

# **REPOSITARIOS DIGITALES INSTITUCIONALES**

**Diseño, implementación y optimización de un  
recurso estratégico para las Universidades**

## **Bloque 3.2: Descripción de objetos digitales**

***Capacitadoras:***

***Patricia Testa***

***Paula Ceriotto***

Sistema Integrado de Documentación (SID)

Universidad Nacional de Cuyo

Apuntes del curso dictado el 8 de mayo de 2013

Curso organizado en el marco del Proyecto PICTO 2010-149 *“Investigación y desarrollo en repositorios institucionales. Aplicaciones y experiencias en universidades de la región bonaerense”* cofinanciado por la Agencia Nacional de Promoción de la Ciencia y la Tecnología y el Consejo Interuniversitario Nacional.

Universidades que integran el proyecto:

Universidad Nacional de La Plata (UNLP),  
Universidad Nacional de Mar del Plata (UNMDP),  
Universidad Nacional de Luján (UNLu),  
Universidad Nacional del Centro de la Provincia de Buenos Aires (UNICEN),  
Universidad Nacional del Sur (UNS),  
Universidad Tecnológica Nacional, Facultad Regional Bahía Blanca (UTN-FRBB)

Investigadora responsable: Dra. Sandra Miguel



**Atribución – No Comercial – Compartir Igual** (*by-nc-sa*): No se permite un uso comercial de la obra original ni de las posibles obras derivadas, la distribución de las cuales se debe hacer con una licencia igual a la que regula la obra original.

## ***Repositorios digitales***

### **Definición**

Archivo digital accesible vía Internet, que reúne la producción intelectual de una disciplina o de una institución (SPARC).

Conjunto de servicios prestados por una institución o comunidad determinada para poner en valor su propia producción documental por medio de las TIC (Rodríguez Palchevich).

Sistema complejo de gestión, acceso y preservación de objetos digitales que responde a la producción científica de una comunidad (Necobelac).

Un repositorio digital no es simplemente un conjunto de enlaces que se dirigen hacia objetos digitales más o menos ordenados. Para asegurar el acceso universal a esa información, estos deben ser gestionados con un software, que permita su ubicación coherente dentro de la Web y que facilite su recuperación a través de una descripción estandarizada por medio de metadatos asociados a cada objeto digital.

### **Objeto digital**

Los objetos digitales son lo que realmente se almacena y gestiona en un repositorio digital.

La definición comúnmente utilizada en el campo de las bibliotecas digitales para un objeto digital, es que es un archivo electrónico que posee, además del contenido, un identificador único y metadatos que lo caracterizan unívocamente.

Según el Diccionario de Datos PREMIS un objeto digital “es una unidad de información específica en formato digital”.

Los repositorios institucionales contienen información nacida en soporte físico (textos, imágenes fijas, videos, grabaciones sonoras) que ha sido digitalizada y también documentos nacidos en formato electrónico. Todos ellos poseen datos que pueden ser interpretados y un identificador único formado por un nombre y una especificación que hace referencia a la norma ISO 2108:2005. Se puede hacer un paralelismo entre un documento en soporte papel, que es identificado con un número único e irrepetible, el ISBN, otorgado por una agencia internacional y el DOI (Digital Object Identifier), otorgado por la Internacional DOI Foundation.

Para que estos objetos no se mezclen con los recursos disponibles en la Web y puedan ser recuperados con eficiencia y rapidez, es indispensable que estén organizados. Esa organización, implica una selección previa, la formación de una colección y su descripción mediante la asignación de etiquetas o metadatos a cada uno de ellos.

## **Metadatos**

La palabra metadato proviene del prefijo griego “meta”, que significa: “entre”, “con”, “después de” o “cambio”, que generalmente se le da el significado de “después de” o “más allá” “al lado”, y del latín “datum” que significa “dato”. En un latín más actual, así también como en inglés, el término “meta” indica algo “trascendental” o “sobrenatural”.

La primera acepción que se le dio (y actualmente la más extendida) fue la de “dato sobre datos”, ya que proporcionaban la información mínima necesaria para identificar un recurso.

La National Information Standards Organization (NISO) dice que “referirse en estos momentos a metadatos es hacerlo con el fin de significar las estructuras de los archivos, conjuntos de datos u otra entidad de información que aseguran la descripción técnica que se necesita para representar las partes del objeto digital”.

Desde el punto de vista de las Ciencias de la Documentación y la Información, Ana Pavani (2009) dice que los metadatos es un conjunto de atributos de catalogación de los documentos, que permiten su identificación sin tener que ejecutarlos.

Pero, ¿por qué surgen los metadatos? El aumento incontrolable de los recursos en la web, generó grandes inconvenientes para la localización y recuperación de los mismos. Por eso es que se debió pensar en alguna forma de etiquetarlos y describirlos. El mecanismo que se creó fueron los metadatos, que no son otra cosa que etiquetas XML, que contienen información sobre un objeto digital y que están asociadas a él. Son datos estructurados sobre la información que hay en la web que pueden ser comprendidos por la computadora por medio de los motores de búsqueda para poder recuperarla. La función de los metadatos es: identificar, describir, clasificar y localizar el contenido de un documento web.

Antes de seguir hablando de metadatos, es importante conocer qué es XML. Víctor García González, en su página Web ***Metadatos y documentos XML/RDF para recuperación***, lo explica muy claramente: “XML surgió como un lenguaje de marcado para sustituir a HTML. Ambos lenguajes son herederos de SGML, el lenguaje de marcas estándar para la descripción formal y de contenido de los documentos, no solamente para la presentación de dichos documentos. HTML era, principalmente, un lenguaje de presentación que definía un conjunto de etiquetas y atributos válidos y que ofrecía un significado visual para cada elemento del lenguaje, por el contrario, XML no define las etiquetas ni cómo se utilizan, sino que ofrece un escaso número de reglas sintácticas para poder crear documentos. Así pues, XML no es un lenguaje, sino un metalenguaje o lenguaje para definir otros lenguajes. XML no sustituye a HTML puesto que sirven para cosas distintas: HTML para presentar la información, y XML para representar e intercambiar datos de forma independiente a su presentación. XML permite

gestionar los datos aunque procedan de diversas fuentes y también permite el intercambio de documentos entre distintas aplicaciones ya estén en un único ordenador o en una red. Como la información en XML está mucho más estructurada, esto facilitará el trabajo a los buscadores y los robots, que accederán a los datos de manera más precisa”.

Ejemplo de XML:

```
<?xml version="1.0" encoding='ISO-8859-1'?>
```

```
<metadata>
```

```
<dc:title>Los catálogos bibliográficos y el multilingüismo: parte I</dc:title>
```

```
<dc:creator>Martín Goikoetxea, Alasne</dc:creator>
```

```
<dc:description>Experiencias nacionales e internacionales llevadas a cabo en la creación de los catálogos bibliográficos multilingües</dc:description>
```

```
<dc:publisher>baratz</dc:publisher>
```

```
<dc:date>2001-09-01</dc:date>
```

```
<dc:identifier>htt://www.absysnet.com/tema/tema6.html</dc:identifier>
```

```
</metadata>
```

Los metadatos pueden incluir información descriptiva sobre el **contexto, calidad y condición o características del dato**<sup>1</sup>. Su finalidad es la de facilitar su recuperación, autenticación, evaluación, preservación o interoperabilidad.

El contexto, está relacionado con quién, qué, por qué, dónde, cuándo y cómo fue creado el objeto digital.

En síntesis, son descripciones estructuradas representadas en etiquetas XML asociadas a un recurso de información, cuya función es hacer útiles los datos para que puedan ser visibles y recuperables.

La gran ventaja de los metadatos, es que permiten normalizar la información contenida dentro de un documento textual, y además recuperar documentos con cualquier otra morfología: audio, video e imágenes.

Los bibliotecarios, desde siempre, han organizado los documentos a través de la descripción bibliográfica y el análisis de contenido, generando registros bibliográficos. Si se toma la definición más amplia de “datos sobre los datos”, estos registros bibliográficos también pueden considerarse metadatos, ya que contienen información sobre documentos. Existe una corriente generalista que utiliza el término metadatos para referirse a la descripción tanto de objetos físicos, como digitales.

Otro grupo, más restrictivo, emplea el término sólo para los recursos Web.

---

<sup>1</sup> Senso, José A. ; De la Rosa Piñero, A. (2003). *El concepto de metadato. Algo más que descripción de recursos electrónicos*. Recuperado de <http://www.scielo.br/pdf/ci/v32n2/17038.pdf>

En el entorno digital, asignar metadatos, consiste en:

- Describir recursos
- Asignar puntos de acceso
- Analizar el contenido
- Localizar recursos

La diferencia importante entre el viejo proceso de catalogación y la asignación de metadatos, es que con los objetos físicos, la descripción se hace manualmente y con posterioridad a la creación del mismo, mientras que la descripción del objeto digital, una parte es realizada automáticamente por un software en el momento de la creación del objeto.

Una biblioteca digital sin metadatos descriptivos, es como una biblioteca tradicional sin catálogo.

