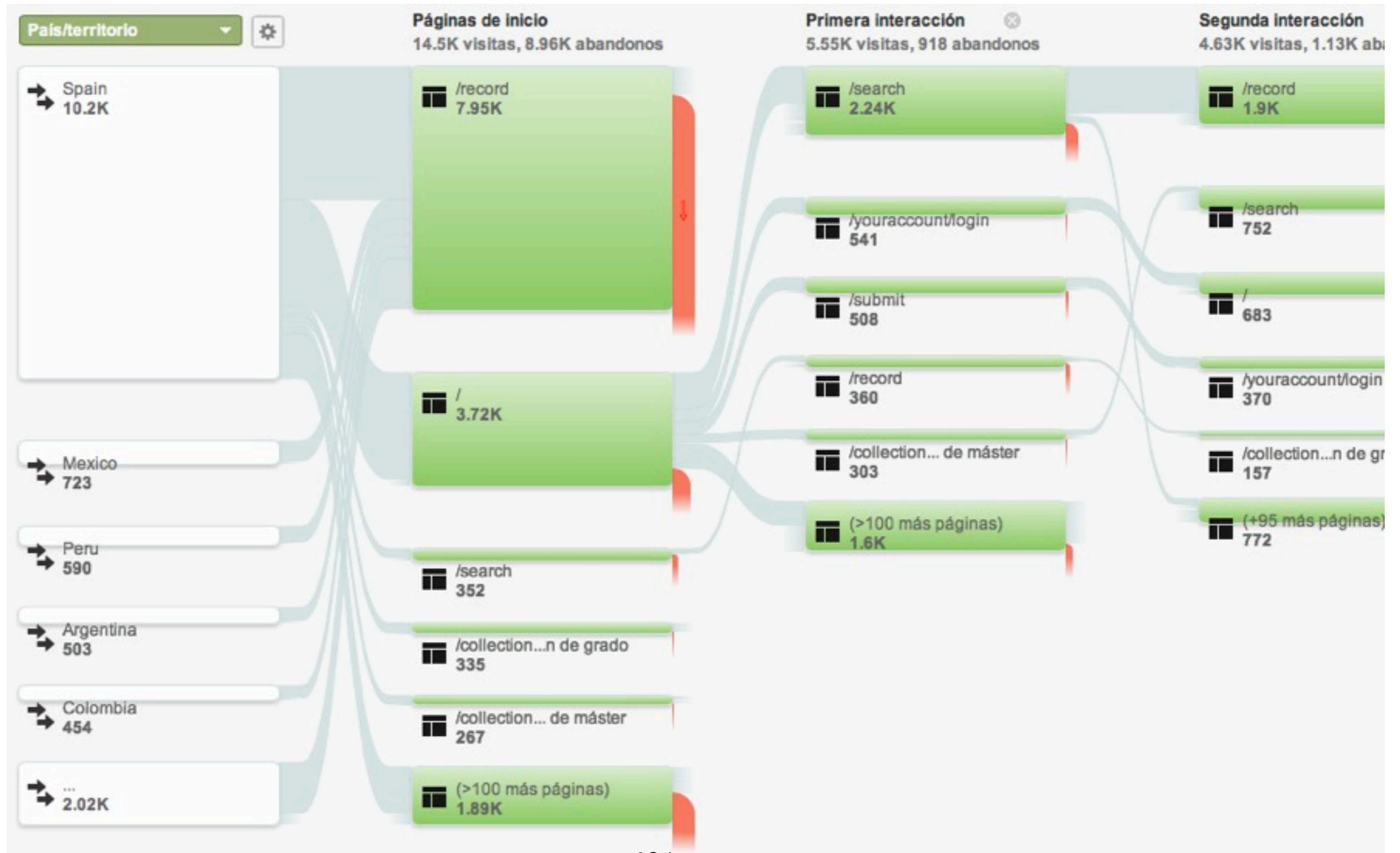
## Cómo se usa

Resolución de pantalla	Visitas	% Visitas
1. 768x1024	256	 24,45%
2. 320x480	187	 17,86%
3. 480x800	60	 5,73%
4. 720x1280	49	 4,68%
5. 320x568	41	 3,92%
6. 1280x800	36	 3,44%
7. 800x1003	31	 2,96%
8. 540x960	16	 1,53%
9. 800x1220	16	 1,53%
10. 800x480	16	 1,53%

