

Cataloging for the 21st Century -- Course 2

Metadata Standards & Applications

Trainee Manual

Original course design by
Diane I. Hillmann
Cornell University Library
Revised by Rebecca Guenther and
Allene Hayes, Library of Congress

For
The Library of Congress
And the
Association for Library Collections & Technical Services

Washington, DC
August 2008

*THIS PAGE INTENTIONALLY LEFT BLANK
FOR DOUBLE SIDED COPY*

Metadata Standards and Applications

Outline

<p>1. Introduction to Digital Libraries and Metadata</p> <ul style="list-style-type: none"> • Discuss similarities and differences between traditional and digital libraries • Understand how the environment where metadata is developing is different from the library automation environment • Explore different types and functions of metadata (administrative, technical, administrative, etc.) <p><u>Exercise:</u> Examine three digital library instances, discuss differences in user approach and experience, and look for examples of metadata use</p>
<p>2. Descriptive Metadata Standards</p> <ul style="list-style-type: none"> • Understand the categories of descriptive metadata standards (e.g., data content standards, data value standards, data structure standards, relationship models) • Learn about the various descriptive metadata standards and the communities that use them • Evaluate the efficacy of a standard for a particular community • Understand how relationship models are used <p><u>Exercise:</u> Create a brief descriptive metadata record using the standard assigned.</p>
<p>3. Technical and Administrative Metadata Standards</p> <ul style="list-style-type: none"> • Understand the different types of administrative metadata • Learn about the metadata needed for supporting digital preservation activities • Understand the importance of technical, structural and rights metadata in digital libraries <p><u>Exercise:</u> Provide technical metadata for the same resource used in the descriptive exercise.</p>
<p>4. Metadata Syntaxes and Containers</p> <ul style="list-style-type: none"> • Overview of syntaxes, including HTML/XHTML, XML, RDF/XML • Overview of containers, including METS and MPEG-21 DID • Discover how container formats are used for managing digital resources and their metadata <p><u>Exercise:</u> Encode a simple resource description in Dublin Core, MARC, and MODS using XML</p>
<p>5. Applying Metadata Standards: Application Profiles</p> <ul style="list-style-type: none"> • Learn about the concept and use of application profiles, including examining specific types (e.g., Dublin Core Application Profiles, METS Profiles) • Learn how different metadata standards are used together in digital library applications <p><u>Exercise:</u> Analyze descriptive metadata application profiles</p>
<p>6. Controlled Vocabularies</p> <ul style="list-style-type: none"> • Understand how different controlled vocabularies are used in metadata • Introduction to vocabulary encodings, including MARC 21 and the Simple Knowledge Organization System (SKOS), an evolving encoding for thesauri
<p>7. Approaches to Metadata Creation, Storage and Retrieval in the Digital Environment</p> <ul style="list-style-type: none"> • Understand the differences between traditional vs. digital library metadata creation • Examine metadata creation management models • Investigate standards and methods of information retrieval and discovery <p><u>Exercise:</u> Using example websites, determine the underlying models for each</p>
<p>8. Metadata Interoperability and Quality Issues</p> <ul style="list-style-type: none"> • Understand interoperability protocols (OAI-PMH for harvesting, OpenURLs for references) • Introduction to crosswalking and mapping • Discuss the criteria that can be used to determine quality in metadata <p><u>Exercise:</u> Evaluate a small set of human and machine-created metadata OR Examine the MARC to DC crosswalk and DC to MARC crosswalk and discuss where metadata loss occurs.</p>

*THIS PAGE INTENTIONALLY LEFT BLANK
FOR DOUBLE SIDED COPY*



Metadata Standards and Applications

Introduction: Background, Goals, and Course Outline

Original course design by Diane I. Hillmann,
Cornell University Library
Revised by Rebecca Guenther and
Allene Hayes, Library of Congress

1



Version 2, August 2008



Cataloging for the 21st Century

Background for this course:

- The first of five courses developed as part of:
 - Bibliographic Control of Web Resources: A Library of Congress Action Plan
 - Action Item 5.3: Continuing Education (CE)
 - Continuing Education Implementation Group (CEIG)
- See course Bibliography for citations



Cataloging for the 21st Century: The five CE course components

- 1. Rules and Tools for Cataloging Internet resources
- 2. Metadata Standards and Applications
- 3. Principles of Controlled Vocabulary and Thesaurus Design
- 4. Metadata and Digital Library Development
- 5. Digital Project Planning and Management Basics



Cataloging for the 21st Century: CE Course Series Objectives

- To equip catalogers to deal with new types of resources and to recognize their unique characteristics
- To equip catalogers to evaluate competing approaches to and standards for providing access to resources
- To equip catalogers to think creatively and work collaboratively with others inside and outside their home institutions
- To ensure that catalogers have a broad enough understanding of the current environment to be able to make their local efforts compatible and interoperable with other efforts
- To prepare catalogers to be comfortable with ambiguity and being less than perfect
- To enable practicing catalogers to put themselves into the emerging digital information environment and to continue to play a significant role in shaping library services



Goals for this Course

- Understand similarities and differences between traditional and digital libraries
- Explore different types and functions of metadata (administrative, descriptive, technical, etc.)
- Understand metadata standards: schemas, data content standards, and data value standards
- Learn how various metadata standards are applied in digital projects, including use of application profiles
- Understand how different controlled vocabularies are used in digital libraries
- Approaches to metadata creation, storage and retrieval
- Learn about metadata interoperability and quality issues



Course objectives

- Increase catalogers' understanding of metadata for digital resources
- Evaluate competing approaches and standards for managing and providing access to resources
- Enable catalogers to think creatively and work collaboratively
- Increase understanding of current environment to allow for compatibility among applications
- Increase flexibility in utilizing different kinds of metadata standards
- Allow catalogers to use expertise to contribute to the emerging digital information environment



Outline of this course

- Session 1. Introduction to Digital Libraries and Metadata
- Session 2. Descriptive Metadata Standards
 - Data content standards, data value standards, data structure standards
 - Specific descriptive metadata formats
 - Relationship models



Outline of this course cont.

- Session 3. Technical and Administrative Metadata Standards
- Session 4. Metadata Syntaxes and Containers
- Session 5. Application Profiles and how they are used in digital libraries



Outline of this course cont.

- Session 6. Controlled Vocabularies
- Session 7. Metadata Creation, Storage and Retrieval
- Session 8. Metadata Interoperability and Quality Issues

*THIS PAGE INTENTIONALLY LEFT BLANK
FOR DOUBLE SIDED COPY*



1. Introduction to Digital Libraries and Metadata

Metadata Standards and Applications Workshop



Goals of Session

- Understand similarities and differences between traditional and digital libraries focusing on metadata
- Explore different types and functions of metadata (descriptive, administrative, structural, etc.)



Traditional vs. Digital Libraries



Traditional library characteristics

Digital library characteristics?



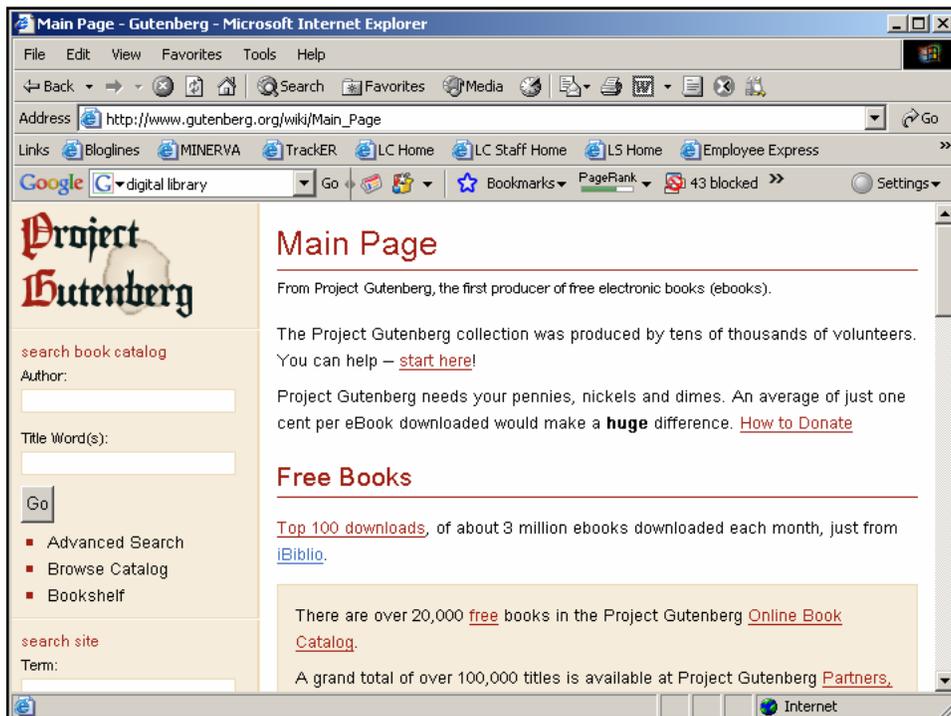
What is a digital library?

- a library in which collections are stored in digital formats and accessed by computers. The digital content may be stored locally, or accessed remotely via computer networks.
- a type of information retrieval system.



Digital Library Federation (DLF)

- “Digital libraries are organizations that provide the resources, including the specialized staff, to select, structure, offer intellectual access to, interpret, distribute, preserve the integrity of, and ensure the persistence over time of collections of digital works so that they are readily and economically available for use by a defined community or set of communities.”
- <http://www.diglib.org/>



Project Gutenberg

search book catalog

Author:

Title Word(s):

Go

- Advanced Search
- Browse Catalog
- Bookshelf

search site

Term:

Main Page

From Project Gutenberg, the first producer of free electronic books (ebooks).

The Project Gutenberg collection was produced by tens of thousands of volunteers. You can help – [start here!](#)

Project Gutenberg needs your pennies, nickels and dimes. An average of just one cent per eBook downloaded would make a **huge** difference. [How to Donate](#)

Free Books

[Top 100 downloads](#), of about 3 million ebooks downloaded each month, just from [iBiblio](#).

There are over 20,000 [free](#) books in the Project Gutenberg [Online Book Catalog](#).

A grand total of over 100,000 titles is available at Project Gutenberg [Partners](#).

The screenshot shows the homepage of the World Digital Library. At the top left is the logo for the World Digital Library, featuring a globe icon and the text "WORLD DIGITAL LIBRARY". To the right of the logo is the URL <http://www.worlddigitallibrary.org/project/english/index.html>. Below the logo is the text "World Digital Library Project Home".

On the left side, there is a navigation menu with the following items: PROJECT HOME, ABOUT THE PROJECT, NEWS, EVENTS, PARTNERS, HOW TO GET INVOLVED, and CONTACT.

The main content area features a section titled "World Digital Library" with a small icon. Below the title is a paragraph of text: "The World Digital Library will make available on the Internet, free of charge and in multilingual format, significant primary materials from cultures around the world, including manuscripts, maps, rare books, musical scores, recordings, films, prints, photographs, architectural drawings, and other significant cultural materials. The objectives of the World Digital Library are to promote international and inter-cultural understanding and awareness, provide resources to educators, expand non-English and non-Western content on the Internet, and to contribute to scholarly research." Below this text is a "Read More" button.

Below the "Read More" button is another paragraph of text: "At the UNESCO General Conference in Paris on October 17, the Library of Congress, the Bibliotheca Alexandrina, the National Library of Brazil, the National Library and Archives of Egypt, the National Library of Russia, and the Russian State Library presented a prototype of the future World Digital Library. The prototype features books, manuscripts, maps, films, prints and photographs, and sound recordings contributed by the partner institutions. It functions in Arabic, Chinese, English, French, Portuguese, Russian and Spanish, and includes content in additional languages. Other features include search and browse by place, time, topic, type of item, and contributing institution; a 'Memory of' section devoted to in-depth exploration of the culture and history of individual countries; and videos by curators that explain why particular primary source documents are important and what they tell us about a culture." Below this text is an "Experience the Prototype" button.

Below the "Experience the Prototype" button is a section titled "Experience the Vision" with a video player. The video player has a "PLAY VIDEO" button and shows a thumbnail image of a person's hands holding a book.

Below the video player is a section titled "Latest News and Events" with a list of links:

- Le prototype d'une Bibliothèque numérique mondiale présenté à l'Unesco (French) 17-10-2007 (Lemonde.fr)
- Checking Out Tomorrow's Library 17-10-2007 (Washington Post)
- U.S. Library of Congress introduces plans for world digital collection 17-10-2007 (International Herald Tribune)
- Library of Congress Advances 2 Digital Projects Abroad

The slide features the Library of Congress logo in the top left corner. The logo is circular and contains the text "THE LIBRARY OF CONGRESS" around the perimeter, with a central emblem of an eagle with wings spread, holding an olive branch and arrows.

Below the logo, there is a horizontal line, followed by the question: "How does the environment affect the creation of metadata?"



Traditional Libraries

- Firm commitment to standards
 - Specifications for metadata content (e.g., AACR2)
 - Specifications for metadata encoding (e.g., MARC)
 - A variety of syntaxes can be used
- Agreements on quality expectations
- Tradition of sharing, facilitated by bibliographic utilities
- Available documentation and training



Digital Libraries

- No dominant content standard
- A variety of “formats” (or “schemas” or “element sets”)
- Some emerging “federated” agreements, mostly in the world of digital libraries attached to traditional libraries
- Variable quality expectations
- Emerging basis for sharing (OAI-PMH)
- Some documentation and training is becoming available



Environmental Factors

■ Differences:

- *Players*: New world of metadata not necessarily led by librarians
- *Goals*: Competition for users critical for sustainability
- *Resources*: No real basis for understanding non-technical needs (including metadata creation and maintenance)
- Many levels of content responsibility (or none)



Environmental Factors

■ Similarities

- It's about discovery (and access, and use and meeting user needs)
- Pressure for fast, cheap and “good enough” (also rich, scalable, and re-usable--is that a contradiction?)
- Wide variety of materials and services
- Maintenance needs often overlooked



What *IS* Metadata?

- Some possibilities:
 - Data about data (or data about resources)
 - Structured information that describes, explains, locates, and otherwise makes it easier to retrieve and use an information resource.”
 - A management tool
 - Computer-processible, human-interpretable information about digital and non-digital objects



“In moving from dispersed digital collections to interoperable digital libraries, the most important activity we need to focus on is standards... most important is the wide variety of metadata standards [including] descriptive metadata... administrative metadata..., structural metadata, and terms and conditions metadata...”

Howard Besser, NYU



Metadata standards in digital libraries

- Interoperability and object exchange requires the use of established standards
- Many digital objects are complex and are comprised of multiple files
- XML is the de-facto standard syntax for metadata descriptions on the Internet
- Complex digital objects require many more forms of metadata than analog for their management and use
 - Descriptive
 - Administrative
 - Technical
 - Digital provenance/events
 - Rights/Terms and conditions
 - Structural



Functions of Metadata

Discover resources	Manage documents	Control IP Rights
Identify versions	Certify authenticity	Indicate status
Mark content structure	Situate geospatially	Describe processes



Types of metadata

- **Descriptive**
- **Administrative**
 - Technical
 - Digital provenance
 - Rights/Access
- Preservation
- **Structural**
- Meta-metadata
- Other?



Cataloging and Metadata

- Cataloging early form of descriptive metadata
 - Underlying models for cataloging based on AACR2 and MARC 21
 - Some new metadata models are emerging (e.g, DC Abstract Model and RDA in development)
- Most metadata models roughly based on attribute/value pairs:
 - <property> = <value>



Some differences between traditional and digital libraries

- Metadata only vs. actual object
- Need to understand Web technologies
- Types of media
- Granularity
- User needs
- Web services
- Digitized vs. born digital

Slide by Brian Surratt



One BIG Difference ...

- Catalogers most often are attempting to fit new items into an already existing world of materials--
 - The structure already exists, as do the rules for describing
- Metadata practitioners are generally working with aggregated “stuff,” attempting to find a way to make it accessible
 - Involves broad understanding, ability to work with others to make decisions that work for whole projects or domains

**Thanks to Marty Kurth for these insights*



Questions to ask when selecting metadata standards

- What type of material will be digitized?
- How rich does the metadata need to be?
- Is there information already available?
- Is there a Community of practice developed for this resource type(s)?
- What is the purpose of digital project?
- Who will be the audience and how they would use the content?
- Are there pre-existing digital projects with which this one needs to function? Is there a need to interact with any existing records?
- What tools or systems options are available?



Exercise

- Examine the digital library sites below, and be prepared to discuss differences in user approach and experience. Look for how metadata is used.
 - Alsos: Digital Library for Nuclear Issues (<http://alsos.wlu.edu/default.aspx>)
 - CSUN Oviatt Library: Digital Collections (<http://library.csun.edu/Collections/SCA/digicoll.html>)
 - Birdsource (<http://www.birdsource.org/>)

*THIS PAGE INTENTIONALLY LEFT BLANK
FOR DOUBLE SIDED COPY*



2. Specific metadata standards: descriptive

Metadata Standards and Applications Workshop



Session 2 Objectives

- Understand the categories of descriptive metadata standards (e.g., data content standards vs. data value standards)
- Learn about the various descriptive metadata standards and the community that developed and use them
- Learn about some relationship models used in descriptive metadata standards



Outline of Session 2: descriptive metadata

- Types of descriptive metadata standards (e.g. element sets, content standards)
- Specific descriptive metadata standards (e.g. MARC, DC, MODS, EAD...)
- Relationship models



Descriptive metadata

- Most standardized and well understood type of metadata
- Major focus of library catalog
- Increased number of descriptive metadata standards for different needs and communities
- Importance for resource discovery
- May support various user tasks



Aspects of descriptive metadata

- Data content standards (e.g., rules: AACR2R/RDA, CCO)
- Data value standards (e.g., values/controlled vocabularies: LCNAF, LCSH, MeSH, AAT)
- Data structure standards (e.g., formats/schemes: DC, MODS, MARC 21, VRA Core)
- Relationship models
- Data exchange/syntax standards (e.g. MARC 21 (ISO 2709), MARCXML, DC/RDF or DC/XML)



Content Standards: Rules

- **AACR2** functions as the content standard for traditional cataloging
- **RDA** (*Resource Description and Access*) is the successor to AACR2 that aspires to be independent of a particular syntax
- **DACS** (*Describing Archives: a Content Standard*)
- **CCO** (*Cataloging Cultural Objects*) new standard developed by visual arts and cultural heritage community
- **CSDGM** (*Content Standards for Digital Geospatial Metadata*)
- Best practices, Guidelines, policies-- less formal content standards



Content Standards: Value Standards/Controlled Vocabularies

- Examples of thesauri
 - Library of Congress Subject Headings
 - Art and Architecture Thesaurus
 - Thesaurus of Geographical Names
- Examples of value lists
 - ISO 639-2 Language codes
 - MARC Geographic Area codes
 - Other enumerated lists (e.g. MARC/008 lists)
 - Dublin Core Resource Types



Data structure standards (element sets and formats)

- Facilitates database creation and record retrieval
- Flexibility because not tied to a particular syntax
- May provide a minimum of agreed upon elements that facilitate record sharing and minimal consistency
- Different user communities develop their own standard data element sets
- May differ in complexity and granularity of fields
- Some data element sets become formats/schemes by adding rules such as repeatability, controlled vocabularies used, etc.



Key descriptive metadata element sets/formats

- MARC 21
- MODS
- Dublin Core
- ONIX
- EAD
- Special resource types
 - CDWA/VRA Core
 - IEEE-LOM
 - PBCore
 - IPTC Core



What is MARC 21?

- A syntax defined by an international standard and was developed in the late 60s
- As a syntax it has 2 expressions:
 - Classic MARC (MARC 2709)
 - MARCXML
- A data element set defined by content designation and semantics
- Institutions do not store “MARC 21”, as it is a communications format
- Many data elements are defined by external content rules; a common misperception is that it is tied to AACR2

02158cam 2200349la
 450000100130000000300060001300500170001900600190003600600190005500700150007400800
 41000890400020001300200015001500430021001650490009001862450119001952460025
 003142600065003395380030004045060038004345360153004725200764006255050094013895000
 086014836000049015696500040016186510039016586510023016977000026017208560050017469
 94001201796 ocm56835268 OCoLC 20060118051017.0 m d szx w s 0 2 cr mn---
 ----- 041028m20049999vau st 000 0 eng d aVA@ cVA@ dOCLCQ a0813922917
 an-us--- an-us-va aVA@@ 04 aThe Dolley Madison digital edition h[electronic resource]
 : bletters 1788-June 1836 / cedited by Holly C. Shulman. 1 iAlso known as: aDMDE
 aCharlottesville, Va. : bUniversity of Virginia Press, c2004- aMode of access: Internet.
 aSubscription required for access. aRotunda editions are made possible by generous grants
 from the Andrew W. Mellon Foundation and the President's Office of the University of Virginia.
 aDolley Payne Madison was the most important First Lady of the nineteenth century. The DMDE will
 be the first-ever complete edition of all of her known correspondence, gathered in an XML-based
 archive. It will ultimately include close to 2,500 letters. From the scattered correspondence were
 gathered letters that have never been previously published. The range and scope of the collection
 makes this edition an important scholarly contribution to the literature of the early republic, women's
 history, and the institution of the First Lady. These letters present Dolley Madison's trials and triumphs
 and make it possible to gain admittance to her mind and her private emotions and to understand the
 importance of her role as the national capital's First Lady. 0 aGeneral introduction -- Biographical
 introduction -- Introduction to the digital edition. aTitle from the opening screen; description based
 on the display of Oct. 21, 2004. 10 aMadison, Dolley, d1768-1849 vCorrespondence.
 0 aPresidents' spouses zUnited States. 0 aUnited States xHistory y1801-1809.
 0 aVirginia xHistory. 1 aShulman, Holly
 Cowan. 40 uhttp://rotunda.upress.virginia.edu:8100/dmde/ aC0 bVA@



MARC 21 Scope

- **Bibliographic Data**
 - books, serials, computer files, maps, music, visual materials, mixed material
- **Authority Data**
 - names, titles, name/title combinations, subjects, series
- **Holdings Data**
 - physical holdings, digital access, location
- **Classification Data**
 - classification numbers, associated captions, hierarchies
- **Community Information**
 - events, programs, services, people, organizations



MARC 21 implementation

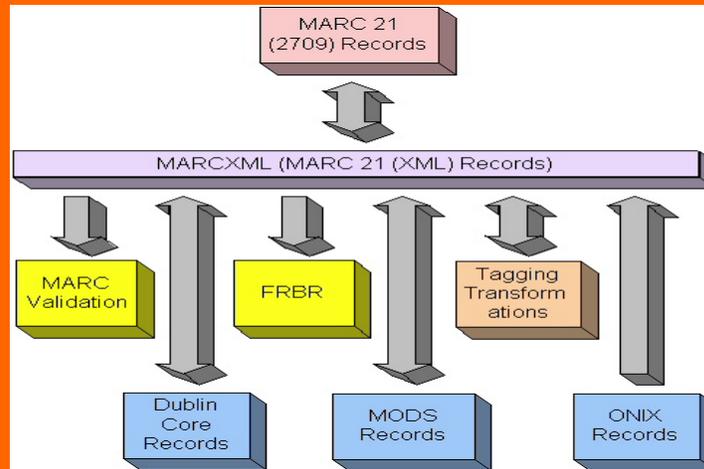
- National formats were once common and there were different flavors of MARC
- Now most have harmonized with MARC 21 (e.g. CANMARC, UKMARC, MAB)
- Billions of records world wide
- Integrated library systems that support MARC bibliographic, authority and holdings format
- Wide sharing of records for 30+ years
- OCLC is a major source of MARC records



Streamlining MARC 21 into the future

- Take advantage of XML
 - Establish standard MARC 21 in an XML structure
 - Take advantage of freely available XML tools
- Develop simpler (but compatible) alternatives
 - MODS
- Allow for interoperability with different XML metadata schemas
 - Assemble coordinated set of tools
- Provide continuity with current data
 - Provide flexible transition options

MARC 21 evolution to XML



MARC 21 in XML – MARCXML

- MARCXML record
 - XML exact equivalent of MARC (2709) record
 - Lossless/roundtrip conversion to/from MARC 21 record
 - Simple flexible XML schema, no need to change when MARC 21 changes
 - Presentations using XML stylesheets
 - LC provides converters (open source)
- <http://www.loc.gov/standards/marcxml>



Example: MARC and MARCXML

- [Music record in MARC](#)
- [Music record in MARCXML](#)

[The Library of Congress](#)[>> Go to Library of Congress Authorities](#)

LIBRARY OF CONGRESS ONLINE CATALOG



Help	New Search	Search History	Headings List	Titles List	Request an Item	Account Status	Start Over
----------------------	----------------------------	--------------------------------	-------------------------------	-----------------------------	---------------------------------	--------------------------------	----------------------------

DATABASE: Library of Congress Online Catalog

YOU SEARCHED: Command = 010a 85753651

SEARCH RESULTS: Displaying 1 of 1.

