## VI Encuentro Nacional de Catalogadores

"Teoría vs. Práctica en la organización y el tratamiento de la información"

# 15, 16 y 17 de noviembre de 2017 Biblioteca Nacional Mariano Moreno Buenos Aires, Argentina

Autores: Lic. Pamela Chatelain<sup>1</sup> - Mter. Mónica Muela <sup>2</sup>

**Institución:** Biblioteca Mayor - Universidad Nacional de Córdoba.

Dirección: Obispo Trejo 242, 1º piso. CP 5000 - Córdoba / Argentina

**Título:** "Evolución y vicisitudes en el establecimiento de metadatos para material antiguo de Biblioteca Mayor de la Universidad Nacional de Córdoba"

#### Resumen:

En el año 2009 se comienza a trabajar en la digitalización de las colecciones históricas pertenecientes a la Universidad Nacional de Córdoba, siendo la Biblioteca Mayor la pionera.

El presente trabajo pretende describir la experiencia transcurrida en cuanto al desarrollo, definición e implementación de metadatos para material antiguo, describiendo la evolución y vicisitudes que se sucedieron en cuanto al establecimiento de políticas de carga de metadatos descriptivos, estructurales y de preservación,

desde el momento en el que se definió la colección a digitalizar hasta el desarrollo definitivo de la plataforma de visualización de las colecciones digitales llamada "*Biblioteca Mayor Digital*".

<sup>1 -</sup> **Lic. Pamela Chatelain.** Responsable del Departamento de Procesos Técnicos. Biblioteca Mayor de la Universidad Nacional de Córdoba. Contacto: <a href="mailto:pamela@bmayor.unc.edu.ar">pamela@bmayor.unc.edu.ar</a> / (54-351) 157533000

<sup>2 -</sup> Mter. Mónica Muela. Responsable del Departamento de Digitalización. Biblioteca Mayor de la Universidad Nacional de Córdoba. Contacto: <a href="mailto:monimuela@bmayor.unc.edu.ar">monimuela@bmayor.unc.edu.ar</a> / (54-351) 156181771

### 1a. parte: los inicios

En el año 2009 se comienza a trabajar en el proyecto de digitalización de la Colección Jesuítica del fondo de Biblioteca Mayor, en el marco de un proyecto macro de cooperación entre la Universidad Nacional de Córdoba y la Universidad Complutense de Madrid. El proyecto llevado a cabo consistió en la selección, acondicionamiento, digitalización y carga de metadatos de unos 500 volúmenes pertenecientes a dicha colección.

Para determinar el set de metadatos que se utilizarían, fue necesario hacer un estudio previo. Al no haber un modelo estándar de metadatos para material antiguo, lo que se hizo fue relevar en las Bibliotecas Digitales de materiales antiguos existentes en ese momento qué metadatos se utilizaban. Se tomaron como principales referentes la Biblioteca Nacional Argentina, la Universidad Complutense de Madrid, la Biblioteca Histórica de Santa Cruz de la Universidad de Valladolid. En base a esta información recabada se tomó la decisión de crear una grilla de metadatos propia, es decir, sin recurrir a ningún esquema de metadatos estandarizado.

Los metadatos definitivos se definieron en tres grupos:

<u>Metadatos Descriptivos</u>: describen e identifican la información contenida en el recurso, contienen atributos físicos y bibliográficos (título, autor/ creador, idioma, palabras claves), los cuales son "levantados", en su mayor parte, de la base de datos de la biblioteca, cargados en el Software Koha (Sistema Integrado de Gestión de Bibliotecas), en formato Marc 21, bajo las normas AACR2 de catalogación.

<u>Metadatos Estructurale</u>s: éstos facilitan la navegación y presentación de las distintas partes que configuran a las obras digitalizadas proporcionando información sobre su estructura interna (página, sección, capítulo, partes, índices, tabla de contenido, etc.) y los distintos archivos que conforman la obra. También registra el tipo de encuadernación, presencia de exlibris y de qué tipo: manuscritos, sello o impresos, presencia de ilustraciones o grabados y de qué tipo, xilográficos o calcográficos:

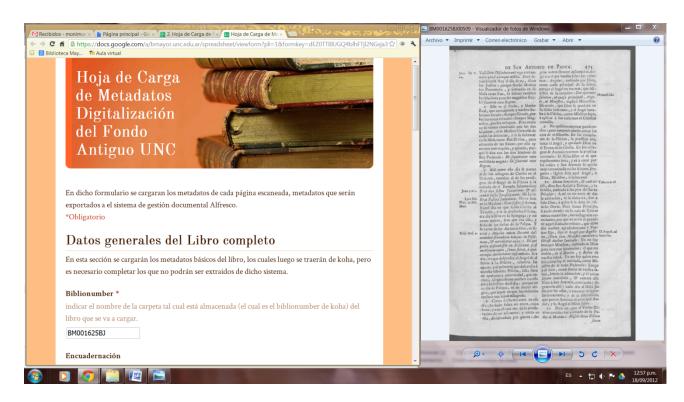
<u>Metadatos de Preservación</u>: éstos tienen por objetivo facilitar el uso de los archivos en el futuro, describiendo las especificaciones técnicas en su creación y las características particulares de cada imagen.

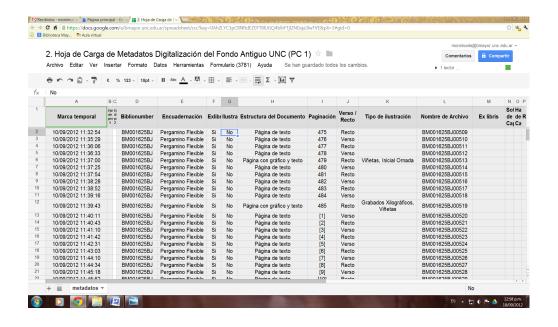
#### La carga de los metadatos

Antes de ser digitalizados en marco del proyecto arriba mencionado,

los libros seleccionados pasaban por el departamento de procesos técnicos para realizar una recatalogación exhaustiva de estos ítems, los metadatos descriptivos del libro ya digitalizado se tomarían posteriormente de esta catalogación provenientes de su correspondiente registro en el catálogo de la Biblioteca.

La carga de metadatos estructurales se realizaba en una planilla de cálculo llamada "Hoja de Carga de Metadatos Digitalización Fondo Antiguo UNC", alojada en la plataforma de Google Apps, que permite la edición de un mismo archivo por varias personas al mismo tiempo. Cada miembro del equipo podía acceder a esta planilla a través de su correo electrónico oficial. En ella se cargaba el conjunto de datos que describen cada una de las imágenes que forman parte del libro digitalizado. Para la carga se trabajó con el libro en mano, de esta manera también se realizaba un segundo control de la digitalización de cada imagen que componía el libro en su totalidad.





Una vez concluido todo el proceso de digitalización, es decir, una vez terminados los controles de calidad de las imágenes, la carga de metadatos y la puesta a disposición de los usuarios del material en el nuevo formato, el original es llevado al Departamento de Preservación para su última revisión. Finalmente, el original es devuelto a su estantería.

### Detalle de metadatos propios

- 1- Biblionumber [BM + Inventario (6 dígitos) + N° de imagen (5 dígitos)]
- 2- Encuadernación
  - 2.1- Pergamino:
    - 2.1.1-Flexible
    - 2.1.2- Rígido
  - 2.2- Cuero:
    - 2.2.1-Flexible
    - 2.2.2- Rígido
  - 2.3- Ornamentada
  - 2.4- Rígida en cartón
- 3- No Encuadernado
- 4- Estructura del documento
  - 4.1- Tapa

- 4.2- Lomo
- 4.3- Hoja de guarda
- 4.4-Frontispicio
- 4.5- Anteportada
- 4.6- Portada
- 4.7- Cuerpo de la obra:
  - 4.7.1- Página en blanco
  - 4.7.2- Página con gráfico y texto
  - 4.7.3- Página de texto
  - 4.7.4- Lámina
- 4.7- Colofón
- 4.8- Contratapa
- 5- Paginación
  - 5.1- (Número o letra, tal como aparece en el libro)
  - 5.2- Recto
  - 5.3- Verso
- 6- Ex-libris (repetible)
  - 6.1-Sello
  - 6.2- Impreso
  - 6.3- Manuscrito
- 7- Ilustración (repetible)
  - 7.1- Grabado Xilográfico
  - 7.2- Grabado Calcográfico
  - 7.3- Viñeta
  - 7.4- Inicial Ornada

### Metadatos de preservación

- 8- Preservación Digital
  - 8.1- Software y versión de captura
  - 8.2- Hardware de Captura
  - 8.3- Resolución
  - 8.4- Profundidad de Bit
  - 8.5- Formato
  - 8.6- Fecha de digitalización: dd/mm/aaaa
  - 8.7- Cantidad de Imágenes que conforman la versión digital