Sistema operativo	Visitas	% Visitas
1. Android	519	 49,57%
2. iOS	464	 44,32%
3. BlackBerry	28	 2,67%
4. Windows Phone	13	 1,24%

# Google Analytics

## Cómo se usa



# AWStats

Statistics for:  
Zaguan

Last Update: 05 Jul 2013 - 00:10



Reported period: Jul 2013 OK

Summary

**When:**

Monthly history

Days of month

Days of week

Hours

**Who:**

Countries

Full list

Cities

Hosts

Full list

Last visit

Unresolved IP Address

Robots/Spiders visitors

Full list

Last visit

**Navigation:**

Visits duration

File type

Downloads

Full list

Viewed

Full list

Entry

Exit

Operating Systems

Versions

Unknown

Browsers

Versions

Unknown

**Referrers:**

Origin

Referring search engines

Referring sites

Search

Search Keyphrases

Search Keywords

**Others:**

Miscellaneous

HTTP Status codes

Pages not found

## Summary

Reported period Month Jul 2013

First visit 01 Jul 2013 - 00:00

Last visit 05 Jul 2013 - 00:01

	Unique visitors	Number of visits	Pages	Hits	Bandwidth
Viewed traffic *	<b>2,038</b>	<b>2,869</b> (1.4 visits/visitor)	<b>36,389</b> (12.68 Pages/Visit)	<b>81,649</b> (28.45 Hits/Visit)	<b>89.26 GB</b> (32623.23 KB/Visit)
Not viewed traffic *			<b>50,733</b>	<b>67,855</b>	<b>68.96 GB</b>

\* Not viewed traffic includes traffic generated by robots, worms, or replies with special HTTP status codes.

## Monthly history



Month	Unique visitors	Number of visits	Pages	Hits	Bandwidth
Jan 2013	9,895	14,181	108,425	339,909	405.02 GB
Feb 2013	11,332	16,552	166,278	455,268	482.37 GB
Mar 2013	12,363	17,230	119,802	439,848	1116.30 GB
Apr 2013	13,372	18,755	129,226	474,225	483.06 GB
May 2013	15,040	21,962	209,714	693,480	751.76 GB
Jun 2013	13,390	20,793	216,783	543,303	515.88 GB
<b>Jul 2013</b>	<b>2,038</b>	<b>2,869</b>	<b>36,389</b>	<b>81,649</b>	<b>89.26 GB</b>
Aug 2013	0	0	0	0	0
Sep 2013	0	0	0	0	0
Oct 2013	0	0	0	0	0
Nov 2013	0	0	0	0	0

# Google Analytics vs AWStats

AWStats **analiza los logs del servidor**, incluyendo accesos de bots (crawlers) y humanos.  
AWStats **analiza todas las peticiones al servidor**. (incluye, p ejemplo, peticiones OAI)

Google Analytics **analiza accesos web**: solo el acceso humano con javascript habilitado (utiliza método de cookie javascript). GA **analiza** únicamente accesos a aquellas **páginas** que tengan el **tracking code** insertado... (NO CONTABILIZA descargas de PDF por enlace directo, por ejemplo...)

AWStats se instala de **forma local** (en nuestros servidores): software libre bajo licencia GNU GPL.

Google analytics es un **servicio de terceros**.

AWStats muestra peticiones fallidas (códigos de error 500 o 404): muy útil para optimizar.

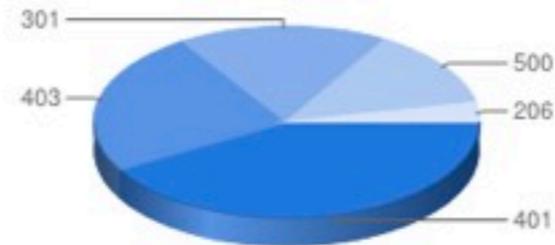
AWStats “cuenta” de forma distinta a GAnalytics.

