

JIIDE  
2012

Madrid, del 17 al 19 de octubre

# De las cartotecas a las IDE: puesta en servicio de una pasarela entre los estándares de catalogación MARC21 e ISO19115

Marta Criado, María Crespo, Joan Capdevila, José M. Agudo, F.Javier Zarazaga-Soria, Jesús Barrera, Alejandra Sánchez, Carolina Soteres

Madrid, 18 de octubre de 2012



IAAA  
Sistemas de Información  
Avanzados  
Universidad Zaragoza



# Índice

## □ Introducción

- ❖ Bibliotecas y archivos

## □ Convergencia entre los centros de información y las IDEs

- ❖ Beneficios de la integración

## □ Subgrupo de Trabajo de Patrimonio Cartográfico en la IDE

- ❖ Pasarela MARC21 - ISO 19115

## □ Conclusiones



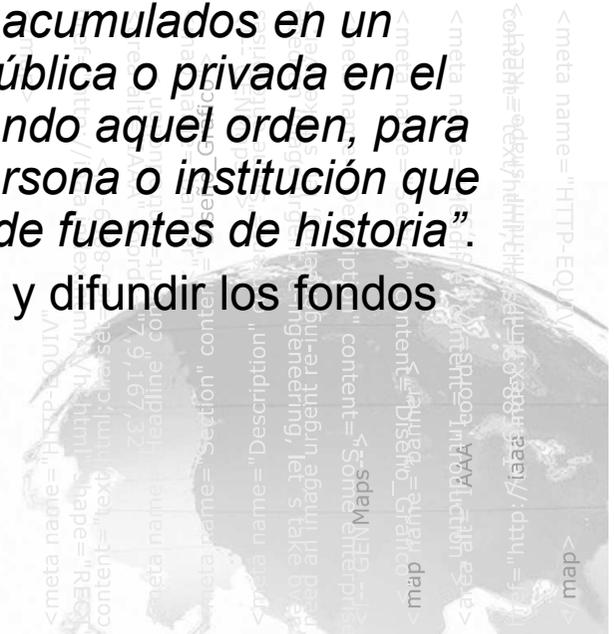
# Introducción

## □ Bibliotecas:

- ❖ Según la ALA (American Library Association) una biblioteca es *una colección de material de información organizada para que pueda acceder a ella un grupo de usuarios*.
- ❖ Misión: conservar los recursos informativos y prestar servicios mediante diversos medios con el fin de cubrir las necesidades de sus usuarios

## □ Archivos:

- ❖ Según Heredia Herrera es *“uno o más conjuntos de documentos, sea cual sea su fecha, su forma y soporte material, acumulados en un proceso natural por una persona o institución pública o privada en el transcurso de su gestión, conservados, respetando aquel orden, para servir como testimonio e información para la persona o institución que los produce, para los ciudadanos o para servir de fuentes de historia”*.
- ❖ Misión: recoger, organizar, clasificar, conservar y difundir los fondos documentales



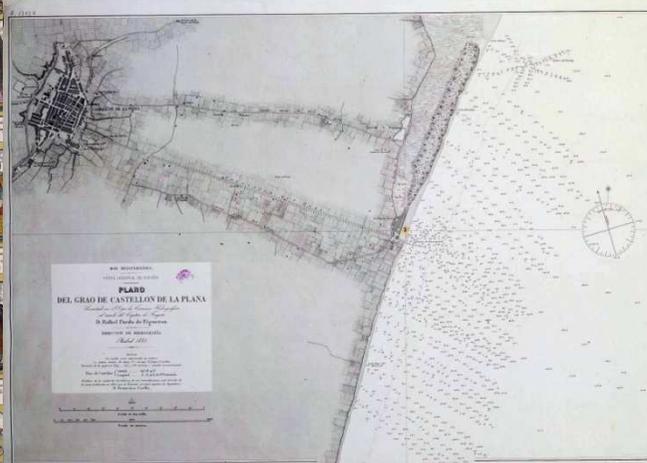
# Fondos en bibliotecas y archivos

- ❑ Tanto las bibliotecas como los archivos ofrecen una amplia diversidad:
  - ❖ de tipologías:
    - bibliotecas nacionales, públicas, escolares, universitarias, especializadas;
    - archivos municipales, provinciales, históricos, etc.
  - ❖ de fondos: libros, revistas, legislación, recursos electrónicos, materiales especiales



# El material cartográfico en bibliotecas y archivos

- Material cartográfico: *“todos los documentos que representan, en su totalidad o en parte, la tierra o los cuerpos celestes a cualquier escala, como los mapas y planos en dos y tres dimensiones, las cartas aeronáuticas, marinas y celestes, los globos, los bloques-diagramas, los cortes de terreno, las fotografías aéreas, los atlas, las vistas a vuelo de pájaro, etc”* (Reglas de Catalogación Españolas, 1999)



# Convergencia entre los centros de información y las IDEs

- Semejanzas entre centros de información y las IDEs
  - ❖ Materiales gráficos e información geográfica identificados a través de normas y estándares
  - ❖ Misión de recopilar, clasificar, ordenar, compartir y difundir la información
- Posibilidad de acercar posiciones



