

5-21-2008

# Metadata Made Simple

Eleta Exline

*University of New Hampshire - Main Campus, [eleta.exline@unh.edu](mailto:eleta.exline@unh.edu)*

Follow this and additional works at: [https://scholars.unh.edu/library\\_pub](https://scholars.unh.edu/library_pub)

 Part of the [Cataloging and Metadata Commons](#), and the [Scholarly Communication Commons](#)

---

## Recommended Citation

Exline, Eleta, "Metadata Made Simple" (2008). *University Library Scholarship*. 60.  
[https://scholars.unh.edu/library\\_pub/60](https://scholars.unh.edu/library_pub/60)

This Presentation is brought to you for free and open access by the University Library at University of New Hampshire Scholars' Repository. It has been accepted for inclusion in University Library Scholarship by an authorized administrator of University of New Hampshire Scholars' Repository. For more information, please contact [nicole.hentz@unh.edu](mailto:nicole.hentz@unh.edu).

<record>

<title> Metadata Made Simple </title>

<creator> Eleta Exline </creator>

<date>May 21, 2008</date>

</record>



UNIVERSITY *of* NEW HAMPSHIRE

# Definitions: NISO

- **Structured** information that **describes**, explains, locates, or otherwise makes it easier to retrieve, use, or **manage** an information resource.

(NISO, 2004. *Understanding Metadata*. Bethesda, MD: NISO Press.  
[http://www.niso.org/standards/resources/  
UnderstandingMetadata.pdf](http://www.niso.org/standards/resources/UnderstandingMetadata.pdf))

# Definitions: ALA

- **Structured**, encoded data that **describe** characteristics of information-bearing entities to aid in the identification, discovery, assessment, and **management** of the described entities.

(Source: CC:DA Task Force on Metadata:

*Final Report* (CC:DA/TF/Metadata/5), June 16, 2000

<http://www.libraries.psu.edu/tas/jca/ccda/tf-meta6.html>)

# A Working Definition

- **Structured** data that **describes** some aspect of a resource so that it can be **managed**.

# Metadata Functions

- Discovery – how do I find it?
- Identification - what is it, exactly?
- Organization – how's it arranged?
- Preservation – what do I need to know to keep it viable?
- Interoperability – can it “play well” with other similar things?

# Categories

1. Descriptive - describes a resource in order to enable identification and discovery.
2. Administrative – provides information about a resource that supports its management.
3. Structural – indicates how the parts of a resource are related, organized, or sequenced.

# Parrot Metadata:



"I'm Mr. Yosuke Nakamura," the bird told the veterinarian. The parrot also provided his full home address, down to the street number.

- Associated Press, May 21, 2008.



# Candy Metadata:



# Candy Metadata:



# Categories

Descriptive – describes a resources in order to enable identification and discovery.

---

MARC

Dublin Core (DC) [Example 1](#), [Example 2](#)

Encoded Archival Description (EAD) [Example](#)

Metadata Object Description Scheme (MODS)

VRA Core

# Categories

Administrative - provides information about a resources that supports its management.

- Rights

- Preservation

---

Preservation Metadata: Implementation Strategies  
(PREMIS)

Metadata for Images in XML (MIX)

# Categories

Structural - indicates how the parts of a resource are related, organized, or sequenced.

---

Metadata Encoding Transmission Standard (METS)

[Example 1](#)

Encoded Archival Description (EAD)

# Metadata Standards

1. Scheme
2. Syntax
3. Content Values
4. Content Rules

# 1. Schemes

Specify names and definitions of elements, along with their possible attributes or qualifiers.

Indicate which elements are required or optional.

Sometimes include formatting rules and requirements or recommendations for the use of controlled vocabularies.

# 1. Schemes

From VRA Core 3.0

## CREATOR

### Qualifiers

Creator.Role

Creator.Attribution

Creator.Personal name

Creator.Corporate name

Description: The names, appellations, or other identifiers assigned to an individual, group, corporate body, or other entity that has contributed to the design, creation, production, manufacture, or alteration of the work or image.