Ejemplo1: Acceso a una misma imagen varias veces seguidas. Si se ha habilitado el cacheo de imagenes en el navegador del cliente, solo la primera petición será anotada en AWStats (las siguientes serán servidas por la cache del navegador) sin embargo todas las peticiones (cacheadas o no) serán apuntadas en GA.

Ejemplo2: Atantes chinos guardando todos los PDF's de Zagan (consumiento bastante ancho de banda): contabilizados como una petición en GA... pero como 10.000 peticiones en AWS.

# AWStats

## HTTP Status codes



### HTTP Status codes\*

HTTP Status codes*		Hits	Percent	Bandwidth
401	Unauthorized	6,220	38.2 %	60.29 MB
403	Forbidden	3,670	22.5 %	793.91 KB
301	Moved permanently (redirect)	2,661	16.3 %	656.62 KB
500	Internal server Error	2,071	12.7 %	25.14 GB
206	Partial Content	601	3.6 %	2.25 GB
404	Document Not Found (hits on favicon excluded)	545	3.3 %	5.51 MB
302	Moved temporarily (redirect)	157	0.9 %	10.78 KB
405	Method not allowed	142	0.8 %	0
300	Multiple documents available	93	0.5 %	52.10 KB
307	Moved temporarily	42	0.2 %	5.94 KB
503	Server busy	39	0.2 %	826 Bytes
400	Bad Request	12	0 %	20.95 KB
416	Requested range not valid	4	0 %	1.23 KB

## Required but not found URLs (HTTP code 404)

URL (334)	Error Hits
/apple-touch-icon-precomposed.png	51
/apple-touch-icon.png	38
/apple-touch-icon-72x72.png	9
/apple-touch-icon-72x72-precomposed.png	7

<http://www.leccionespracticas.com/cds-invenio/apple-touch-icon-404-errors/>

# Google Analytics vs AWStats

En resumen:

**GA como herramienta de marketing y decisiones de negocio.**

**AWStats como herramienta para el administrador del servidor**

# Índice

Origen del Repositorio

CDS Invenio: arquitectura y módulos

Formatos de archivos a texto completo

Flujos de incorporación de información

Colecciones del Repositorio

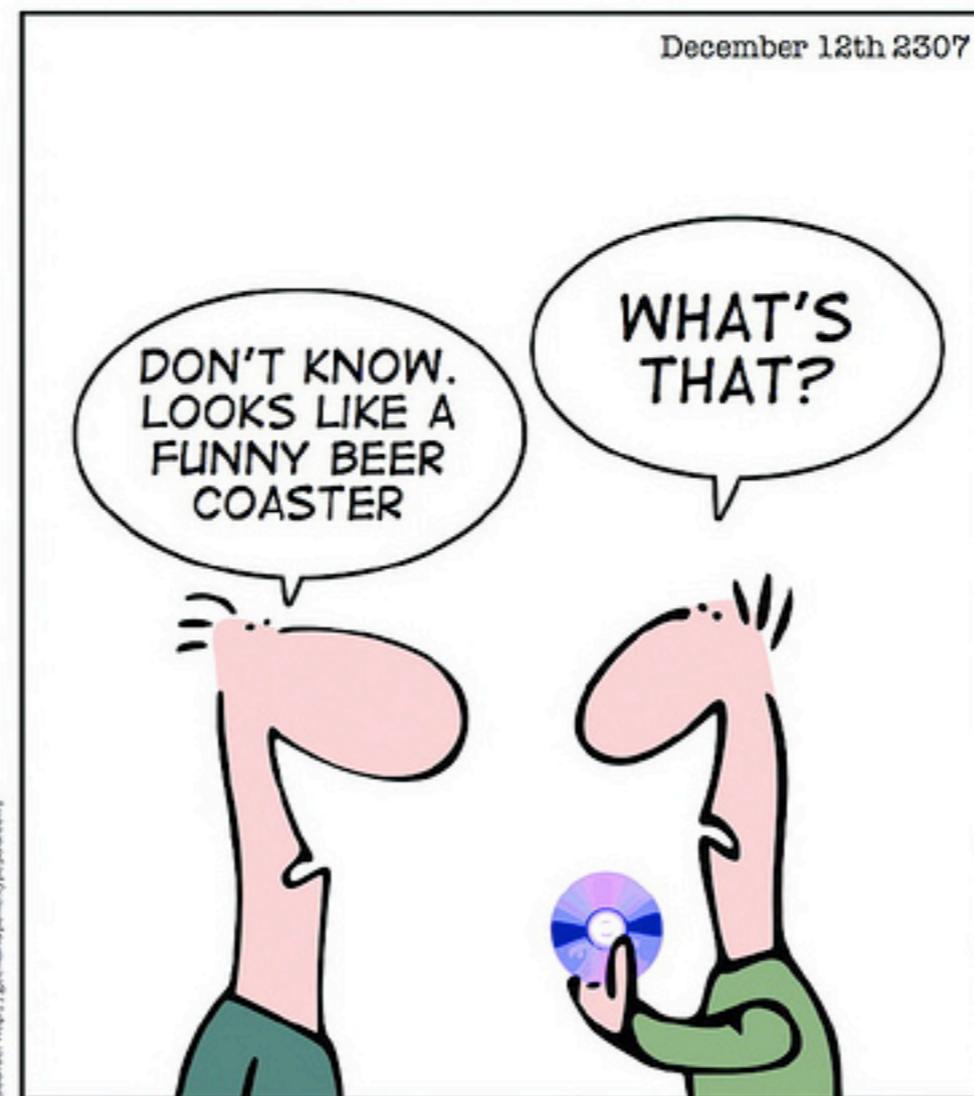
Feedback: Google Analytics + AWStats

Plan de Preservación: la realidad de las Instituciones



# Preservación: la realidad en las Instituciones

HOW TO SAVE YOUR DIGITAL WORK FOR THE POSTERITY?



# Preservación: un poco de historia...

Babilonia (hoy, Iraq) - Confiaban en (Nabu) *los dioses, autores de los alfabetos y autores del conocimiento [...]*

Hindúes - Ganesh (dios con cabeza de elefante) *“is a keeper of wisdom and knowledge”*

Monasterios cristianos - “La maldición del libro” (en la última página del mismo) - *“For him that stealeth a book from this library, may it change into a serpent in his hand and rend him. May he be struck with palsy, and all his members blasted. May bookworms gnaw his entrails ...”*

[The Challenges of Preservation of Archives and Records in the Electronic Age](http://unllib.unl.edu/LPP/PNLA%20Quarterly/asogwa-ezema76-3.htm) by Brendan E. Asogwa, Ifeany J. Ezema <http://unllib.unl.edu/LPP/PNLA%20Quarterly/asogwa-ezema76-3.htm>

Wikipedia EN for “Preservation” [20130702] [http://en.wikipedia.org/wiki/Preservation\\_%28library\\_and\\_archival\\_science%29](http://en.wikipedia.org/wiki/Preservation_%28library_and_archival_science%29)

# Preservar

(Del lat. *praeservāre*).

**1. tr.** Proteger, resguardar anticipadamente a una persona, animal o cosa, de algún daño o peligro. **U. t. c. pml.**

*Real Academia Española © Todos los derechos reservados*

**CARO** (€/página)

**DIFÍCIL** (¿qué preservar? ¿cómo hacerlo?)

# Preservar es CARO

Ejemplo real: 13TB de espacio de almacenamiento NAS

## **ADQUISICIÓN (coste inicial)**

Cabina VNXe 3300 (actual NAS en unizar) - coste repercutido a la Institución (“gratis” para la Biblioteca)

13TB netos sobre una nueva bandeja y RAID6 = Necesidad de 10 discos NL-SAS de 7200rpm de 2tb para tener 9+1 HotSpare

TOTAL: 9721€ (+iva)

**MANTENIMIENTO:** 12% del coste del hardware anual

TOTAL: 1000€

**TOTAL en 3 años: 12.721 euros...**

**¡¡solo por el disco para preservar 1891 registros  
de fondo histórico!!**

# Preservar es CARO

Por eso es necesario PLANEAR y PENSAR con sumo cuidado.