[◀ Previous](#) [Next ▶](#)[Brief Record](#)[Subjects/Content](#)[Full Record](#)[MARC Tags](#)*3 Viennese arias : for soprano, obligato clarinet in B flat, and piano /...*

LC Control No.: 85753651

LCCN Permalink: <http://lcn.loc.gov/85753651>

000 01917ccm a2200409 a 450

001 5594130

005 19950601141653.9

008 850813s1984 enkopa z n ita

035 __ |9 (DLC) 85753651

906 __ |a 7 |b cbc |c orignew |d 3 |e ncip |f 19 |g y-genmusic

010 __ |a 85753651

028 32 |a N.M. 275 |b Nova Music

040 __ |a DLC |c DLC |d DLC

041 1_ |a ita |e itaeng |h ita

048 __ |b va01 |a wc01 |a ka01

050 00 |a M1506 |b .A14 1984

245 00 |a 3 Viennese arias : |b for soprano, obligato clarinet in B flat, and piano / |c G.B. Bononcini and Emperor Joseph I ; edited by Colin Lawson.

260 __ |a London : |b Nova Music, |c c1984.

300 __ |a 1 score (12 p.) + 2 parts ; |c 31 cm.

440 _0 |a Music for voice and instrument

500 __ |a Opera excerpts.

500 __ |a Acc. arr. for piano; obligato for the 2nd-3rd excerpts originally for chalumeau.

500 __ |a Italian words.

500 __ |a Cover title.

500 __ |a The 1st excerpt composed for inclusion in M.A. Ziani's Chilonida.

500 __ |a Texts with English translations on cover p. [2].

505 0_ |a Tutto in pianto il cor struggete / Emperor Joseph I -- E sempre inquieto quel

core infelice : from Endimione / G. Bononcini -- L'adorata genitrice : from Muzio [i.e. Mutio] Scevola / G. Bononcini.

650 _0 |a Operas |x Excerpts, Arranged |x Scores and parts.

650 _0 |a Songs (High voice) with instrumental ensemble |x Scores and parts.

700 1_ |a Lawson, Colin |q (Colin James)

700 02 |a Joseph |b I, |c Holy Roman Emperor, |d 1678-1711. |t Tutto in pianto il cor struggete; |o arr. |f 1984.

700 12 |a Bononcini, Giovanni, |d 1670-1747. |t Endimione. |p E sempre inquieto quel core infelice; |o arr. |f 1984.

700 12 |a Bononcini, Giovanni, |d 1670-1747. |t Mutio Scevola. |p Adorata genitrice; |o arr. |f 1984.

740 0_ |a Three Viennese arias.

740 0_ |a Viennese arias.

953 __ |a TA28

991 __ |b c-Music |h M1506 |i .A14 1984 |t Copy 1 |w MUSIC

CALL NUMBER: [M1506 .A14 1984](#)

Copy 1

-- **Request in:** Performing Arts Reading Room (Madison, LM113)

-- **Status:** Not Charged

◀ Previous Next ▶

Save, Print or Email Records (View Help)	
Select Download Format:	Text (Brief Information) <input type="button" value="Press to SAVE or PRINT"/>
Email Text (Full Info) to:	<input type="text"/> <input type="button" value="Press to SEND EMAIL"/>

[Help](#) - [Search](#) - [History](#) - [Headings](#) - [Titles](#) - [Request](#) - [Account](#) - [Exit](#)



The Library of Congress

URL: <http://www.loc.gov/>
Mailing Address:
101 Independence

Ave, S.E.
Washington, DC 20540

Catalog/authority record errors?
[Use our Error Report Form](#)
Questions about searching?
[Ask a Librarian](#)

Library of Congress Online Catalog
URL: <http://catalog.loc.gov/>
Library of Congress Authorities
URL: <http://authorities.loc.gov/>

```

<?xml version="1.0" ?>
- <zs:searchRetrieveResponse xmlns:zs="http://www.loc.gov/zing/srw/">
  <zs:version>1.1</zs:version>
  <zs:numberOfRecords>1</zs:numberOfRecords>
- <zs:records>
  - <zs:record>
    <zs:recordSchema>info:srw/schema/1/marcxml-v1.1</zs:recordSchema>
    <zs:recordPacking>xml</zs:recordPacking>
  - <zs:recordData>
    - <record xmlns="http://www.loc.gov/MARC21/slim">
      <leader>01917ccm 2200409 a 4500</leader>
      <controlfield tag="001">5594130</controlfield>
      <controlfield tag="005">19950601141653.9</controlfield>
      <controlfield tag="008">850813s1984 enkopa z n ita</controlfield>
    - <datafield tag="035" ind1="" ind2="">
      <subfield code="9">(DLC) 85753651</subfield>
    </datafield>
    - <datafield tag="906" ind1="" ind2="">
      <subfield code="a">7</subfield>
      <subfield code="b">cbc</subfield>
      <subfield code="c">orignew</subfield>
      <subfield code="d">3</subfield>
      <subfield code="e">ncip</subfield>
      <subfield code="f">19</subfield>
      <subfield code="g">y-genmusic</subfield>
    </datafield>
    - <datafield tag="010" ind1="" ind2="">
      <subfield code="a">85753651</subfield>
    </datafield>
    - <datafield tag="028" ind1="3" ind2="2">
      <subfield code="a">N.M. 275</subfield>
      <subfield code="b">Nova Music</subfield>
    </datafield>
    - <datafield tag="040" ind1="" ind2="">
      <subfield code="a">DLC</subfield>
      <subfield code="c">DLC</subfield>
      <subfield code="d">DLC</subfield>
    </datafield>
    - <datafield tag="041" ind1="1" ind2="">
      <subfield code="a">ita</subfield>
      <subfield code="e">itaeng</subfield>
      <subfield code="h">ita</subfield>
    </datafield>
    - <datafield tag="048" ind1="" ind2="">
      <subfield code="b">va01</subfield>
      <subfield code="a">wc01</subfield>
      <subfield code="a">ka01</subfield>
    </datafield>
    - <datafield tag="050" ind1="0" ind2="0">
      <subfield code="a">M1506</subfield>
      <subfield code="b">.A14 1984</subfield>
    </datafield>
  </record>
  </zs:recordData>
  </zs:record>
</zs:records>
</zs:searchRetrieveResponse>

```

- <datafield tag="245" ind1="0" ind2="0">
 <subfield code="a">3 Viennese arias :</subfield>
 <subfield code="b">for soprano, obbligato clarinet in B flat,
 and piano /</subfield>
 <subfield code="c">G.B. Bononcini and Emperor Joseph I ;
 edited by Colin Lawson.</subfield>
 </datafield>
- <datafield tag="260" ind1="" ind2="">
 <subfield code="a">London :</subfield>
 <subfield code="b">Nova Music,</subfield>
 <subfield code="c">c1984.</subfield>
 </datafield>
- <datafield tag="300" ind1="" ind2="">
 <subfield code="a">1 score (12 p.) + 2 parts ;</subfield>
 <subfield code="c">31 cm.</subfield>
 </datafield>
- <datafield tag="440" ind1="" ind2="0">
 <subfield code="a">Music for voice and instrument</subfield>
 </datafield>
- <datafield tag="500" ind1="" ind2="">
 <subfield code="a">Opera excerpts.</subfield>
 </datafield>
- <datafield tag="500" ind1="" ind2="">
 <subfield code="a">Acc. arr. for piano; obbligato for the 2nd-
 3rd excerpts originally for chalumeau.</subfield>
 </datafield>
- <datafield tag="500" ind1="" ind2="">
 <subfield code="a">Italian words.</subfield>
 </datafield>
- <datafield tag="500" ind1="" ind2="">
 <subfield code="a">Cover title.</subfield>
 </datafield>
- <datafield tag="500" ind1="" ind2="">
 <subfield code="a">The 1st excerpt composed for inclusion in
 M.A. Ziani's Chilonida.</subfield>
 </datafield>
- <datafield tag="500" ind1="" ind2="">
 <subfield code="a">Texts with English translations on cover p.
 [2].</subfield>
 </datafield>
- <datafield tag="505" ind1="0" ind2="">
 <subfield code="a">Tutto in pianto il cor struggete / Emperor
 Joseph I -- E sempre inquieto quel core infelice : from
 Endimione / G. Bononcini -- L'adorata genitrice : from
 Muzio [i.e. Mutio] Scevola / G. Bononcini.</subfield>
 </datafield>
- <datafield tag="650" ind1="" ind2="0">
 <subfield code="a">Operas</subfield>
 <subfield code="x">Excerpts, Arranged</subfield>
 <subfield code="x">Scores and parts.</subfield>
 </datafield>
- <datafield tag="650" ind1="" ind2="0">

```

    <subfield code="a">Songs (High voice) with instrumental
      ensemble</subfield>
    <subfield code="x">Scores and parts.</subfield>
  </datafield>
- <datafield tag="700" ind1="1" ind2="">
  <subfield code="a">Lawson, Colin</subfield>
  <subfield code="q">(Colin James)</subfield>
</datafield>
- <datafield tag="700" ind1="0" ind2="2">
  <subfield code="a">Joseph</subfield>
  <subfield code="b">I,</subfield>
  <subfield code="c">Holy Roman Emperor,</subfield>
  <subfield code="d">1678-1711.</subfield>
  <subfield code="t">Tutto in pianto il cor struggete;</subfield>
  <subfield code="o">arr.</subfield>
  <subfield code="f">1984.</subfield>
</datafield>
- <datafield tag="700" ind1="1" ind2="2">
  <subfield code="a">Bononcini, Giovanni,</subfield>
  <subfield code="d">1670-1747.</subfield>
  <subfield code="t">Endimione.</subfield>
  <subfield code="p">E sempre inquieto quel core
    infelice;</subfield>
  <subfield code="o">arr.</subfield>
  <subfield code="f">1984.</subfield>
</datafield>
- <datafield tag="700" ind1="1" ind2="2">
  <subfield code="a">Bononcini, Giovanni,</subfield>
  <subfield code="d">1670-1747.</subfield>
  <subfield code="t">Mutio Scevola.</subfield>
  <subfield code="p">Adorata genitrice;</subfield>
  <subfield code="o">arr.</subfield>
  <subfield code="f">1984.</subfield>
</datafield>
- <datafield tag="740" ind1="0" ind2="">
  <subfield code="a">Three Viennese arias.</subfield>
</datafield>
- <datafield tag="740" ind1="0" ind2="">
  <subfield code="a">Viennese arias.</subfield>
</datafield>
- <datafield tag="953" ind1="" ind2="">
  <subfield code="a">TA28</subfield>
</datafield>
- <datafield tag="991" ind1="" ind2="">
  <subfield code="b">c-Music</subfield>
  <subfield code="h">M1506</subfield>
  <subfield code="i">.A14 1984</subfield>
  <subfield code="t">Copy 1</subfield>
  <subfield code="w">MUSIC</subfield>
</datafield>
</record>
</zs:recordData>

```

```
<zs:recordPosition>1</zs:recordPosition>  
</zs:record>  
</zs:records>  
</zs:searchRetrieveResponse>
```



What is MODS?

- Metadata Object Description Schema
- An XML descriptive metadata standard
- A derivative of MARC
 - Uses language based tags
 - Contains a subset of MARC data elements
 - Repackages elements to eliminate redundancies
- MODS does not assume the use of any specific rules for description
- Element set is particularly applicable to digital resources



MODS high-level elements

- | | |
|------------------------|---------------------|
| ■ Title Info | ■ Note |
| ■ Name | ■ Subject |
| ■ Type of resource | ■ Classification |
| ■ Genre | ■ Related item |
| ■ Origin Info | ■ Identifier |
| ■ Language | ■ Location |
| ■ Physical description | ■ Access conditions |
| ■ Abstract | ■ Part |
| ■ Table of contents | ■ Extension |
| ■ Target audience | ■ Record Info |



Advantages of MODS

- Element set is compatible with existing descriptions in large library databases
- Element set is richer than Dublin Core but simpler than full MARC
- Language tags are more user-friendly than MARC numeric tags
- Hierarchy allows for rich description, especially of complex digital objects
- Rich description that works well with hierarchical METS objects



Uses of MODS

- Extension schema to METS
 - Rich description works well with hierarchical METS objects
- To represent metadata for harvesting (OAI)
 - Language based tags are more user friendly
- As a specified XML format for SRU
- As a core element set for convergence between MARC and non-MARC XML descriptions
- For original resource description in XML syntax that is simpler than full MARC



Example: MODS

- Music record in MODS

```

<?xml version="1.0" ?>
- <zs:searchRetrieveResponse xmlns:zs="http://www.loc.gov/zing/srw/">
  <zs:version>1.1</zs:version>
  <zs:numberOfRecords>1</zs:numberOfRecords>
- <zs:records>
  - <zs:record>
    <zs:recordSchema>info:srw/schema/1/mods-v3.0</zs:recordSchema>
    <zs:recordPacking>xml</zs:recordPacking>
  - <zs:recordData>
    - <mods xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
      xmlns="http://www.loc.gov/mods/v3" version="3.0"
      xsi:schemaLocation="http://www.loc.gov/mods/v3
      http://www.loc.gov/standards/mods/v3/mods-3-0.xsd">
      - <titleInfo>
        <title>3 Viennese arias</title>
        <subTitle>for soprano, obbligato clarinet in B flat, and
          piano</subTitle>
        </titleInfo>
      - <name type="personal">
        <namePart>Lawson, Colin (Colin James)</namePart>
        </name>
        <typeOfResource>notated music</typeOfResource>
      - <originInfo>
        - <place>
          <placeTerm type="code"
            authority="marccountry">enk</placeTerm>
          </place>
        - <place>
          <placeTerm type="text">London</placeTerm>
          </place>
          <publisher>Nova Music</publisher>
          <dateIssued>c1984</dateIssued>
          <dateIssued encoding="marc">1984</dateIssued>
          <issuance>monographic</issuance>
        </originInfo>
      - <language>
        <languageTerm authority="iso639-2b"
          type="code">ita</languageTerm>
        </language>
      - <language>
        <languageTerm authority="iso639-2b"
          type="code">eng</languageTerm>
        </language>
      - <physicalDescription>
        <form authority="marcform">print</form>
        <extent>1 score (12 p.) + 2 parts ; 31 cm.</extent>
        </physicalDescription>
        <tableOfContents>Tutto in pianto il cor struggete / Emperor
          Joseph I -- E sempre inquieto quel core infelice : from
          Endimione / G. Bononcini -- L'adorata genitrice : from Muzio
          [i.e. Mutio] Scevola / G. Bononcini.</tableOfContents>
        <note type="statement of responsibility">G.B. Bononcini and
          Emperor Joseph I ; edited by Colin Lawson.</note>
    </mods>
  </zs:recordData>
  </zs:record>
</zs:records>
</zs:searchRetrieveResponse>

```

```

<note>Opera excerpts.</note>
<note>Acc. arr. for piano; obbligato for the 2nd-3rd excerpts
  originally for chalumeau.</note>
<note>Italian words.</note>
<note>Cover title.</note>
<note>The 1st excerpt composed for inclusion in M.A. Ziani's
  Chilonida.</note>
<note>Texts with English translations on cover p. [2].</note>
- <subject authority="Icsh">
  <topic>Operas</topic>
  <topic>Excerpts, Arranged</topic>
  <topic>Scores and parts</topic>
</subject>
- <subject authority="Icsh">
  <topic>Songs (High voice) with instrumental
    ensemble</topic>
  <topic>Scores and parts</topic>
</subject>
<classification authority="Icc">M1506 .A14 1984</classification>
- <relatedItem type="series">
  - <titleInfo>
    <title>Music for voice and instrument</title>
  </titleInfo>
</relatedItem>
- <relatedItem type="constituent">
  - <titleInfo>
    <title>Tutto in pianto il cor struggete; arr. 1984</title>
  </titleInfo>
  - <name type="personal">
    <namePart>Joseph</namePart>
    <namePart type="termsOfAddress">I, Holy Roman
      Emperor</namePart>
    <namePart type="date">1678-1711</namePart>
  </name>
</relatedItem>
- <relatedItem type="constituent">
  - <titleInfo>
    <title>Endimione. arr. 1984</title>
    <partName>E sempre inquieto quel core infelice; arr.
      1984</partName>
  </titleInfo>
  - <name type="personal">
    <namePart>Bononcini, Giovanni,</namePart>
    <namePart type="date">1670-1747</namePart>
  </name>
</relatedItem>
- <relatedItem type="constituent">
  - <titleInfo>
    <title>Mutio Scevola. arr. 1984</title>
    <partName>Adorata genitrice; arr. 1984</partName>
  </titleInfo>
  - <name type="personal">

```

```
      <namePart>Bononcini, Giovanni,</namePart>
      <namePart type="date">1670-1747</namePart>
    </name>
  </relatedItem>
  <identifier type="lccn">85753651</identifier>
  <identifier type="music publisher">N.M. 275 Nova
    Music</identifier>
- <recordInfo>
  <recordContentSource
    authority="marcorg">DLC</recordContentSource>
  <recordCreationDate
    encoding="marc">850813</recordCreationDate>
  <recordChangeDate
    encoding="iso8601">19950601141653.9</recordChangeDate>
  <recordIdentifier>5594130</recordIdentifier>
</recordInfo>
</mods>
</zs:recordData>
<zs:recordPosition>1</zs:recordPosition>
</zs:record>
</zs:records>
</zs:searchRetrieveResponse>
```



Status of MODS



- Open listserv collaboration of possible implementers, LC coordinated (1st half 2002)
- First comment and use period: 2nd half 2002
- Now in MODS version 3.3
- Companion for authority metadata (MADS) in version 1.0
- Endorsed as METS extension schema for descMD
- Many expose records as MODS in OAI
- MODS Editorial Committee being formed



A selection of MODS projects

- LC uses of MODS
 - LC [web archives](#)
 - Digital library METS projects
- University of Chicago Library
 - [Chopin early editions](#)
 - Finding aid discovery
- Digital Library Federation Aquifer initiative
- National Library of Australia
 - MusicAustralia: MODS as exchange format between National Library of Australia and ScreenSoundAustralia
 - Australian national bibliographic database metadata project
- See: [MODS Implementation registry](http://www.loc.gov/mods/registry.php)
<http://www.loc.gov/mods/registry.php>



What is MADS?

- Metadata Authority Description Schema
- A companion to MODS for authority data using XML
- Defines a subset of MARC authority elements using language-based tags
- Elements have same definitions as equivalent MODS
- Metadata about people, organizations, events, subjects, time periods, genres, geographics, occupations



MADS elements

- | | |
|---|---|
| <ul style="list-style-type: none"> ■ authority <ul style="list-style-type: none"> ■ name ■ titleInfo ■ topic ■ temporal ■ genre ■ geographic ■ hierarchicalGeographic ■ occupation ■ related <ul style="list-style-type: none"> ■ same subelements ■ variant <ul style="list-style-type: none"> ■ same subelements | <ul style="list-style-type: none"> ■ note ■ affiliation ■ url ■ identifier ■ fieldOfActivity ■ extension ■ recordInfo |
|---|---|



Uses of MADS

- As an XML format for information about people, organizations, titles, events, places, concepts
 - To expose library metadata in authority files
 - To allow for linking to an authoritative form and fuller description of the entity from a MODS record
- For a simpler authority record than full MARC 21 authorities
- To integrate bibliographic/authority information for presentation

Example: MADS Name Record

```
<mads xsi:schemaLocation="http://www.loc.gov/mads/
http://www.loc.gov/mads/mads.xsd">
  <authority>
    <name type="personal">
      <namePart>Smith,John</namePart>
      <namePart type="date">1995-</namePart>
    </name>
  </authority>
  <variant type="other">
    <name>
      <namePart>Smith, J</namePart>
    </name>
  </variant>
  <variant type="other">
    <name>
      <namePart>Smith, John J</namePart>
    </name>
  </variant>
  <note type="history">Biographical note about John Smith.</note>
  <affiliation>
    <organization>Lawrence Livermore Laboratory</organization>
    <dateValid>1987</dateValid>
  </affiliation>
</mads>
```

Example: MADS Organization Record

```
<mads xsi:schemaLocation="http://www.loc.gov/mads/
http://www.loc.gov/mads/mads.xsd">
  <authority>
    <name type="corporate">
      <namePart>Unesco</namePart>
    </name>
  </authority>
  <related type="parentOrg">
    <name>
      <namePart>United Nations</namePart>
    </name>
  </related>
  <variant type="expansion">
    <name>
      <namePart>United Nations Educational, Cultural, and Scientific
      Organization</namePart>
    </name>
  </variant>
</mads>
```



Some MADS implementations

- [Irish Virtual Research Library and Archive Repository Prototype](#)
- Perseus Digital Library (Tufts)
- Mark Twain Papers (University of California)
- Library of Congress/National Library of Egypt



Dublin Core: Simple

- Fifteen elements; one namespace
- Controlled vocabulary values may be expressed, but not the sources of the values
- Minimal standard for OAI-PMH
- Used also as:
 - core element set in some other schemas
 - switching vocabulary for more complex schemas



Dublin Core Metadata Element Set (DCMES) 1996

The 15 Dublin Core elements can be divided into three categories:

CONTENT	INTELLECTUAL PROPERTY	INSTANTIATION
Title	Creator	Date
Description	Contributor	Language
Subject	Publisher	Identifier
Relation	Rights	Format
Source		
Coverage		
Type		



Advantages

- International and cross-domain
- Increase efficiency of the discovery/retrieval of digital objects
- Provide a framework of elements which will aid the management of information
- Promote collaboration of cultural/educational information



Dublin Core Characteristics

- Simplicity
- Supports resource discovery
- All elements are optional/repeatable
- No order of elements prescribed
- Extensible
- Interdisciplinary/International
- Semantic interoperability



Dublin Core metadata

- Original 15 included in DC simple
- Elements defined subsequently and all refinements/encoding schemes under dcterms (qualified)
- DCMI Type values for high level resource types
- Simple DC widely implemented and required for metadata harvesting using OAIPMH (until current version)
- Application profiles developing to document usage and additional elements needed
- <http://dublincore.org>



Dublin Core: Qualified

- ‘Qualified’ includes element refinements and encoding schemes
 - More specific properties
 - Two namespaces
 - Explicit vocabularies
- Additional elements, including ‘Audience,’ ‘InstructionalMethod,’ ‘RightsHolder’ and ‘Provenance’

Ex.: Simple Dublin Core

```

<metadata>
  <dc:title>3 Viennese arias: for soprano, obbligato clarinet in B flat, and
  piano.</dc:title>
  <dc:contributor>Lawson, Colin (Colin James)</dc:contributor>
  <dc:contributor>Bononcini, Giovanni, 1670-1747.</dc:contributor>
  <dc:contributor>Joseph I, Holy Roman Emperor, 1678-1711.</dc:contributor>
  <dc:subject>Operas--Excerpts, Arranged--Scores and parts</dc:subject>
  <dc:subject>Songs (High voice) with instrumental ensemble--Scores and
  parts</dc:subject>
  <dc:subject>M1506 .A14 1984</dc:subject>
  <dc:subject></dc:subject>
  <dc:subject></dc:subject>
  <dc:date>1984</dc:date>
  <dc:format>1 score (12 p.) + 2 parts ; 31 cm.</dc:format>
  <dc:type>text</dc:type>
  <dc:identifier>85753651</dc:identifier>
  <dc:language>it</dc:language>
  <dc:language>en</dc:language>
  <dc:publisher>Nova Music</dc:publisher></metadata>

```

Ex.: Qualified Dublin Core

```

<metadata>
  <dc:title xml:lang="en">3 Viennese arias: for soprano, obbligato clarinet in B
  flat, and piano.</dc:title>
  <dc:contributor>Lawson, Colin (Colin James)</dc:contributor>
  <dc:contributor>Bononcini, Giovanni, 1670-1747.</dc:contributor>
  <dc:contributor>Joseph I, Holy Roman Emperor, 1678-1711.</dc:contributor>
  <dc:subject xsitype="LCSH">Operas--Excerpts, Arranged--Scores and
  parts</dc:subject>
  <dc:subject xsitype="LCSH">Songs (High voice) with instrumental
  ensemble--Scores and parts</dc:subject>
  <dc:subject xsitype="LCC">M1506 .A14 1984</dc:subject>
  <dc:date xsitype="W3CDTF">1984</dc:date>
  <dcterms:extent>1 score (12 p.) + 2 parts ; 31 cm.</dcterms:extent>
  <dc:type xsitype="DCMIType">Sound</dc:type>
  <dc:identifier>85753651</dc:identifier>
  <dc:language xsitype="RFC3066">it</dc:language>
  <dc:language xsitype="RFC3066">en</dc:language>
  <dc:publisher>Nova Music</dc:publisher>
</metadata>

```



Encoded Archival Description (EAD)

- Standard for electronic encoding of finding aids for archival and manuscript collections
- Expressed as an SGML/XML DTD
- Supports archival descriptive practices and standards
- Supports discovery, exchange and use of data
- Developed and maintained by Society of American Archivists; LC hosts the website



EAD, continued

- Based on the needs of the archival community
- Good at describing blocks of information, poor at providing granular information
- Some uptake by museum community
- Not a content standard
- EAC is a companion for information about creators of archival material
- Example:
<http://purl.dlib.indiana.edu/iudl/findingaids/lilly/InU-Li-VAA1292>



Benefits of an EAD finding aid

- Documents the interrelated descriptive information of an archival finding aid
- Preserves the hierarchical relationships existing between levels of description
- Represents descriptive information that is inherited by one hierarchical level from another
- Supports element-specific indexing and retrieval of descriptive information



Text Encoding Initiative (TEI)

- Consortium of institutions and research projects which collectively maintains and develops guidelines for the representation of texts in digital form.
- Includes representation of title pages, chapter breaks, tables of contents, as well as poetry, plays, charts, etc.
- The TEI file contains a “header” that holds metadata about the digital file & about the original source.

