



LA IUPI  
UPRPP



# Metadatos. Conceptual

Taller del Certificado edición de revistas electrónicas

*Joel Blanco, Ph.D.*

Escuela Graduada de Ciencias y Tecnologías de la Información  
(EGCTI)

*Snejanka Penkova, Ph.D.*

Biblioteca de Administración de Empresas, Sistema de Bibliotecas

# Conceptos básicos

- **Metainformación**

“Metainformation is information about information. For example, if a document is considered to be information, its title, location, and subject are examples of metainformation. This term is sometimes used interchangeably with the term metadata.”

- **Metaetiquetas (*metatags*)**

“Conjunto de datos de identificación y/o descripción de un recurso de información, disponible en Internet, para su recuperación. Basado en algún sistema de o norma reconocida, por ejemplo “Dublin Core”.



# Conceptos básicos...

- **Metadatos**

“Metadata is structured information that describes, explains, locates, or otherwise makes it easier to retrieve, use, or manage an information resource. Metadata is often called data about data or information about information.”

...metadata is commonly used for any formal scheme of resource description, applying to any type of object, digital or non-digital.

- Metadata can describe resources at any level of aggregation. (Ex. Collection, single resource, a component part of a larger resource (photograph in an article)).



# Tipos de metadatos

- **Descriptivos**

“describes a resource for purposes such as discovery and identification. It can include elements such as title, abstract, author, and keywords.” (*FRONT en XML-JATS*)

- **Estructurales**

“indicates how compound objects are put together, for example, how pages are ordered to form chapters.”

(*BODY en XML-JATS*)

- **Administrativos**

“provides information to help manage a resource, such as when and how it was created, file type and other technical information, and who can access it. There are several subsets of administrative data; two that sometimes are listed as separate metadata types are:

- *Rights management meta- data*, which deals with intellectual property rights, and

- *Preservation metadata*, which contains information needed to archive and preserve a resource.” (*BACK and Float Group en XML-JATS*)

<http://marciazeng.slis.kent.edu/metadatabasics/types.htm>



An important reason for creating descriptive metadata is to facilitate discovery of relevant information. In addition to resource discovery, metadata can help organize electronic resources, facilitate interoperability and legacy resource integration, provide digital identification, and support archiving and preservation.



# Funciones de los metadatos

- ***Resource Discovery***
  - allowing resources to be found by relevant criteria;
  - identifying resources;
  - bringing similar resources together;
  - distinguishing dissimilar resources; and
  - giving location information

- ***Organizing Electronic Resources***

As the number of Web-based resources grows exponentially, aggregate sites or portals are increasingly useful in organizing links to resources based on audience or topic.



# Funciones de los metadatos...

- ***Interoperability***

Describing a resource with metadata allows it to be understood by both humans and machines in ways that promote interoperability. **Interoperability is the ability of multiple systems with different hardware and software platforms, data structures, and interfaces to exchange data with minimal loss of content and functionality.** Using defined metadata schemes, shared transfer protocols, and crosswalks between schemes, resources across the network can be searched more seamlessly.

- Two approaches to interoperability are **cross-system search** and **metadata harvesting**.
- **Protocolo Z39.50** es mayormente utilizado para “*cross-system search*”
- **Open Archives Initiative** – permite a todos los proveedores de datos transferir sus metadatos nativos a un “*common core set of elements*” y exponerlos para recuperación o cosecha (harvesting).



# Funciones de los metadatos...

- ***Digital Identification***

Most metadata schemes include elements such as standard numbers to uniquely identify the work or object to which the metadata refers. The location of a digital object may also be given using a file name, URL (Uniform Resource Locator), or some more persistent identifier such as a PURL (Persistent URL) or DOI (Digital Object Identifier). Persistent identifiers are preferred because object locations often change, making the standard URL (and therefore the metadata record) invalid.





# Funciones de los metadatos...

- ***Archiving and Preservation***

Digital information is fragile; it can be corrupted or altered, intentionally or unintentionally. It may become unusable as storage media and hardware and software technologies change. Format migration and perhaps emulation of current hardware and software behavior in future hardware and software platforms are strategies for overcoming these challenges.

Metadata is key to ensuring that resources will survive and continue to be accessible into the future.

Archiving and preservation require special elements to track the lineage of a digital object (where it came from and how it has changed over time), to detail its physical characteristics, and to document its behavior in order to emulate it on future technologies.