La Guía de uso del Dublin Core dice que la relación entre un registro de metadatos y el recurso al que se describe puede darse de dos formas:

- 1- Los elementos pueden estar en un registro separado del documento, como por ejemplo el caso de los registros de un catálogo de biblioteca.
- 2- Los metadatos pueden estar incluidos o incrustados en el propio recurso.

## **Clasificación de los metadatos**

Los metadatos pueden describir tanto colecciones de objetos como sus partes componentes y toda clase de recursos digitales: texto, audio, video e imágenes. También incluyen datos sobre los procesos en los que están involucrados y definen las relaciones entre los objetos.

Según la complejidad de los metadatos, pueden clasificarse de distintas formas:

- **Metadatos desestructurados:** son aquellos que extraen automáticamente los datos de los recursos electrónicos, para lo cual se emplean, regularmente, robots e índices que localizan la información a partir del texto HTML

- **Metadatos semiestructurados:** tienen una estructura semántica sencilla, como Dublin Core

- **Metadatos estructurados:** son esquemas muy ricos y complejos. Comúnmente utilizan algunos datos creados automáticamente y otros en forma manual. Por ejemplo MODS.

Según su **función**, existen diversas clasificaciones, pero la más generalizada es la de dividirlos en 3 amplias categorías: **administrativos, estructurales y descriptivos.**

- **Metadatos administrativos:** ayudan a gestionar y rastrear los recursos. Se refieren a sus características y propiedades, facilitando la gestión y procesamiento tecnológico y físico de las colecciones digitales, tanto a corto como a largo plazo.

Pueden subdividirse en:

- **Metadatos de gestión de derechos:** ayudan a administrar las limitaciones de acceso a los recursos y registrar toda la información relacionada con la propiedad intelectual y los derechos (fecha del registro de derecho de autor, limitaciones en cuanto al copiado y distribución, información sobre licencia...)
- **Metadatos de preservación:** definen las acciones que garantizan que los objetos digitales permanezcan viables a largo plazo (que los soportes se puedan leer –migrar-) y recuperables (que puedan visualizarse, ejecutarse o representarse de alguna manera por una aplicación de software –emular-). Aseguran, además, que los recursos digitales no se han alterado inadvertidamente y que se han documentado los cambios legítimos de los objetos (seguridad en el almacenamiento).
- **Metadatos técnicos:** informan sobre el funcionamiento del sistema (requisitos de hardware y software, resolución, profundidad de bit, espacio de color, formato de archivo, compresión, fuente de luz, propietario, fecha del registro...).

- **Metadatos estructurales:** indican cómo reunir objetos digitales complejos para que se puedan visualizar o utilizar de algún modo. Proporcionan información sobre la estructura interna de los recursos electrónicos, como página, sección, capítulo, partes, índices, tabla de contenido, etc. y describen la relación entre los materiales. Facilitan la navegación y presentación de los recursos y relacionan las diferentes partes que lo componen.

- **Metadatos descriptivos:** ayudan a identificar y recuperar los recursos. Se utilizan para la descripción e identificación de la información contenida en el recurso. Contienen atributos físicos (medios, dimensiones) y atributos bibliográficos (título, autor/ creador, idioma, palabras claves).

En la siguiente figura, se observan las distintas clases de metadatos:



No. de OD: 779

**Título: Huellas : búsquedas en artes y diseño / Nº 5 (2006)**  
**Editor: Universidad Nacional de Cuyo. Facultad de Artes y Diseño**

#### Contenido

**Hacia una nueva semiótica**, Guembe, Gabriela, p. 9 –16  
**Obras para piano y orquesta en la Argentina**, De Marinis, Dora, p. 17-40  
**Cuicani**, Plana, Beatriz, p. 41- 52  
**Renovación de la música en Argentina**, Loyola, María Enriqueta, p. 53-66  
**Forma móvil: Intermission 6, de Morton Feldman**, Solare, Juan María, p. 67-74  
**El teatro mendocino en el siglo XIX**, Navarrete, José Francisco, p. 75-84  
**Homenaje al Prof. Carlos Agustín Gómez**, Coll, Roxana, p. 85-92

Formato

PDF

83 Kb

124 Kb

579 Kb

3,4 Mb

501 Kb

92,9 Kb

1,3 Mb

◆ **Metadatos administrativos**    ◆ **Metadatos estructurales**    ◆ **Metadatos descriptivos**

Todos estos tipos de metadatos conviven con el objeto digital y existen diferentes modelos para cada uno. Los administrativos y estructurales en su mayoría son gestionados automáticamente por el software cuando se crea el registro. Los metadatos descriptivos son ingresados por el catalogador.

### ***Esquemas de metadatos***

Según el uso o el objetivo de los metadatos, se agrupan en esquemas que son un conjunto de metadatos para una aplicación específica.

Un esquema de metadatos es un plan lógico que muestra las relaciones entre los elementos de metadatos (campos), incorporando normalmente un conjunto de reglas, incluidas aquellas relacionadas con la semántica y la sintaxis, que permiten la gestión de metadatos. (Norma UNE-ISO 23081).

- o **Reglas de contenido:** indican qué datos se incluyen en cada elemento, y en qué forma. Por ejemplo AACR2r, ISBD, RDA, ISAD, APA, códigos ISO para



lenguas, escrituras, países y fechas, listas normalizadas para tipos de recurso, audiencia, formato, CDU, CDD, tesauros, listas de encabezamientos, etc.

- o **Semántica:** indica las características del elemento, por ejemplo, si es obligatorio, si puede ser repetible, etc.
- o **Sintaxis:** establece la manera en que los elementos y su contenido deben ser codificados, para ofrecer un formato de intercambio. Por ejemplo, XML

El esquema define qué elementos de metadatos puede contener un documento, cómo están organizados, y qué atributos y de qué tipo son.

Básicamente, todos los esquemas constan de una serie de elementos y atributos normalizados y el valor de los mismos es el que sirve para recuperar la información. A modo de ejemplo, podemos decir que el esquema nos da el nombre de un elemento, qué dato debe contener y la forma adecuada de consignarlo. Por ejemplo, el elemento que el esquema denomina <creator> es repetible y obligatorio (semántica), debe tener el nombre del autor del documento colocado de una manera específica (sintaxis). El dato colocado es la información recuperable.

**<creator>**Borges, Jorge Luis**<creator>**

En el año 2001 se realizó una convención en Ottawa, donde se establecieron los principios comunes y que pueden servir de base para el diseño de cualquier esquema de metadatos o aplicación:

- **Modularidad:** la arquitectura de los metadatos debe ser flexible de manera tal que se puedan combinar de manera interoperable, sintáctica (por ej. XML) y semánticamente, elementos provenientes de esquemas preestablecidos diferentes. Cada componente del esquema debe incluir la funcionalidad y los requisitos específicos para una determinada aplicación.

- **Extensibilidad:** los esquemas de metadatos deben ser lo suficientemente flexibles como para poder ajustarse a las necesidades específicas del repositorio, sin perder de vista la interoperabilidad proporcionada por el esquema base. Así pues, el elemento *creador* va estar presente en todos los esquemas, como así también el *identificador*, mientras que otros serán específicos para una aplicación particular.

- **Refinamiento:** el esquema de metadatos debe permitir un nivel de detalle más exhaustivo mediante, por ejemplo, la adición de calificadores a los elementos. Esto hará más específico su significado. Por ejemplo en el elemento *colaborador*, el calificador que aclare su función: ilustrador, prologuista, editor, compilador, director, escenógrafo, etc. También el uso de vocabularios controlados ayuda al refinamiento dado que mejora la precisión de las

descripciones. Este principio es muy importante para lograr una recuperación más profunda y precisa de la información contenida en el repositorio.

- **Plurilingüismo:** los esquemas de metadatos deben respetar la diversidad lingüística y cultural y la internacionalización de la Web, poniendo a disposición del usuario los recursos en su idioma nativo.

Si bien existen muchos esquemas de metadatos descriptivos, cada institución propone una aplicación diferente según su objetivo. Las propuestas son múltiples y van desde lo más simple hasta lo más complejo. Al momento de interactuar resulta difícil compatibilizar los diferentes modelos adoptados. Suele pasar que los elementos no se llamen igual, no tengan un significado unívoco y que los valores sean diferentes.

Para garantizar la interoperabilidad de los repositorios y asegurar los mejores resultados de búsqueda, es imprescindible que todos adopten un mismo criterio semántico a nivel de metadatos.

## **Clasificación de los esquemas de metadatos**

Existen diferentes formas de clasificar la gran variedad de esquemas de metadatos existentes. Jenn Riley analiza más de 100 estándares de metadatos utilizados en el sector del patrimonio cultural y los clasifica mediante el gráfico “Seeing Standards: A Visualization of the Metadata Universe”, según su propósito, su función, su dominio y la comunidad por la cual puede ser usado.