TEI

```

<fileDesc>
  <titleStmt>
    <title type="main">A chronicle of the conquest of
      Granada</title>
    <author>
      <name type="last">Irving</name>
      <name type="first">Washington</name>
      <dateRange from="1783" to="1859">1783-
        1859</dateRange>
    </author>
  </titleStmt>
  <extent>455 kilobytes</extent>
  <publicationStmt>
    <publisher>University of Virginia Library</publisher>
    <pubPlace>Charlottesville, Virginia</pubPlace>
    <date value="2006">2006</date>
  </publicationStmt>
  <availability status="public">
    <p n="copyright">Copyright &copy; 2006 by the Rector
      and Visitors of the University of Virginia</p>
    <p n="access">Publicly accessible</p>
  </availability>

```

MORE TEI

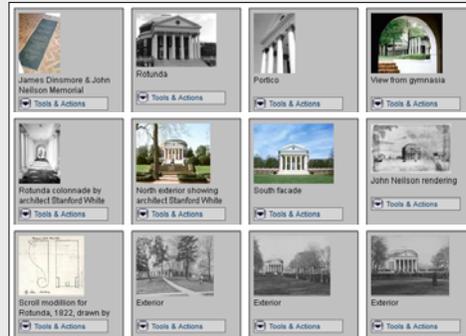
```

<sourceDesc>
  <titleStmt>
    <title type="main">A chronicle of the conquest of
      Granada</title>
    <author>
      <name type="last">Irving</name>
      <name type="first">Washington</name>
      <dateRange from="1783" to="1859">1783-
        1859</dateRange>
    </author>
  </titleStmt>
  <extent>345 p. ; 21 cm.</extent>
  <publicationStmt>
    <publisher>Carey, Lea & Carey</publisher>
    <pubPlace>Philadelphia</pubPlace>
    <date value="1829">1829</date>
    <idno type="LC call number">DP122 .I7 1829a</idno>
    <idno type="UVa Title Control Number">a1599744</idno>
  </publicationStmt>
</sourceDesc>

```

VRA Core

- Maintained by the Visual Resources Association. Version 4 is currently in beta release.
- A categorical organization for the description of works of visual culture as well as the images that document them.
- Consists of a metadata element set and an initial blueprint for how those elements can be hierarchically structured.



credit: K. Edward Lay



Work, Collection or Image

- | | |
|-----------------------------|----------------|
| ■ work, collection or image | ■ relation |
| ■ agent | ■ rights |
| ■ culturalContext | ■ source |
| ■ date | ■ stateEdition |
| ■ description | ■ stylePeriod |
| ■ inscription | ■ subject |
| ■ location | ■ technique |
| ■ Material | ■ textRef |
| ■ Measurements | ■ title |
| | ■ workType |



Advantages of VRA

- Allows description of original and digital object
- Level of granularity greater than Dublin Core, less than MARC and supports specific discipline
- Now content rules have been developed (CCO)

VRA Core

```

<work>
<titleSet>
  <title pref="true" source="LC NAF">Rotunda</title>
</titleSet>
<agentSet><agent>
  <name type="personal" vocab="LC NAF" refid="n 79089957">
    Jefferson, Thomas</name>
  <dates type="life">
    <earliestDate>1743</earliestDate><latestDate>1826</latestDate>
  </dates>
  <role>architect</role>
  <culture>American</culture>
</agent></agentSet>
<agentSet><agent>
  <name type="personal" vocab="LC NAF" refid="n 50020242">
    White, Stanford</name>
  <dates type="life">
    <earliestDate>1853</earliestDate><latestDate>1906</latestDate>
  </dates>
  <role>architect</role>
  <culture>American</culture>
  <notes>Architect of 1896-1897 renovation</notes>
</agent></agentSet>

```

```

<dateSet>
  <date type="construction">
    <earliestDate>1822</earliestDate><latestDate>1826</latestDate>
  </date>
  <notes>Construction begun October, 1822, completed September, 1826.</notes>
</dateSet>
<dateSet>
  <date type="destruction">
    <earliestDate>1895</earliestDate>
  </date>
  <notes>Burned October 27, 1895.</notes>
</dateSet>
<dateSet>
  <date type="renovation">
    <earliestDate>1896</earliestDate><latestDate>1897</latestDate>
  </date>
  <notes>Rebuilt to designs of Stanford White, 1896-1897.</notes>
</dateSet>
<locationSet><location type="site">
  <name type="geographic" vocab="TGN" refid="2002201">
    Charlottesville, Virginia</name>
</location></locationSet>
</work>

```

More VRA Core

```

<image>
<titleSet>
  <title type="descriptive">general view</title>
</titleSet>
<agentSet><agent>
  <name type="personal" vocab="LC NAF"
    refid="n 82111472">Lay, K. Edward</name>
  <culture>American</culture>
  <role>photographer</role>
</agent></agentSet>
<dateSet><date type="creation">
  <earliestDate>1990</earliestDate>
  <latestDate>2000</latestDate>
</date></dateSet>
<locationSet><location type="repository">
  <name type="corporate">University of Virginia Library</name>
  <name type="geographic" vocab="TGN" refid="2002201">
    Charlottesville</name>
</location></locationSet>
<rightsSet>
  <rights type="credit">K. Edward Lay</rights>
  <rights type="access">Publicly accessible</rights>
</rightsSet>
</image>

```



credit: K. Edward Lay

More VRA Core

```

<image>
<titleSet>
  <title type="descriptive">View from gymnasium</title>
</titleSet>
<agentSet><agent>
  <name type="personal" vocab="LC NAF"
    refid="n 82111472">Lay, K. Edward</name>
  <culture>American</culture>
  <role>photographer</role>
</agent></agentSet>
<dateSet><date type="creation">
  <earliestDate>1995</earliestDate>
  <latestDate>2000</latestDate>
</date></dateSet>
<locationSet><location type="repository">
  <name type="corporate">University of Virginia Library</name>
  <name type="geographic" vocab="TGN" refid="2002201">
    Charlottesville</name>
</location></locationSet>
<rightsSet>
  <rights type="credit">K. Edward Lay</rights>
  <rights type="access">Publicly accessible</rights>
</rightsSet>
</image>

```



credit: K. Edward Lay



Other schema

- ONIX (ONline Information eXchange)
 - Originally devised to simplify the provision of book product information to online retailers
 - Includes marketing & shipping oriented information: book jacket blurb and photos, full size, weight info, etc.
 - Uses extensive codelists rather than free text
- IEEE LOM (Learning Object Metadata)
 - An array of related standards for description of 'learning objects' or 'learning resources'
 - One well-documented implementation is CanCore: <http://www.cancore.ca/>



Modeling metadata: why use models?

- To understand what entities you are dealing with
- To understand what metadata are relevant to which entities
- To understand relationships between different entities
- To organize your metadata to make it more predictable (and be able to use automated tools)



Descriptive metadata models

- Conceptual models for bibliographic and authority data
 - Functional Requirements for Bibliographic Records (FRBR)
 - Functional Requirement for Authority Data (FRAD)
- Dublin Core Abstract Model (DCAM)
- Some other models:
 - CIDOC Conceptual Reference Model (emerged from museum community)
 - INDECS (for intellectual property rights)
- There are many conceptual models intended for different purposes



Bibliographic relationships (pre-FRBR)

- Tillett's Taxonomy (1987)
 - Equivalence
 - Derivative
 - Descriptive
 - Whole-part
 - Accompanying
 - Sequential
 - Shared-characteristic



Bibliographic relationships in MARC/MODS

- MARC Linking entry fields
- MARC relationships by specific encoding format
 - Authority vs bibliographic vs holdings
- MODS relationships
 - relatedItem types
 - Relationship to METS document



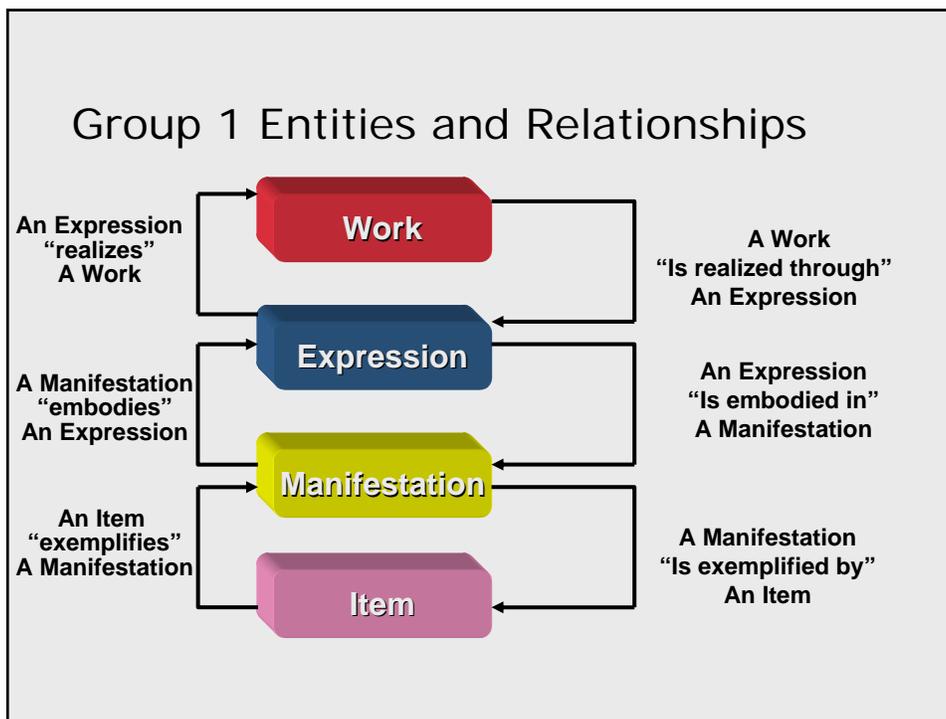
FRBR (1996)

- IFLA Study Group on the Functional Requirements for Bibliographic Records
- Focused on the bibliographic record rather than the catalog
- Used an entity relationship model, rather than descriptive analysis without a structural model
- Broader in scope than previous studies



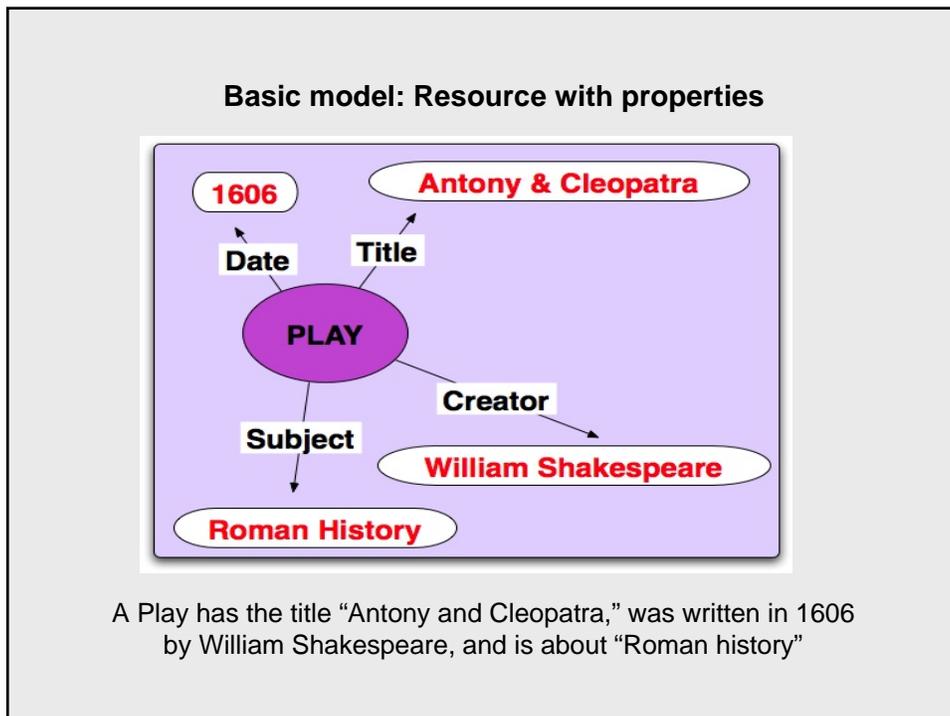
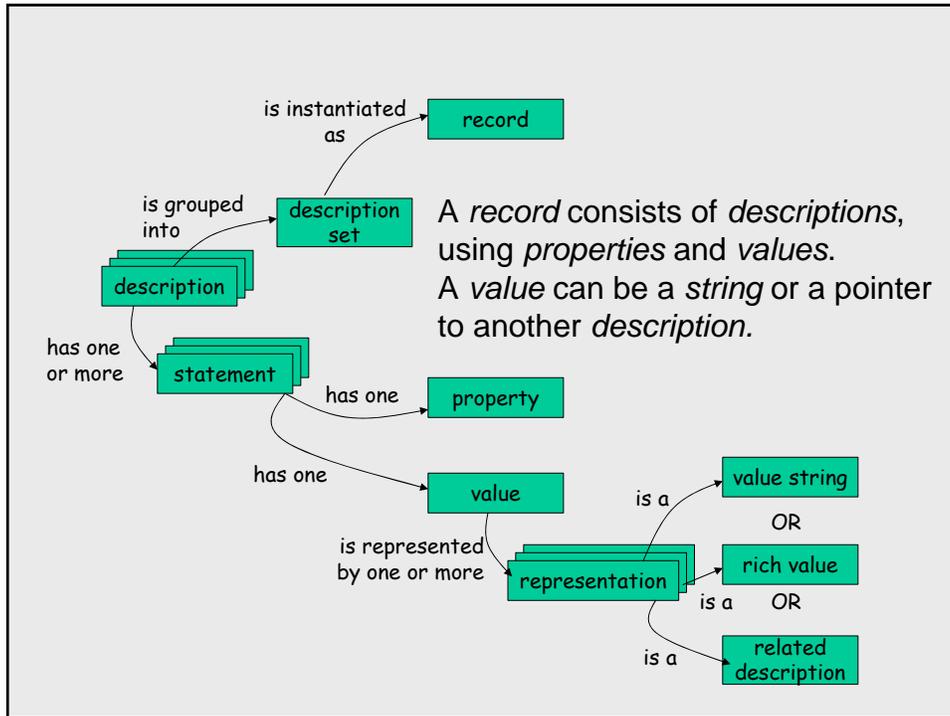
FRBR Entities

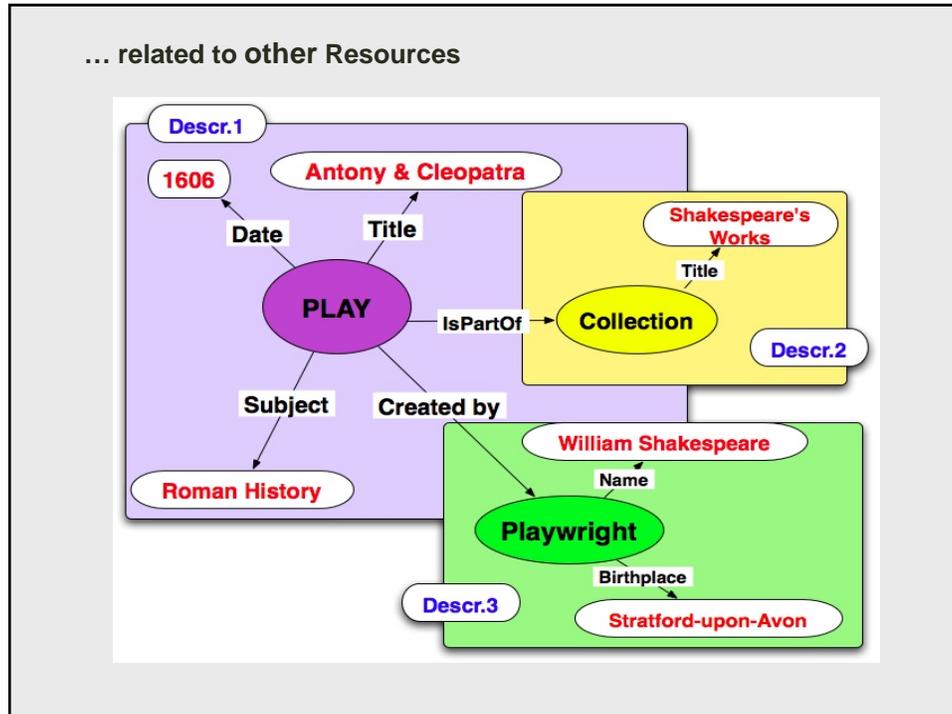
- Bibliographic entities: *works, expressions, manifestations, items*
- Responsible parties: *persons, corporate bodies*
- Subject entities: *concepts, objects, events, places*



DC Abstract Model

- Reaffirms the One-to-One Principle
- Defines 'statement' as the atomic level
- Distinguishes between "description" and "description set":
 - **Description:** "A description is made up of one or more statements about one, and only one, resource."
 - **Description Set:** "A description set is a set of one or more descriptions about one or more resources."
- Work is being done to understand RDA in terms of the DCAM





An Exercise

- *Each group will be given a printout of a digital object*
- *Create a brief metadata record based on the standard assigned to your group*
- *Take notes about the issues and decisions made*
- *Appoint a spokesperson to present the metadata record created & the issues involved (5-10 minutes)*

Exercise for Session 2: Descriptive metadata
Student instructions and Template

Example 1:

America's pinch hit march

Sheet music; 3 pages (2 pages music with cover)

<http://lcweb2.loc.gov/diglib/ihas/loc.natlib.ihas.200033287/default.html>

(images are in the packet)

Example 2:

52nd Street, New York, N.Y.