## IMPORTANCIA vs CALIDAD

*"Importance" relates to the collection's role as a record, and "quality" covers comprehensiveness, depth, uniqueness, authenticity and reputation of the collection*

¿Es necesario preservarlo todo?

¿El material es único?

¿Para quién tiene valor? ¿En cuánto se estima la cuantía de la pérdida?

¿Hay que preservar masters y copias ligeras? ¿Solo masters? ¿Algunos masters y otros copias ligeras?

¿Cuál es la vida útil del documento?

¿Tienen los masters una calidad de digitalización suficiente? ¿Hay que re-digitalizar algo?

# Preservar es DIFÍCIL

Lo más difícil es convencer de la necesidad y una estrategia (o varias)

**Refreshing** *Refreshing* is the transfer of data between two types of the same storage medium so there are no [bitrot](#) changes or alteration of data.[\[15\]](#) For example, transferring [census](#) data from an old preservation [CD](#) to a new one. This strategy may need to be combined with migration when the [software](#) or [hardware](#) required to read the data is no longer available or is unable to understand the format of the data. Refreshing will likely always be necessary due to the deterioration of physical media.

**Migration** *Migration* is the transferring of data to newer system environments (Garrett et al., 1996). This may include conversion of resources from one [file format](#) to another (e.g., conversion of [Microsoft Word](#) to [PDF](#) or [OpenDocument](#)) or from one [operating system](#) to another (e.g., [Windows](#) to [GNU/Linux](#)) so the resource remains fully accessible and functional. Two significant problems face migration as a plausible method of digital preservation in the long terms. Due to the fact that digital objects are subject to a state of near continuous change, migration may cause problems in relation to authenticity and migration has proven to be time-consuming and expensive for "large collections of heterogeneous objects, which would need constant monitoring and intervention."[\[2\]](#)

**Replication** Creating duplicate copies of data on one or more systems is called *replication*. Data that exists as a single copy in only one location is highly vulnerable to software or hardware failure, intentional or accidental alteration, and environmental catastrophes like fire, flooding, etc. Digital data is more likely to survive if it is replicated in several locations. Replicated data may introduce difficulties in refreshing, migration, versioning, and [access control](#) since the data is located in multiple places.

**Emulation** *Emulation* is the replicating of functionality of an obsolete system. According to van der Hoeven, "Emulation does not focus on the digital object, but on the hard- and software environment in which the object is rendered. It aims at (re)creating the environment in which the digital object was originally created."[\[23\]](#) Examples are having the ability to replicate or imitate another operating system.[\[24\]](#) Examples include emulating an [Atari 2600](#) on a [Windows](#) system or emulating [WordPerfect 1.0](#) on a [Macintosh](#). [Emulators](#) may be built for applications, operating systems, or hardware platforms. Emulation has been a popular strategy for retaining the functionality of old video game systems, such as with the [MAME](#) project. The feasibility of emulation as a catch-all solution has been debated in the academic community. (Granger, 2000) Raymond A. Lorie has suggested a [Universal Virtual Computer](#) (UVC) could be used to run any software in the future on a yet unknown platform.[\[25\]](#) The UVC strategy uses a combination of emulation and migration. The UVC strategy has not yet been widely adopted by the digital preservation community. Jeff Rothenberg, a major proponent of Emulation for digital preservation in libraries, working in partnership with Koninklijke Bibliotheek and National Archief of the Netherlands, developed a software program called Dioscuri, a modular emulator that succeeds in running MS-DOS, WordPerfect 5.1, DOS games, and more.[\[26\]](#) Another example of emulation as a form of digital preservation can be seen in the example of Emory University and the Salman Rushdie's papers. Rushdie donated an outdated computer to the Emory University library, which was so old that the library was unable to extract papers from the harddrive. In order to procure the papers, the library emulated the old software system and was able to take the papers off his old computer.[\[27\]](#)