- **Función:** se refiere al papel que tiene el estándar en la creación y almacenamiento de los metadatos. Los divide en: modelo conceptual (ej. FRBR, FRAD, RDF), estándar de contenido (ej. AACR2r, ISBD, RDA, CCO), vocabularios controlados (listas de encabezamientos y tesauros), tecnologías marco (ej. XML, OAI-PMH, RDF, Z39.50), lenguajes de marcado (XML, SGML, TEI, EAD), formatos de registro (ej. MARC, MARCXML) y estándares de estructura (ej. DC, MODS). Nos referiremos solamente a los estándares más utilizados en los repositorios digitales del ámbito de las bibliotecas y archivos.

- **Propósito:** se refiere al tipo de registro para el cual ha sido diseñado el estándar de metadatos. En general, un estándar va a estar más enfocado hacia un fin específico, pero puede incluir elementos de datos para otros fines. Estos propósitos pueden ser: descriptivos, de preservación, contenedores, de gestión de derechos, estructurales y/o técnicos.

- **Dominio:** se refiere a los tipos de materiales para los cuales está destinada la norma o podría potencialmente ser útil. No es una lista exhaustiva, y cada estándar puede utilizarse para varios materiales a la vez. Se incluyen: objetos culturales, conjunto o set de datos, datos geoespaciales, imágenes en movimiento, materiales musicales, textos y recursos visuales.

- **Comunidad:** se refiere al ámbito en que estos estándares se utilizan o potencialmente podrían ser utilizados. Se incluyen: bibliotecas (académicas, públicas, escolares, especializadas y/o de instituciones), archivos, museos y la industria de la información.<sup>2</sup>

ESTANDAR	PROPOSITO	DOMINIO	FUNCION	COMUNIDAD
<b>DC</b> Dublin Core Metadata Initiative	Metadatos: <ul style="list-style-type: none"> <li>• <b>descriptivos</b></li> <li>• técnicos</li> <li>• estructurales</li> <li>• de gestión de derechos</li> </ul>	Imágenes fijas Textos Materiales musicales Imágenes en movimiento Datos geoespaciales Conjunto de datos Objetos culturales	<b>Estándar de            estructura</b> Formato de registro	Archivos <b>Bibliotecas</b> Museos <b>Industria de            la            Información</b>
<b>MODS</b> Metadata Object Description Schema	Metadatos: <ul style="list-style-type: none"> <li>• <b>descriptivos</b></li> <li>• técnicos</li> <li>• estructurales</li> <li>• de gestión de derechos</li> </ul>	Imágenes fijas Textos Materiales musicales Imágenes en movimiento Datos geoespaciales Conjunto de datos Objetos culturales	<b>Estándar de            estructura</b> Formato de registro	Archivos <b>Bibliotecas</b> Museos

<sup>2</sup> Todos los segmentos o actividades relacionadas con la producción de recursos de información -procesamiento y reprocesamiento, organización, almacenamiento y recuperación de la información que se operan con fines lucrativos o no. En: Rebolledo Saavedra, Gustavo (comp.). La era de la información y la sociedad de la información, Recuperado de: [http://b3.bibliotecologia.cl/ar-era\\_de\\_la\\_informacion.htm](http://b3.bibliotecologia.cl/ar-era_de_la_informacion.htm)

<p><b>FGDC/CSDGM</b> Content Standard for Digital Geospatial Metadata ISO 19115</p>	<p>Metadatos:</p> <ul style="list-style-type: none"> <li>• <b>descriptivos</b></li> <li>• técnicos</li> <li>• estructurales</li> <li>• de gestión de derechos</li> </ul>	<p><b>Datos geoespaciales</b></p>	<p><b>Estándar de estructura</b>  Lenguaje de marcado</p>	<p>Archivos <b>Bibliotecas</b>  Museos Industria de la Información</p>
<p><b>IEEE/LOM</b> Learning Object Metadata</p>	<p>Metadatos:</p> <ul style="list-style-type: none"> <li>• <b>descriptivos</b></li> <li>• técnicos</li> <li>• estructurales</li> <li>• de gestión de derechos</li> </ul>	<p><b>Imágenes fijas</b>  Textos Materiales musicales Imágenes en movimiento Datos geoespaciales Conjunto de datos Objetos culturales</p>	<p><b>Estándar de estructura</b>  <b>Formato de registro</b>  Modelo conceptual</p>	<p>Archivos Bibliotecas Museos <b>Industria de la Información</b></p>
<p><b>MIX</b> Metadata for Images in XML</p>	<p>Metadatos:</p> <ul style="list-style-type: none"> <li>• <b>de preservación</b></li> <li>• técnicos</li> </ul>	<p><b>Imágenes fijas</b>  Textos Materiales musicales <b>Objetos culturales</b></p>	<p><b>Estándar de estructura</b>  Formato de registro</p>	<p><b>Archivos Bibliotecas</b>  Museos Industria de la Información</p>
<p><b>PREMIS</b> PREservation Metadata: Implementation Strategies</p>	<p>Metadatos:</p> <ul style="list-style-type: none"> <li>• <b>de preservación</b></li> <li>• técnicos</li> </ul>	<p>Imágenes fijas Textos Materiales musicales Imágenes en movimiento Datos</p>	<p><b>Estándar de estructura</b>  Formato de registro</p>	<p><b>Archivos Bibliotecas</b>  Museos Industria de la Información</p>

		geoespaciales Conjunto de datos Objetos culturales		
<b>METS</b> Metadata Encoding and Transmission Standard	Metadatos: <ul style="list-style-type: none"> <li>• <b>contenedores</b></li> <li>• estructurales</li> </ul>	Imágenes fijas Textos Materiales musicales Imágenes en movimiento Datos geoespaciales Conjunto de datos Objetos culturales	Estándar de estructura Formato de registro	Archivos <b>Bibliotecas</b> Museos Industria de la Información
<b>EAD</b> Encoded Archival Description	Metadatos: <ul style="list-style-type: none"> <li>• <b>descriptivos</b></li> <li>• estructurales</li> <li>• de gestión de derechos</li> </ul>	Imágenes fijas Textos Materiales musicales Imágenes en movimiento Objetos culturales	<b>Lenguaje de marcado</b> Formato de registro Estándar de estructura	<b>Archivos</b> Bibliotecas Museos
<b>TEI</b> Text Encoding Initiative	Metadatos: <ul style="list-style-type: none"> <li>• <b>descriptivos</b></li> <li>• <b>técnicos</b></li> <li>• <b>estructurales</b></li> <li>• de gestión de derechos</li> </ul>	<b>Textos</b>	<b>Lenguaje de marcado</b> Formato de registro	Archivos <b>Bibliotecas</b> Museos Industria de la Información
<b>RDF</b> Resource Description Framework	Metadatos: <ul style="list-style-type: none"> <li>• <b>descriptivos</b></li> <li>• técnicos</li> <li>• estructurales</li> <li>• de gestión de derechos</li> </ul>	Imágenes fijas Textos Materiales musicales Imágenes en movimiento	<b>Modelo conceptual</b> <b>Tecnología marco</b>	Archivos <b>Bibliotecas</b> Museos <b>Industria de la Información</b>

		Datos geoespaciales Conjunto de datos Objetos culturales		
--	--	---	--	--

## ***Breve descripción de esquemas de metadatos***

### **DC: Dublin Core Metadata Initiative**

( <http://dublincore.org/>)

En 1995 una reunión, realizada en Dublin (Ohio, Estados Unidos), de bibliotecarios, investigadores en bibliotecas digitales, expertos en computación y lenguaje de marcado de texto electrónico, pactaron un grupo de 15 elementos que fueron concebidos originalmente como un conjunto de etiquetas XML, que debían ser generadas por el autor del documento, con la finalidad de facilitar su identificación y posterior recuperación en Internet. Es mantenido y actualizado por The Dublin Core Metadata Iniciativa (DCMI).

Los elementos se pueden dividir en tres grupos de datos relacionados con el contenido, la propiedad intelectual y el uso, instalación y manipulación.

<b>Contenido</b>	<b>Propiedad intelectual</b>	<b>Uso, instalación y manipulación</b>
<b>Título</b> <b>&lt;Title&gt;</b> Normalmente el título será el nombre por el que se conoce formalmente el recurso.	<b>Creador</b> <b>&lt;Creator&gt;</b> Autor personal o entidad responsable del contenido del recurso.	<b>Fecha</b> <b>&lt;Date&gt;</b> Fecha clave del recurso.
<b>Materia</b> <b>&lt;Subject&gt;</b> Descriptores o palabras clave que definen el tema del que trata el recurso.	<b>Editor</b> <b>&lt;Publisher&gt;</b> Editor del recurso.	<b>Tipo</b> <b>&lt;Type&gt;</b> Naturaleza del contenido (artículo, tesis, ponencia, etc.).