Data Values (controlled): recommend ULAN and AAAF (LC authority files).



# 1. Schemes

MODS

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

MARC XML

<http://www.loc.gov/marcxml>

DC

<http://dublincore.org/>

TEI

<http://www.tei-c.org/>

MIX

<http://www.loc.gov/standards/mix/>

## 2. Syntax

The formal structure in which the scheme is expressed.

MARC

XML

# MARC

=100 1\\$\aFrost, Robert,\$d1874-1963.  
=245 10\$aNorth of Boston /\$cby Robert Frost.  
=250 \\\a1st ed.  
=260 \\\aLondon :\$bD. Nutt,\$c1914.  
=300 \\\a143 p. ;\$c20 cm.  
=910 \\\a4397967 lb 851212

# MARC XML

```
<marc:datafield tag="100" ind1="1" ind2=" " >  
<marc:subfield code="a">Frost, Robert,</marc:subfield>  
<marc:subfield code="d">1874-1963.</marc:subfield>  
</marc:datafield>  
<marc:datafield tag="245" ind1="1" ind2="0">  
<marc:subfield code="a">North of Boston /</  
  marc:subfield><  
marc:subfield code="c">by Robert Frost.</marc:subfield>  
</marc:datafield>  
<marc:datafield tag="250" ind1=" " ind2=" " >  
<marc:subfield code="a">1st ed.</marc:subfield>  
</marc:datafield>  
<marc:datafield tag="260" ind1=" " ind2=" " >  
<marc:subfield code="a">London :</marc:subfield>  
<marc:subfield code="b">D. Nutt,</marc:subfield>  
<marc:subfield code="c">1914.</marc:subfield>  
</marc:datafield>
```

# 3. Content Rules

Anglo-American Cataloguing Rules (AACR2)

Archival Moving Image Materials: a Cataloging Manual (AMIM)

Archives, Personal Papers, and Manuscripts : a Cataloging Manual for Archival Repositories, Historical Societies, and Manuscript Libraries (APPM)

Describing Archives: A Content Standard (DACS)

Categories for the Description of Works of ART (CDWA)

[http://www.getty.edu/research/conducting\\_research/standards/cdwa/](http://www.getty.edu/research/conducting_research/standards/cdwa/)

## 4. Content Values: Standardized Vocabularies

DCMI Type Vocabulary

<http://dublincore.org/documents/dcmi-type-vocabulary/>

[MIME] Internet Media Types

<http://www.isi.edu/in-notes/iana/assignments/media-types/media-types>

ISO 639 - Codes for the representation of names of languages.

<http://www.oasis-open.org/cover/iso639a.html>

## 4. Content Values: Thesauri & Classification Schemes

Thesaurus for Graphic Materials I: Subject Terms (TGM I)

[http://www.loc.gov/lexico/servlet/lexico?usr=pub-375:0&op=frames&db=TGM\\_I](http://www.loc.gov/lexico/servlet/lexico?usr=pub-375:0&op=frames&db=TGM_I)

Art and Architecture Thesaurus (AAT)

<http://www.getty.edu/research/tools/vocabulary/aat/index.html>

Library of Congress Subject Headings (LCSH)

<http://www.loc.gov/cds/classweb.html>

Library of Congress Classification (LCC)

<http://www.loc.gov/catdir/cpsolcco/>

## 4. Content Values: Authority lists

*The Union List of Artist Names (ULAN)*

<http://www.getty.edu/research/tools/vocabulary/ulan/>

*The Getty Thesaurus of Geographic Names (TGN)*

<http://www.getty.edu/research/tools/vocabulary/tgn/>

LC Name Authority file = LC/NACO Authority File (LC/NAF)

<http://authorities.loc.gov/>

# Crosswalks

Function as maps from one scheme to another:

	Dublin Core	EAD	MARC 21
<b>Title Element</b>	Title	<titleproper>	245 00\$a (Title Statement/Title proper)
<b>Author Element</b>	Creator	<author>	700 1#\$a (Added Entry--Personal Name) (with \$e=author) 720\$a (Added Entry--Uncontrolled Name/Name) (with \$e=author)
<b>Date Created Element</b>	Date.Created	<unitdate>	260 ##\$c (Date of publication, distribution, etc.)