# Beneficios de la integración de las cartotecas en una IDE

- Las ventajas de utilizar las IDE: la georeferenciación ...
  - ❖ Búsquedas geográficas
  - ❖ Comparaciones inmediatas
  - ❖ Con capacidad de medida
- ... interoperable
  - ❖ Sobre distintos catálogos
  - ❖ Superponiendo mapas de diferentes fuentes
  - ❖ Midiendo lo mismo



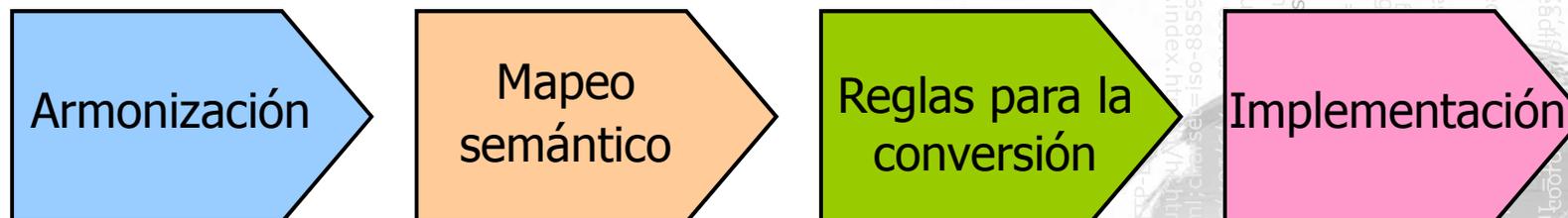
# Subgrupo de Trabajo de Patrimonio Cartográfico en la IDE

- ❑ Constituido en 2008
- ❑ Grupo interdisciplinar coordinado por el Instituto Geográfico Nacional y compuesto por personal que trabaja en bibliotecas, institutos cartográficos, universidades y centros de investigación
- ❑ Objetivo: estimular la publicación de cartografía histórica a través de las IDEs
- ❑ Primer gran proyecto : pasarela entre el formato de catalogación MARC21 y el formato de metadatos ISO 19115



# Proyecto GTI PC-IDE: interoperabilidad de registros MARC21 e ISO19115

- ❑ Formato MARC (MACHINE Readable Cataloguing):
  - ❖ muy extendido en el ámbito bibliotecario para intercambiar registros
  - ❖ el patrimonio cartográfico se conserva, describe y facilita a través de este formato
- ❑ ISO19115 Geographic Information – Metadata:
  - ❖ respaldado por la International Organization for Standardization
  - ❖ recomendado por INSPIRE
  - ❖ el más extendido y recomendado para describir información geográfica
- ❑ Metodología
  - ❖ Documentada y aplicada previamente en otros contextos (\*)
    - Pasarelas Dublin Core - ISO19115, CSDGM – ISO19115
  - ❖ Fases en la ejecución de la pasarela



(\*) J. Nogueras-Iso, F.J. Zarazaga-Soria, J. Lacasta, R. Béjar, P.R. Muro-Medrano, 2004. "Metadata standard interoperability: application in the geographic information domain". Computers, Environment and Urban Systems, Vol. 28/6, pp. 611-634.



# Fase 1: armonización

## ❑ MARC21 (estudio realizado por el Instituto Cartográfico de Cataluña)

- ❖ Perfil mínimo de MARC 21 para la catalogación de documentos cartográficos del Consorci de Biblioteques Universitàries de Catalunya (CBUC).
- ❖ Los campos específicos que contiene el MARC 21 para documentos cartográficos

## ❑ ISO19115 (estudio realizado por el Instituto Geográfico Nacional)

### ❖ Perfil NEM

Etiqueta	Ind. 1	Ind. 2	Nombre elemento	Restric (3)	Multiplicidad (4)	Definición semántica	Tipo de datos (5)	Valores por defecto (6)
<b>Cabecera</b>								
pos. 00-04			Longitud del registro	M	NR	Cadena de 5 caracteres numéricos generada automáticamente que especifica la longitud del registro.	Numérico	
			Estado del registro catalogado	M	NR	Indica la relación del registro con un fichero.	a - Incremento del nivel de codificación / c - corregido o revisado / d - borrado / n - nuevo / p - incremento del nivel de codificación de prepublicación.	
			Tipo de registro	M	NR	Indica las características del registro y define sus componentes.	e - Material cartográfico (mapas, atlas, globos, mapas digitalizados y otros documentos cartográficos) / f - material cartográfico manuscrito	e - Material cartográfico (mapas, atlas, globos, mapas digitalizados y otros documentos cartográficos)
			Nivel bibliográfico del registro	M	NR		a - Parte componente monográfico / b - parte componente seriada / c - Colección / d - subunidad / i - Recurso integrado / m - Monografía/documento / s - Publicación en serie	Manual