<b>Descripción</b> <b>&lt;Description&gt;</b> Resumen del recurso.	<b>Colaborador</b> <b>&lt;Contributor&gt;</b> Personas o entidades colaboradoras en el contenido del recurso.	<b>Formato</b> <b>&lt;Format&gt;</b> Formato del recurso (tamaño, duración, software, hardware necesario para su reproducción, etc.).
<b>Fuente</b> <b>&lt;Source&gt;</b> Recurso relacionado del cual se deriva el recurso que se describe.	<b>Derechos</b> <b>&lt;Rights&gt;</b> Derechos de autor que afectan al recurso (copyright, IPR, etc.).	<b>Identificador</b> <b>&lt;Identifier&gt;</b> Referencia unívoca de identificador del recurso (ISBN, URL, DOI, etc.).
<b>Idioma</b> <b>&lt;Language&gt;</b> Idioma en el que está expresado el recurso.		
<b>Relación</b> <b>&lt;Relation&gt;</b> Otro recurso relacionado con el que se describe.		
<b>Cobertura</b> <b>&lt;Coverage&gt;</b> Periodo de tiempo o lugar geográfico sobre los que trata el recurso.		

Todos los elementos son opcionales, repetibles, no poseen subcampos y no requieren control de los valores colocados en ellos.

Se usan para describir documentos en cualquier ámbito o materia. Es concebido como una forma muy simple de generar metadatos y de permitir una catalogación mucho más sencilla, ya que se aparta de las normas bibliotecológicas.

El esquema, además, permite usar **calificadores** opcionales para cada elemento que aumentan la especificidad y precisión de los metadatos y/o indican la normativa empleada. Se debe utilizar solamente el conjunto de calificadores aprobados por Dublin Core, desarrollados por esa comunidad, de manera tal que se mantenga la compatibilidad con otros usuarios del estándar.

Por ejemplo:

**Calificador que refina el elemento <Title> (Título)**

- **Alternativa**

Nombre: Alternative

Etiqueta: Alternativa

Definición: Cualquier alternativa al título usada para sustituir al título formal del recurso (puede incluir títulos abreviados, variantes y/o traducciones)

### Calificador que indica esquema de codificación del elemento <Subject> (Materia)

- **LCSH**

Nombre: LCSH

Etiqueta: LCSH

Definición: Library of Congress Subject Headings (Encabezamientos de materia de la Library of Congress)

### MODS: Metadata Object Description Schema

(<http://www.loc.gov/standards/mods/>)

La Biblioteca del Congreso de los Estados Unidos, en el año 2002, crea un esquema de metadatos orientado a bibliotecas digitales, que permite tanto la descripción original de objetos digitales, como tomar descripciones ya existentes en MARC21. Dentro de ese ámbito bibliotecario, fueron seleccionados 20 elementos principales, subelementos y atributos que pueden ser usados en uno o más elementos.

Tiene una complejidad intermedia. Está pensado para ser utilizado por especialistas del ámbito de las Ciencias de la Información.

Elemento	Subelementos	Atributos
Título <TitleInfo>	Título, subtítulo, número de la parte, nombre de la parte, caracteres a no alfabetizar	Lengua, transliteración, abreviado, traducido, alternativo, uniforme
Nombre <Name>	Nombre, forma de visualización, afiliación, rol ( <i>atrib.:</i> código o texto), descripción	Lengua, transliteración, personal, corporativo, conferencia
Tipo de recurso <TypeResource>	No posee. Incluye los siguientes valores: Texto Cartografía Música anotada Grabación sonora (musical y no musical) Imágenes fijas	Colección, manuscrito



	<p>Imágenes en movimiento</p> <p>Objetos tridimensionales</p> <p>Software, multimedia</p> <p>Material mixto</p>	
<p>Género</p> <p><b>&lt;Genre&gt;</b></p>	<p>No posee.</p> <p>Incluye los siguientes valores:</p> <p>Categoría que caracteriza un estilo particular, la forma o contenido, como artísticas, musicales, composición literaria, etc.</p>	No posee
<p>Información de origen</p> <p><b>&lt;OriginInfo&gt;</b></p>	<p>Lugar (<i>atrib.</i>: código o texto), editor, fecha de edición, fecha de creación o fabricación, fecha de digitalización, fecha de validez del contenido del recurso, fecha de copyright, otras fechas, edición, emisión (recurso continuado, monografía), frecuencia</p>	<p>códigos, fechas inicio-fin, fecha aproximada, inferida o cuestionable</p>
<p>Idioma</p> <p><b>&lt;Language&gt;</b></p>	Idioma	Código o texto y lista de autoridad
<p>Descripción física</p> <p><b>&lt;PhysicalDescription&gt;</b></p>	<p>Forma (<i>atrib.</i>: lista de autoridad y tipo), calidad del reformato, tipo de formato electrónico, extensión del recurso, origen digital, nota (<i>atrib.</i>: idioma, transliteración, visualización y tipo)</p>	
<p>Resumen</p> <p><b>&lt;Abstract&gt;</b></p>	No posee	Idioma, transliteración, visualización y tipo.
<p>Tabla de contenido</p> <p><b>&lt;TableOfContents&gt;</b></p>	No posee	Idioma, transliteración, visualización y tipo.
<p>Audiencia</p> <p><b>&lt;TargetAudience&gt;</b></p>	No posee	Idioma, transliteración, lista de autoridad

Nota <b>&lt;Note&gt;</b>	Idioma, transliteración, visualización y tipo.	
Materia <b>&lt;Subject&gt;</b>	Materia general, término geográfico, término cronológico ( <i>atrib.:</i> código, inicio-fin, fecha aproximada, inferida o cuestionable), título como materia, nombre como materia, código geográfico ( <i>atrib.:</i> lista de autoridad), género como materia, jerarquía geográfica, ocupación como materia, término cartográfico	Idioma, transliteración, lista de autoridad.
clasificación <b>&lt;Classification&gt;</b>	No posee	Idioma, transliteración, lista de autoridad, no. de edición, visualización.
Identificador <b>&lt;Identifier&gt;</b>	No posee	Idioma, transliteración, tipo (doi, isbn, issn, etc.), visualización.
Ubicación <b>&lt;Location&gt;</b>	Ubicación física, signatura topográfica, url ( <i>atrib.:</i> fecha del último acceso, visualización, nota, acceso, uso), existencias.	
Condiciones de acceso <b>&lt;AccessCondition&gt;</b>	No posee	Idioma, transliteración, visualización, tipo ( <i>atrib.:</i> restricciones de acceso, uso y reproducción)
Parte <b>&lt;Part&gt;</b>	Detalle, extensión, fecha, texto	Tipo (volumen, edición, capítulo, sección, párrafo, pista), orden
Extensión <b>&lt;Extension&gt;</b>	No posee	

Información del registro <b>&lt;RecordInfo&gt;</b>	Organización que crea o modifica el registro original, fecha de creación del registro, fecha de modificación del registro, no. identificador del registro, origen del registro MODS, idioma de la catalogación, normas de descripción.	Idioma, transliteración
---	--	-------------------------

Todos los elementos y atributos son opcionales. Los elementos son repetibles y los atributos no repetibles. Los subelementos contienen atributos. En ciertos elementos se puede hacer control de autoridades.

Se utilizan en cualquier ámbito o materia. Tiene un nivel de extensibilidad que permite tanto una descripción muy sencilla como una más compleja. Para el control de los valores de los datos colocados en algunos elementos sugiere esquemas de codificación, tales como normas ISO, reglas de catalogación, sistemas de clasificación y tesauros. Permite exportar a esquemas más sencillos como Dublin Core y a otros más complejos.

Es mantenido y actualizado por The Network Development and MARC Standards Office, de la Library of Congress.

**FGDC/CSDGM: Content Standard for Digital Geospatial Metadata**  
**(<http://www.fgdc.gov/metadata/geospatial-metadata-standards>)**

Desarrollado por The Federal Geographic Data Committee (FGDC)<sup>3</sup>. Este estándar básico puede ser aplicado por todas las áreas de datos geoespaciales, pero también se puede adaptar para otras disciplinas. La primera versión (CSDGM) surgió en 1994, pero con el uso se fue modificando y extendiendo, y esas sugerencias fueron ratificadas por el FGDC. La versión vigente actualmente, es la 2.0 del año 1998.

Tiene una estructura jerárquica y con 336 elementos organizados en 7 secciones principales y 3 de apoyo (Cita, Período temporal, Contacto).

---

<sup>3</sup> Comité de agencias que promueve el desarrollo coordinado, el uso, el intercambio y la diseminación de datos básicos geoespaciales nacionales de Estados Unidos.

## Secciones principales:

**Información de identificación:** contiene información básica del recurso geográfico como el nombre del set de datos, quién lo ha desarrollado, la cobertura geográfica, el tipo de información que incluye, actualidad de los datos, restricciones de acceso y uso del set, etc.

**Información de calidad de datos:** contiene información relativa a las fuentes utilizadas y proceso de producción, consistencia lógica, completitud del set de datos, informes acerca de la corrección posicional, tanto vertical como horizontal, etc.

**Información de organización de datos espaciales:** contiene información sobre los mecanismos utilizados para representar la información espacial en el set de datos (directa o indirecta), número y tipo de cada objeto si es vectorial, filas y columnas si es raster, modelo de datos espaciales usado para codificar los datos espaciales, etc.

**Información de referencia espacial:** contiene información sobre el sistema de referencia utilizado para codificar las coordenadas de los objetos geométricos del set de datos (Horizontal y Vertical), tales como detalles sobre resolución de la latitud y la longitud, detalles sobre el sistema de proyección o cuadrícula, qué parámetros deberían usarse para convertir el dato a otro sistema de coordenadas, etc.