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

<http://www.niso.org/standards/resources/UnderstandingMetadata.pdf>)




# Dublin Core

```
<record xmlns:xsi="http://www.w3.org/
  2001/XMLSchema-instance"
  xmlns:dc="http://purl.org/dc/elements/
  1.1/" xmlns:dcterms="http://purl.org/dc/
  terms/">
  <dc:title>North of Boston </dc:title>
  <dc:creator>Frost, Robert,1874-1963.</
  dc:creator>
  <dc:type>text</dc:type>
  <dc:publisher>London : D. Nutt,</
  dc:publisher>
  <dc:date>1914.</dc:date>
  <dc:language>eng</dc:language>
</record>
```

# MODS

```
<mods version="3.0">
  <titleInfo>
    <title>North of Boston</title>
  </titleInfo>
  <name type="personal">
    <namePart>Frost, Robert</namePart>
    <namePart type="date">1874-1963</namePart>
    <role>
      <roleTerm authority="marcrelator" type="text">creator</
      roleTerm>
    </role>
  </name>
  <typeOfResource>text</typeOfResource>
  <originInfo>
    <place>
      <placeTerm type="code" authority="marccountry">enk</
      placeTerm>
    </place>
    <place>
      <placeTerm type="text">London</placeTerm>
    </place>
    <publisher>D. Nutt</publisher>
    <dateIssued>1914</dateIssued>
    <edition>1st ed.</edition>
    <issuance>monographic</issuance>
  </originInfo>
```

# Example: California Digital Library



## CALIFORNIA DIGITAL LIBRARY

Supporting Scholarship • Building Collections & Services • Fostering Innovation & Collaboration

[Site Map](#)

### ABOUT THE CDL

- [At a Glance](#)
- [Programs & Services](#)
- [Partners & Research](#)
- [News & Publications](#)
- [Contact Us](#)

### ESPECIALLY FOR

- [Vendors & Content Providers](#)
- [UC Library Staff & CDL Partners: Inside CDL](#)


### POWERED BY CDL

- [Calisphere](#)
- [Counting California](#)
- [eScholarship Editions](#)
- [eScholarship Repository](#)
- [Mark Twain Project Online](#)
- [Melvyl Catalog](#)
- [Online Archive of California \(OAC\)](#)

## California Digital Library

The California Digital Library supports the assembly and creative use of the world's scholarship and knowledge for the University of California libraries and the communities they serve.

In addition, the CDL provides tools that support the construction of online information services for research, teaching, and learning, including services that enable the UC libraries to effectively share their materials and provide greater access to digital content.



### News

- University of California is testing new Google software that will allow greater access to digitized content available online. March 13, [2008 press release](#) [PDF]
- University of California eScholarship® Repository Exceeds 5 Million Full-Text Downloads; 20,000 Papers: [January 2008 CDLINFO](#)
- University of California launches Mark Twain Project Online: [November 2007 CDLINFO](#)

### University of California Libraries

The CDL was established as a UC library in 1997. Find out more about the UC libraries' shared collections, strategic planning initiatives, and efforts to reshape scholarly communications at the [University of California Libraries](#) website.