Nº (1)	Nombre elemento	Abreviatura (2)	Restricciones (3)	Multiplicidad (4)	Definición semántica	Tipo de datos
2	Identificador del fichero	mdFileID	Op	1	Identificador único para el fichero de metadatos	Cadena de caracteres
3	Idioma de los metadatos	mdLang	C si no está definido en <i>Encoding</i>	1	Idioma usado para documentar metadatos	Cadena de caracteres
4	Conjunto de caracteres de los metadatos	mdChar	C si ISO 10646-1 no la usó y no está definida	1	Nombre completo de la norma de codificación de caracteres usada en el conjunto de metadatos	Clase
6	Nivel jerárquico	mdHrLv	C si "Nivel jerárquico" no es igual para todo el conjunto de datos	N	Subconjunto de datos al que se refieren estos metadatos	Clase
8	Punto de contacto para los metadatos					374
374	CI_ResponsibleParty					375-376-377-378-379
375	Nombre individual	rplndname	C si "organisationName" y "positionName" no está documentado	1	Nombre de la persona responsable dando apellido, nombre y cargo, separados por un delimitador	Cadena de caracteres

# Fase 2: mapeo semántico

Elemento INSPIRE/ISO/NEM [C1]	Elemento MARC21 [C2]	Valores de los elementos de MARC21 [C3]	Observaciones [C4]
Conjunto de caracteres de los metadatos [Nombre completo de la norma de codificación de caracteres de conjunto]	No se ha		Propuesta DMSGroup: Este elemento es condicional según ISO/NEM. Si no se puede cruzar con ningún elemento de MARC21, podría

trabajo

Elemento INSPIRE/ISO/NEM [C1]	Propuesta DMSGroup	Respuesta BNE	Resp	Elemento ISO/NEM v.1.1/INSPIRE	Cruce	Respuesta CICTEX	Porcentaje de respuestas	Cruce
<p>Lista controlada</p> <p>001=Ucs2 002=Ucs4 003=Utf7 (Utf8) 005=Utf16 006=8859 007=8859 008=8859 009=8859 010=8859 011=8859 012=8859 013=8859 014=8859 015=8859 016=8859</p> <p>Título [Nombre por el que se conoce el recurso mencionado]</p>	<p>Propuesta DMSGroup: El título que aparezca en INSPIRE/ISO/NEM deberá componerse con los 3 elementos seleccionados en C2 (245\$a – 130\$a – 130\$f). De esta forma se podrá asegurar mejor las búsquedas de los mapas por el campo Título.</p>	<p>NO El 245 describe fielmente el título de un mapa, no tiene que añadir nada más</p>		<p>Título</p> <p>245\$a + 008 pos.07-10, 130\$f</p>				
				<p>Título alternativo</p> <p>246 \$a</p>				
				<p>Forma de presentación</p> <p>Lista controlada</p> <p>001= Documento digital 002= Copia impresa del documento 003= Imagen digital 004= Copia impresa de la imagen 005= Mapa digital 006= Copia impresa del mapa 007= Modelo digital 008= Copia impresa del modelo 009= Perfil digital 010= Copia impresa del perfil 011= Tabla digital 012= Copia impresa de la tabla 013= Video digital 014= Video en cinta 015= Material cartográfico manuscrito 016= Material manuscrito textual 017= Recurso electrónico 018= Globo</p>	<p>Cabecera pos.06: a(material impreso textual)=002, e(material cartográfico)=006, f(material cartográfico manuscrito)=015 (introducir como nuevo valor), t(material manuscrito textual)=016 (introducir como nuevo valor)</p> <p>Campo 006-Mapas pos.00: e(material cartográfico)=006, f(material cartográfico manuscrito)=015 (introducir como nuevo valor)</p> <p>Campo 007-Mapas pos.00: a(mapa)=006, c(recurso electrónico)=017 (introducir como nuevo valor en la lista), d(globo)=018 (introducir como nuevo valor), r(imágenes con sensores remotos)=003</p>		<p>33 %</p>	
<p>Título alternativo</p> <p>[Acrónimo o nombre en otro idioma por el que se conoce el recurso mencionado]</p>	<p>Propuesta DMSGroup: El elemento 246 (\$a) de C2 es completado de forma minoritaria por las Bibliotecas. En este caso, se ha optado por realizar el cruce al encontrar un elemento de ISO de forma directa.</p>	<p>SI</p>				<p>SI</p>	<p>Porcentaje de respuestas afirmativas (SI): 100 %</p>	<p>Título alternativo = 246 \$a</p>



# Fase 3: reglas adicionales para la conversión



Grupo de Trabajo Interdisciplinar de Patrimonio Cartográfico  
Pasarela entre formatos de catalogación MARC21 – INSPIRE, ISO19115,  
NEM