**Información de entidad y atributo:** contiene información geográfica sobre calles, casas, elevaciones, temperatura, etc., cómo y dónde está codificada y significado de los códigos.

**Información de distribución:** contiene información sobre a través de quién se pueden obtener los datos, en qué formato están disponibles, si están disponibles en línea y su precio, etc.

**Información de referencia de metadatos:** contiene información sobre control del registro de metadatos, como el autor del registro, fecha de creación y última modificación, nombre del estándar de metadatos y versión.

El esquema FGDC/CSDGM es el estándar estadounidense para documentar los datos geoespaciales creados a partir de enero de 1995. Si embargo se ha aplicado mucho más allá de las fronteras de ese país.

A través de la ISO, y la colaboración de otras entidades, entre ellas la FGDC, se ha elaborado y aprobado un estándar internacional de metadatos para información geográfica, la **ISO 19115**, hacia la cual están comenzando a migrar los diversos sistemas, por lo tanto, se está en una etapa de transición.

Algunas características de esta norma son:

- Contiene menos elementos obligatorios y más elementos opcionales
- Se han ampliado algunos elementos y se han creado otros nuevos para registrar información más específica
- Tiene una estructura jerárquica que crea "paquetes o set" de metadatos que pueden ser reutilizados y combinarse para formar nuevos registros de metadatos
- Incorpora documentación de apoyo como bases de datos geográficas, aplicaciones web de cartografía, modelos de datos, portales de datos, ontologías, etc.
- Sugiere las mejores prácticas para la creación de registros de metadatos a fin de mejorar la calidad y utilidad de los metadatos.

### **IEEE/LOM: Learning Object Metadata (<http://ltsc.ieee.org/wg12/>)**

Más conocido como LOM, se trata del estándar por excelencia para describir objetos educativos. Está patrocinado por el Comité de Estandarización de Tecnologías Educativas del Institute of Electrical and Electronic Engineers (IEEE)<sup>4</sup>.

Su objetivo principal es facilitar la búsqueda, la evaluación, la adquisición y el uso de los recursos educativos ya sea por parte de los instructores, como de los alumnos y facilitar el intercambio de objetos educacionales, teniendo en cuenta la diversidad cultural donde estos recursos y metadatos pueden ser usados<sup>5</sup>.

Según el estándar los objetos educativos son toda entidad, digital o no digital, que puede ser utilizada, reutilizada o referenciada durante cualquier actividad de aprendizaje basada en la tecnología (por ej. sistemas de educación a distancia, ambientes de aprendizaje colaborativos, sistemas de aprendizaje interactivo, etc.). Los objetos de aprendizaje pueden

---

<sup>4</sup> Asociación técnico-profesional mundial dedicada a la estandarización entre otras cosas. Es la mayor asociación internacional sin fines de lucro formada por profesionales de las nuevas tecnologías, como ingenieros en electricidad, electrónica, informática, biomédica, telecomunicaciones y mecatrónica, científicos de la computación y matemáticos aplicados.

<sup>5</sup> [http://ltsc.ieee.org/wg12/files/LOM\\_1484\\_12\\_1\\_v1\\_Final\\_Draft.pdf](http://ltsc.ieee.org/wg12/files/LOM_1484_12_1_v1_Final_Draft.pdf)

incluir contenidos multimedia, objetivos de aprendizaje, software educativo, herramientas de software<sup>6</sup>.

La norma especifica la sintaxis y la semántica de los metadatos para objetos de aprendizaje y los atributos necesarios para describirlos adecuadamente.

Los metadatos LOM poseen una estructura jerárquica (arbórea), donde el nodo raíz corresponde al documento que se está describiendo y suele recibir el nombre de "lom". En el siguiente nivel encontramos sub-elementos (ramas), que pueden contener a su vez otros sub-elementos (hojas). Para cada elemento de la jerarquía se especifica la definición, el tipo de datos, los valores permitidos y si se permite su repetibilidad.

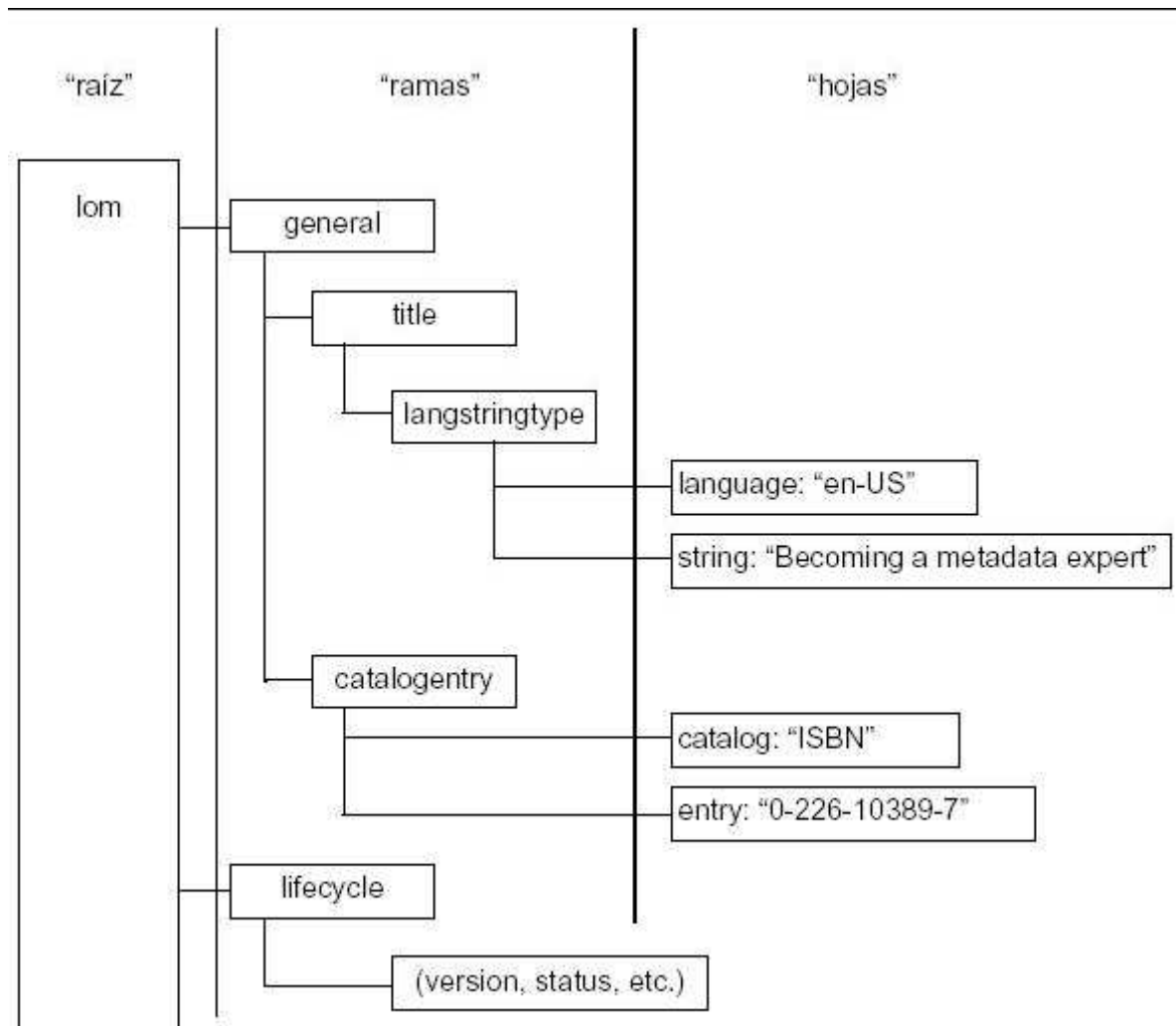
El Esquema base de LOM se compone de 9 categorías y 47 elementos:

Categorías	Elementos
1. General	Identificador, Título, Entrada de catálogo, Lengua, Descripción, Descriptor, Cobertura, Estructura, Nivel de agregación
2. Ciclo de vida	Versión, Estatus, Otros colaboradores
3. Meta-metainformación	Identificador, Entrada de catálogo, Otros colaboradores, Esquema de metadatos, Lengua
4. Técnica	Formato, Tamaño, Ubicación, Requisitos, Comentarios sobre la instalación, Otros requisitos para plataformas, Duración
5. Uso educativo	Tipo de interactividad, Tipo de recurso de aprendizaje, Nivel de interactividad, Densidad semántica, Usuario principal, Contexto [Nivel educativo], Edad, Dificultad, Tiempo previsto de aprendizaje, Descripción, Lengua
6. Derechos	Coste, Copyright y otras restricciones, Descripción

<sup>6</sup> <http://ltsc.ieee.org/wg12/>

7. Relación [con otros recursos]	Tipo [naturaleza de la relación con el recurso principal], Recurso [recurso principal al que se refiere esta relación]
8. Observaciones	Persona, Fecha, Descripción
9. Clasificación	Finalidad, Nivel táxon (taxonómico), Descripción, Descriptor

En el gráfico siguiente se muestran los sub-elementos que pertenecen al elemento "general".