<http://lcweb2.loc.gov/diglib/ihas/loc.natlib.gottlieb.02771/default.html>

(image is in the packet)

Fill in the following metadata elements in the metadata scheme assigned:

Title/subtitle

Creator/name (with role defined if possible)

Type of resource

Publication/origin information with place, publisher, date

Physical description

Subject

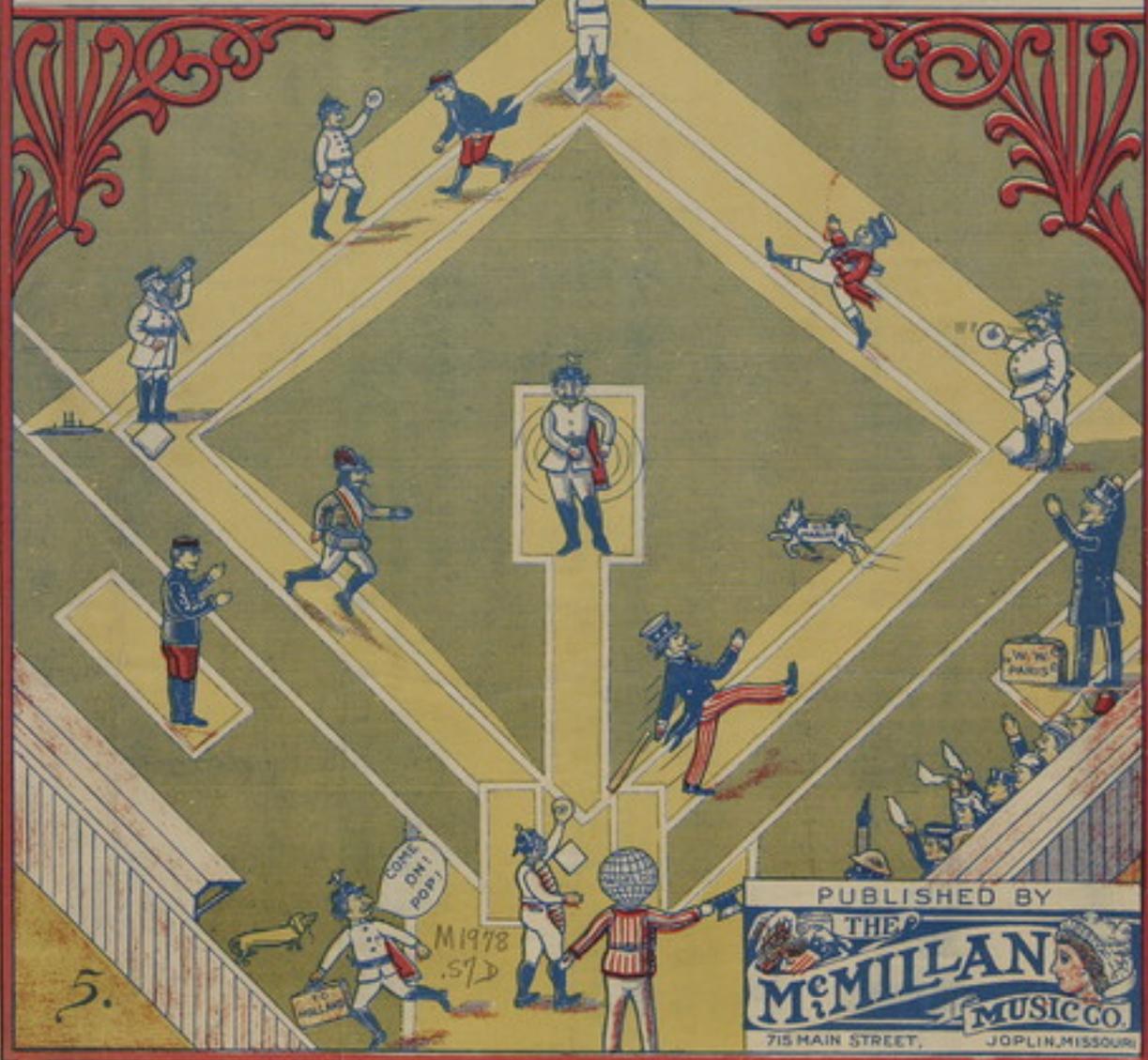
URL

*THIS PAGE INTENTIONALLY LEFT BLANK
FOR DOUBLE SIDED COPY*

“AMERICA'S PINCH HIT” MARCH

THE **HIT** THAT ENDED THE WORLDS GREATEST WAR

COMPOSED BY Music Division, Reserve Storage
BERTHA STANFIELD DEMPSEY



PUBLISHED BY
THE McMILLAN MUSIC CO.
 715 MAIN STREET, JOPLIN, MISSOURI

America's Pinch Hit March

(THE HIT THAT ENDED THE WORLD'S GREATEST WAR)

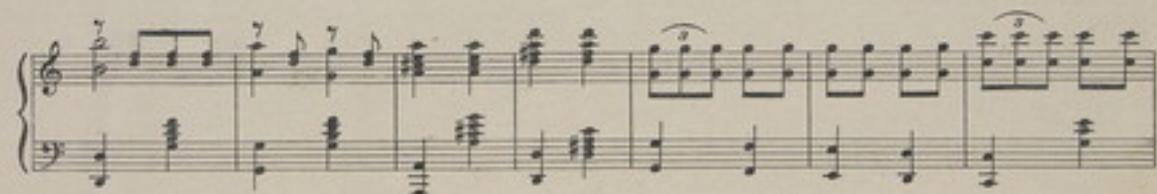
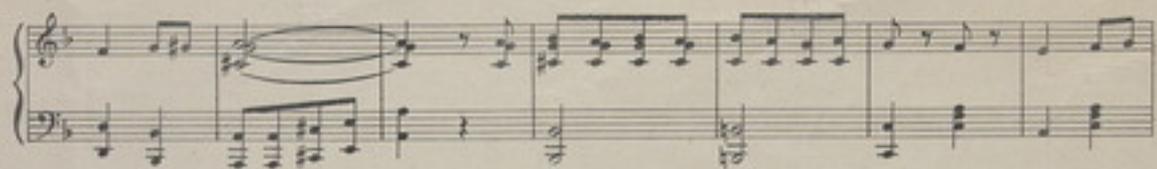
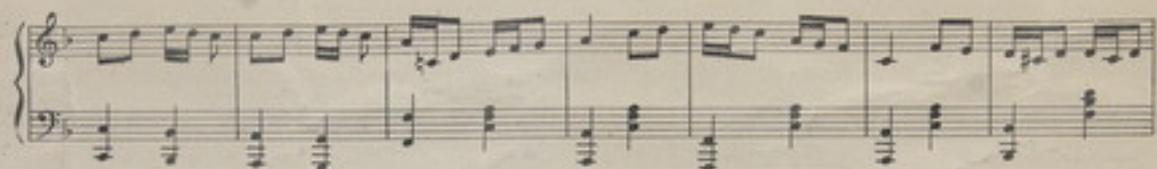
By BERTHA STANFIELD DEMPSEY

Intro.

The introduction consists of six systems of piano music. Each system has a treble and bass staff. The music is in 2/4 time and features a steady, rhythmic accompaniment with various melodic lines and chords. The first system includes a dynamic marking of *f* (forte) in the bass staff. The piece concludes with a final chord in the bass staff.

TRIO

The Trio section consists of two systems of piano music. Each system has a treble and bass staff. The music is in 2/4 time and features a steady, rhythmic accompaniment with various melodic lines and chords. The first system includes a dynamic marking of *p* (piano) in the bass staff. The piece concludes with a final chord in the bass staff.



*THIS PAGE INTENTIONALLY LEFT BLANK
FOR DOUBLE SIDED COPY*

[52nd Street, New York, N.Y., ca. 1948] / William P. Gottlieb [photograph][default view](#) | [640](#) | [1024](#)[default view](#) | [640](#) | [1024](#)

*THIS PAGE INTENTIONALLY LEFT BLANK
FOR DOUBLE SIDED COPY*

*THIS PAGE INTENTIONALLY LEFT BLANK
FOR DOUBLE SIDED COPY*

*THIS PAGE INTENTIONALLY LEFT BLANK
FOR DOUBLE SIDED COPY*

*THIS PAGE INTENTIONALLY LEFT BLANK
FOR DOUBLE SIDED COPY*



Dublin Core Metadata Initiative

Making it easier to find information

Dublin Core Metadata Element Set, Version 1.1

Identifier: <http://dublincore.org/documents/2008/01/14/dces/>

Supersedes: <http://dublincore.org/documents/2006/12/18/dces/>

Latest version: <http://dublincore.org/documents/dces/>

Date 2008-01-14

Issued:

Status of document: This is a DCMI [Recommendation](#).

Description of document: This document provides ready reference for the Dublin Core Metadata Element Set, of Version 1.1. For more detailed documentation and links to historical versioning information, see the document "[DCMI Metadata Terms](#)".

Introduction

The Dublin Core Metadata Element Set is a vocabulary of fifteen properties for use in resource description. The name "Dublin" is due to its origin at a 1995 invitational workshop in Dublin, Ohio; "core" because its elements are broad and generic, usable for describing a wide range of resources.

The fifteen element "Dublin Core" described in this standard is part of a larger set of metadata vocabularies and technical specifications maintained by the Dublin Core Metadata Initiative (DCMI). The full set of vocabularies, DCMI Metadata Terms [DCMI-TERMS], also includes sets of resource classes (including the DCMI Type Vocabulary [DCMI-TYPE]), vocabulary encoding schemes, and syntax encoding schemes. The terms in DCMI vocabularies are intended to be used in combination with terms from other, compatible vocabularies in the context of application profiles and on the basis of the DCMI Abstract Model [DCAM].

All changes made to terms of the Dublin Core Metadata Element Set since 2001 have been reviewed by a DCMI Usage Board in the context of a DCMI Namespace Policy [DCMI-NAMESPACE]. The namespace policy describes how DCMI terms are assigned Uniform Resource Identifiers (URIs) and sets limits on the range of editorial changes that may allowably be made to the labels, definitions, and usage comments associated with existing DCMI terms.

This document, an excerpt from the more comprehensive document "DCMI Metadata Terms" [[DCTERMS](#)] provides an abbreviated reference version of the fifteen element descriptions that have been formally endorsed in the following standards:

- ISO Standard 15836-2003 of February 2003 [[ISO15836](#)]
- ANSI/NISO Standard Z39.85-2007 of May 2007 [[NISOZ3985](#)]
- IETF RFC 5013 of August 2007 [[RFC5013](#)]

Since 1998, when these fifteen elements entered into a standardization track, notions of best

practice in the Semantic Web have evolved to include the assignment of formal domains and ranges in addition to definitions in natural language. Domains and ranges specify what kind of described resources and value resources are associated with a given property. Domains and ranges express the meanings implicit in natural-language definitions in an explicit form that is usable for the automatic processing of logical inferences. When a given property is encountered, an inferencing application may use information about the domains and ranges assigned to a property in order to make inferences about the resources described thereby.

Since January 2008, therefore, DCMI includes formal domains and ranges in the definitions of its properties. So as not to affect the conformance of existing implementations of "simple Dublin Core" in RDF, domains and ranges have not been specified for the fifteen properties of the dc: namespace (<http://purl.org/dc/elements/1.1/>). Rather, fifteen new properties with "names" identical to those of the Dublin Core Metadata Element Set Version 1.1 have been created in the dcterms: namespace (<http://purl.org/dc/terms/>). These fifteen new properties have been defined as subproperties of the corresponding properties of DCMES Version 1.1 and assigned domains and ranges as specified in the more comprehensive document "DCMI Metadata Terms" [[DCTERMS](#)].

Implementers may freely choose to use these fifteen properties either in their legacy dc: variant (e.g., <http://purl.org/dc/elements/1.1/creator>) or in the dcterms: variant (e.g., <http://purl.org/dc/terms/creator>) depending on application requirements. The RDF schemas of the DCMI namespaces describe the subproperty relation of dcterms:creator to dc:creator for use by Semantic Web-aware applications. Over time, however, implementers are encouraged to use the semantically more precise dcterms: properties, as they more fully follow emerging notions of best practice for machine-processable metadata.

References

[RFC5013]	http://www.ietf.org/rfc/rfc5013.txt
[NISOZ3985]	http://www.niso.org/standards/z39-85-2007/
[ISO15836]	http://www.iso.org/iso/search.htm?qt=15836&searchSubmit=Search&sort=rel&type=simple&published=on
[TRANSLATIONS]	http://dublincore.org/resources/translations/
[DCTERMS]	http://dublincore.org/documents/dcmi-terms/

The Elements

Term Name: contributor	
URI:	http://purl.org/dc/elements/1.1/contributor
Label:	Contributor
Definition:	An entity responsible for making contributions to the resource.
Comment:	Examples of a Contributor include a person, an organization, or a service. Typically, the name of a Contributor should be used to indicate the entity.
Term Name: coverage	
URI:	http://purl.org/dc/elements/1.1/coverage
Label:	Coverage

Definition:	The spatial or temporal topic of the resource, the spatial applicability of the resource, or the jurisdiction under which the resource is relevant.
Comment:	Spatial topic and spatial applicability may be a named place or a location specified by its geographic coordinates. Temporal topic may be a named period, date, or date range. A jurisdiction may be a named administrative entity or a geographic place to which the resource applies. Recommended best practice is to use a controlled vocabulary such as the Thesaurus of Geographic Names [TGN]. Where appropriate, named places or time periods can be used in preference to numeric identifiers such as sets of coordinates or date ranges.
References:	[TGN] http://www.getty.edu/research/tools/vocabulary/tgn/index.html
Term Name: creator	
URI:	http://purl.org/dc/elements/1.1/creator
Label:	Creator
Definition:	An entity primarily responsible for making the resource.
Comment:	Examples of a Creator include a person, an organization, or a service. Typically, the name of a Creator should be used to indicate the entity.
Term Name: date	
URI:	http://purl.org/dc/elements/1.1/date
Label:	Date
Definition:	A point or period of time associated with an event in the lifecycle of the resource.
Comment:	Date may be used to express temporal information at any level of granularity. Recommended best practice is to use an encoding scheme, such as the W3CDTF profile of ISO 8601 [W3CDTF].
References:	[W3CDTF] http://www.w3.org/TR/NOTE-datetime
Term Name: description	
URI:	http://purl.org/dc/elements/1.1/description
Label:	Description
Definition:	An account of the resource.
Comment:	Description may include but is not limited to: an abstract, a table of contents, a graphical representation, or a free-text account of the resource.
Term Name: format	
URI:	http://purl.org/dc/elements/1.1/format
Label:	Format
Definition:	The file format, physical medium, or dimensions of the resource.

Comment:	Examples of dimensions include size and duration. Recommended best practice is to use a controlled vocabulary such as the list of Internet Media Types [MIME].
References:	[MIME] http://www.iana.org/assignments/media-types/
Term Name: identifier	
URI:	http://purl.org/dc/elements/1.1/identifier
Label:	Identifier
Definition:	An unambiguous reference to the resource within a given context.
Comment:	Recommended best practice is to identify the resource by means of a string conforming to a formal identification system.
Term Name: language	
URI:	http://purl.org/dc/elements/1.1/language
Label:	Language
Definition:	A language of the resource.
Comment:	Recommended best practice is to use a controlled vocabulary such as RFC 4646 [RFC4646].
References:	[RFC4646] http://www.ietf.org/rfc/rfc4646.txt
Term Name: publisher	
URI:	http://purl.org/dc/elements/1.1/publisher
Label:	Publisher
Definition:	An entity responsible for making the resource available.
Comment:	Examples of a Publisher include a person, an organization, or a service. Typically, the name of a Publisher should be used to indicate the entity.
Term Name: relation	
URI:	http://purl.org/dc/elements/1.1/relation
Label:	Relation
Definition:	A related resource.
Comment:	Recommended best practice is to identify the related resource by means of a string conforming to a formal identification system.
Term Name: rights	
URI:	http://purl.org/dc/elements/1.1/rights
Label:	Rights

Definition:	Information about rights held in and over the resource.
Comment:	Typically, rights information includes a statement about various property rights associated with the resource, including intellectual property rights.
Term Name: source	
URI:	http://purl.org/dc/elements/1.1/source
Label:	Source
Definition:	A related resource from which the described resource is derived.
Comment:	The described resource may be derived from the related resource in whole or in part. Recommended best practice is to identify the related resource by means of a string conforming to a formal identification system.
Term Name: subject	
URI:	http://purl.org/dc/elements/1.1/subject
Label:	Subject
Definition:	The topic of the resource.
Comment:	Typically, the subject will be represented using keywords, key phrases, or classification codes. Recommended best practice is to use a controlled vocabulary. To describe the spatial or temporal topic of the resource, use the Coverage element.
Term Name: title	
URI:	http://purl.org/dc/elements/1.1/title
Label:	Title
Definition:	A name given to the resource.
Comment:	Typically, a Title will be a name by which the resource is formally known.
Term Name: type	
URI:	http://purl.org/dc/elements/1.1/type
Label:	Type
Definition:	The nature or genre of the resource.
Comment:	Recommended best practice is to use a controlled vocabulary such as the DCMI Type Vocabulary [DCMITYPE]. To describe the file format, physical medium, or dimensions of the resource, use the Format element.
References:	[DCMITYPE] http://dublincore.org/documents/dcmi-type-vocabulary/

Errata

2008-06-09. Updated URIs for ANSI/NISO Z39.85 and ISO 15836.

*THIS PAGE INTENTIONALLY LEFT BLANK
FOR DOUBLE SIDED COPY*



Dublin Core Metadata Initiative
Making it easier to find information

DCMI Type Vocabulary

Title: DCMI Type Vocabulary

Creator: [DCMI Usage Board](#)

Identifier: <http://dublincore.org/documents/2008/01/14/dcmi-type-vocabulary/>

Date Issued: 2008-01-14

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

Replaces: <http://dublincore.org/documents/2006/08/28/dcmi-type-vocabulary/>

Replaced By: Not applicable

Translations: <http://dublincore.org/resources/translations/>

Document Status: This is a DCMI Recommendation.

Status:

Description: The DCMI Type Vocabulary provides a general, cross-domain list of approved terms that may be used as values for the Resource Type element to identify the genre of a resource. The terms documented here are also included in the more comprehensive document "DCMI Metadata Terms" at <http://dublincore.org/documents/dcmi-terms/>.

Term Name: Collection	
URI:	http://purl.org/dc/dcmitype/Collection
Label:	Collection
Definition:	An aggregation of resources.
Comment:	A collection is described as a group; its parts may also be separately described.
Type of Term:	Class

Member Of:	http://purl.org/dc/terms/DCMIType
Version:	http://dublincore.org/usage/terms/history/#Collection-003
Term Name: Dataset	
URI:	http://purl.org/dc/dcmitype/Dataset
Label:	Dataset
Definition:	Data encoded in a defined structure.
Comment:	Examples include lists, tables, and databases. A dataset may be useful for direct machine processing.
Type of Term:	<u>Class</u>
Member Of:	http://purl.org/dc/terms/DCMIType
Version:	http://dublincore.org/usage/terms/history/#Dataset-003
Term Name: Event	
URI:	http://purl.org/dc/dcmitype/Event
Label:	Event
Definition:	A non-persistent, time-based occurrence.
Comment:	Metadata for an event provides descriptive information that is the basis for discovery of the purpose, location, duration, and responsible agents associated with an event. Examples include an exhibition, webcast, conference, workshop, open day, performance, battle, trial, wedding, tea party, conflagration.
Type of Term:	<u>Class</u>
Member Of:	http://purl.org/dc/terms/DCMIType
Version:	http://dublincore.org/usage/terms/history/#Event-003
Term Name: Image	

URI:	http://purl.org/dc/dcmitype/Image
Label:	Image
Definition:	A visual representation other than text.
Comment:	Examples include images and photographs of physical objects, paintings, prints, drawings, other images and graphics, animations and moving pictures, film, diagrams, maps, musical notation. Note that Image may include both electronic and physical representations.
Type of Term:	<u>Class</u>
Broader Than:	http://purl.org/dc/dcmitype/StillImage
Broader Than:	http://purl.org/dc/dcmitype/MovingImage
Member Of:	http://purl.org/dc/terms/DCMIType
Version:	http://dublincore.org/usage/terms/history/#Image-004
Term Name: InteractiveResource	
URI:	http://purl.org/dc/dcmitype/InteractiveResource
Label:	Interactive Resource
Definition:	A resource requiring interaction from the user to be understood, executed, or experienced.
Comment:	Examples include forms on Web pages, applets, multimedia learning objects, chat services, or virtual reality environments.
Type of Term:	<u>Class</u>
Member Of:	http://purl.org/dc/terms/DCMIType
Version:	http://dublincore.org/usage/terms/history/#InteractiveResource-003
Term Name: MovingImage	

URI:	http://purl.org/dc/dcmitype/MovingImage
Label:	Moving Image
Definition:	A series of visual representations imparting an impression of motion when shown in succession.
Comment:	Examples include animations, movies, television programs, videos, zoetropes, or visual output from a simulation. Instances of the type Moving Image must also be describable as instances of the broader type Image.
Type of Term:	<u>Class</u>
Narrower Than:	http://purl.org/dc/dcmitype/Image
Member Of:	http://purl.org/dc/terms/DCMIType
Version:	http://dublincore.org/usage/terms/history/#MovingImage-003
Term Name: PhysicalObject	
URI:	http://purl.org/dc/dcmitype/PhysicalObject
Label:	Physical Object
Definition:	An inanimate, three-dimensional object or substance.
Comment:	Note that digital representations of, or surrogates for, these objects should use Image, Text or one of the other types.
Type of Term:	<u>Class</u>
Member Of:	http://purl.org/dc/terms/DCMIType
Version:	http://dublincore.org/usage/terms/history/#PhysicalObject-003
Term Name: Service	
URI:	http://purl.org/dc/dcmitype/Service
Label:	Service

Definition:	A system that provides one or more functions.
Comment:	Examples include a photocopying service, a banking service, an authentication service, interlibrary loans, a Z39.50 or Web server.
Type of Term:	<u>Class</u>
Member Of:	http://purl.org/dc/terms/DCMIType
Version:	http://dublincore.org/usage/terms/history/#Service-003
Term Name: Software	
URI:	http://purl.org/dc/dcmitype/Software
Label:	Software
Definition:	A computer program in source or compiled form.
Comment:	Examples include a C source file, MS-Windows .exe executable, or Perl script.
Type of Term:	<u>Class</u>
Member Of:	http://purl.org/dc/terms/DCMIType
Version:	http://dublincore.org/usage/terms/history/#Software-003
Term Name: Sound	
URI:	http://purl.org/dc/dcmitype/Sound
Label:	Sound
Definition:	A resource primarily intended to be heard.
Comment:	Examples include a music playback file format, an audio compact disc, and recorded speech or sounds.
Type of Term:	<u>Class</u>
Member Of:	http://purl.org/dc/terms/DCMIType

Version:	http://dublincore.org/usage/terms/history/#Sound-003
Term Name: StillImage	
URI:	http://purl.org/dc/dcmitype/StillImage
Label:	Still Image
Definition:	A static visual representation.
Comment:	Examples include paintings, drawings, graphic designs, plans and maps. Recommended best practice is to assign the type Text to images of textual materials. Instances of the type Still Image must also be describable as instances of the broader type Image.
Type of Term:	<u>Class</u>
Narrower Than:	http://purl.org/dc/dcmitype/Image
Member Of:	http://purl.org/dc/terms/DCMIType
Version:	http://dublincore.org/usage/terms/history/#StillImage-003
Term Name: Text	
URI:	http://purl.org/dc/dcmitype/Text
Label:	Text
Definition:	A resource consisting primarily of words for reading.
Comment:	Examples include books, letters, dissertations, poems, newspapers, articles, archives of mailing lists. Note that facsimiles or images of texts are still of the genre Text.
Type of Term:	<u>Class</u>
Member Of:	http://purl.org/dc/terms/DCMIType
Version:	http://dublincore.org/usage/terms/history/#Text-003

[The Library of Congress](#) >> [Standards](#) >> [MODS](#)

Metadata Object Description Schema (MODS) [Official Web Site](#)



[HOME](#) >> [Schemas](#) >> [Outline of Elements and Attributes](#)

Outline of Elements and Attributes in MODS Version 3.3

This document contains a listing of elements and their related attributes in MODS Version 3.3 with values or value sources where applicable. It is an "outline" of the schema. Items highlighted in **red** indicate changes made to MODS in Version 3.3.

All top-level elements and all attributes are optional, but you must have at least one element. Subelements are optional, although in some cases you may not have empty containers. Attributes are not in a mandated sequence and not repeatable (per XML rules). "Ordered" below means the subelements must occur in the order given. Elements are repeatable unless otherwise noted.

"Authority" attributes are either followed by codes for authority lists (e.g., iso639-2b) or "see" references that link to documents that contain codes for identifying authority lists.

For additional information about any MODS elements (except for new 3.3 elements), please see the [MODS User Guidelines](#).