Documento de definición de reglas de conversión

Elemento ISO/NEM v.1.1/INSPIRE	Cruce	Reglas de conversión
<b>Título</b>	245\$a + 008 pos.07-10, 130\$f	<p>*El elemento de ISO "Título" se completará con la unión de varios elementos de MARC21.</p> <p>*La fecha del elemento "Título" de ISO se completará primero con el elemento 008 pos.07-10 de MARC21. Si no hay información en este elemento la pasarela debe leer la información del elemento 130\$f.</p>
<b>Título alternativo</b>	246 \$a	*El cruce de elementos es directo
<b>Forma de presentación</b> <b>Lista controlada</b> 001= Documento digital 002= Copia impresa del documento 003= Imagen digital 004= Copia impresa de la imagen 005= Mapa digital 006= Copia impresa del mapa 007= Modelo digital 008= Copia impresa del modelo 009= Perfil digital 010= Copia impresa del perfil 011= Tabla digital 012= Copia impresa de la tabla 013= Video digital 014= Video en cinta 015= Material cartográfico manuscrito 016= Material manuscrito textual 017= Recurso electrónico 018= Globo	<b>Cabecera pos.06:</b> a(material impreso textual)=002, e(material cartográfico)=006, f(material cartográfico manuscrito)=015(introducir como nuevo valor), t(material manuscrito textual)=016(introducir como nuevo valor) <b>Campo 006-Mapas pos.00:</b> e(material cartográfico)=006, f(material cartográfico manuscrito)=015 (introducir como nuevo valor) <b>Campo 007-Mapas pos.00:</b> a(mapa)=006, c(recurso electrónico)=017 (introducir como nuevo valor en la lista), d(globo)=018 (introducir como nuevo valor), r(imágenes con sensores remotos)=003	<p>*El elemento de ISO "Forma de presentación" se completará con la información de tres elementos de MARC21.</p> <p>*Si el valor de MARC21 se repite la pasarela debe detectarlo y no duplicar el valor.</p> <p>*Primer paso: cruzar los valores de MARC21 con los valores de la lista controlada.</p> <p>*Segundo paso: ampliar la lista controlada de ISO con los valores de MARC21 que no se han podido cruzar en El Primer paso.</p>
<b>Fecha de referencia del conjunto de datos</b> <b>Lista controlada</b> Fecha de creación Fecha de publicación Fecha de revisión	<b>Fecha de creación</b> = 008 pos. 07-10; 130\$f  <b>Fecha de publicación</b> = 260 \$c	<p>*El elemento de ISO "Fecha de referencia del conjunto de datos" se cruzará primero con el elemento de 008 pos.07-10 de MARC21.</p> <p>*Si no hay información en este elemento de MARC21 la pasarela debe leer la información del elemento 130\$f.</p> <p>*Este elemento sólo se puede completar en ISO según la</p>



## Fase 4: implementación

- ❑ Incorporación de la pasarela a una herramienta OpenSource de creación de metadatos: CatMDEdit



**CatMDEdit** OpenSource Project

<http://catmdeedit.sourceforge.net/>

About CatMDEdit

Download

Manuals

License

Credits

Related publications

Related projects

Contact

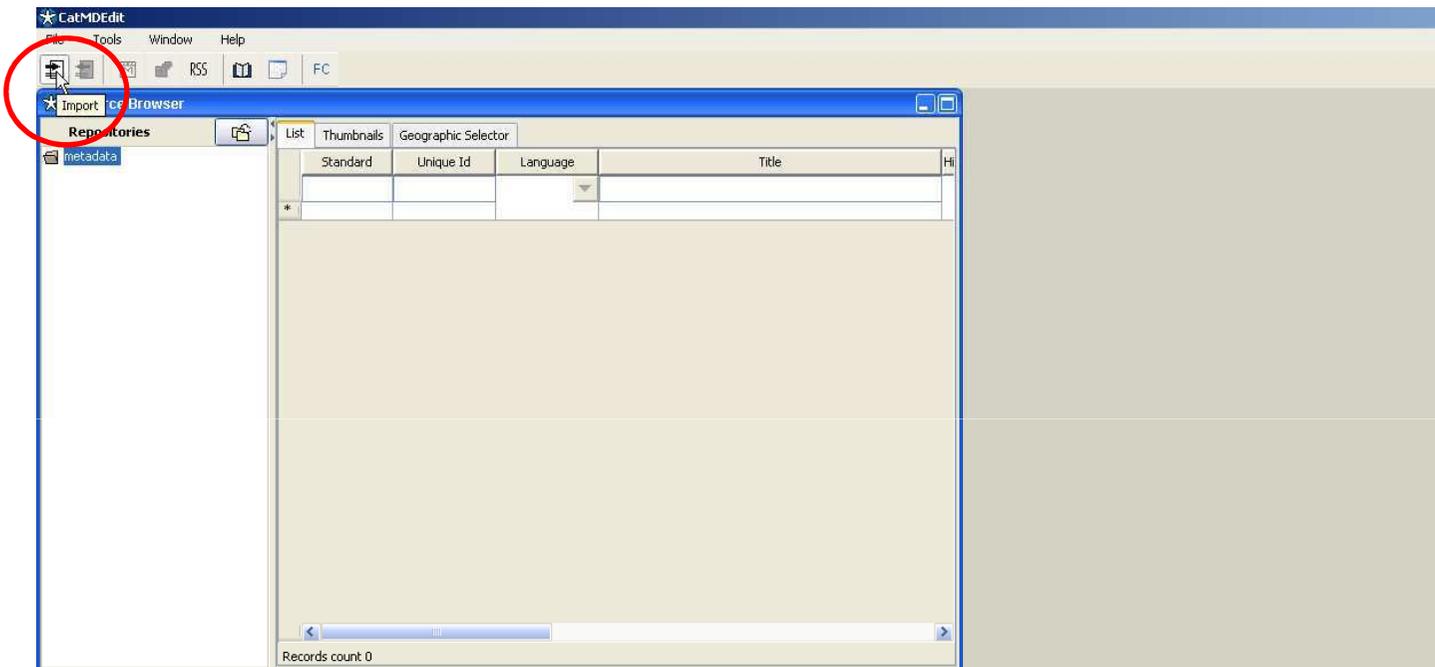
CatMDEdit is a metadata editor tool that facilitates the documentation of resources, with special focus on the description of geographic information resources. It is an initiative of the National Geographic Institute of Spain (IGN), which is the result of the scientific and technical collaboration between IGN and the Advanced Information Systems Group (IAAA) of the University of Zaragoza with the technical support of GeoSpatiumLab (GSL).