**Encapsulation** This method maintains that preserved objects should be self-describing, virtually "linking content with all of the information required for it to be deciphered and understood".[\[2\]](#) The files associated with the digital object would have details of how to interpret that object by using "logical structures called "containers" or "wrappers" to provide a relationship between all information components[\[28\]](#) that could be used in future development of emulators, viewers or converters through machine readable specifications.[\[29\]](#) The method of encapsulation is usually applied to collections that will go unused for long periods of time[\[29\]](#)

¿Qué hardware usar? ¿Qué software usar? ¿Estrategia? (refresco, migración, replicación, emulación, encapsulamiento) ¿Qué metadatos mínimos deben añadirse?

¿Qué preservar? ¿Datos+formato? ¿Datos solo?

¿Subcontratar?

# Ejemplo: FH en unizar...

Pleno desarrollo de plan de Preservación... a falta de conocer el presupuesto concedido a tal fin.

FH **Antes**: discos duros USB I.I en cajas externas con los master en TIF etiquetados por firmas. Discos duros metidos en cajas de cartón...

FH **Futuro**: infraestructura de disco NAS y/o cintas de backup + máquina virtual VMWare (estrategia *Refresh*) en RAID6, inventariados, con CRC, con los formatos claramente definidos.

Mantener los otros discos y “moverlos” cada cierto tiempo.

# Ejemplo: FH en unizar..

	<b>NIVEL 1 (proteger datos)</b>	<b>NIVEL 2 (conocer datos)</b>	<b>NIVEL 3 (monitorizar datos)</b>	<b>NIVEL 4 (reparar datos)</b>
<b>Almacenamiento</b>	<p><b>Sacar datos discos duros / DVD's y volcarlos a sistema común</b></p> <p>Dos copias separadas físicamente</p>	<p><b>Tres copias separadas físicamente</b></p> <p>Sistema de almacenamiento documentado y necesidades del mismo</p>	<p>Al menos dos planes ante desastres distintos</p> <p>Monitorización de obsolescencia de los sistemas de almacenamientos empleados</p>	<p>Al menos 3 copias, separadas geográficamente y con distintos planes de desastre</p>
<b>Integridad</b>	<p>Comprobar integridad de datos volcados</p> <p><b>Generación de CRC's (si no existen)</b></p>	<p>Verificar contenidos "peligrosos" (con posibilidad de virus)</p> <p>Usar protección de escritura en sistemas con los datos originales</p>	<p>Mantener logs de chequeos periódicos de los CRC's</p> <p>Verificar virus (todos los contenidos)</p>	<p>Comprobaciones de CRC al producirse determinados eventos</p> <p>Asegurar posibilidad de reconstruir datos corruptos</p> <p>Distintas personas, distintas copias.</p>
<b>Seguridad</b>	<p>Identificar quién tiene permiso para leer, escribir, mover y borrar</p> <p>Restringir dichas autorizaciones</p>	<p>Documentar las restricciones al acceso de los documentos.</p>	<p>Loggear todas las acciones sobre los documentos (incluyendo borrados y acciones específicas de preservación)</p>	<p>Auditar todos los logs</p>
<b>Metadatos</b>	<p><b>Inventariar contenido y localización</b></p> <p>Realizar backup de metadatos</p>	<p>Almacenar metadatos administrativos</p> <p>Loggear transformaciones y eventos</p>	<p>Almacenar metadatos técnicos y descriptivos</p>	<p>Almacenar información de preservación</p>
<b>Formatos de fichero</b>	<p>Limitar el número de formatos</p>	<p>Inventariar los formatos usados</p>	<p>Monitorizar y controlar obsolescencia en los formatos de ficheros empleados.</p>	<p>Realizar migraciones de formatos, emulación o actividades análogas cuando sean necesarias</p>



*That's all Folks!*