Top Level Elements:

titleInfo	note
name	subject
typeOfResource	classification
genre	relatedItem
originInfo	identifier
language	location
physicalDescription	accessCondition
abstract	part
tableOfContents	extension
targetAudience	recordInfo

Root Elements:

[mods](#)
[modsCollection](#)

Top Level Elements

1. titleInfo

Subelements:

title
subTitle
partNumber

partName
nonSort

Attributes:

ID; xlink; lang; xml:lang; script; transliteration
type (enumerated: abbreviated, translated, alternative, uniform)
authority (see: www.loc.gov/marc/sourcecode/authorityfile/authorityfilesources.html)
displayLabel

2. name

Subelements:

namePart

Attribute: type (date, family, given, termsOfAddress)

displayForm

affiliation

role

roleTerm

Attributes: type (code, text); authority
(see: www.loc.gov/marc/sourcecode/relator/relatorsources.html)

description

Attributes:

ID; xlink; lang; xml:lang; script; transliteration
type (enumerated: personal, corporate, conference)
authority (see: www.loc.gov/marc/sourcecode/authorityfile/authorityfilesources.html)

3. typeOfResource

Enumerated values:

text
cartographic
notated music
sound recording-musical
sound recording-nonmusical
sound recording
still image
moving image
three dimensional object
software, multimedia
mixed material

Subelements:

[none]

Attributes:

collection (yes)
manuscript (yes)

4. genre

Subelements:

[none]

Attributes:

lang; xml:lang; script; transliteration
authority (see: www.loc.gov/marc/sourcecode/genre/genresources.html)

type (examples: class, work type, or style)

5. originInfo

Subelement:

place

placeTerm

Attributes: type (code, text); authority (marcgac, marccountry, iso3166)

publisher

dateIssued

Attributes: encoding (w3cdtf, iso8601, marc); point (start, end); keyDate (yes);
qualifier (approximate, inferred, questionable)

dateCreated

Attributes: encoding (w3cdtf, iso8601, marc); point (start, end); keyDate (yes);
qualifier (approximate, inferred, questionable)

dateCaptured

Attributes: encoding (w3cdtf, iso8601, marc); point (start, end); keyDate (yes);
qualifier (approximate, inferred, questionable)

dateValid

Attributes: encoding (w3cdtf, iso8601, marc); point (start, end); keyDate (yes);
qualifier (approximate, inferred, questionable)

dateModified

Attributes: encoding (w3cdtf, iso8601, marc); point (start, end); keyDate (yes);
qualifier (approximate, inferred, questionable)

copyrightDate

Attributes: encoding (w3cdtf, iso8601, marc); point (start, end); keyDate (yes);
qualifier (approximate, inferred, questionable)

dateOther

Attributes: encoding (w3cdtf, iso8601, marc); point (start, end); keyDate (yes);
qualifier (approximate, inferred, questionable); type

edition

issuance (continuing, monographic)

frequency

Attribute: authority

(see: www.loc.gov/marc/sourcecode/frequency/frequencyhome.html)

Attributes:

lang; xml:lang; script; transliteration

6. language

Subelements:

languageTerm

Attributes: type (code, text); authority (iso639-2b, rfc3066, [iso639-3](#), [rfc4646](#))

Attributes:

objectPart

7. physicalDescription

Subelements:

form

Attribute: authority (see: www.loc.gov/marc/sourcecode/form/formsource.html); type
(Examples: material, technique)

reformattingQuality (access, preservation, replacement)
internetMediaType
extent
digitalOrigin (born digital, reformatted digital, digitized microfilm, digitized other analog)
note

Attributes: xlink; lang; xml:lang; script; transliteration; displayLabel; type
(For a list of implemented note types, see:
www.loc.gov/standards/mods/mods-notes.html)

Attributes:

lang
xml:lang
script
transliteration

8. abstract

Subelements:

[none]

Attributes:

xlink; lang; xml:lang; script; transliteration
displayLabel
type (Examples: review, scope and content)

9. tableOfContents

Subelements:

[none]

Attributes:

xlink; lang; xml:lang; script; transliteration
displayLabel
type (Examples: incomplete contents, partial contents)

10. targetAudience

Subelements:

[none]

Attributes:

lang; xml:lang; script; transliteration
authority (see: www.loc.gov/marc/sourcecode/target/targetsource.html)

11. note

Subelements:

[none]

Attributes:

ID; xlink; lang; xml:lang; script; transliteration
displayLabel
type (For a list of implemented note types,
see: www.loc.gov/standards/mods/mods-notes.html)

12. subject

Subelements:

topic
 geographic
 temporal
Attributes: encoding (w3cdtf, iso8601, marc); point (start,end); keyDate (yes);
 qualifier (approximate, inferred, questionable)
 titleInfo (see: [titleInfo](#))
 name (see: [name](#))
 geographicCode
Attribute: authority (marcgac, marccountry, iso3166)
 genre
 hierarchicalGeographic
 continent
 country
 province
 region
 state
 territory
 county
 city
 island
 area
 extraterrestrialArea
 citySection
 cartographics [ordered]
 scale
 projection
 coordinates
 occupation
Attributes:
 ID; xlink; lang; xml:lang; script; transliteration
 authority (see: www.loc.gov/marc/sourcecode/subject/subjectsourc.html)

13. classification

Subelements:

[none]

Attributes:

lang; xml:lang; script; transliteration
 authority (see: www.loc.gov/marc/sourcecode/classification/classificationsourc.html)
 edition
 displayLabel

14. relatedItem

Subelements:

(Any MODS element may be used as defined in the schema with appropriate subelements.)

[titleInfo](#)

[name](#)

[typeOfResource](#)
[genre](#)
[originInfo](#)
[language](#)
[physicalDescription](#)
[abstract](#)
[tableOfContents](#)
[targetAudience](#)
[note](#)
[subject](#)
[classification](#)
[relatedItem](#)
[identifier](#)
[location](#)
[accessCondition](#)
[part](#)
[extension](#)
[recordInfo](#)

Attributes:

ID; xlink
 displayLabel
 type (enumerated: preceding, succeeding, original, host, constituent, series, otherVersion, otherFormat, isReferencedBy)

15. identifier*Subelements:*

[none]

Attributes:

lang; xml:lang; script; transliteration
 type
 (suggested values: hdl, doi, isbn, isrc, ismn, issn, issue number, istc, lccn, local, matrix number, music publisher, music plate, sici, uri, upc, videorecording identifier, stock number)
 invalid (yes)

16. location*Subelements:*

physicalLocation

Attributes: authority

(see: www.loc.gov/marc/sourcecode/organization/organizationsource.html);
 displayLabel; type (Examples: current, discovery, former, creation); lang;
 xml:lang; script; transliteration; **xlink**

shelfLocator

url

Attributes:

dateLastAccessed
 displayLabel
 note

access (preview, raw object, object in context)

usage (primary display)

holdingSimple (not repeatable)

copyInformation

form (not repeatable)

Attribute: authority

sublocation

shelfLocator

electronicLocator

note

Attributes: displayLabel, type

enumerationAndChronology

Attributes: unitType (1,2,3)

Note: 1=basic bibliographic unit; 2=supplement; 3=index

holdingExternal (not repeatable)

(Extensible to use other holdings schemas)

17. accessCondition

(Extensible to allow for other more detailed rights schemas.)

Subelements:

[none]

Attributes:

xlink; lang; xml:lang; script; transliteration

displayLabel

type (suggested values: restriction on access; use and reproduction)

18. part

Subelements:

detail

number

caption

title

Attributes: type (suggested values: part, volume, issue, chapter, section, paragraph, track) ; level

extent [ordered]

start

end

total

list

Attribute: unit (suggested values: pages, minutes)

date

Attributes: encoding (w3cdf, iso8601, marc); point (start,end); qualifier (approximate, inferred, questionable)

text

Attributes: xlink; lang; xml:lang; transliteration; script; displayLabel; type

Attributes:

ID

type (suggested values: volume, issue, chapter, section, paragraph, track)

order

19. extension

Subelements:

[none]

Attributes:

[none]

20. recordInfo

Subelements:

recordContentSource

Attributes: authority

(see: www.loc.gov/marc/sourcecode/organization/organizationsource.html);

lang; xml:lang; script; transliteration

recordCreationDate

Attributes: encoding (w3cdtf, iso8601, marc); point (start, end); keyDate (yes);

qualifier (approximate, inferred, questionable)

recordChangeDate

Attributes: encoding (w3cdtf, iso8601, marc); point (start, end); keyDate (yes);

qualifier (approximate, inferred, questionable)

recordIdentifier

Attribute: source

recordOrigin

languageOf Cataloging

languageTerm

Attributes: type (code, text); authority (iso639-2b, rfc3066)

descriptionStandard (see: <http://www.loc.gov/marc/relators/reladesc.html#rela040b>)

Attributes:

lang; xml:lang; script; transliteration

Root Elements

1. mods (A single MODS record)

Subelements:

See: [Top Level Elements](#)

Attributes:

ID

version

2. modsCollection (A collection of MODS records)

Subelements:

mods

Attributes:

[none]

[HOME](#) >> [Schemas](#) >> **Outline of Elements and Attributes**

Questions and comments:

[Contact Us](#) (*September 4, 2008*)



*THIS PAGE INTENTIONALLY LEFT BLANK
FOR DOUBLE SIDED COPY*

VRA Core 4.0 Outline

Global Attributes

- dataDate
- extent
- href
- pref
- refid
- rules
- source
- vocab
- xml:lang

Syntax (using date element as example)

```
<work id="">
  <dateSet>
    <display></display>
    <notes></notes>
    <date type="">
      <earliestDate></earliestDate>
      <latestDate></latestDate>
    </date>
  </dateSet>
</work>
```

ELEMENTS

- **work, collection, or image** (*id*)
- **agent**
 - attribution
 - culture
 - dates (*type*)
 - earliestDate (*circa*)
 - latestDate (*circa*)
 - name (*type*)
 - role
- **culturalContext**
- **date** (*type*)
 - earliestDate (*circa*)
 - latestDate (*circa*)
- **description**
- **inscription**
 - author
 - position
 - text (*type*)
- **location** (*type*)
 - name (*type*)
 - refid (*type*)
- **material** (*type*)
- **measurements** (*type, unit*)
- **relation** (*type, relids*)
- **rights** (*type*)
 - rightsHolder
 - text
- **source**
 - name (*type*)
 - refid (*type*)
- **stateEdition** (*count, num, type*)
 - description
 - name
- **stylePeriod**
- **subject**
 - term (*type*)
- **technique**
- **textref**
 - name (*type*)
 - refid (*type*)
- **title** (*type*)
- **worktype**

Appendix 1: Related Data Standards (when online versions are available they are listed below)

Data element sets:

Categories for the Description of Works of Art
http://www.getty.edu/research/conducting_research/standards/cdwa/

Data content:

Cataloging Cultural Objects: A Guide to Describing Cultural Works and Their Images/
editors, Murtha Baca ... [et al.]. on behalf of the Visual Resources Association. Chicago:
American Library Association, 2006. See also <http://vraweb.org/ccoweb/cco/index.html>

[Anglo-American cataloguing rules](#) / prepared under the direction of the Joint Steering
Committee for Revision of AACR, a committee of the American Library Association ... let
al. 2nd ed., 2002 revision. Ottawa: Canadian Library Association; Chicago: American
Library Association, 2002- (AACR2)

Data values:

Library of Congress Subject Headings (LCSH)

Library of Congress Thesaurus for Graphic Materials (LCTGM - Parts I and II)
<http://www.loc.gov/rr/prinUtgm1/>
<http://www.loc.gov/rr/prinUtgm2/>

Getty Art and Architecture Thesaurus (AAT)
<http://www.getty.edu/research/conductingresearch/vocabularies/aaUindex.html>

Getty Thesaurus of Geographic Names (TGN)
<http://www.getty.edu/research/conductingresearch/vocabularies/tgn/index.html>

Getty Union List of Artist Names (ULAN)
<http://www.getty.edu/research/conductingresearch/vocabularies/ulan/index.html>

Getty Editorial Guidelines for ULAN: Appendix G: Nationalities and Places
http://www.getty.edu/research/conducting_research/vocabularies/guidelines/ulan_4_7_appendix_9_nationality_place.pdf

Appendix 2: Recommended XML introductory resources

XML tutorial put out by w3Schools
<http://www.w3schools.com/xml/default.asp>

Eric Lease Morgan's Getting Started with XML
<http://www.infomotions.com/musings/getting-started/>

Help files within software programs- For example, in Access search for "XML for the uninitiated"

Gilmour, Rom. XML: a Guide for Librarians. LITA Guide #11. American Library
Association, 2003.



3. Technical and administrative metadata standards

Metadata Standards and Applications Workshop



Goals of session

- To understand the different types of administrative metadata standards
- To learn what types of metadata are needed for digital preservation
- To learn the importance of technical, structural and rights metadata in digital libraries



Types of administrative metadata

- Provides information to help manage a resource
 - Preservation metadata
 - Technical characteristics
 - Information about actions on an object
 - Structural metadata may be considered administrative; indicates how compound objects are put together
 - Rights metadata
 - Access rights and restrictions
 - Preservation rights and restrictions

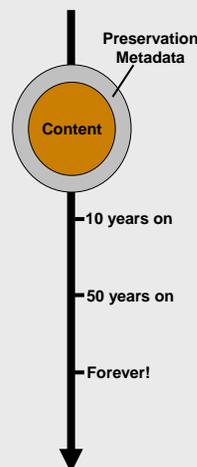


PREMIS: introduction

- Preservation metadata that includes subcategories:
 - Technical metadata
 - Relationships (structural and derivative)
 - Digital provenance (what actions performed on objects)
 - Rights

Preservation metadata includes:

- Provenance:
 - *Who has had custody/ownership of the digital object?*
- Authenticity:
 - *Is the digital object what it purports to be?*
- Preservation Activity:
 - *What has been done to preserve it?*
- Technical Environment:
 - *What is needed to render and use it?*
- Rights Management:
 - *What IPR must be observed?*



➤ **Makes digital objects self-documenting across time**



PREMIS Data Dictionary

- **May 2005:**
Data Dictionary for Preservation Metadata: Final Report of the PREMIS Working Group
- 237-page report includes:
 - PREMIS Data Dictionary 1.0
 - Accompanying report
 - Special topics, glossary, usage examples
- **Data Dictionary:** comprehensive, practical resource for implementing preservation metadata in digital archiving systems
 - Used *Framework* as starting point
 - Detailed description of metadata elements
 - Guidelines to support implementation, use, management
 - Based on deep pool of institutional experiences in setting up and managing operational capacity for digital preservation
- Set of **XML schema** developed to support use of Data Dictionary





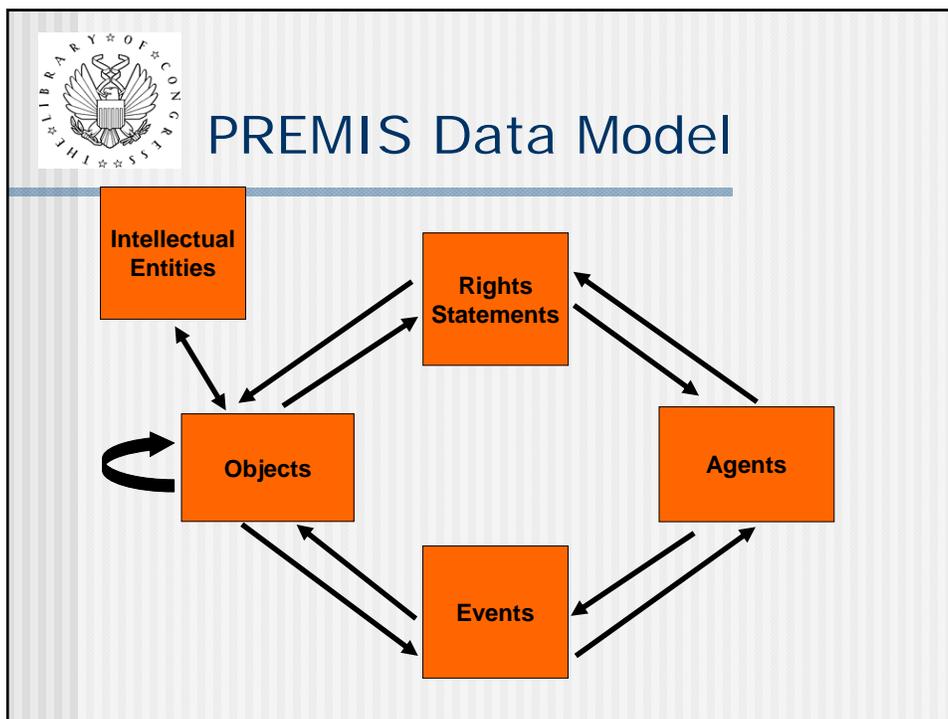
Scope of data dictionary

- Implementation independent
- Descriptive metadata out of scope
- Technical metadata applying to all or most format types
- Media or hardware details are limited
- Business rules are essential for working repositories, but not covered
- Rights information for preservation actions, not access



What PREMIS is and is not

- What PREMIS is:
 - Common data model for organizing/thinking about preservation metadata
 - Guidance for local implementations
 - Standard for exchanging information packages between repositories
- What PREMIS is not:
 - Out-of-the-box solution: need to instantiate as metadata elements in repository system
 - All needed metadata: excludes business rules, format-specific technical metadata, descriptive metadata for access, non-core preservation metadata
 - Lifecycle management of objects outside repository
 - Rights management: limited to permissions regarding actions taken within repository



-
- The diagram lists the types of information covered in PREMIS, categorized by entity type. The information is organized into two columns:
- Object**
 - Object ID
 - Preservation level
 - Object characteristics (format, size, etc.)
 - Storage
 - Environment
 - Digital signatures
 - Relationships
 - Linking identifiers
 - Event**
 - Event ID
 - Event type
 - Event date/time
 - Event outcomes
 - Linking identifiers
 - Agent**
 - Agent ID
 - Agent name
 - Rights**
 - Rights statement
 - Granting agent
 - Permission granted



PREMIS Maintenance Activity

Permanent Web presence,
hosted by Library of Congress

Centralized destination for
information, announcements,
and other PREMIS-related
resources

Discussion list for PREMIS
implementers (PIG list)

Coordinate future revisions of Data Dictionary and XML schema

Editorial committee guides development and revisions



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



Current activities

- PREMIS Implementers' Registry
 - <http://www.loc.gov/standards/premis/premis-registry.html>
- Revision of data dictionary and schemas (March 2008)
- Guidelines for use of PREMIS within METS have been developed
- PREMIS tutorials
 - One or one and a half day tutorials have been given in several locations: Glasgow, Boston, Stockholm, Albuquerque, Washington, San Diego, Berlin
 - Training materials available from LC



Why is PREMIS important to catalogers?

- As we take responsibility for more digital materials, we need to ensure that they can be used in the future
- Most preservation metadata will be generated from the object, but catalogers may need to verify its accuracy
- Catalogers may need to play a role in assessing and organizing digital materials
 - Understanding the structure of complex digital objects
 - Determining significant properties that need to be preserved



Technical metadata for images

- NISO Z39.87 and MIX
- Adobe and XMP
- Exif
- IPTC (International Press Telecommunications Council)/XMP
- Some of these deal with embedded metadata in images



Metadata For Images in XML (MIX)

- An XML Schema designed for expressing technical metadata for digital still images
- Based on the NISO Z39.87 Data Dictionary – Technical Metadata for Digital Still Images
- Can be used standalone or as an extension schema with METS/PREMIS



Using MIX

- Includes
 - Characteristics that apply to all or most object types, e.g. size, format (elements also in PREMIS)
 - Format specific metadata for images
- Some examples of format specific metadata elements in MIX:
 - Image width
 - Color space, color profile
 - Scanner metadata
 - Digital camera settings
- Most well developed of format specific technical metadata standards



Technical metadata for textual objects

- [textMD](#) is an XML Schema designed for expressing technical metadata for textual objects
- Developed at New York University; maintenance transferred to LC
- Includes format specific technical metadata for text, e.g. byte order, character set encoding, font script



Technical metadata for audio and video

- Not as well developed as other technical metadata
- Complexities of file formats requires expertise to develop these
- LC developed XML technical metadata schemas in 2003/2004 for LC Audiovisual Prototype Project used with METS; these were widely implemented because of the lack of other schemas
- Audio and video technical metadata schemas under development by expert organizations
- [Moving Image Catalog](#) (MIC) project is also experimenting with these



Structural metadata

- Supports the intended presentation and use and navigation of an object
- Binds the parts together; expresses relationships between parts of a multipart object
- Examples of structural metadata expressions:
 - METS structMap
 - PREMIS relationship elements
 - EAD hierarchical structure



Rights metadata

- Rights schemas with limited scope
- Full blown Rights Expression Languages (REL) for managing intellectual property rights, particularly by rights owners
- Rights information is not well understood
 - Different laws in different jurisdictions
 - Machine actionable vs. human understandable
- Rights take different forms
 - legal statutes, e.g. copyright
 - contractual rights, e.g. licenses



Rights schemas with limited scope

- METS Rights
 - Access rights for use with METS objects
 - Rights declarations
 - Rights holder
 - Context
- CDL copyright schema
 - Specifically copyrights, not other intellectual property rights
 - Information you need to know to assess copyright status (e.g. creators, rights holders, dates, jurisdiction)
- Note that a new field 542 has been added to MARC 21 with information about copyright to help the cataloger assess the status of the item (based on the CDL work)



Rights schemas with limited scope cont.

- PREMIS Rights
 - Focused on rights for preservation rather than access
 - Revision of PREMIS data dictionary expanded this area
 - Allows for extensibility, i.e. inserting another rights schema
- Creative commons
 - Allows creators to choose a license for their work
 - Simple rights statements that fit a lot of situations
 - An example: [MIC catalog](#)



Rights metadata for specific object types

- PLUS for images
- MPEG-21 REL for moving images, etc.
- ONIX for licensing terms
- Full Rights Expression Languages
 - XRML/MPEG 21
 - ODRL



Exercise

- Provide administrative/technical metadata for the object used in the descriptive metadata exercise

Session 3: Administrative metadata exercise (PREMIS)
Information for students needed to fill in template

Example 1 (sheet music):

This is a digitized version of 3 pages of sheet music. The first page is the cover. There are 3 files, one for each page. The files use the JPEG format.

America's pinch hit march

<http://lcweb2.loc.gov/diglib/ihas/loc.natlib.ihas.200033287/default.html>

Sheet music; 3 pages (2 pages music with cover)

File 1 (cover):

FileID: FN10057

Full path:

<http://lcweb2.loc.gov/natlib/ihas/service/encyclopedia/200033287/0001v.jpg>

File format: image/jpeg

Size: 629507

Software used to access: Macromedia Fireworks MX Version 6.0

File created: 2 January 2008

Application used for creation: Adobe photoshop version CS3

File 2 (1st page of music):

FileID: FN10075

Full path:

<http://lcweb2.loc.gov/natlib/ihas/service/encyclopedia/200033287/0002v.jpg>

File format: image/jpeg

Size: 399565

Software used to access: Macromedia Fireworks MX Version 6.0

File created: 2 January 2008

Application used for creation: Adobe photoshop version CS3

Example 2 (photograph):

This item was digitized in the TIFF format, which is a high quality master format used as a preservation copy. A derivative was made in the JPEG format, which is considered a "service" copy (used for retrieval). There are 2 files: the TIFF master and the JPEG derivative.

52nd Street, New York, N.Y., ca. 1948

<http://lcweb2.loc.gov/natlib/ihas/warehouse/gottlieb/02771/ver01/0001.tif>

File 1 (master TIFF)

FileID: masterd1e30196

Full path:

<http://lcweb2.loc.gov/natlib/ihas/warehouse/gottlieb/02771/ver01/0001.tif>

File format: image/tiff

Size: 60158210

Software used to access: Firefox version 5.0
File created: Feb. 15, 2003
Application used for creation: Image Alchemy PS v1.11
Date created: 20030215

File 2 (derivative JPEG)
FileID: serviced1e30196
Full path:
<http://lcweb2.loc.gov/natlib/ihas/service/gottlieb/02771/ver01/0001v.jpg>
File format: image/jpeg
Size: 167367
Software used to access: Firefox version 5.0
File created: Feb. 15, 2003
Application used for creation: Image Alchemy PS v1.11
Use metadata template provided.