The tool has been implemented in Java and has the following features:

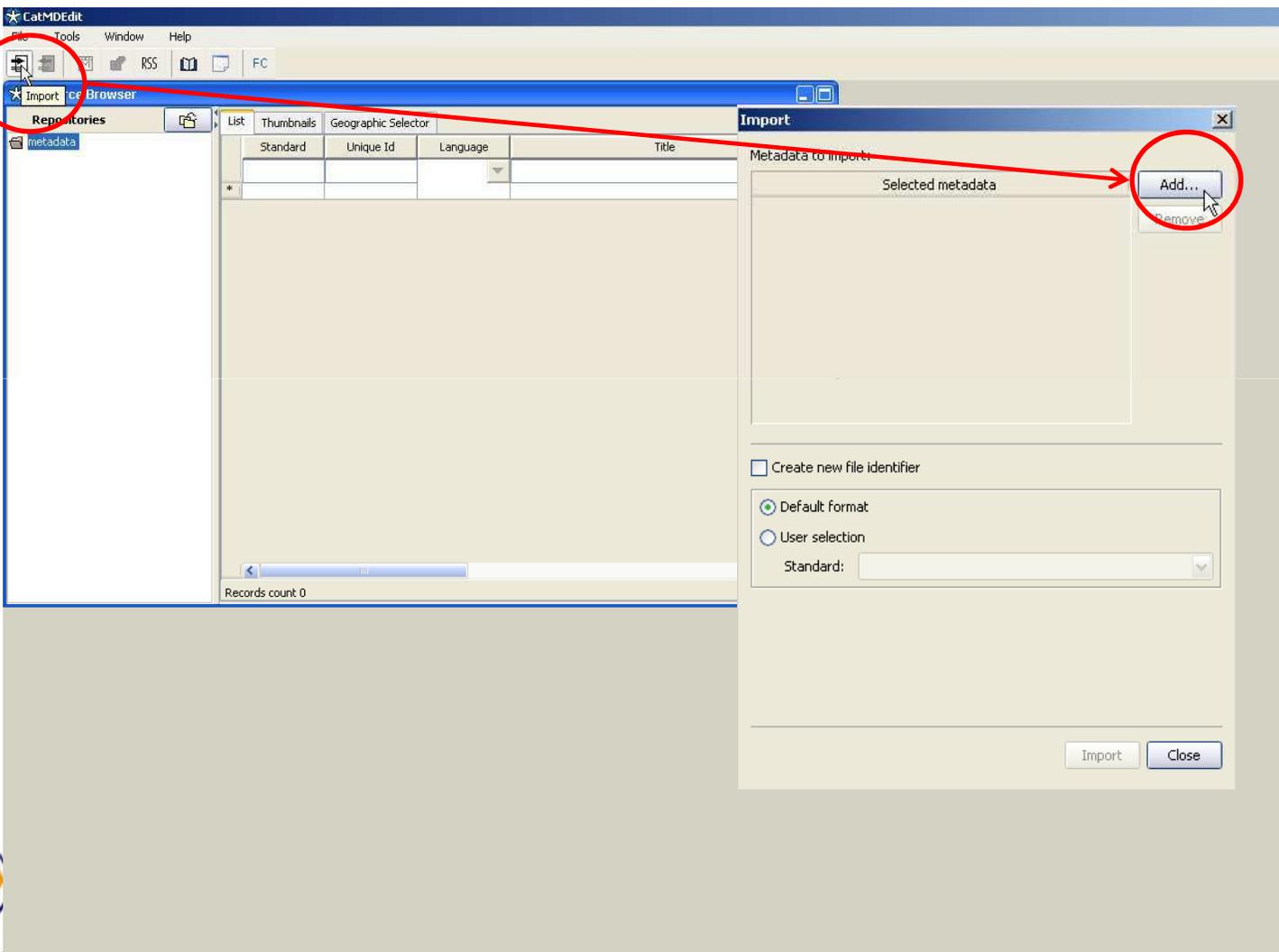
- Multi-platform (Windows, Unix). As it has been developed in Java and the storage of metadata records is managed directly through the file system, the application can be deployed in any platform with the minimum requirement of having installed a Java virtual machine.
- Multilingual. The application has been developed following the Java internationalization methodology. Nowadays, there is a Spanish, English, French, German, Polish, Portuguese and Czech version. [Contributors](#) are welcome for customization to other languages.
- Definition and management of different metadata repositories (repositories may also contain data files), including the selection and filtering of metadata records stored in each local metadata repository.
- Metadata edition in conformance with "ISO 19115. Geographic Information - Metadata" standard ([ISO 19115:2003/Cor 1 2006](#), ISO/TS 19139:2007 XML encoding). Edition interfaces adapted to different metadata profiles:



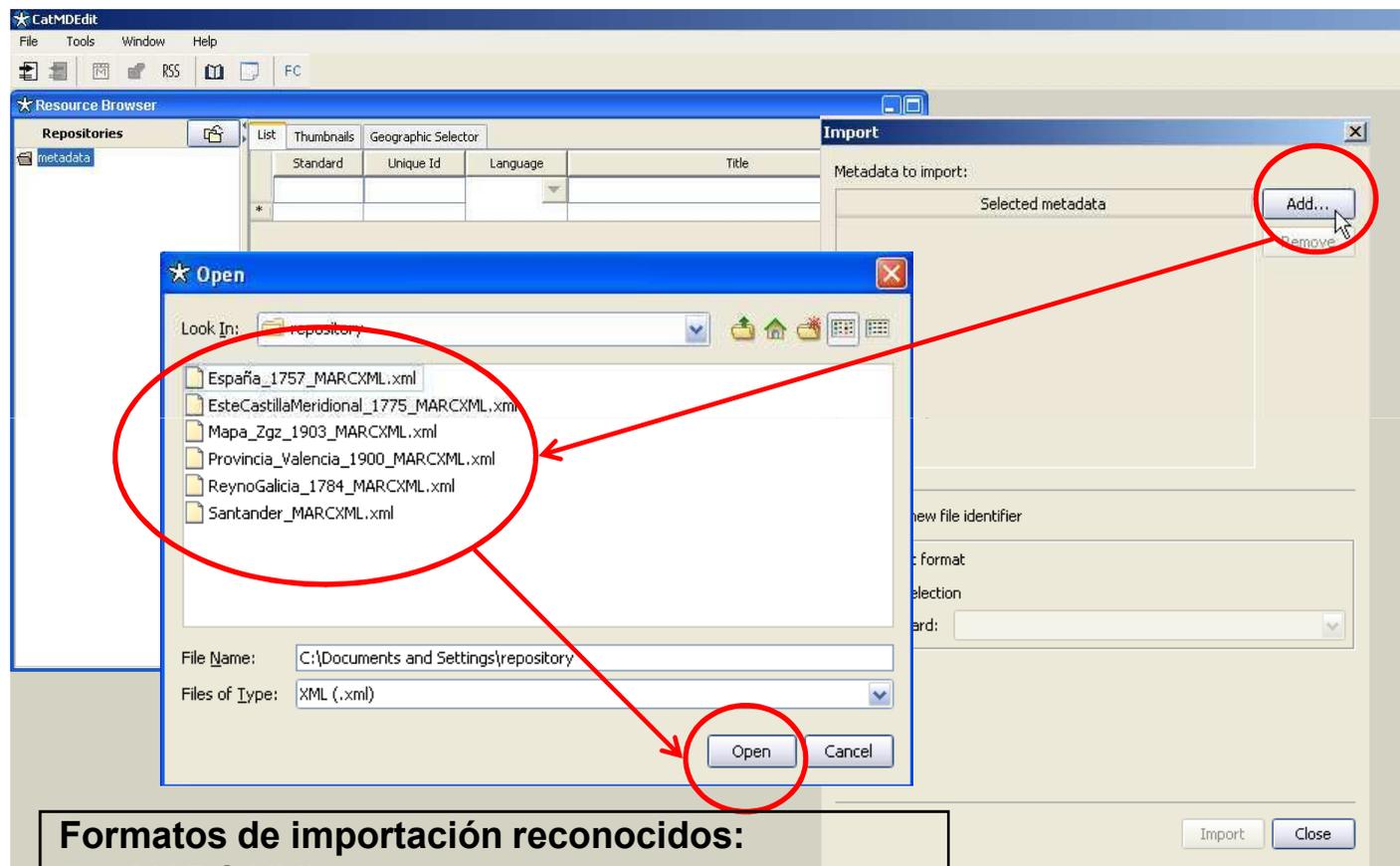
# Ejemplo práctico: importación de registros MARC21



# Ejemplo práctico: importación de registros MARC21

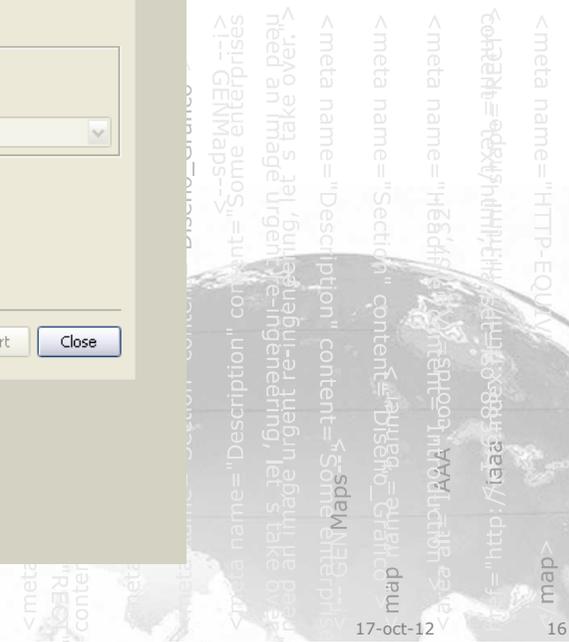


# Ejemplo práctico: importación de registros MARC21



## Formatos de importación reconocidos:

- MARC XML
- MARC21 communications format – ISO 2709
- MARC21 etiquetado



# Ejemplo práctico: importación de registros MARC21

The screenshot shows the CatMDEdit application with the following components:

- Resource Browser:** Shows a list of repositories. The 'metadata' repository is selected.
- Open Dialog:** Shows the 'Look In' field set to 'repository'. A list of XML files is displayed:
  - España\_1757\_MARCXML.xml
  - EsteCastillaMeridional\_1775\_MARCXML.xml
  - Mapa\_Zgz\_1903\_MARCXML.xml
  - Provincia\_Valencia\_1900\_MARCXML.xml
  - ReynoGalicia\_1784\_MARCXML.xml
  - Santander\_MARCXML.xmlThe 'File Name' field is set to 'C:\Documents and Settings\repository' and 'Files of Type' is set to 'XML (.xml)'. The 'Open' button is circled in red.
- Import Dialog:** Shows the 'Metadata to import:' section with a list of selected metadata files:
  - España\_1757\_MARCXML.xml
  - EsteCastillaMeridional\_1775\_MARCXML.xml
  - Mapa\_Zgz\_1903\_MARCXML.xml
  - Provincia\_Valencia\_1900\_MARCXML.xml
  - ReynoGalicia\_1784\_MARCXML.xml
  - Santander\_MARCXML.xmlThe 'Import' button is circled in red.



# Ejemplo práctico: importación de registros MARC21

The screenshot shows the CatMDEdit application interface. The main window displays a table of records with the following columns: Standard, Unique Id, Language, and Title. A red circle highlights a row in the table. An 'Import' dialog box is open, showing a list of selected metadata files and a list of files that have been successfully imported.

Standard	Unique Id	Language	Title
ISO19115	esp-M-RAH-Esp...	Spanish	España, Dedicada al Re
ISO19115	fra-GR-GR-BVA...	Spanish	Partie meridionale des
ISO19115	esp-ARGMM-ZA...	Spanish	Zaragoza - 1903
ISO19115	esp-M-RAH-Val...	Spanish	[Mapa de la provincia
ISO19115	esp-M-RAH-Gali...	Spanish	Mapa geográfico del Re
ISO19115	esp-ARGMM-Sa...	Spanish	Santander - 1903

Import dialog box details:

- Metadata to import: España\_1757\_MARCXML.xml, EsteCastillaMeridional\_1775\_MARCXML.xml, Mapa\_Zgz\_1903\_MARCXML.xml, Provincia\_Valencia\_1900\_MARCXML.xml, ReynoGalicia\_1784\_MARCXML.xml, Santander\_MARCXML.xml
- Standard: ISO 19115
- Imported files: Mapa\_Zgz\_1903\_MARCXML.xml [OK], Provincia\_Valencia\_1900\_MARCXML.xml [OK], ReynoGalicia\_1784\_MARCXML.xml [OK], Santander\_MARCXML.xml [OK]



# Ejemplo práctico: importación de registros MARC21

The screenshot shows the CatMDEdit application interface. The main window is titled "Edition: esp-M-RAH-España-Mapas generales-20101013112331.0-RAH20101000577 - España. Dedicada al Rey N. S. D. Fer...". The interface includes a menu bar (File, Tools, Window, Help) and a toolbar with icons for File, RSS, and FC. A "Resource Browser" window is open, showing a "Repositories" list with "metadata" selected. The main editing area is divided into two panes. The left pane shows a map of Spain with a scale bar (0.0 to 5.0) and a red circle around a mouse cursor pointing to a specific location. The right pane displays the metadata editor for the selected record. The editor has tabs for "HTML", "ISO19115", "NEM Profile", "CORE Profile", "INSPIRE Profile", and "WISE Profile". The "ISO HTML (en)" tab is active, showing the title "España. Dedicada al Rey N. S. D. Fernando VI, que Dios guar.de -" and an "Index" of metadata fields: Metadata file identifier, Language, Character set, Contact, Metadata date stamp, Metadata standard name, Metadata standard version, Reference System information, Identification Information, and Data quality information. The "Metadata file identifier" field is highlighted in green and contains the value "esp-M-RAH-España--Mapas generales-20101013112331.0-RAH20101000577". At the bottom of the editor, there are buttons for "Edit", "Save", "Cancel", and "Validate".