# Funciones de los metadatos

## 3.3 Getty's definitions on types of metadata

Type	Definition	Examples
Administrative	Metadata used in managing and administering collections and information resources	<ul style="list-style-type: none"> <li>• Acquisition information</li> <li>• Rights and reproduction tracking</li> <li>• Documentation of legal access requirements</li> <li>• Location information</li> <li>• Selection criteria for digitization</li> </ul>
Descriptive	Metadata used to identify and describe collections and related information resources	<ul style="list-style-type: none"> <li>• Cataloging records</li> <li>• Finding aids</li> <li>• Differentiations between versions</li> <li>• Specialized indexes</li> <li>• Curatorial information</li> <li>• Hyperlinked relationships between resources</li> <li>• Annotations by creators and users</li> </ul>
Preservation	Metadata related to the preservation management of collections and information resources	<ul style="list-style-type: none"> <li>• Documentation of physical condition of resources</li> <li>• Documentation of actions taken to preserve physical and digital versions of resources, e.g., data refreshing and migration</li> <li>• Documentation of any changes occurring during digitization or preservation</li> </ul>
Technical	Metadata related to how a system functions or metadata behaves	<ul style="list-style-type: none"> <li>• Hardware and software documentation</li> <li>• Technical digitization information, e.g., formats, compression ratios, scaling routines</li> <li>• Tracking of system response times</li> <li>• Authentication and security data, e.g., encryption keys, passwords</li> </ul>
Use	Metadata related to the level and type of use of collections and information resources	<ul style="list-style-type: none"> <li>• Circulation records</li> <li>• Physical and digital exhibition records</li> <li>• Use and user tracking</li> <li>• Content reuse and multiversioning information</li> <li>• Search logs</li> <li>• Rights metadata</li> </ul>

Source:

Murtha Baca ed.(2008).

*Introduction to Metadata Version 3.*

Getty Information Institute, Table 2.

## 3.4 Dublin Core's elements outline



# Dublin Core

- *Ohio College Library Center (OCLC) and National Center for Supercomputing Applications (NCSA) – 1995*
- The original objective of the Dublin Core was to define a **set of elements** that could be used by authors to describe their own Web resources.
- The Dublin Core was developed to be simple and concise, and to describe Web-based documents. However, Dublin Core has been used with other types of materials...
- All Dublin Core elements are optional and all are repeatable. The elements may be presented in any order.



# Ejemplo

Source:

Murtha Baca ed.(2008),

*Introduction to Metadata Version 3.*

Getty Information Institute, Table 2.

## 3.4 Dublin Core's elements outline

DUBLIN CORE Metadata Element Set (= "simple Dublin Core")

Website: <http://dublincore.org/>

DC Element set

Content	Intellectual Property	Instantiation
Coverage	Contributor	Date
Description	Creator	Format
Type	Publisher	Identifier
Relation	Rights	Language
Source		
Subject		
Title		

## Additional notes

Majority of the descriptive metadata are manually created (refer to Section 2) using various tools (refer to Section 5).

*Technical metadata* may be automatically captured by a software.

- E.g., what is the resolution of this [satellite image](#)? What is its size? Will it be too large to download to a smart phone or too small to zoom-in? -- Adobe Bridge can help to provide technical metadata for each [digital file](#) like this.

*Use metadata* can be learned.

- Publishers, social media, and marketing services have been using such data collected based on usage
  - e.g., about a published or posted item: number of times it has been viewed, downloaded, discussed, reviewed, recommended, shared, or cited, and [viewed or bought together](#) with another thing.

Descriptive metadata can also be crawled through [computer programming](#) or [APIs](#).

<-- [Back to Table of Contents](#) |||| [Go to Next Section](#) -->



- Características revistas electrónicas LATINDEX (34-36)

<http://www.redalyc.org/infope.oa?page=/politica-editorial/metodologiaevalua.html>

- Glosario LATINDEX

[http://www.latindex.org/lat/documentos/GlosarioLatindex\\_esp.pdf](http://www.latindex.org/lat/documentos/GlosarioLatindex_esp.pdf)

- Metodología REDALYC (Criterios 56-60)

<http://www.redalyc.org/infope.oa?page=/politica-editorial/metodologiaevalua.html>



# Ejercicio de práctica



# XML

- ¿Qué es XML?
  - Extensible Markup Language
  - Lenguaje de meta-marcado
  - Provee un método uniforme para intercambiar metadatos.

```
▼<titlestmt>  
  ▼<titleproper encodinganalog="title">  
    Corporación de Promoción y Defensa de los Derechos del Pueblo - CODEPU  
  </titleproper>  
  <author encodinganalog="creator">Jorge E.E. Vivar y Graciela Karababikian</author>  
</titlestmt>
```

# XML y HTML

- HTML es un lenguaje para presentar páginas web; XML no lo es.
- HTML tiene unas etiquetas establecidas (ej. <p> para párrafo); XML permite definir etiquetas propias.

```
▼<titlestmt>  
  ▼<titleproper encodinganalog="title">  
    Corporación de Promoción y Defensa de los Derechos del Pueblo - CODEPU  
  </titleproper>  
  <author encodinganalog="creator">Jorge E.E. Vivar y Graciela Karababikian</author>  
</titlestmt>
```





# Estructura básica de XML

- Elementos
- Etiquetas
- Atributos

```
▼<titlestmt>  
  ▼<titleproper encodinganalog="title">  
    Corporación de Promoción y Defensa de los Derechos del Pueblo - CODEPU  
  </titleproper>  
  <author encodinganalog="creator">Jorge E.E. Vivar y Graciela Karababikian</author>  
</titlestmt>
```



# Ejemplo

**Title**

Repertorio histórico de Puerto-Rico.