Selected PREMIS elements

- objectIdentifier
 - objectIdentifierType
 - objectIdentifierValue
- objectCharacteristics
 - size
 - format
 - formatDesignation
 - formatName
 - formatVersion
- significantProperties
- creatingApplication
 - creatingApplicationName
 - creatingApplicationVersion
 - dateCreatedByApplication
- environment
 - software
 - swName
 - swVersion
 - swType
- relationship
 - relationshipType
 - relationshipSubType
 - relatedObjectIdentification
 - relatedObjectIdentifierType
 - relatedObjectIdentifierValue
 - relatedObjectSequence

Controlled vocabularies:

RelationshipType:

structural = a relationship between parts of an object

derivation = a relationship where one object is the result of a transformation performed on the related object

RelationshipSubType:

is child of = the object is directly subordinate in a hierarchy to the related object (Note that this is semantically equivalent to "Has parent," which may be preferred by some implementations.

is parent of = the object is directly superior in a hierarchy to the related object (Note that this is semantically equivalent to "Has child," which may be preferred by some implementations.

has sibling = the object shares a common parent with the related object

is part of = the object is contained by the related object

has part = the object contains the related object

source of = the related object is a version of this object created by a transformation

has root = for a representation only, the related object is the file that must be processed first in order to render the representation



4. Metadata syntaxes and containers

Metadata Standards and Applications Workshop



Goals of session

- Understand syntaxes used for encoding information, including HTML, XML and RDF
- Discover how container formats are used for managing digital resources and their metadata



Overview of Syntaxes

- HTML, XHTML: Hypertext Markup Language; eXtensible Hypertext Markup Language
- XML: Extensible Markup Language
- RDF/XML: Resource Description Framework



HTML

- HyperText Markup Language
- HTML 4 is the current standard
- HTML is an SGML (Standard Generalized Markup Language) application conforming to International Standard ISO 8879
- Widely regarded as the standard publishing language of the World Wide Web
- HTML addressed the problem of SGML complexity by specifying a small set of structural and semantic tags suitable for authoring relatively simple documents



XHTML

- XML-ized version of HTML 4.0, tightens up HTML to match XML syntax
 - Requires ending tags, quoted attributes, lower case, etc., to conform to XML requirements
- XHTML is a W3C specification, redefining HTML as an XML implementation, rather than an SGML implementation
- Imposes requirements that are intended to lead to more well-formed, valid XML, easier for browsers to handle

```

<link rel="schema.DC" href="http://purl.org/dc/elements/1.1/" />
<link rel="schema.DCTERMS" href="http://purl.org/dc/terms/" />
<meta name="DC.title" content="Using Dublin Core" />
<meta name="DC.creator" content="Diane Hillmann" />
<meta name="DC.subject" content="documents; Bibliography; Model; meta; Glossary; mark; matching;
refinements; XHTML; Controlled; Qualifiers; Hillmann; mixing; encoding; Diane; Issues; Appendix; elements;
Simple; Special; element; trademark/service; DCMI; Dublin; pages; Section; Resource; Grammatical; Qualified;
XML; Using; Principles; Documents; licensing; OCLC; formal; Usageguide; Roles; Implementing; Contents;
Guidelines; Expressing; Table; Syntax; Content; Element; DC.dot; Home; document; Metadata; RDF/XML;
Website; metadata; privacy; schemes; liability; profiles; Elements; Copyright; Localization; schemas;
HTML/XHTML; Core; Guide; registry; Research; contact; Scope; Projects; languages; Maintenance; Application;
available; Internationalization; HTML; Recommended; link; Purpose; Abstract; AskDCMI; Vocabularies; software;
Storage; Introduction" />
<meta name="DC.description" content="This document is intended as an entry point for users of Dublin Core. For
non-specialists, it will assist them in creating simple descriptive records for information resources (for example,
electronic documents). Specialists may find the document a useful point of reference to the documentation of
Dublin Core, as it changes and grows." />
<meta name="DC.publisher" content="Dublin Core Metadata Initiative" />
<meta name="DC.type" scheme="DCTERMS.DCMIType" content="Text" />
<meta name="DC.format" content="text/html" />
<meta name="DC.format" content="31250 bytes" />
<meta name="DC.identifier" scheme="DCTERMS.URI" content="http://dublincore.org/documents/usageguide/" />

```

An XHTML Example



XML

- Extensible Markup Language
- A '*metamarkup*' language: has no fixed tags or elements
- Strict grammar imposes structure designed to be read by machines
- Two levels of conformance:
 - well-formed--conforms to general grammar rules
 - valid--conforms to particular XML schema or DTD (document type definition)



XML: Extensible Markup Language

- A technical approach to convey meaning with data
- Not a natural language, although uses natural languages
 - <姓名>Louis Armstrong</姓名>
 - <name>Louis Armstrong</name>
- Not a programming language
- Language in the sense of:
 - A limited set of tags defines the elements that can be used to markup data
 - The set of tags and their relationships need to be explicitly defined (e.g., in XML schema)
 - We can build software that uses XML as input and processes them in a meaningful way
- You can define your own markups and schemas



XML is the *lingua franca* of the Web

- Web pages increasingly use at least XHTML
- Business use for data exchange/ messaging
- Family of technologies can be leveraged
 - XML Schema, XSLT, XPath, and XQuery
- Software tools widely available (open source)
 - Storage, editing, parsing, validating, transforming and publishing XML
- *Microsoft Office 2003* supports XML as document format (WordML and ExcelML)
- *Web 2.0* applications based on XML (AJAX, Semantic Web, Web Services, etc.)



An XML Schema Defines:

- What elements may be used
- Of which types
- Any attributes
- In which order
- Optional or compulsory
- Repeatability
- Subelements
- Logic

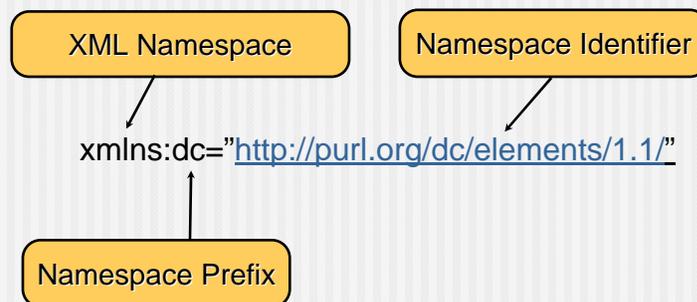


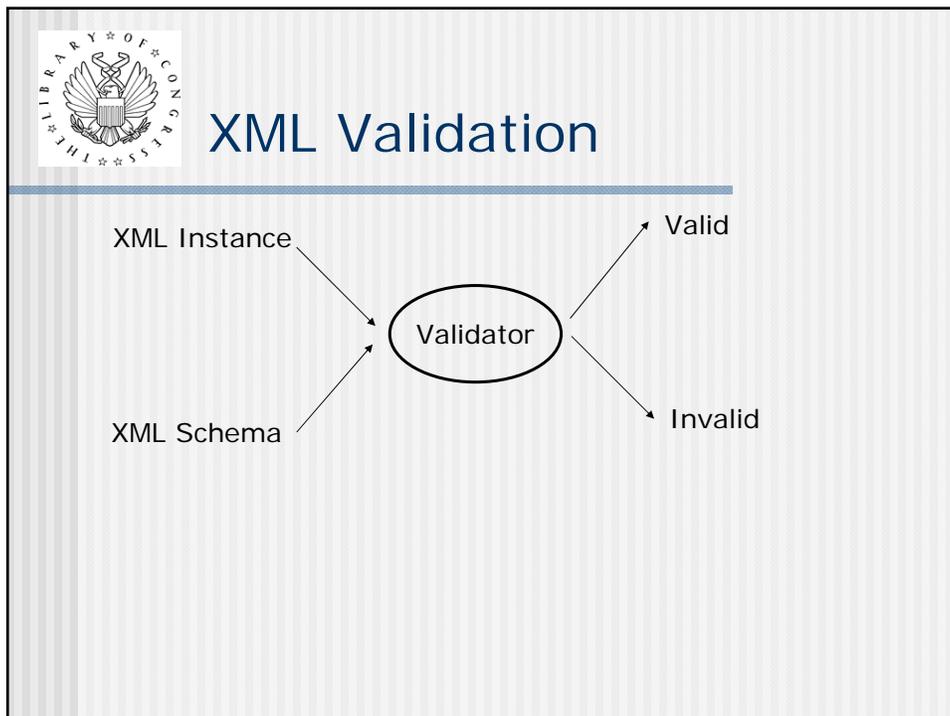
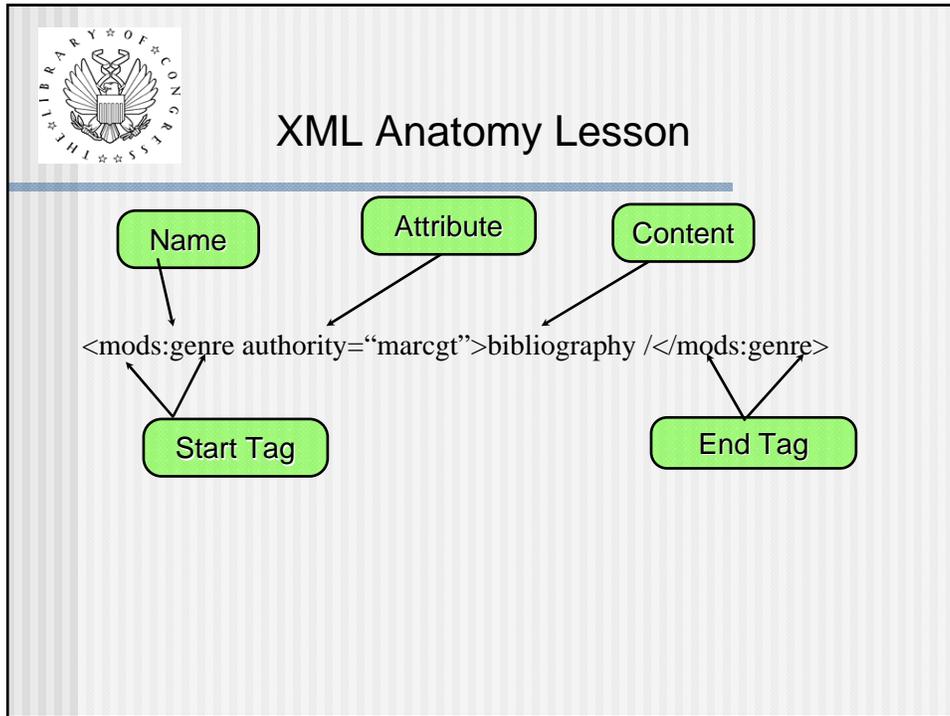
Anatomy of an XML Record

- XML declaration--prepares the processor to work with the document and states the XML version
- Namespaces (uses `xmlns:prefix` and a URI to attach a prefix to each element and attribute)
 - Distinguishes between elements and attributes from different vocabularies that might share a name (but not necessarily a definition) using association with URIs
 - Groups all related elements from an application so software can deal with them
 - The URIs are the standardized bit, not the prefix, and they don't necessarily lead anywhere useful, even if they look like URLs



XML Namespaces





XML Schema Example

```
<xs:element name="software" minOccurs="0"
  maxOccurs="unbounded">
  <xs:complexType>
    <xs:sequence>
      <xs:element name="swName" minOccurs="1"
        maxOccurs="1" type="xs:string"></xs:element>
      <xs:element name="swVersion" minOccurs="0"
        maxOccurs="1" type="xs:string"></xs:element>
      <xs:element name="swType" minOccurs="1"
        maxOccurs="1" type="xs:string"></xs:element>
      <xs:element name="swOtherInformation"
        minOccurs="0" maxOccurs="unbounded" type="xs:string">
      </xs:element>
      <xs:element name="swDependency" minOccurs="0"
        maxOccurs="unbounded" type="xs:string"> </xs:element>
    </xs:sequence>
  </xs:complexType>
</xs:element>
```



Will the following XML instance validate?

```
<software>
  <swName>Windows</swName>
  <swVersion>2000</swVersion>
  <swType>Operating System</swType>
</software>
```

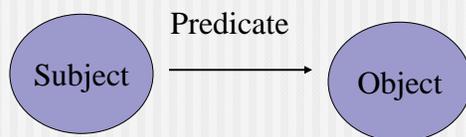
How about this?

```
<swVersion>2000</swVersion>
```



Resource Description Framework

- Application for describing and exchanging metadata
- Structure consists of "triples"
- Designed to be read by computers, not humans
- An ontology language to support semantic interoperability—understanding meanings
- Considered an essential part of the Semantic Web
- Can be expressed using XML



<http://www.w3.org/RDF>



Some RDF Concepts

- A **Resource** is anything you want to describe
- A **Class** is a category; it is a set that comprises individuals
- A **Property** is a Resource that has a name, such as "creator" or "homepage"
- A **Property value** is the value of a Property, such as "Barack Obama" or "<http://dublincore.org>" (note that a property value can be another resource)



RDF Statements

- The combination of a **Resource**, a **Property**, and a **Property value** forms a *Statement* (known also as the subject, predicate and object of a Statement), also known as “triples”
- An example *Statement*: "The editor of <http://dublincore.org/documents/usageguide/> is Diane Hillmann"
- The subject of the statement above is: <http://dublincore.org/documents/usageguide/>
- The predicate is: editor
- The object is: Diane Hillmann



RDF and OWL

- RDF does not have the language to specify all relationships
- Web Ontology Language (OWL) can specify richer relationships, such as equivalence, inverse, unique
- RDF and OWL may be used together
- RDFS: a syntax for expressing relationships between elements

An RDF/XML Example (with DC)

```

<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://www.dlib.org">
    <dc:title>D-Lib Program - Research in Digital Libraries</dc:title>
    <dc:description>The D-Lib program supports the community of people
    with research interests in digital libraries and electronic
    publishing.</dc:description>
    <dc:publisher>Corporation For National Research Initiatives</dc:publisher>
    <dc:date>1995-01-07</dc:date>
    <dc:subject>
      <rdf:Bag>
        <rdf:li>Research; statistical methods</rdf:li>
        <rdf:li>Education, research, related topics</rdf:li>
        <rdf:li>Library use Studies</rdf:li>
      </rdf:Bag>
    </dc:subject>
    <dc:type>World Wide Web Home Page</dc:type>
    <dc:format>text/html</dc:format>
    <dc:language>en</dc:language>
  </rdf:Description>
</rdf:RDF>

```

Note
unordered
list



Overview of container formats

- A container format is needed to package together all forms of metadata and digital content
- Use of a container is compatible with and an implementation of the OAIS information package concept
- METS: packages metadata with objects or links to objects and defines structural relationships
- MPEG 21 DID: represents digital objects using a flexible and expressive model



Metadata Encoding & Transmission Standard *METS*

- Developed by the Digital Library Federation, maintained by the Library of Congress
- "... an XML document format for encoding metadata necessary for both management of digital library objects within a repository and exchange of such objects between repositories (or between repositories and their users)."
- Records the (possibly hierarchical) structure of digital objects, the names and locations of the files that comprise those objects, and the associated metadata

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



What is METS used for?

- To package metadata with digital object in XML syntax
- For retrieving, storing, preserving, serving resource
- For interchange of digital objects with metadata
- As information package in a digital repository (may be a unit of storage or a transmission format)



Characteristics of METS

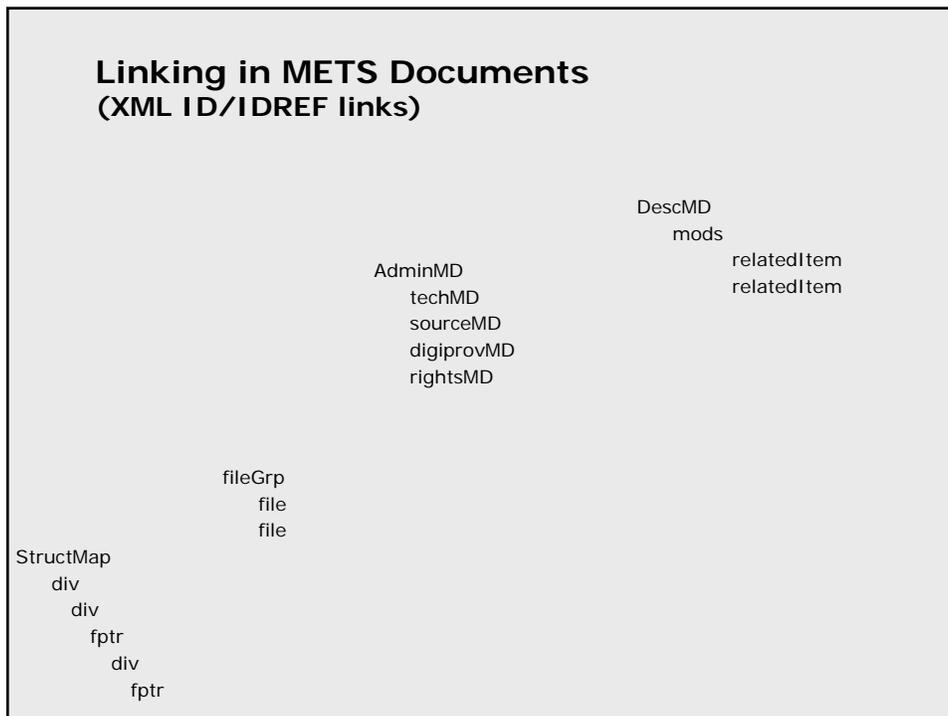
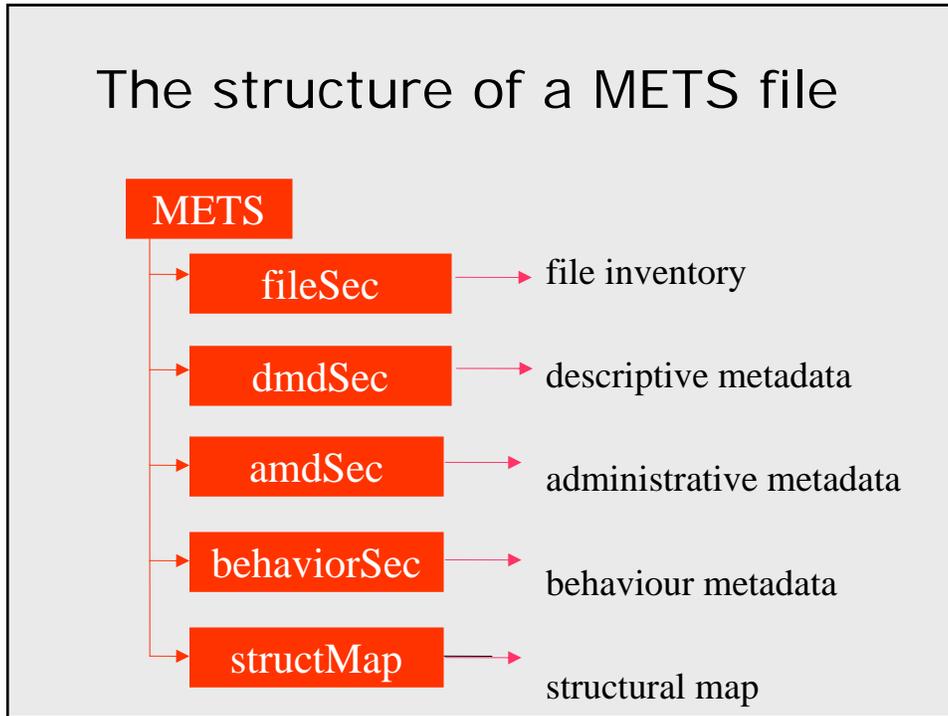
- Open non-proprietary standard
- Extensible
- Modular
- Developed by the digital library community

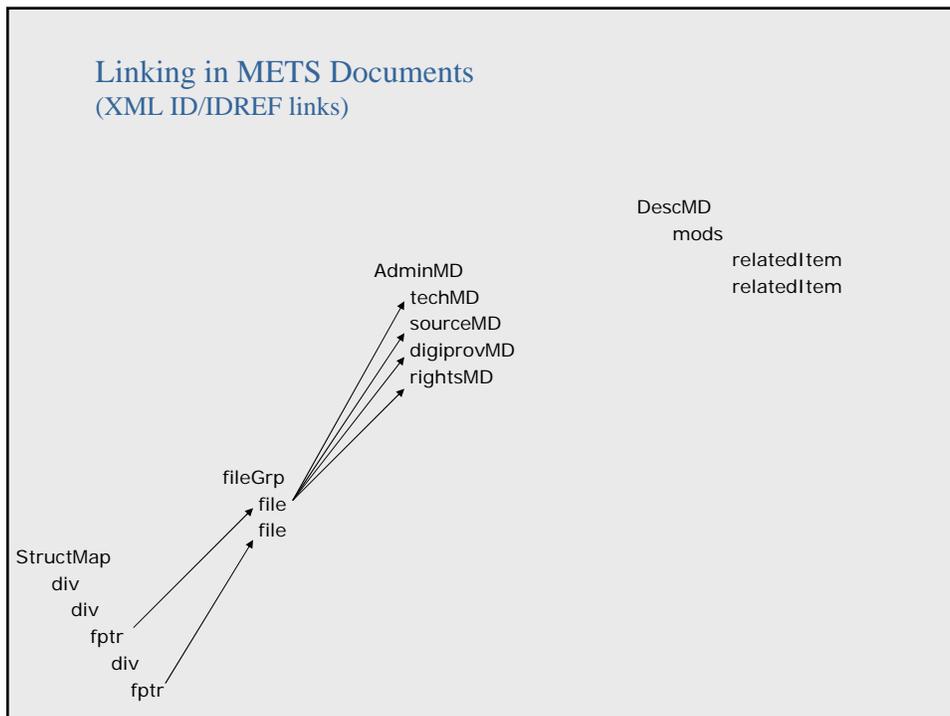
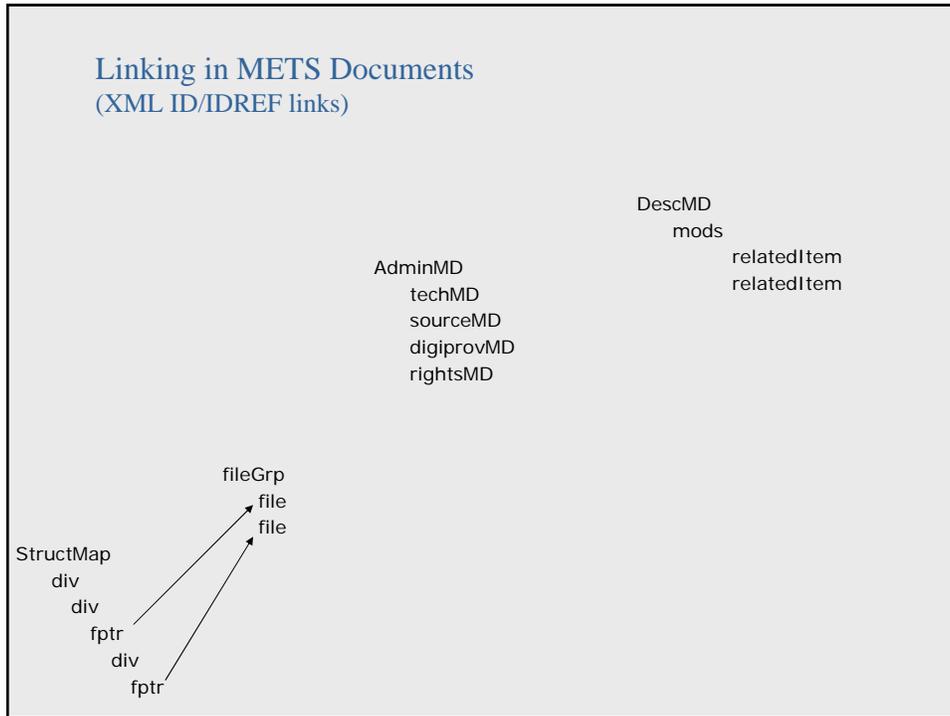