UNIVERSITY OF CALIFORNIA

# Example: Internet Archive



[Web](#) | [Moving Images](#) | [Texts](#) | [Audio](#) | [Software](#) | [Education](#) | [Patron Info](#) | [About IA](#)

[Home](#) | [American Libraries](#) | [Canadian Libraries](#) | [Open Source Books](#) | [Project Gutenberg](#) | [Biodiversity Heritage Library](#) | [Children's Library](#) | [Additional Collections](#)

Search:  Texts  [Advanced Search](#)

[Upload](#)

## Spotlight Item



[The complete works of Friedrich Nietzsche : the first complete and authorized English translation](#)

500. MAKING USE OF EBB AND FLOW. For the purpose of knowledge we must know how to make use of the inward current which draws us towards a thing, and also of the current which after a time draws us away from it. (Human, All Too Human)

## Welcome to Text Archive

414,080 items

Welcome to the Open-Access Text Archive. This collection is open to the community for the contribution of any type of text, many licensed using [Creative Commons licenses](#). Please feel free to [contribute your texts!](#) (Uploaders, please try to set a Creative Commons license as part of the upload process, so people know what they can do with your texts - thanks!)

[See recent additions](#) - [RSS](#)

## Sub-Collections

### [American Libraries](#)

191,599 items

Colorful Favorites Goody Two-Shoes (details | read it) Publisher's chromolithographed pictorial wrappers c1888 Caw caw; or, the chronicle of crows (details | read it) The house that Jack built, a game of forfeits : to which is added, The entertaining fable of "The Magpie" (details | read...

### [Canadian Libraries](#)

114,684 items

Historical Texts De civitate Dei (1475) (details | pdf) A native of France, Nicolas Jenson was one of the most important printers operating in Venice in the fifteenth century. Between c.1470 and 1480, Jenson produced around 150 books including the 1475 printing of St. Augustine's "De...

### [Open Source Books](#)

14,681 items

Open source books are books contributed by the community. [Click here to contribute your book!](#)

# Example: OAIster



Home Search Help • About Using OAIster News

## About OAIster

OAIster currently provides access to **16,167,543** records from **969** contributors.

OAIster is a **union catalog of digital resources**. We provide access to these digital resources by "harvesting" their descriptive metadata (records) using **OAI-PMH** (the Open Archives Initiative Protocol for Metadata Harvesting).

➤ [More...](#)

[View Data Contributors](#)

[Collection Development Policy](#)

[Improvements](#)

[Presentations & Publications](#)

[Statistics on our Growth](#)

[Staff](#)

## Using OAIster

OAIster can be searched by Title, Author/Creator, Subject, Language or Entire Record. Searches can also be limited by resource type (text, image, audio, video, dataset) and sorted by title, author, date and hit frequency. Results allow further limiting by data contributor (i.e., where the record was harvested from).

➤ [More...](#)

SEARCH

[Search Help](#)

[View Data Contributors](#)

[How-to Become a Data Contributor](#)

[Using OAIster Data Outside this Interface](#)

## News & Updates

2007 Dec 12

UM has developed an [OAI toolkit](#) for both data providers and harvesters.

➤ [More...](#)

.....  
Last update: 14 May 2008

Next update: 21 May 2008  
.....

New data contributors harvested:

ARROW@DIT

Atatürk Üniversitesi Açık Arşiv Sistemi

Centro Internacional de la Papa (CIP) Publications

DEA @ DEENK

Frattura ed Integrità Strutturale

IUScholarWorks Journals

KOPS-Datenbank OPUS

Northwestern University: World War II

Poster Collection

University of Nevada, Las Vegas Digital Collections



Product of [DLPS/DLXS](#) at the [University of Michigan](#)  
for more information contact [oaiSTER at umich dot edu](#)  
copyright © 2002-2008

# Tools

JH  VE

*JSTOR/Harvard Object Validation Environment*



# Resources:

- Metadata Basics. Library of Congress Cataloger's Learning Workshop.  
<http://www.loc.gov/catworkshop/readings/metadatabasics/index.html>
- NISO, 2004. *Understanding Metadata*. Bethesda, MD: NISO Press. <http://www.niso.org/standards/resources/UnderstandingMetadata.pdf>
- Moving Theory into Practice: Digital Imaging Tutorial. Cornell University.  
<http://www.library.cornell.edu/preservation/tutorial/contents.html>