## **MIX: Metadata for Images in XML**

**(<http://www.loc.gov/standards/mix/>)**

El esquema MIX fue desarrollado por la Network Development and MARC Standards Office de la Library of Congress, junto con la NISO Technical Metadata for Digital Still Images Standards Committee, para administrar colecciones de imágenes y facilitar el intercambio y el archivo a largo plazo. Está mantenido por la Biblioteca del Congreso y contiene una descripción de las características técnicas de una imagen.

El estándar es la representación en XML del diccionario de elementos definido en la norma: *NISO Draft Standard Data Dictionary: Technical Metadata for Digital Still Images*. El modelo proporciona un formato para el intercambio y / o almacenamiento de los datos especificados en el diccionario.

Este diccionario ha sido diseñado para facilitar la interoperabilidad entre los sistemas, servicios y software, así como para apoyar la gestión a largo plazo y el acceso continuo a las colecciones de imágenes digitales.

Consta de una lista de elementos de datos técnicos, necesarios para la gestión de imágenes fijas digitales. En este contexto, la "gestión" se refiere a las tareas y operaciones necesarias para apoyar la evaluación de la calidad de la imagen, el procesamiento de los datos de las imágenes y el mantenimiento a largo plazo, durante todo el ciclo de vida de la imagen.

Este esquema contiene específicamente metadatos de preservación y técnicos. No contempla elementos descriptivos ni sobre gestión de derechos, por lo que puede ser utilizado junto con otros estándares más generales, como DC, MODS, METS, PREMIS, etc.

Agrupar los elementos en cinco categorías principales:

1. **Basic digital object information:** Información básica del objeto digital, como tamaño del archivo, información sobre el formato.
2. **Basic image information:** Información básica de la imagen, como la compresión, dimensiones de la imagen.
3. **Image capture metadata:** Metadatos de captura de la imagen, incluye información sobre el proceso de captura, tal como el dispositivo de escaneo, ajustes y software utilizado en el proceso.
4. **Image assessment metadata:** Metadatos de evaluación de la imagen. Importantes para el mantenimiento de la calidad de la imagen. Esta sección es necesaria para evaluar la precisión de salida. Incluye información de color (tales como puntos blancos y mapas de colores) y sobre la resolución de la imagen.



5. **Change history:** Metadatos de historial de cambios. Tienen como función documentar los procesos aplicados a los datos de la imagen durante su ciclo de vida, como cualquier edición o transformación (migración) del archivo.

Cabe aclarar que los metadatos de la sección **Image capture metadata**, se utilizan para documentar la creación de una imagen a partir de una fuente analógica. Los metadatos de la sección **Change history** se utilizan para documentar todas las operaciones digitales posteriores a la captura.

**PREMIS: PREservation Metadata: Implementation Strategies**  
**(<http://www.loc.gov/standards/premis/>)**  
**([www.loc.gov/standards/premis/v2/premis-2-0.pdf](http://www.loc.gov/standards/premis/v2/premis-2-0.pdf))**

PREMIS son las siglas de "PREservation Metadata: Implementation Strategies" (Metadatos de preservación: estrategias de implementación) que es el nombre de un grupo de trabajo internacional patrocinado por OCLC y RLG desde 2003-2005. Este grupo de trabajo elaboró un informe denominado PREMIS Data Dictionary for Preservation Metadata (Diccionario de datos PREMIS de metadatos de preservación) que incluye un diccionario de datos e información sobre los metadatos de preservación. En marzo de 2008 se publicó la segunda versión actualizada. La Library of Congress mantiene un esquema de representación de PREMIS en XML.

El Diccionario de datos PREMIS define unidades semánticas, no elementos de metadatos. Esta diferencia es importante ya que una unidad semántica es una pieza de información o de conocimiento. Un elemento de metadatos es una manera definida de representar esa información en un registro de metadatos, en un esquema o en una base de datos. PREMIS no especifica cómo deben representarse los metadatos en un sistema, únicamente define lo que necesita entender el sistema y lo que debe poder exportarse a otros sistemas.

PREMIS solamente define los metadatos que se necesitan, por lo general, para llevar a cabo las funciones de preservación de todos los materiales. No incluye metadatos descriptivos. Para este objetivo, el esquema debe complementarse con otro estándar tal como MARC21, MODS, Dublin Core, etc. Tampoco contempla la información sobre derechos y permisos, excepto los que afectan directamente a las funciones de preservación.

PREMIS se utiliza fundamentalmente para el diseño de los repositorios, para su evaluación y para el intercambio de los paquetes de información archivada entre los repositorios de preservación.

La mayor parte de sus elementos se han diseñado para que los proporcione

automáticamente la aplicación que administra el repositorio de preservación. Esto no quiere decir, que todas las aplicaciones actualmente disponibles lo permitan.

Las unidades semánticas de PREMIS están organizadas en cinco entidades diferentes: Entidades Intelectuales, Objetos, Acontecimientos (o eventos), Derechos y Agentes.

## **METS: Metadata Encoding and Transmission Standard**

([http://www.loc.gov/standards/mets/METSOverview\\_spa.html](http://www.loc.gov/standards/mets/METSOverview_spa.html))

Se trata de un esquema para la gestión de objetos de bibliotecas digitales complejas y para su intercambio entre repositorios, que utiliza el lenguaje XML y que empaqueta todos los metadatos asociados con un recurso digital: estructurales, administrativos y descriptivos, necesarios para recuperar, preservar y ofrecer recursos digitales. Está pensado principalmente para el envío de los archivos, imágenes y objetos multimedia de una biblioteca digital.

El estándar es mantenido por la Network Development and MARC Standards Office de la Biblioteca del Congreso.

Permite una gestión integral y jerárquica de manifestaciones de una misma obra: por ejemplo, texto, grabación sonora y video de una conferencia.

Un documento METS consta de siete secciones:

1. **Cabecera METS (metsHdr):** contiene metadatos que describen el propio documento METS, e incluye datos como su creador, editor, fecha y hora de su creación, etc.
2. **Metadatos Descriptivos (dmdSec):** contiene la descripción del objeto al que se hace referencia en el documento METS. Esta sección puede:
  - a) apuntar a metadatos descriptivos externos al documento METS (por ejemplo una URN o URL de un registro MARC en un OPAC o un documento EAD disponible en un servidor web)
  - b) contener internamente los metadatos descriptivos, o
  - c) combinar ambas aproximaciones. En la sección Metadatos Descriptivos se pueden incluir múltiples metadatos descriptivos, tanto internos como externos.

Ejemplo de un documento METS:

```
<mets:dmdSec ID=dmd001>
```

```
  <mets:mdWrap mimetype="text/xml" mdtype="DC" Label="Dublin Core  
  Metadata">
```

```
    <dc:title>El caballero del jubón amarillo</dc:title>
```

```
    <dc:creator>Pérez-Reverte, Arturo</dc:creator>
```

```
    <dc:date>2003</dc:date>
```

```
    <dc:publisher>Alfaguara</dc:publisher>
```

```
    <dc:type>text</dc:type>
```

```
  </mets:mdWrap>
```

```
</mets:dmdSec>
```

3. **Metadatos Administrativos (admSec):** ofrece información sobre cómo se crearon y almacenaron los archivos que conforman el objeto digital.

En los documentos METS hay cuatro tipos de metadatos administrativos:

1. Metadatos técnicos (información relativa a la creación del archivo, su formato y características de uso)

2. Metadatos sobre derechos de propiedad intelectual (copyright e información sobre licencias)

3. Metadatos sobre el origen (metadatos descriptivos y administrativos sobre el documento origen a partir del cual se ha generado el objeto digital)

4. Metadatos sobre la procedencia digital (información sobre la relación entre el documento original y su representación digital, incluyendo la relación entre copias maestras y derivadas, migraciones y transformaciones realizadas sobre los archivos desde su digitalización inicial).

Al igual que sucede con los metadatos descriptivos, los metadatos administrativos pueden ser externos o codificarse dentro del propio documento METS.

4. **Sección Archivo (fileGrp):** contiene uno o más elementos <fileGrp>. Estos agrupan archivos relacionados entre sí. Un <fileGrp> reúne todos los archivos que conforman una misma versión electrónica del objeto digital. Por ejemplo, puede haber elementos

<fileGrp> para las miniaturas, las copias maestras (alta resolución) de las imágenes, la versión en pdf, la versión codificada en TEI, etc.

5. **Mapa Estructural (StructMap):** es la parte principal de un documento METS. Recoge la estructura jerárquica del objeto digital, y enlaza sus secciones con los archivos de contenido y los metadatos correspondientes a cada una de ellas. Esta estructura puede presentarse a los usuarios para navegar a través del objeto digital.
6. **Enlaces Estructurales (SmLink):** permite registrar la existencia de hiperenlaces entre las secciones del mapa estructural. Tiene gran valor cuando se usa METS para archivar sitios web y mantener un registro de su estructura hipertextual aparte de la que se establecen mediante los hiperenlaces de las propias páginas HTML.
7. **Comportamientos:** Se utiliza para asociar comportamientos ejecutables con los contenidos del objeto METS.