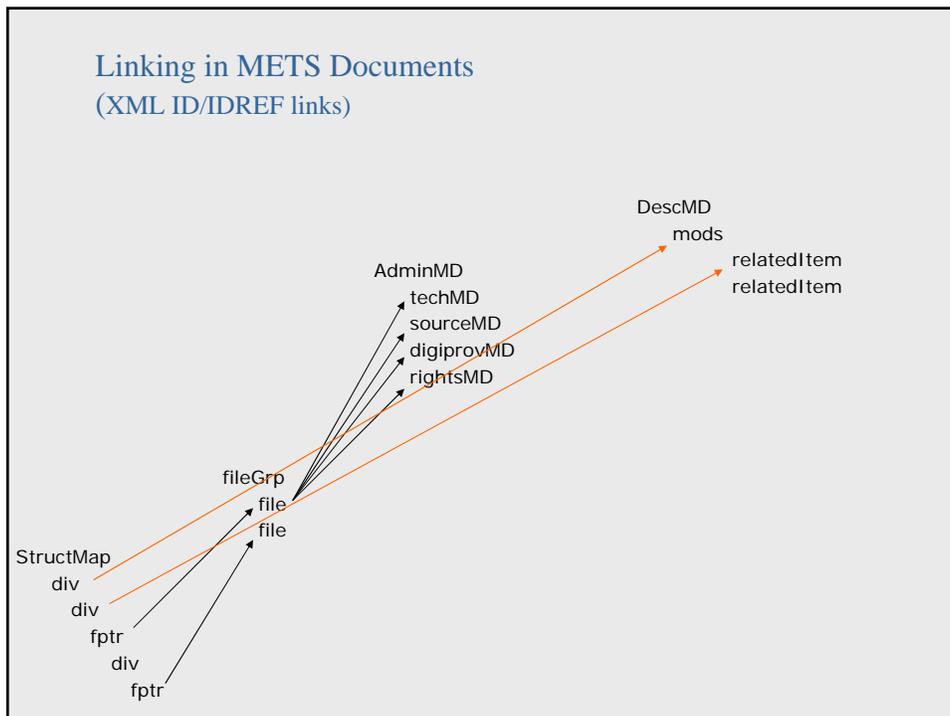
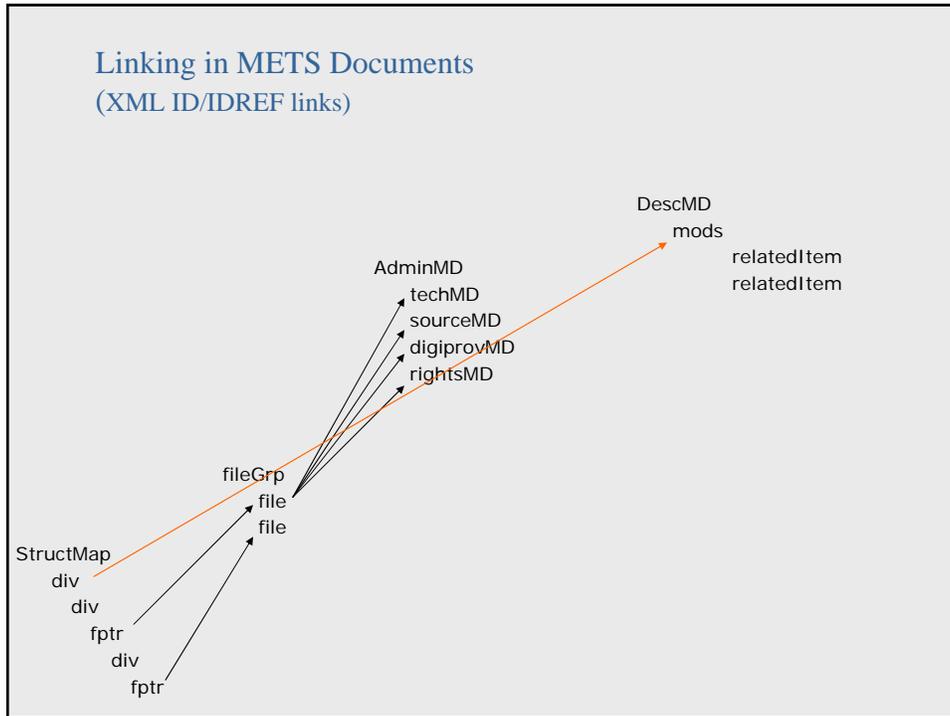
METS Sections

Defined in METS schema for navigation & browsing

1. Header (XML Namespaces)
2. File inventory,
3. Structural Map & Links
4. Descriptive Metadata (not part of METS but uses an externally developed descriptive metadata standard, e.g. MODS)
5. Administrative Metadata (points to external schemas):
 1. Technical, Source
 2. Digital Provenance
 3. Rights









METS extension schemas

- “wrappers” or “sockets” where elements from other schemas can be plugged in
- Provides extensibility
- Uses the XML Schema facility for combining vocabularies from different Namespaces
- Endorsed extension schemas:
 - Descriptive: MODS, DC, MARCXML
 - Technical metadata: MIX (image); textMD (text)
 - Preservation related: PREMIS



Descriptive Metadata Section (dmdSec)

Two methods: Reference and Wrap

```
<mets>  
  <dmdSec></dmdSec>  
  <fileSec></fileSec>  
  <structMap></structMap>  
</mets>
```



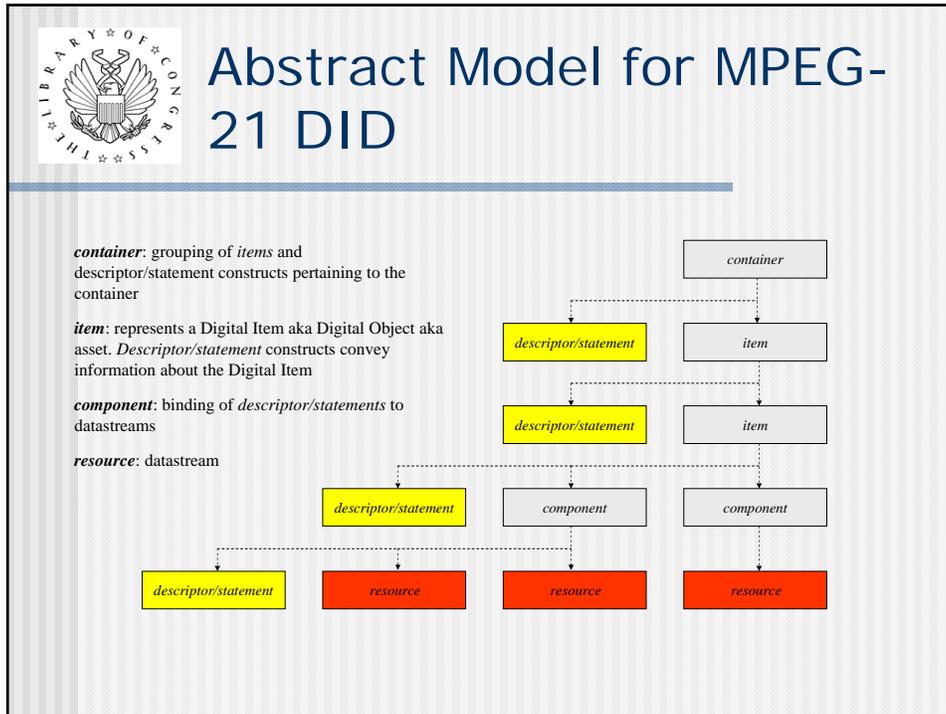
METS examples

- METS with MODS
 - [Recorded event](#)
 - [Patriotic melodies](#)
- METS with MODS, PREMIS and MIX
 - [Portrait of Louis Armstrong](#) (XML)
 - [Portrait of Louis Armstrong](#) (presentation)



MPEG-21 Digital Item Declaration (DID)

- ISO/IEC 21000-2: Digital Item Declaration
- An alternative to represent Digital Objects
- Starting to get supported by some repositories, e.g., aDORe, DSpace, Fedora
- A flexible and expressive model that easily represents compound objects (recursive “item”)
- MPEG DID is an ISO standard and has industry support, but is often implemented in a proprietary way and standards development is closed; METS is open source and developed by open discussion, mainly cultural heritage community



Exercise

- Encode your resource in DC and MODS using XML
- Use the template forms provided

*THIS PAGE INTENTIONALLY LEFT BLANK
FOR DOUBLE SIDED COPY*

Exercise for Session 4: XML Syntax
DC Template

```
<?xml version="1.0" encoding="UTF-8"?>
<metadata
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  http://dublincore.org/schemas/xmls/qdc/2003/04/02/dcterms.xsd
  xmlns:dc="http://purl.org/dc/elements/1.1/"
  xmlns:dcterms="http://purl.org/dc/terms/"
  xmlns:dcmitype="http://purl.org/dc/dcmitype/">

  <dc:title xml:lang="  ">

</dc:title>

<dc:creator>

</dc:creator>

<dc:type>

</dc:type>

<dc:publisher>

</dc:publisher>

<dc:date>

</dc:date>

<dc:format>

</dc:format>

<dc:identifier>

</dc:identifier>

<dc:subject xsitype="dcterms:LCSH">

</dc:subject>

<dc:relation>
```

```
</dc:relation>
```

```
</metadata>
```

Exercise for Session 4: XML syntax
MODS Template

```
<?xml version="1.0" encoding="UTF-8"?>
<mods:mods version="3.2" ID="MODS1"
xsi:schemaLocation="http://www.loc.gov/mods/v3
http://www.loc.gov/standards/mods/mods.xsd">

  <mods:titleInfo>
    <mods:title>

    </mods:title>
    <mods:subTitle>

    </mods:subTitle>
  </mods:titleInfo>

  <mods:name type="personal">
    <mods:namePart>

    </mods:namePart>
    <mods:role>
      <mods:roleTerm authority="marcrelator" type="text">

      </mods:roleTerm>
    </mods:role>
  </mods:name>

  <mods:typeOfResource>

  </mods:typeOfResource>

  <mods:originInfo>
    <mods:place>
      <mods:placeTerm>

      </mods:placeTerm>
    </mods:place>

    <mods:publisher>

    </mods:publisher>

    <mods:dateIssued>
```

```
</mods:dateIssued>
<mods:dateCreated>

</mods:dateCreated>
</mods:originInfo>
<mods:physicalDescription>
  <mods:form authority=" " >sheet music</mods:form >
  <mods:extent>

</mods:extent>
</mods:physicalDescription>
<mods:accessCondition>

</mods:accessCondition>
<mods:subject authority="lcsch">
  <mods:topic>

  </mods:topic>
  <mods:genre>

</mods:genre>
</mods:subject>
<mods:subject authority=" " >
  <mods:name>
    <mods:namePart>

    </mods:namePart>
  </mods:name>
</mods:subject>
<mods:physicalLocation>
  <mods:url>

</mods:url>
</mods:physicalLocation>
<mods:relatedItem type="host">
  <mods:titleInfo>
    <mods:title>
```

```
    </mods: title>  
  </mods: titleInfo>  
</mods: relatedItem>  
  
</mods: mods>
```

```
<?xml version="1.0" encoding="UTF-8"?>
<mods:mods version="3.2" ID="MODS1"
xsi:schemaLocation="http://www.loc.gov/mods/v3
http://www.loc.gov/standards/mods/mods.xsd">

  <mods:titleInfo>
    <mods:title>

      </mods:title>
    <mods:subTitle>

      </mods:subTitle>
    </mods:titleInfo>

    <mods:name type="personal">
      <mods:namePart>

        </mods:namePart>
      <mods:role>
        <mods:roleTerm authority="marcrelator" type="text">

          </mods:roleTerm>
        </mods:role>
      </mods:name>

      <mods:typeOfResource>

        </mods:typeOfResource>

      <mods:originInfo>
        <mods:place>
          <mods:placeTerm>

            </mods:placeTerm>
          </mods:place>

          <mods:publisher>

            </mods:publisher>

          <mods:dateIssued>

            </mods:dateIssued>
```

```
<mods:dateCreated>

</mods:dateCreated>
</mods:originInfo>

<mods:physicalDescription>
  <mods:form authority=" " >sheet music</mods:form >

  <mods:extent>

</mods:extent>
</mods:physicalDescription>

<mods:accessCondition>

</mods:accessCondition>

<mods:subject authority="lcsch">
  <mods:topic>

  |</mods:topic>
  <mods:genre>

</mods:genre>
</mods:subject>

<mods:subject authority=" " >
  <mods:name>
    <mods:namePart>

      </mods:namePart>
    </mods:name>
  </mods:subject>

<mods:physicalLocation>
  <mods:url>

</mods:url>
</mods:physicalLocation>

<mods:relatedItem type="host">
  <mods:titleInfo>
    <mods:title>
```

```
    </mods: title>  
  </mods: titleInfo>  
</mods: relatedItem>  
  
</mods: mods>
```

```
<?xml version="1.0" encoding="UTF-8"?>
<metadata
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  http://dublincore.org/schemas/xmls/qdc/2003/04/02/dcterms.xsd
  xmlns:dc="http://purl.org/dc/elements/1.1/"
  xmlns:dcterms="http://purl.org/dc/terms/"
  xmlns:dcmitype="http://purl.org/dc/dcmitype/">

  <dc:title xml:lang="  ">

</dc:title>

<dc:creator>

</dc:creator>

<dc:type>

</dc:type>

<dc:publisher>

</dc:publisher>

<dc:date>

</dc:date>

<dc:format>

</dc:format>

<dc:identifier>

</dc:identifier>

<dc:subject xsitype="dcterms:LCSH">

</dc:subject>

<dc:relation>

</dc:relation>
```

</metadata>



5. Applying metadata standards: Application profiles

Metadata Standards and Applications Workshop



Goals of Session

- Learn how metadata standards are applied and used:
 - Learn about the concept and use of application profiles
 - Learn about how different metadata standards may be used together in digital library applications



Overview of session

- Use of Application profiles
 - Dublin Core
 - METS
 - MODS
- Case study: using metadata standards together based on an application profile



Application profiles

- The set of metadata elements, policies, and guidelines defined for a particular application, implementation, or object type
- A declaration of the metadata terms an organization, information resource, application, or user community uses in its metadata
- Documents metadata standards used in instances, including schemas and controlled vocabularies, policies, required elements, etc.
- Called “application profile” or just “profile”



Function of Application Profiles

- **Many metadata standards are sufficiently flexible that they need a mechanism to impose some constraints**
- **Express the decisions made for a project in machine-readable form (XML or RDF)**
- **Refining**
 - A narrower interpretation of a standard to suit your project
- **Combining**
 - Mixing tags from various different standards
- **Benefit**
 - Machine-readable
 - Enforceable



Dublin Core application profiles: term usages

- *Identifying attributes*
 - **Term URI, Name, Label, Defined By**
- *Definitional attributes*
 - **Definition, Comments, Type of Term**
- *Relational attributes*
 - **Refines, Refined By, Encoding Scheme For,**
 - **Uses Encoding Scheme, Similar To**
- *Constraints*
 - **Obligation, Condition, Datatype, Occurrence**



DCAP: mixing and matching metadata

- Currently a great deal of contention and discussion about the technical issues around re-use of properties
- Determining reusability of terms in a DCAP
 - Is the term a real “property” and defined as such within the source schema?
 - Is the term declared properly, with a URI and adequate documentation and support?
 - Properties whose meaning is determined by its place in a hierarchy are not appropriate for reuse without reference to the hierarchy.
- Example: [Collection Description AP](#)



Singapore Framework

- A Framework for designing metadata applications for maximum interoperability
- Defines a set of descriptive components that are necessary for documenting an Application Profile
- Forms a basis for reviewing Dublin Core application profiles
- Relates APs to standard domain models and Semantic Web standards
- <http://dublincore.org/documents/singapore-framework/>



Application profile development for RDA

- A DCMI/RDA Work Group has been formed to define RDA entities as an RDF vocabulary
- This will constitute an application profile of RDA according to the DCMI application profile framework
- RDA elements will be given URIs
- Controlled vocabularies in RDA will be represented in SKOS/RDF



What is a METS Profile?

- Description of a class of METS documents
 - provides document authors and programmers guidance to create and process conformant METS documents
- XML document using a schema
 - Expresses the requirements that a METS document must satisfy
- “Data standard” in its own right
 - A sufficiently explicit METS Profile may be considered a “data standard”
- METS Profiles are output in human-readable prose and not intended to be “machine actionable” (but they use a standard XML schema)



Components of a METS Profile

1. Unique URI
2. Short Title
3. Abstract
4. Date and time of creation
5. Contact Information
6. Related profiles
7. Extension schemas
8. Rules of description
9. Controlled vocabularies
10. Structural requirements
11. Technical requirements
12. Tools and applications
13. Sample document



Case study of a METS profile

- [LC Audio Compact Disc Profile](#)
- Features:
 - Specifies MODS for descriptive metadata
 - Specifies description rules as AACR2
 - Specifies controlled vocabularies used in various elements
 - dmdSec requirements 2 and 3 specify use of relatedItem type="constituent" if there are multiple works on the CD
 - Specifies how to detail the physical structure, whether multiple CDs or multiple tracks on a CD (structMap requirements 2 and 3)



MODS profiles

- Some applications are establishing MODS profiles to document usage, required elements, controlled vocabularies used, etc.

Some examples:

- [DLF Aquifer MODS profile](#): to establish implementation guidelines for rich shared metadata for cultural heritage materials
- British Library electronic journal MODS profile



Using metadata standards together: a case study

- METS can be used to package together the metadata with the objects
 - METS allows for use of any XML metadata schema in its extensions
 - MODS works well with METS for descriptive metadata and can be associated with any level of the description
 - Technical metadata can be inserted and associated with specific files

<dmdSec> with MODS Extension Schema

```

<mets:mets>
...
<mets:dmdSec>
  <mets:mdWrap>
    <mets:xmlData>
      <mods:mods></mods:mods>
    </mets:xmlData>
  </mets:mdWrap>
</mets:dmdSec>
...
</mets:mets>
    
```

Descriptive metadata section

MODS data contained inside the metadata wrap section

Use of prefixes before element names to identify schema

<dmdSec> with <mods:relatedItem>

```

<mets:mets>
...
<mets:dmdSec>
  <mets:mdWrap>
    <mets:xmlData>
      <mods:mods>
        <mods:relatedItem type="constituent">
          <mods:relatedItem type="constituent"></mods:relatedItem>
        </mods:relatedItem>
      </mods:mods>
    </mets:xmlData>
  </mets:mdWrap>
</mets:dmdSec>
...
</mets:mets>
    
```

The MODS relatedItem element can be nested and can be used to express a hierarchy.

```

<mods:mods>
  <mods:titleInfo>
    <mods:title>Bernstein conducts Beethoven </mods:title>
  </mods:titleInfo>
  <mods:name>
    <mods:namePart>Bernstein, Leonard</mods:namePart>
  </mods:name>
  <mods:relatedItem type="constituent">
    <mods:titleInfo>
      <mods:title>Symphony No. 5</mods:title>
    </mods:titleInfo>
    <mods:name>
      <mods:namePart>Beethoven, Ludwig van</mods:namePart>
    </mods:name>
    <mods:relatedItem type="constituent">
      <mods:titleInfo>
        <mods:partName>Allegro con moto</mods:partName>
      </mods:titleInfo>
    </mods:relatedItem>
    <mods:relatedItem type="constituent">
      <mods:titleInfo>
        <mods:partName>Adagio</mods:partName>
      </mods:titleInfo>
    </mods:relatedItem>
  </mods:relatedItem>
</mods:mods>

```




Use of MODS relatedItem type="constituent"

- A first level child element to MODS
- **relatedItem** element uses MODS content model
 - titleInfo, name, subject, physicalDescription, note, etc.
- Makes it possible to create rich analytics for contained works within a MODS record
- Repeatable and nestable recursively
 - Making it possible to build a hierarchical tree structure
- Makes it possible to associate descriptive data with any structural element

METS 2 Hierarchies: Logical & Physical

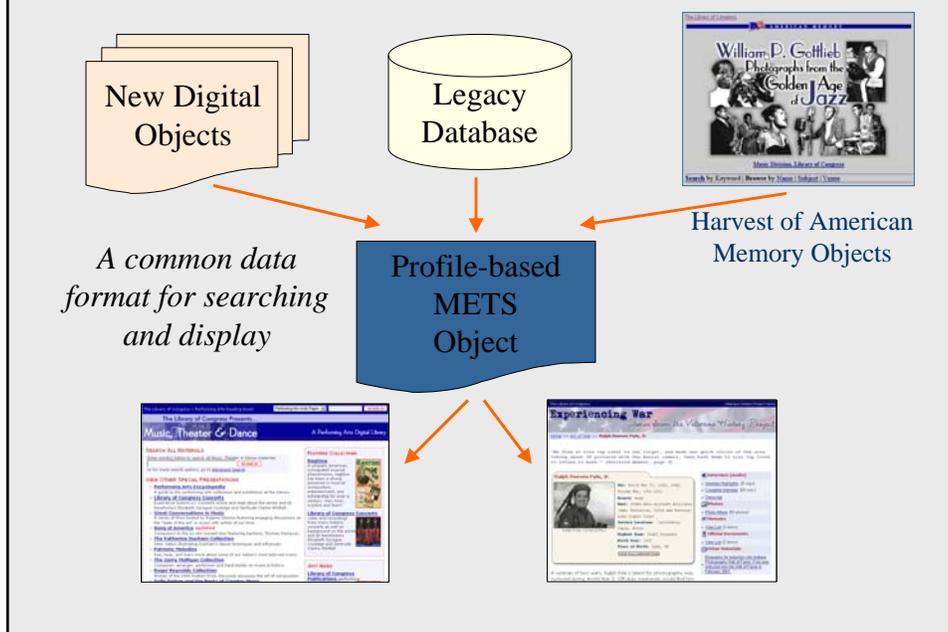
```

<mets:mets>
  <mets:dmdSec>
    <mets:mdWrap>
      <mets:xmlData>
        <mods:mods>
          <mods:relatedItem>
            <mods:relatedItem></mods:relatedItem>
          </mods:relatedItem>
        </mods:mods>
      </mets:xmlData>
    </mets:mdWrap>
  </mets:dmdSec>
  <mets:fileSec></mets:fileSec>
  <mets:structMap>
    <mets:div>
      <mets:div></mets:div>
    </mets:div>
  </mets:structMap>
</mets:mets>
    
```

Hierarchy to represent "logical" structure (nested relatedItems)

Hierarchy to represent "physical" structure (nested div elements)

Multiple Inputs to Common Data Format





Example: Using an application profile

- METS Photograph Profile
- William P. Gottlieb Collection
Portrait of Louis Armstrong
- Photographic object

Convert file of 1600 MARC records, using *marc4j*, to XML modsCollection (single file).

Used XSLT stylesheet to create 1600 records conforming to the METS photograph profile.



Logical & Physical Relationships

Logical (MODS)

```
<mods:mods ID="ver01">
  <mods:titleInfo>
    <mods:title>Original Work</mods:title>
  </mods:titleInfo>
  <mods:relatedItem type="otherVersion" ID="ver02">
    <mods:titleInfo>
      <mods:title>Derivative Work</mods:title>
    </mods:titleInfo>
  </mods:relatedItem>
  <mods:relatedItem type="otherVersion" ID="ver03">
    <mods:titleInfo>
      <mods:title>Derivative Work</mods:title>
    </mods:titleInfo>
  </mods:relatedItem>
</mods:mods>
```

mods:mods and *mods:relatedItem type="otherVersion"* elements create a sequence of 3 nodes

div TYPE="photo:version" elements correspond to the 3 nodes using a logical sequence of ID to DMDID relationships

Physical (METS structMap)

```
<structMap>
  <div TYPE="photo:photoObject" DMDID="MODS1">
    <div TYPE="photo:version" DMDID="ver01">
      <div TYPE="photo:image">
        <ptr FILEID="FN10081"/>
      </div>
    </div>
    <div TYPE="photo:version" DMDID="ver02">
      <div TYPE="photo:image">
        <ptr FILEID="FN10090"/>
      </div>
    </div>
    <div TYPE="photo:version" DMDID="ver03">
      <div TYPE="photo:image">
        <ptr FILEID="FN1009F"/>
      </div>
    </div>
  </div>
</structMap>
```



Using a METS profile-based approach

- Ability to model complex library objects
- Use of modern, non-proprietary software tools
- Use of XSLT for...
 - Legacy data conversion
 - Batch METS creation and editing
 - Web displays and behaviors
- Use of a common syntax – XML
 - For data creation, editing, storage and searching

continued...