#### Software de gestión de documentos

La selección del software para la gestión de los documentos digitalizados ha sido muy compleja. En una primera instancia los informáticos se inclinaron por ICA - AtoM, un software de código abierto para la descripción archivística, pero cuando comenzamos a ponerlo a prueba no respondía a nuestras necesidades, ya que los libros digitalizados en algunos casos (los más grandes) contaban con hasta 1300 imágenes y un peso total de casi 65 Gb, para ello debíamos cargar imagen por imagen o en lotes de 50 imágenes y luego seleccionar de un listado cada imagen para cargar los metadatos, esto hacía que el trabajo fuese muy engorroso y por ende demasiado lento.

Como el tiempo para la selección y adaptación de un nuevo software para la gestión volvía a foja cero, lo que implicaba varios meses de demora, se optó por generar una planilla de cálculo, para poder comenzar a cargar los metadatos. Esta planilla, como se ha mencionado con anterioridad se realizó en la plataforma de Google Drive (antes Google Doc), haciéndole periódicamente el backup correspondiente en un disco duro de una computadora de la biblioteca.

Mientras tanto los informáticos seguían trabajando en un nuevo software de gestión de documentos: Alfresco. Se realizaron pruebas, se migraron metadatos, pero el sistema resultó ser muy engorroso, y solo se utilizó para subir las imagenes al servidor.

#### 2a. parte: Plataforma Biblioteca Mayor Digital

Siete años después, en el año 2015 se comienza a desarrollar una plataforma propia para la carga y visualización de los libros digitalizados.

El trabajo en las bibliotecas digitales en este momento había evolucionado, y debíamos adaptar la estructura de metadatos propia pensada con anterioridad debido en gran parte, a la estandarización necesaria para poder trabajar cooperativamente con otras instituciones. Poder compartir archivos y su descripción es fundamental, y para eso debíamos hablar el mismo lenguaje.

Razón por la que decidimos reservar los metadatos propios estructurales y de preservación, pero migrar al Dublin Core Calificado los metadatos descriptivos.

Las razones por las cuales se seleccionó este esquema fueron:

- Reconocimiento internacional y amplia utilización.
- Utilizados por la plataforma digital Trapalanda de la Biblioteca Nacional Argentina

- Simplicidad: fáciles de entender para personas que no sean necesariamente catalogadores.
- Se aplican a todo tipo de materiales

Estos metadatos se tomaron, en su mayoría, de los registros bibliográficos MARC del catálogo correspondientes al material digitalizado de la biblioteca, se exportaron esos registros en XML para importarlos posteriormente a la plataforma en los campos Dublin Core correspondientes.

```
Registro Marc importado en
                                                                                                                                                                                                                                                                                                                    formato XML
                    <leader>02095cam a2200385 4500</leader>
                  <controlfield tag="001">0000002</controlfield>
<controlfield tag="003">AR-CdUBM</controlfield>
        datafield tag="080" ind1=" " ind2=
                  <subfield code="a">929:1</subfield>
</datafield>

ind1="1" ind2=
                      <subfield code="a">\text{subfield code="a">\text{subfield code="a">\text{subfield code="d">\text{subfield}\\
<subfield code="d">\text{subfield}\\
<subfield code="9">\text{24973</subfield}\\
<subfield code="9">\text{subfield}\\
<subfield code="9"</subfield>\text{subfield}\\
</subfield code="9"</subfield code="9"</subfield>\text{subfield}\\
</subfield code="9"</subfield code="9
                  </datafield>
         <subfield code="c">Compuesta por ... Manuel Abad Illana ...</subfield</pre>

<
                      <subfield code="a">Salamanca :</subfield>
<subfield code="b">Eugenio Garcia de Honorato y S.Miguel ...,</subfield>
<subfield code="c">1755.</subfield>
                   </datafield>
         datafield tag="500" ind1=" " ind2=" ">
                       <subfield code="a">Tomo primero, primer y segundo libros de los varones ilustres</subfield>
         <datafield tag="500" ind1=" "</pre>
                        <subfield code="a">1 v. de 31 cm.</subfield>
                  </datafield>
        <datafield tag="500" ind1=" " ind2=" ";</pre>
                  <subfield code="a">Sig:: []2,[calderon]2, 3[calderon]-5[calderon]2, 6[calderon]5, [parágrafo]2, *3, A-502, 5P3, *-4*2, 5*1</subfield>
</datafield>
         datafield tag="500" ind1=" " ind2=" ";
                        <subfield code="a">Texto a dos columnas</subfield>
                  </datafield>
```