### **EAD: Encoded Archival Description (<http://www.loc.gov/ead/>)**

Este esquema se inicia en la Biblioteca de la Universidad de California, Berkeley, en 1993. La primera versión surge en 1998. Es un esquema no-propietario, que proporciona los elementos administrativos y descriptivos para la organización y descripción de material digitalizado de archivo y manuscritos.

Un documento codificado utilizando EAD, básicamente consta de tres segmentos:

**<eadheader>:** Proporciona información sobre el documento en sí mismo (título, compilador, fecha de compilación).

**<frontmatter>:** Incluye las cuestiones preliminares necesarias para la publicación formal del documento

**<findaid>:** Proporciona la descripción del material archivístico en sí misma, además de la información contextual y administrativa asociada.

Utiliza la sintaxis XML y la última versión (2002) contiene 146 elementos (distribuidos en los 3 segmentos). Se adapta a una amplia gama de prácticas descriptivas internacionalmente divergentes. Es una estructura de datos y no una norma de contenido de datos.

### **TEI: Text Encoding Initiative (<http://www.tei-c.org/>)**

Se trata de un proyecto internacional que desarrolla pautas para el marcado de textos electrónicos y se enfoca principalmente al campo de las humanidades, ciencias sociales y lingüística, con el objetivo de permitir a los investigadores de todas las disciplinas intercambiar y reutilizar recursos, independientemente del software y hardware que utilicen y sin tener en cuenta dónde están localizados.

Todos los textos TEI constan de una cabecera (**teiHeader**) y un cuerpo (**text**). En la cabecera se codifican todos los datos bibliográficos del recurso electrónico y en el cuerpo, va el texto en sí mismo, donde se marcan las distintas secciones en lenguaje XML. La cabecera puede almacenarse como parte separada del documento al que se refiere o ir unida intrínsecamente a él.

La cabecera TEI consta de 4 secciones:

1. La descripción del archivo **<fileDesc>**, contiene la descripción bibliográfica completa que permite citarlo.
2. La descripción sobre la codificación **<encodingDesc>**, en la que se precisan las incidencias en el momento de la transcripción.
3. El perfil del texto **<profileDesc>**, aporta información adicional contextual y no bibliográfica como el idioma, colaboradores, materias, descriptores, etc.
4. El historial de revisiones **<revisionDesc>**, permite llevar un registro de los cambios realizados sobre la versión electrónica.

Las descripciones que se proporcionan para cumplimentar los elementos más estructurados, aplican las AACR2 y las ISBD. Los elementos no estructurados contienen texto libre.

Su gran flexibilidad facilita su adaptación a cualquier usuario, ya que permite un mayor o menor nivel de detalle en la descripción, según las necesidades.

## **RDF Resource Description Framework ([www.w3.org/RDF](http://www.w3.org/RDF))**

Desarrollado por el W3C<sup>7</sup> para la descripción de recursos de la Web. Provee un mecanismo para integrar múltiples esquemas de metadatos. Mientras que XML es un lenguaje para modelar datos, RDF es un lenguaje para especificar metadatos. Está particularmente indicado para representar metadatos sobre recursos Web, tales como el título, autor, modificaciones de los datos de la página, copyright y otras licencias de información sobre documentos *Web*, así como la disponibilidad para algunos recursos compartidos.

Es un modelo estándar para intercambiar datos en Internet. Permite la interoperabilidad entre aplicaciones que intercambian información comprensible por la Web, proporcionando una infraestructura que soporta acciones sobre los metadatos.

El lenguaje RDF es muy útil en situaciones en las que la información necesita ser procesada por aplicaciones que intercambian datos legibles por máquina, más que por

---

<sup>7</sup> El Consorcio World Wide Web (W3C) es una comunidad internacional donde las organizaciones miembros trabajan conjuntamente para desarrollar estándares para la Web. Liderado por el inventor de la Web Tim Berners-Lee y el Director Ejecutivo (CEO) Jeffrey Jaffe. La misión del W3C es aprovechar al máximo el potencial de Internet.

humanos. RDF también provee una sintaxis basada en XML, llamada RDF/XML, que define cómo debe ser etiquetada la información a ser intercambiada.

Este estándar permite que en la descripción de un recurso puedan ser combinados elementos de diferentes esquemas de metadatos. De esta forma pueden enlazarse a otras múltiples descripciones creadas en diferentes momentos y con propósitos diferentes.

Ejemplo de RDF:

**<rdf:RDF**

xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"

xmlns:dc="http://purl.org/dc/elements/1.1/">

**<rdf:Description** rdf:about="http://en.wikipedia.org/Tony\_Benn">

**<dc:title>**Tony Benn**</dc:title>**

**<dc:publisher>**Wikipedia**</dc:publisher>**

**</rdf:Description>**

**</rdf:RDF>**

## ***Elección del esquema de metadatos para un repositorio institucional***

### ***¿Creación de un esquema propio o de un perfil de aplicación?***

Cuando la institución está decidida a crear un repositorio digital, una de las primeras preguntas que debe realizarse es: ¿Es necesario crear un nuevo esquema de metadatos o existe alguno que pueda servir para cumplir con el objetivo del repositorio?

La respuesta surgirá, una vez que se hayan resuelto los siguientes temas:

- Tipo de institución (académica, gubernamental, científica, etc.)
- El tipo de información que contendrá el repositorio (general, especializada, artículos, tesis, ponencias, etc.)
- Los tipos de objetos digitales que se van a incluir (texto, video, audio, etc.).
- El objetivo del repositorio (acceso, difusión, preservación, etc.).
- Profundidad que se desea dar a la descripción.
- Qué han hecho otras instituciones similares, tanto de la región como de otros países.
- Relevar y analizar los esquemas de metadatos existentes.
- Evaluar si estos esquemas cubren las necesidades y objetivos del repositorio.
- Analizar el cumplimiento de los requisitos de interoperabilidad para la integración a redes.

Es importante tener en cuenta, que se pueden utilizar varios esquemas complementarios para describir el mismo recurso y para múltiples propósitos. Por ejemplo, un artículo de una revista podría tener el conjunto de metadatos MARC en el catálogo en línea de una biblioteca, pero a la vez, podría contener una descripción MODS si forma parte de una biblioteca digital, y tener embebidos elementos del conjunto de metadatos Dublin Core para la interoperabilidad. O, también, si una de las finalidades del repositorio es la preservación, se pueden utilizar elementos de un esquema de metadatos descriptivo (DC, MODS u otro) y además el conjunto de elementos PREMIS.

#### **Creación de un esquema de metadatos propio:**

Es indispensable tener en cuenta que la creación de un esquema propio implica gestionarlo y mantenerlo durante el lapso de vida de los documentos, lo que incluye su actualización, su compatibilidad en el pasado y en el futuro, su registro en organismos

relevantes y la infraestructura necesaria para su mantenimiento (tiempo, recursos humanos especializados, software y hardware, documentación, etc.).

Si la institución decide crearlo, es requisito imprescindible estudiar el grupo de normas **UNE-ISO 23081**:

- 23081-1 (2008): Información y documentación. Procesos de gestión de documentos. Metadatos para la gestión de documentos. Parte 1: Principios.
- 23081-2 (2011): Información y documentación. Procesos de gestión de documentos. Metadatos para la gestión de documentos. Parte 2: Elementos de implementación y conceptuales.
- 23081-3 (2012): Información y documentación. Metadatos para la gestión de documentos. Parte 3: Método de auto-evaluación.

### **Creación de un perfil de aplicación:**

Un perfil de aplicación es el uso específico que se le da a un esquema de metadatos existente, en una entidad en particular. Incluye las políticas propias y las directrices en el uso de los elementos. Determina su obligatoriedad y limitaciones, explica cómo deben usarse y da ejemplos aclaratorios. Para poder cumplir con sus objetivos, un perfil puede integrar elementos provenientes de más de un esquema de metadatos, con la precaución de que éstos no se solapen.

Para este propósito es necesario seguir los siguientes pasos:

- Estudiar el/los esquema/s existente/s:
  - Para qué tipo de documentos ha sido diseñado.
  - Revisar su estructura.
  - Consultar con otras instituciones que ya lo hayan aplicado para crear su propio perfil (selección de los elementos, mejoras (subelementos) y esquemas de codificación usados<sup>8</sup>).
- Estudiar las normas UNE-ISO 23081.
- Determinar el alcance del perfil de aplicación:
  - Qué información contendrá el repositorio
  - Qué objetos digitales se van a describir (texto, imágenes, videos, etc.).
  - Qué interacción tendrá el repositorio con otros (interoperabilidad).

---

<sup>8</sup> “Lista controlada de todos los valores aceptables en lenguaje natural o cadena de caracteres formateada con una sintaxis concreta, diseñados para su procesamiento automatizado”. UNE-ISO 23081-1. Capítulo 3, Términos y definiciones.