**Summary**

Cayetano Coll y Toste was born in Arecibo, Puerto Rico in 1850 and became a doctor of medicine at the University of Barcelona, Spain in 1874. He was an important nineteenth-century Puerto Rican historian who brought the techniques of modern historical methodology to bear on the study of Puerto Rican history. His works corrected many of the errors and misconceptions found in previous histories of the island. His most famous work was Boletín histórico de Puerto Rico, which he edited into fourteen volumes between 1914 and 1927.

**Contributor Names**

Coll y Toste, Cayetano, 1850-1930.

**Created / Published**

Puerto Rico : Sucesion de J.J. Acosta, 1896.

**Subject Headings**

- Puerto Rico--History--Periodicals

**Notes**

- "Director propietario, Cayetano Coll y Toste."
- Also available in microform
- Microfilm. Año 1, núm. 1(Nov. 1896) Washington : Library of Congress Preservation Microfilming Program : Available from Library of Congress Photoduplication Service. 1966. 1 microfilm reel ; 35 mm. d18961897dcccc

**Call Number/Physical Location**

F1951

<https://www.loc.gov/item/2005700060/>



# Ejemplo

```
▼<srw_dc:dc xmlns:srw_dc="info:srw/schema/1/dc-schema" xmlns="http://purl.org/dc/elements/1.1/" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="info:srw/schema/1/dc-schema http://www.loc.gov/standards/sru/resources/dc-schema.xsd">
▼<title>
  Repertorio histórico de Puerto-Rico. [electronic resource].
</title>
<creator>Coll y Toste, Cayetano, 1850-1930.</creator>
<type>text</type>
<publisher>Puerto Rico : Sucesion de J.J. Acosta,</publisher>
<date>1896.</date>
<language>spa</language>
▼<description>
  Cayetano Coll y Toste was born in Arecibo, Puerto Rico in 1850 and became a doctor of medicine at the University of Barcelona, Spain in 1874. He
was an important nineteenth-century Puerto Rican historian who brought the techniques of modern historical methodology to bear on the study of
Puerto Rican history. His works corrected many of the errors and misconceptions found in previous histories of the island. His most famous work
was Boletín histórico de Puerto Rico, which he edited into fourteen volumes between 1914 and 1927.
</description>
▼<description>
  Cayetano Coll y Toste was born in Arecibo, Puerto Rico in 1850 and became a doctor of medicine at the University of Barcelona, Spain in 1874. He
was an important nineteenth-century Puerto Rican historian who brought the techniques of modern historical methodology to bear on the study of
Puerto Rican history. His works corrected many of the errors and misconceptions found in previous histories of the island. His most famous work
was Boletín histórico de Puerto Rico, which he edited into fourteen volumes between 1914 and 1927.
</description>
<coverage>Puerto Rico--History--Periodicals.</coverage>
<relation>Also available in microform (Microfilm 38852).</relation>
<relation>Repertorio histórico de Puerto Rico</relation>
<identifier>http://hdl.loc.gov/loc.gdc/lhbpr.08425</identifier>
</srw_dc:dc>
```

<https://www.loc.gov/item/2005700060/>



# Ejercicio de práctica

- Catalogar artículo de revista académica utilizando la plantilla xml para importar metadatos al Open Journal Systems.



# Marcalyc

- XML-JATS

<https://xmljatsredalyc.org/2016/07/28/que-es-xml-jats/>

Journal Article Tag Suite (JATS)

<https://jats.nlm.nih.gov/>



# Ejemplos de metadatos de revistas

- Journal metadata

<https://jats.nlm.nih.gov/archiving/tag-library/1.1d1/n-qcf0.html>

- Article metadata

<https://dtd.nlm.nih.gov/publishing/tag-library/2.2/n-5wc0.html>



# Referencias

Definition Metainformation. Retrieved from

<http://searchoracle.techtarget.com/definition/metainformation>

Glosario LATINDEX. Recuperado de

[http://www.latindex.org/lat/documentos/Glosario Latindex esp .pdf](http://www.latindex.org/lat/documentos/Glosario_Latindex_esp.pdf)

Méndez Rodríguez, E. (2002). *Metadatos y recuperación de información: Estándares, problemas y aplicabilidad en bibliotecas digitales*. Madrid: Ediciones Trea.

Understanding Metadata. (2004). Bethesda, MD: NISO Press.

Retrieved from

[http://www.niso.org/publications/press/UnderstandingMetadata .pdf](http://www.niso.org/publications/press/UnderstandingMetadata.pdf)



# Gracias!

*Datos de contacto:*

Joel Blanco

[joel.blanco@upr.edu](mailto:joel.blanco@upr.edu)

Snejanka Penkova

[snejanka.penkova@upr.edu](mailto:snejanka.penkova@upr.edu)







UPR RP LA IUPI