# Ejemplo práctico: exportación de ISO19115 a MARC21

The screenshot displays the CatMDEdit software interface. On the left, the 'Resource Browser' shows a table of metadata records. The first record is highlighted with a red dashed box. A red arrow points from this record to the 'Edition' window on the right, which shows the XML structure of the metadata. The XML view includes a 'Citation' section with the following content:

```
<gmd:CI_Citation>
  <gmd:title>
    <gco:CharacterString>Límites hidrológicos de la Cuenca del Ebro</gco:CharacterString>
  </gmd:title>
  <gmd:date>
    <gmd:CI_Date>
      <gmd:date>
        <gco:Date>2007-03-22</gco:Date>
      </gmd:date>
      <gmd:dateType>
        <gmd:CI_DateTypeCode>
          <gco:CharacterString>public</gco:CharacterString>
        </gmd:CI_DateTypeCode>
      </gmd:dateType>
    </gmd:CI_Date>
  </gmd:date>
</gmd:CI_Citation>
```

The 'Edition' window also shows a 'Citation' section with the following content:

```
<gmd:CI_Citation>
  <gmd:title>
    <gco:CharacterString>Límites hidrológicos de la Cuenca del Ebro</gco:CharacterString>
  </gmd:title>
  <gmd:date>
    <gmd:CI_Date>
      <gmd:date>
        <gco:Date>2005-03-08</gco:Date>
      </gmd:date>
      <gmd:dateType>
        <gmd:CI_DateTypeCode>
          <gco:CharacterString>public</gco:CharacterString>
        </gmd:CI_DateTypeCode>
      </gmd:dateType>
    </gmd:CI_Date>
  </gmd:date>
</gmd:CI_Citation>
```

# Ejemplo práctico: exportación de ISO19115 a MARC21

**Export Dialog Configuration:**

- Metadata to export: Selected metadata
- Standard: ISO 19115
- Format: MARC21 ISO2709 (.mrc)
- Output directory: C:\Documents and Settings\...

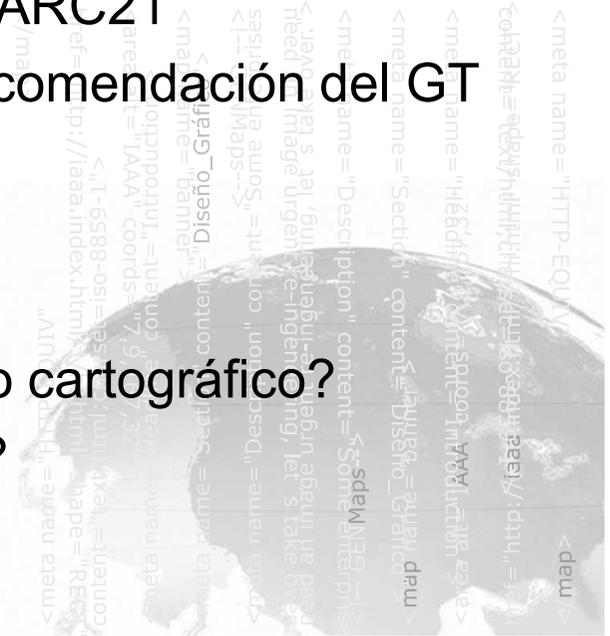
**MARC21 Record:**

```

01163nam #2200205 a
4500001000300000000500090000300600100001200800390002203200410006103400870010204000950
018924500470028424600050033125000120033626000150034835200540036350000970041750603910
0514856005200905R$IdR$20070322R$e bR$121009
es R$ U$a6472068c-cb4d-11e1-93e1-001cc0ae8493R$
U$d-004.405307586015035U$e001.9893912234775344U$f043.17500036089162U$g040.25554780
617321R$ U$besU$aGrupo IAAA. Depto. Informática e Ingeniería de Sistemas.
Universidad de Zaragoza | IGMU$aLímites hidrológicos de la Cuenca del
EbroR$0 U$aR$ U$aUnknownR$ U$c2005-03-08R$ U$bsurfaceU$clU$gJPEG -Joint
Photographic Group FormatR$ U$aCobertura de los límites hidrológicos de la
cuenca hidrográfica del Ebro, a escala 1:50.000.R$1 U$aDebe tenerse en cuenta que
el presente geodataset ha sido creado con la finalidad de servir de cobertura
territorial de base en los trabajos relacionados con el Plan Hidrológico de la
cuenca del Ebro.. La Confederación Hidrográfica del Ebro agradece a todos los
usuario de nuestras coberturas por favor se sirvan citar la fuente de procedencia
de la información gráfica en sus trabajos.R$ U$gARCE - ARC/INFO Export format |
SHP | DBF | SHXR$CS
    
```

# Conclusiones

- ❑ Finalización de un trabajo de más de cuatro años
  - ❖ Pasarela disponible desde junio de 2.012
    - Incluida en versión 4.6.6 de CatMDEdit
- ❑ Pero esto no acaba aquí...
  - ❖ Test de robustez
    - Llevar a cabo un caso de uso en una cartoteca con un fondo importante de registros en MARC21
  - ❖ Proponer la pasarela definida como recomendación del GT IDEE
  - ❖ Otras ideas de cara al futuro
    - TimeML?
    - Catálogo de recursos en patrimonio cartográfico?
    - Geoportal Patrimonio Cartográfico?





```

<meta name="HTTP-EQUIV"
content="text/html; charset=iso-8859-1" >
<area alt="IAAA" coords="167, 32, 179, 167" shape="RECT" />
<area alt="Diseño Gráfico" coords="167, 167, 321, 167" shape="RECT" />
<meta name="Section" content="Diseño Gráfico" />
<meta name="Description" content="Some surprises
and an image urgent re-inventing, let's take over." />
<meta name="Section" content="Diseño Gráfico" />
<meta name="Section" content="Diseño Gráfico" />
<map name="GeoLab" points="100,100,100,100" />
  
```