- Registrar el perfil de aplicación en organismos relevantes.
- Identificar:
  - Los elementos que se van a incluir en el perfil.
  - Los cambios o mejoras (subelementos) que sea necesario realizar.
  - Qué elementos son obligatorios, recomendados y opcionales.
  - Esquemas de codificación para consignar la información.
  - Elementos de otros esquemas reconocidos, que puedan ser útiles para registrar datos específicos y que no estén contemplados en el esquema elegido (por ej. Datos geoespaciales, de preservación).
- Redactar la documentación y manuales de procedimiento.

Diseñar un esquema propio tiene la gran ventaja de que se va a ajustar exactamente a nuestras necesidades, pero no podemos perder de vista dos aspectos fundamentales: por un lado, el costo que implica esta tarea, y por otro, las ventajas de la normalización.

Ante la disyuntiva de **crear un esquema o realizar un perfil de aplicación**, lo más recomendable es **utilizar un esquema existente** con el agregado de subelementos locales específicos (mejoras). Es más fácil y sustentable en el tiempo, adoptar algo que ya existe, que está bien diseñado y que está probado y apoyado globalmente.

El uso de esquemas de metadatos estandarizados permite una comunicación en otro nivel con los cosechadores de datos (por ejemplo iniciativa OAI). En este nuevo nivel, nosotros, como proveedores, definimos con precisión el contenido creando una relación de “confianza” entre cosechador y proveedor. Sin el uso de estándares no se produce una buena comunicación, por lo tanto va en desmedro de una buena recuperación y visibilidad de la información, lo que entorpece el acceso a la misma, principal insumo para promover el desarrollo de la actividad cultural, científica y tecnológica de las instituciones.

## ***EJEMPLOS DE PERFILES DE APLICACIÓN***

### **Aplicaciones basadas en DC:**

#### **DRIVER Project (2008)**

DRIVER, “Digital Repository Infrastructure Vision for European Research” (Visión de infraestructura de repositorios digitales para la investigación europea), es un proyecto realizado por un consorcio financiado por la Unión Europea que creó un marco de trabajo tecnológico y

organizativo para recolectar datos de todas las instituciones del ámbito de la investigación y la educación superior.

La red de repositorios digitales europea es muy heterogénea, en cuanto a los recursos incluidos (texto, datos, multimedia), como a las plataformas tecnológicas y a las distintas políticas de metadatos.

Desarrolló una infraestructura de servicios y una infraestructura de datos. Para ello creó un portal desde donde se puede acceder a toda la información y además generó unas directrices cuyo objetivo fue unificar criterios en el uso de los metadatos para poder interoperar entre repositorios.

La aplicación “DRIVER 2.0 Directrices para proveedores de contenido - Exposición de recursos textuales con el protocolo OAIPMH”, están basadas en la Iniciativa Dublin Core sin calificar, y establece las características de cada elemento según los requerimientos del proyecto.

### **OpenAIRE Guidelines (2010)**

OpenAIRE es un proyecto de la Comunidad Europea, que tiene por objeto proveer un soporte para el libre acceso a la información en Europa. Los recursos recolectados serán los resultados de las investigaciones financiadas por el CE y que además, hayan sido revisados por pares.

Estas directrices, elaboradas por un grupo europeo de expertos, proporcionan orientación a los gestores de repositorios para definir y aplicar sus políticas locales de gestión de datos de acuerdo con las demandas de acceso abierto de la Comisión Europea. Las directrices OpenAIRE complementan y se construyen sobre las Directrices DRIVER.

Para ser compatible con las OpenAIRE, todos los aspectos de las Directrices DRIVER son válidos, con las excepciones siguientes:

- Los recursos pueden estar en acceso abierto, bajo embargo o tener acceso restringido.
- El uso de los campos dc:relation y dc:rights es más específico.
- No es obligatorio para OpenAIRE utilizar el estilo APA en los elementos dc:creator y dc:contributor. Se puede poner el nombre completo cuando esté disponible.

### **Directrices SNRD (2012)**

El Sistema Nacional de Repositorios Digitales (SNRD) es una iniciativa del Ministerio de Ciencia, Tecnología e Innovación Productiva conjuntamente con el Consejo Interinstitucional de Ciencia y Tecnología (CICYT) a través de sus representantes en el Consejo Asesor de la Biblioteca Electrónica de Ciencia y Tecnología.

El Sistema se creó por Resolución Ministerial N. 469/11 del 17 de mayo de 2011 y tiene como propósito conformar una red interoperable de repositorios digitales en ciencia y tecnología, a partir del establecimiento de políticas, estándares y protocolos comunes a todos los integrantes del Sistema.

Las “Directrices SNRD para proveedores de contenido” surgen en 2012 con el propósito de obtener un documento práctico que sirva de guía para la creación de registros de metadatos producidos por organismos académicos y/o científicos del país, a fin de lograr un estándar nacional, que permita participar de iniciativas OAI-PMH y de esta manera, los repositorios digitales creados en el país puedan ser absolutamente compatibles con los del resto del mundo y convertirse en proveedores de contenido para los cosechadores de datos (harvesters), facilitando la visibilidad de su información científico/académica.

Se tomaron como modelos las siguientes directrices basadas en Dublin Core:

- Driver 2.0 (Digital Repository Infrastructure Vision for European Research)
- OpenAire
- BDCol (Bibliotecas Digitales de Colombia)

Cada repositorio determina qué registros aportará al SNRD, los cuales deberán respetar el esquema de metadatos propuestos por el SNRD y prestar especial atención a:

- Aportar una URI accesible (que enlace en forma directa al objeto digital descrito o a la ficha de acceso al mismo).
- Incluir sólo valores de las listas controladas definidas en el esquema de metadatos (dc:type, dc:rights).
- Respetar los esquemas de codificación definidos en las directrices, en especial campos como fechas, derechos e idioma (dc:date, dc:Rights, dc:language).

## **Aplicaciones basadas en MODS:**

### **DLF/Aquifer**

La Federación de Bibliotecas Digitales de los Estados Unidos (Digital Library Federation-DLF), un programa del Consejo de Bibliotecas y Recursos de Información (Council on Library and Information Resources), estableció las Directrices para la Implementación de Registros MODS Compartibles (Digital Library Federation / Aquifer Implementation Guidelines for Shareable MODS Records).

El principal objetivo de la Federación, es permitir que los contenidos distribuidos por las

bibliotecas miembros, puedan ser usados efectivamente por bibliotecarios y académicos para la enseñanza, el aprendizaje y la investigación.

Tal como lo dice su nombre, es una implementación del esquema MODS, versión 3.2. Tratan de proporcionar una estructura tan simple como sea posible para la presentación de metadatos teniendo en cuenta de no perder la interoperabilidad (ya sea a través del protocolo OAI-PMH, o algún otro medio) más allá de su contexto local.

### **Directrices LUCIS**

Las Directrices LUCIS surgen a partir de la necesidad de dotar a la Biblioteca Digital de la Universidad Nacional de Cuyo de una estructura normalizada de metadatos que permitiera su posterior interoperabilidad e intercambio de información a nivel nacional e internacional.

A partir de la comparación de los esquemas de metadatos Dublin Core (DC) y Metadata Object Description Schema (MODS), se propuso una aplicación para la descripción de los objetos digitales producidos por las distintas Unidades Académicas de la Universidad, a fin de lograr un esquema que se adapte a las características propias del repositorio y además que permita participar de la iniciativa OAI-PMH.

Esta Biblioteca Digital había sido creada originalmente con una estructura basada en Dublin Core, la cual resultó insuficiente al momento de comenzar a cargar los metadatos. La complejidad de los recursos (texto, audio y video) y los usuarios a los cuales éstos van dirigidos, obligaron, necesariamente, a realizar una descripción acorde a sus necesidades. Por lo tanto, se eligió como base el esquema de metadatos MODS a partir del cual se generó la aplicación.

Se tomaron como modelos las siguientes directrices:

- Driver 2.0 (Digital Repository Infrastructure Vision for European Research)
- OpenAIRE Guidelines 1.0
- DLF/Aquifer.

Curso organizado en el marco del Proyecto PICTO 2010-149 “Investigación y desarrollo en repositorios institucionales. Aplicaciones y experiencias en universidades de la región bonaerense” cofinanciado por la Agencia Nacional de Promoción de la Ciencia y la Tecnología y el Consejo Interuniversitario Nacional.

Universidades que integran el proyecto:

Universidad Nacional de La Plata (UNLP),  
Universidad Nacional de Mar del Plata (UNMDP),  
Universidad Nacional de Luján (UNLu),  
Universidad Nacional del Centro de la Provincia de Buenos Aires (UNICEN),  
Universidad Nacional del Sur (UNS),  
Universidad Tecnológica Nacional, Facultad Regional Bahía Blanca (UTN-FRBB)

Investigadora responsable: Dra. Sandra Miguel



**Atribución – No Comercial – Compartir Igual** (*by-nc-sa*): No se permite un uso comercial de la obra original ni de las posibles obras derivadas, la distribución de las cuales se debe hacer con una licencia igual a la que regula la obra original.