Para tal fin se realizó una correspondencia de datos entre las etiquetas Dublin Core y campos MARC. A continuación una tabla de los campos Dublin Core utilizados y fuente de dónde se toman los datos:

Etiqueta dc.title (Título)	Registro MARC, campo ^245 subcampos a, b, c
Etiqueta dc.creator (Autor o Creador)	Registro MARC, campo ^1xx subcampo a, d
Etiqueta dc.subject (Materia o palabras clave)	Registro MARC, campos ^6xx subcampo a
Etiqueta dc.publisher (editor)	Registro MARC, ^260 subcampo b

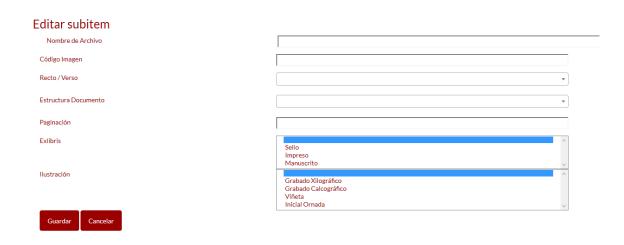
Etiqueta dc.publisher.place (lugar del editor)	Registro MARC, ^260 subcampo a
Etiqueta dc.contributor (Colaborador)	Registro MARC, ^7xx subcampo a
Etiqueta <b>dc.date.available</b> (Fecha en la que el recurso se pone a disponibilidad) (aaaa-mm-dd)	Se coloca aquí la fecha de subida del documento a la plataforma de forma manual
Etiqueta dc.date.created (creación)	Fecha de digitalización: dd/mm/aaaa
Etiqueta dc.date.issued (de edición)	Registro MARC, ^260 subcampo c
Etiqueta <b>dc.type</b> (Tipo de recurso lista predefinida	Se generó una lista desplegable con los tipos de formato para seleccionar el correspondiente.
Etiqueta <b>dc.identifier</b> (Identificador unívoco, puede ser internacional ISBN o local administrativo)	Se coloca en este campo la url generada por la Plataforma BMDigital del objeto.
Etiqueta <b>dc.source</b> (Fuente número o cadena de símbolos para identificar unívocamente el documento del que deriva el recurso)	URL del registro del libro en el catálogo
Etiqueta dc.language (Lengua del recurso)	Se toma de campo 008, posición 35-37, código de idioma

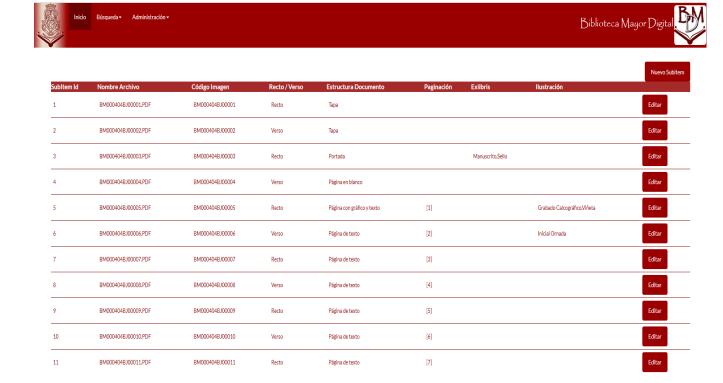
Administración + Biblioteca Mayor Digital
Historia del gran padre y patriarcha San Norberto, fundador del Orden Candido Premonstratense / Compuesta por Manuel Abad Illana
Abad Illana, Manuel, 1716-1780
BIOGRAFIAS
v1
Eugenio Garcia de Honorato y S.Miguel
Salamanca
19/09/2014
1755
application/pdf

La carga de metadatos Dublin Core dentro de la Plataforma se realiza utilizando la siguiente pantalla de carga:



Los metadatos estructurales del libro antiguo anteriormente cargados en planillas de cálculo se importaron también a la nueva plataforma. En las siguientes imágenes se muestra cómo se realiza la carga de estos metadatos dentro de la plataforma y cómo quedan registrados:

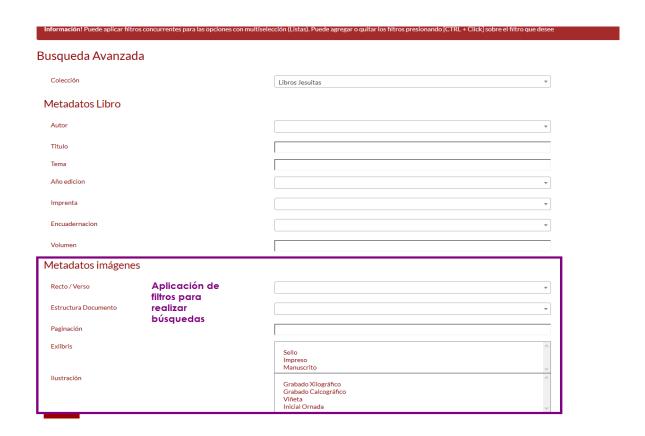




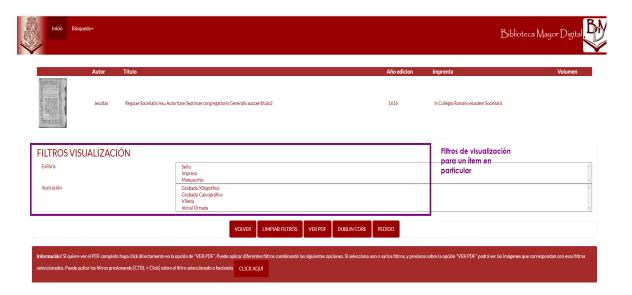
La carga de estos metadatos estructurales nos permite buscar y navegar dentro del documento.

En la pantalla de búsqueda se pueden aplicar estos filtros y restringir la búsqueda sólo a aquellos materiales que poseen exlibris, o un tipo específico de éstos como manuscrito,

impreso o sellos. Como así también aquellos que poseen ilustraciones o un tipo específico de éstas, como xilográficas o calcográficas.



También se pueden realizar búsquedas dentro de cada ítem en base a estos atributos: ilustraciones, exlibris, etc.



#### Controles de calidad

La Biblioteca Mayor certifica todos sus procesos bajo las normas de calidad ISO 9001:2008, por este motivo se controlaron todos los metadatos tanto descriptivos como estructurales y de preservación de cada ítem digitalizado. El área de Procesos Técnicos realizó el control y corrección de los metadatos descriptivos de los ítems importados y realizó la vinculación de los registros del OPAC de la Biblioteca a los ítems digitalizados en la Plataforma *BMDigital*. El área de digitalización realizó en control de los metadatos estructurales y de preservación cargados.

# Hacia adelante ... más allá de libros antiguos.

Hasta el momento sólo se definieron metadatos estructurales para libro antiguo en el marco de la Biblioteca Mayor Digital, sin embargo, la plataforma fue pensada para albergar todo tipo de materiales como vídeo, audio y prensa histórica. El desafío a futuro consiste en continuar con el estudio para la aplicación de metadatos estructurales para estos tipos de materiales con el fin de potenciar la recuperación de información del material digitalizado. Como así también establecer convenios de cooperación con otras instituciones con colecciones patrimoniales digitalizadas.

# Bibliografía.

- Eito Brun, Ricardo. (2008). Lenguajes de marcas para la gestión de recursos digitales : aproximación técnica, especificaciones y referencia. Guijón, Asturias : Ediciones Trea.
- García Camarero, Ernesto, García Melero, Luis Ángel. (2001). *La biblioteca digital*. Madrid: Arco/Libros.
- Herrera Morillas, José Luis. (2003). *Tratamiento y difusión digital del libro antiguo : directrices metodológicas y guía de recursos*. Gijón : Ediciones Trea.
- Morales del Castillo, José Manuel. (2011). *Hacia la biblioteca digital semántica*. Guijón, Asutiras : Ediciones Trea.
- Testa, Patricia, Ceriotto, Paula. (2009). [en línea]. *Descripción de objetos digitales : metadatos*. II Encuentro Nacional de Catalogadores, Biblioteca Nacional Argentina. Revisado el 07 de octubre de 2016 de : http://www.bn.gov.ar/descargas/catalogadores/ponencias/251109 09a.pdf
- *Metadata innovation : Dublin Core Metadata Initiative*. (2017), [en línea]. Revisado el 22 de octubre de 2017 de: <a href="http://dublincore.org/">http://dublincore.org/</a>