Using a METS profile-based approach (cont.)

- Creation of multiple outputs from XML
 - HTML/XHTML for Web display; PDF for printing
- Ease of editing
 - Single records or selected batches of records
- Ability to validate data
- Ability to aggregate disparate data sources
- Ease of data management and publishing
- Excellent positioning for the future
 - New web applications (Web 2.0)
 - Repository submission and OAI harvesting
 - Cooperative projects (to test interoperability)



Closing thoughts on application profiles

- Many metadata standards are sufficiently flexible that profiling is necessary
- Documenting what is used in an application will simplify and enhance data presentation, conversion from other sources, ability to provide different outputs
- Constraining a metadata standard by specifying what is used and how facilitates data exchange and general interoperability



Exercise: critique an application profile

- University of Maryland Descriptive Metadata
<http://www.lib.umd.edu/dcr/publications/taglibrary/umdm.html>
- UVa DescMeta
<http://lib.virginia.edu/digital/metadata/descriptive.html>
- Texas Digital Library profile for electronic theses and dissertations
<http://metallogger.files.wordpress.com/2007/06/tdl-etd-mods-profile.pdf>



Exercise: Questions to address

- Does the profile define its user community and expected uses?
- How usable would the profile be for a potential implementer?
- How (well) does the profile specify element/term usage?
- How (well) does the profile define and manage controlled vocabularies?
- Does the profile use existing metadata standards?
- Are there key anomalies, omissions, or implementation concerns?

University of Maryland Descriptive Metadata Tag Library

Authored by [Jennifer O'Brien Roper](#) Jennifer O'Brien Roper

Presentation designed by Sean Daugherty

This tag library provides current documentation for all descriptive metadata for digital objects in the University of Maryland Digital Collections. The University of Maryland Descriptive Metadata (UMDM) DTD and element set are based on Descriptive Metadata (Descmeta) DTD (Version 1.06) authored by Daniel McShane and Perry Roland at the University of Virginia. UMDM was designed within the local context, yet allows data to be adaptable to comply with national standards such as Dublin Core, VRA Core, or MODS.

The tag library contains definitions, notation of parent and child elements, attributes, input standards, and examples for each element. Wherever possible, links to appropriate external standards are present. UMDM is an evolving standard, and while stable, is not static. As new projects and material types are added to the repository, the standard may be modified.

- [<address>](#)
- [<addressLine>](#)
- [<agent>](#)
- [<availability>](#)
- [<bibRef>](#)
- [<bibScope>](#)
- [<century>](#)
- [<color>](#)
- [<corpName>](#)
- [<covPlace>](#)
- [<covTime>](#)
- [<culture>](#)
- [<date>](#)
- [<dateRange>](#)
- [<decade>](#)
- [<descMeta>](#)
- [<description>](#)
- [<extent>](#)
- [<extPtr>](#)

- [<extRef>](#)
- [<form>](#)
- [<geogName>](#)
- [<identifier>](#)
- [<imprint>](#)
- [<language>](#)
- [<linkGrp>](#)
- [<mediaType>](#)
- [<other>](#)
- [<persName>](#)
- [<physDesc>](#)
- [<pid>](#)
- [<price>](#)
- [<relation>](#)
- [<relationships>](#)
- [<repository>](#)
- [<rights>](#)
- [<series>](#)
- [<size>](#)
- [<style>](#)
- [<subject>](#)
- [<title>](#)
- [<version>](#)

[View Document Type Definition \(DTD\)](#)

Required base elements:

Finite:

- [<covPlace>](#)
- [<covTime>](#)
- [<mediaType>](#)
- [<physDesc>](#)
- [<pid>](#)
- [<relationships>](#)
- [<repository>](#)
- [<rights>](#)

Repeatable:

- [<culture>](#)

- [<description>](#)
- [<subject>](#)
- [<title>](#)

Optional base elements:

Repeatable:

- [<agent>](#)
- [<language>](#)
- [<style>](#)
- [<identifier>](#)

<address>

Definition:

Contains a postal or other address, for example of an organization, individual, or internet resource. Addresses may be encoded either as a sequence of lines, or using any sequence of address component elements.

Must Contain:

Repeatable:

- [<addressLine>](#)

May Occur Within:

- [<imprint>](#)
- [<repository>](#)

Input Standard(s):

Address as found on item, or as found from a reliable source.

Attributes:

id	ID	Optional
label	CDATA	Optional
type	CDATA	Optional
xml:lang	CDATA	Optional

Example:

```
<descMeta>
<relationships>
[... ]
<relation type="isPartOf">Digital Collections @ UM</relation>
```

```
</relationships>
<repository>
<corpName>Historical Manuscripts</corpName>
<address>
<addressLine>http://www.lib.umd.edu/histmss/</addressLine>
</address>
</repository>
[... ]
</descMeta>
```

<addressLine>

Definition:

Individual line of an address, postal or otherwise.

Must Contain:

Finite:

- PCDATA

May Occur Within:

- [<address>](#)

Input Standard(s):

Create as necessary to separate discrete portions of an address.

Attributes:

id	ID	Optional
label	CDATA	Optional
type	CDATA	Optional
xml:lang	CDATA	Optional

Example:

```
<descMeta>
<relationships>
[... ]
<relation type="isPartOf">Performing Arts Digital Videos</relation>
</relationships>
<repository>
<corpName>Nonprint Media Services</corpName>
<address>
<addressLine>Hornbake Library</addressLine>
<addressLine>College Park, MD 20742-7011</addressLine>
</address>
</repository>
[... ]
</descMeta>
```

<agent>

Definition:

An entity primarily responsible for the creation or distribution of the intellectual content of a resource.

Must Contain At Least One of the Following:

Finite:

- [<corpName>](#)
- [<other>](#) *Use of this element is restricted. See the Usage Rules for Specific Collections for this element.*
- [<persName>](#)

May Occur Within:

- [<bibRef>](#)
- [<descMeta>](#)
- [<imprint>](#)

Input Standard(s):

Attribute type must be selected.

Attributes:

id	ID	Optional
label	CDATA	Optional
normal	CDATA	Optional
role	"author", "illustrator", "publisher", "broadcaster", "composer", "conductor", "director", "lyricist", "narrator", "performer", "producer", "speaker", "storyteller"	Optional
type	"creator", "contributor", "provider"	REQUIRED
xml:lang	CDATA	Optional

Specific Usage Rules for Collections

The following elements, attributes, and values may be used with this element in these specific conditions.

Prange Digital Children's Book Collection

- `<agent type="creator" role="author">`
- `<agent type="contributor" role="illustrator">`
- `<agent type="provider" role="publisher">`

Example:

```

<descMeta>
[... ]
<title type="main">Agamemnon</title>
<agent type="creator">
<persName>Aeschylus</persName>
</agent>
<agent type="creator">
<persName>Harrison, Tony, 1937-</persName>
</agent>
<agent type="provider">
<corpName>National Theatre (Great Britain)</corpName>
</agent>
<agent type="provider">
<corpName>Channel Four (Great Britain)</corpName>
</agent>
<covPlace>
<geogName>Greece</geogName>
</covPlace>
[... ]
</descMeta>

```

<availability>

Definition:

Supplies information about the availability of a resource.

Must Contain:

Repeatable:

- [<date type=\[...\]>](#)
- and/or [<price>](#)

May Occur Within:

- [<imprint>](#)

Attributes:

id	ID	Optional
label	CDATA	Optional
type	CDATA	Optional
xml:lang	CDATA	Optional

Example:

```

<descMeta>
[... ]
<relationships>
<relation label="collection" type="isPartOf" >Treasure of World's Fair Art &
Architecture </relation>
<relation label="fair" type="isPartOf">Centennial Exhibition (1876: Philadelphia,
Pa.)</relation>
<relation type="isPartOf">
<bibRef>
<imprint>
<geogName>London</geogName>
<agent type="provider">
<corpName>John Murray</corpName>
</agent>
<address>

```

```
<addressLine>Albemarle-street</addressLine>
</address>
<date era="ad" type="exact">1818</date>
<availability>
<price units="shillings">10</price>
</availability>
</imprint>
</bibRef>
</relation>
</relationships>
[... ]
</descMeta>
```

<bibRef>

Definition:

Bibliographic reference. Contains a loosely-structured bibliographic citation of which the sub-components may or may not be explicitly tagged.

Must Contain:

Any of the following

Finite:

- [<bibScope>](#)
- [<repository>](#)
- [<version>](#)

Repeatable:

- [<agent>](#)
- [<date type="\[...\]">](#)
- [<imprint type="\[...\]">](#)
- [<series>](#)
- [<title type="\[...\]">](#)

May Occur Within:

- [<corpName>](#)
- [<description>](#)
- [<linkGrp>](#)
- [<other>](#)
- [<persName>](#)
- [<relation>](#)

Input Standard(s):

From the resource itself or an accepted authoritative source.

Attributes:

id	ID	Optional
label	CDATA	Optional
type	CDATA	Fixed
xml:lang	CDATA	Optional

Example:

```

<descMeta>
[... ]
<physDesc>
<color>color</color>
<extent units="image">1</extent>
</physDesc>
<relationships>
<relation label="collection" type="isPartOf">University AlbUM</relation>
<relation label="citation" type="isPartOf">UM Call Number: ARCH REF GV 885.43.U535
U54 2004
<bibRef>Ungrady, Dave. Legends of Maryland basketball. Champaign, IL: Sports
Publishing, 2004.</bibRef>
</relation>
</relationships>
<repository>
<corpName>University Archives</corpName>
</repository>
[... ]
</descMeta>

```

<bibScope>

Definition:

Scope of citation. Defines the scope of a bibliographic citation, for example as a list of page numbers, or a named subdivision of a larger work.

Must Contain:

Finite:

- PCDATA

May Occur Within:

- [<bibRef>](#)

Input Standard(s):

From the resource itself or an accepted authoritative source.

Attributes:

id	ID	Optional
label	CDATA	Optional
type	CDATA	Optional
xml:lang	CDATA	Optional

Example:

```

<descMeta>
[... ]
<relationships>
<relation type="isReferencedBy">

```

```

<bibRef>
<agent type="creator">
<persName>Earle Leighton, 1917-</persName>
</agent>
<title type="main">Confederate Broadside Verse</title>
<imprint>
<geogName>Texas</geogName>
<geogName>New Braunfels</geogName>
<agent type="provider">
<corpName>Book Farm</corpName>
</agent>
<date>1950</date>
</imprint>
<bibScope type="citation number">60</bibScope>
</bibRef>
</relation>
</relationships>
<repository>
<corpName>Art & Architecture Libraries</corpName>
</repository>
[... ]
</descMeta>

```

<century>

Definition:

A numerical representation of the century or centuries associated with the coverage time or subject matter of a resource.

Must Contain:

Finite:

- PCDATA

May Occur Within:

- [<covTime>](#)
- [<subject>](#)

Input Standard(s):

YY01-YY00

Attributes:

id	ID	Optional
label	CDATA	Optional
type	CDATA	Optional
normal	CDATA	Optional
certainty	"exact", "circa"	Optional
era	"ad", "bc", "cc", "cd"	REQUIRED
xml:lang	CDATA	Optional

Example:

```

<descMeta>
[... ]

```

```

<covPlace>
<geogName>not captured</geogName>
</covPlace>
<covTime>
<century era="ad">1901-2000</century>
<dateRange era="ad" from="1974" to="1978">1974-1978</date>
</covTime>
<culture>American</culture>
<culture>Maryland</culture>
[... ]
</descMeta>

```

<color>

Definition:

Describes the color quality of image and moving image resources.

Must Contain:

Finite:

- PCDATA

May Occur Within:

- [<physDesc>](#)

Input Standard(s):

Choose from the following list:

- black and white
- color
- monochrome

Attributes:

id	ID	Optional
label	CDATA	Optional
type	CDATA	Optional
xml:lang	CDATA	Optional

Example:

```

<descMeta>
[... ]
<identifier type="handle">hdl:1903.1/2240</identifier>
<physDesc>
<color>monochrome</color>
<extent units="image">1</size>
</physDesc>
<relationships>
<relation label="collection" type="isPartOf">University AlbUM</relation>
<relation label="archivalcollection" type="isPartOf">Records of the Department of
Intercollegiate Athletics</relation>
</relationships>
[... ]
</descMeta>

```

<corpName>

Definition:

Corporate name. Contains the text for the name of a corporate entity.

May Contain:**Finite:**

- PCDATA
- [<bibRef>](#)
- [<extPtr>](#)
- [<extRef>](#)
- [<linkGrp>](#)

May Occur Within:

- [<agent>](#)
- [<repository>](#)
- [<subject>](#)

Input Standard(s):

[Library of Congress Name Authority File](#). For names not found in the LCNAF, enter corporate names in full direct form, omitting initial articles. Enter more than one agent if necessary for access.

Attributes:

id	ID	Optional
label	CDATA	Optional
type	CDATA	Optional
normal	CDATA	Optional
role	CDATA	Optional
xml:lang	CDATA	Optional

Example:

```

<descMeta xml:lang="en">
[... ]
<agent type="contributor">
<persName>Price, Kenneth</persName>
</agent>
<agent type="contributor">
<persName>Cossons, Neil, 1939-</persName>
</agent>
<agent type="provider">
<corpName>HTV West (Firm)</corpName>
</agent>
<agent type="provider">
<corpName>Films for the Humanities (Firm)</corpName>
</agent>
[... ]
</descMeta>

```

<covPlace>

Definition:

Coverage place. The geographical location associated with the production of the content of a resource (e.g. place of publication for a scanned image of a postcard). Provide information from broadest level (i.e. country or continent) to the most specific level known (region, state, city). Each location level should be provided in a separate [<geogName>](#) element.

Must Contain:

Repeatable:

- [<geogName>](#)

May Occur Within:

- [<descMeta>](#)

Input Standard(s):

Materials with an unknown place of origin may contain the phrase "*not captured*". All other materials require place names encoded within [<geogName>](#) sub-elements.

Attributes:

id	ID	Optional
label	CDATA	Optional
type	CDATA, "printing"	Optional
xml:lang	CDATA	Optional

Example:

```
<descMeta xml:lang="en">
[... ]
<covPlace>
<geogName type="continent">North America</geogName>
<geogName type="country">United States</geogName>
<geogName type="region">New Jersey</geogName>
</covPlace>
<covTime>
<century era="ad">1901-2000</century>
<date era="ad">1990</date>
</covTime>
[... ]
</descMeta>
```

<covTime>

Definition:

Coverage time. Time period associated with the production of the content of the resource (e.g. date of publication for a video digitized from videotape). Identification of a century is required, and additional elements describing a specific date or date range are also available.

Must Contain:

Repeatable:

- [<century>](#)

May Contain:

Finite:

- [<dateRange>](#)
- [<date>](#)

May Occur Within:

- [<descMeta>](#)

Attributes:

id	ID	Optional
label	CDATA	Optional
type	CDATA	Optional
xml:lang	CDATA	Optional

Specific Usage Rules for Collections

See the usage rules for the [<date>](#) element.

Example:

```
<descMeta xml:lang="en">
[... ]
<covPlace>
<geogName type="continent">North America</geogName>
<geogName type="country">United States</geogName>
<geogName type="region">New Jersey</geogName>
</covPlace>
<covTime>
<century era="ad">1901-2000</century>
<date era="ad">1990</date>
</covTime>
[... ]
</descMeta>
```

<culture>

Definition:

Contains text that indicates the culture of origin or context for a resource.

Must Contain:

Finite:

- PCDATA

May Occur Within:

- [<descMeta>](#)

Attributes:

id	ID	Optional
label	CDATA	Optional

type	CDATA	Optional
xml:lang	CDATA	Optional

Example:

```
<descMeta xml:lang="en">
[... ]
<covTime>
<century era="ad">1900-2000</century>
<date era="ad">1990</date>
</covTime>
<culture>European</culture>
<culture>British</culture>
<language>eng</language>
[... ]
</descMeta>
```

<date>

Definition:

A date associated with an event in the life cycle of the resource. Most often, date will be associated with the creation or availability of the resource.

Must Contain:**Finite:**

- PCDATA

May Occur Within:

- [<availability>](#)
- [<bibRef>](#)
- [<covTime>](#)
- [<imprint>](#)
- [<rights>](#)
- [<subject>](#)

Input Standard(s):

From the resource itself or an accepted authoritative source. Enter date in the form: YYYY-MM-DD. Range of dates may be expressed by using the ISO 8601 [<dateRange>](#) element. Enter more than one coverage-time if necessary for access.

Attributes:

id	ID	Optional
label	CDATA	Optional
type		Optional
normal	CDATA	Optional
certainty	exact, circa	Optional

era	"ad", "bc", "cc", "cd"	REQUIRED
xml:lang	CDATA	Optional

Specific Usage Rules for Collections

The following elements, attributes, and values may be used for this element in these specific collections.

University Album

- `<covTime><date certainty="exact">` *Required:* When using the `<date>` element in `<covtime>`, the *certainty* attribute is required with a value of "exact" or "circa".

Prange Digital Children's Book Collection

- `<covTime><date era="ad" label="pcbccd">` *Optional*

Example:

```
<descMeta>
[... ]
<covPlace>
<geogName>not captured</geogName>
</covPlace>
<covTime>
<century era="ad">1901-2000</century>
<date era="ad" type="exact">1926-12-26</date>
</covTime>
<culture>Maryland</culture>
<culture>American</culture>
<culture>University of Maryland</culture>
[... ]
</descMeta>
```

<dateRange>

Definition:

A range of dates associated with an event in the life cycle of the resource. Most often, a date range will be associated with the creation or availability of the resource.

Must Contain:

Finite:

- PCDATA

May Occur Within:

- [<covTime>](#)
- [<rights>](#)

Input Standard(s):

Enter dateRange in the form: YYYY-MM-DD to YYYY-MM-DD. Attributes from and to are required, following the same pattern.

Attributes:

id	ID	Optional
----	----	----------

label	CDATA	Optional
type	CDATA	Optional
normal	"exact", "circa"	Optional
certainty	CDATA	Optional
era	"ad", "bc", "cc", "cd"	REQUIRED
from	CDATA	REQUIRED
to	CDATA	REQUIRED
xml:lang	CDATA	Optional

Example:

```

<descMeta>
[... ]
<covPlace>
<geogName>not captured</geogName>
</covPlace>
<covTime>
<century era="ad">1901-2000</century>
<dateRange era="ad" from="1931" to="1933">1930-1933</dateRange>
</covTime>
<culture>Maryland</culture>
<culture>American</culture>
<culture>University of Maryland</culture>
[... ]
</descMeta>

```

<decade>

Definition:

A decade associated with the subject matter of a resource.

Must Contain:**Finite:**

- PCDATA

May Occur Within:

- [<subject>](#)

Input Standard(s):

Enter decade in the form: YYYY-YYYY

Attributes:

id	ID	Optional
label	CDATA	Optional
type	CDATA	Optional
xml:lang	CDATA	Optional

Example:

```

<descMeta>
[... ]

```

```
<description type="summary">Aerial view of University of Maryland, College Park
campus at 1200 feet looking northwest.</description>
<subject scheme="TGM2" type="topical">Aerial views</subject>
<subject scheme="TGM1" type="topical">Buildings</subject>
<subject type="temporal">
<decade>1931-1940</decade>
</subject>
<subject type="geographical">
<geogName type="continent">North America</geogName>
<geogName type="country">United States</geogName>
<geogName type="region">Maryland</geogName>
<geogName type="settlement">College Park</geogName>
</subject>
[... ]
</descMeta>
```

<descMeta>

Definition:

Descriptive metadata. Wrapper element for all elements used to describe a resource.

Must Contain:

Finite:

- [<covTime>](#)
- [<mediaType>](#)
- [<pid>](#)
- [<relationships>](#)
- [<repository>](#)

Repeatable:

- [<covPlace>](#)
- [<culture>](#)
- [<description>](#)
- [<physDesc>](#)
- [<rights>](#)
- [<subject>](#)
- [<title type="\[...\]">](#)

May Contain:

Repeatable:

- [<agent>](#)
- [<identifier>](#)
- [<language>](#)
- [<style>](#)

Attributes:

id	ID	Optional
label	CDATA	Optional
type	CDATA	Optional
xml:lang	CDATA	Optional

Example:

The following elements constitute the minimum set of elements for a descMeta instance (i.e.

those required by the DTD):

```

<descMeta>
<pid>[...]</pid>
<mediaType type="[...]">
[... ]
<form>[...]</form>
</mediaType>
<title type="[...]">[...]</title >
<covPlace>
<geogName>[...]</geogName>
</covPlace>
<covTime>
<century>[...]</century>
</covTime>
<culture>[...]</culture>
<description>[...]</description>
<subject type="[...]">[...]</subject>
<physDesc>[...]</physDesc>
<relationships>
<relation type="[...]">[...]</relation>
</relationships>
<repository>
<corpName>[...]</corpName>
</repository>
<rights>[...]</rights>
</descMeta>

```

<description>

Definition:

A textual description of the content of the resource, including abstracts or summaries for document-like objects, content descriptions for visual resources, or other descriptions not included in other elements.

Must Contain:**Finite:**

- PCDATA

Repeatable:

- [<bibRef>](#)
- [<extPtr>](#)
- [<extRef>](#)
- [<linkGrp>](#)

May Occur Within:

- [<descMeta>](#)
- [<relation>](#)

Input Standard(s):

Enter descriptive text, remarks, and comments about the object. This information can be taken from the item, or supplied by the agency if no structured description or abstract is available.

Attributes:

id	ID	Optional
label	CDATA	Optional
type	CDATA, "summary", "credits"	Optional
lang	CDATA	Optional
xml:lang	CDATA	Optional

Specific Usage Rules for Collections

The following elements, attributes and values may be used with this element in these specific collections.

Films@UM

- `<description type="summary">` *REQUIRED*
- `<description type="credits">` *Optional*

The Jim Henson Works

- `<description type="credits">` *Optional*
- `<description type="summary">` *REQUIRED*

Prange Digital Children's Books Collection

- `<description type="papertype" label="pcbcensor">` *Optional*
- `<description label="pcbcensor">` *Optional*
- `<description label="pcbnotes">` *Optional*

University AlBUM

- `<description type="summary">` *REQUIRED*

World's Fair

- o `<description type="caption">` *Optional; used to denote a caption on an original item NOT a caption created to describe digital content*

Example:

```
<descMeta>
[... ]
<covTime>
```

```
<century>1901-2000</century>
<date>1997</date>
</covTime>
<culture>European</culture>
<language>eng</language>
<description type="summary">Photo of the first official game played by the Baltimore Orioles versus Jersey City in the original Baltimore Stadium after Oriole Park burned down.</description>
[... ]
</descMeta>
```

<extent>

Definition:

The number of a given measurement that comprises a resource (e.g. 3 images)

Must Contain:

Finite:

- PCDATA

May Occur Within:

- [<physDesc>](#)

Input Standard(s):

Record the number of a given measurement for the whole of the resource being described.

Attribute units is required to indicate the unit of measure.

Attributes:

id	ID	Optional
label	CDATA	Optional
type	CDATA	Optional
units	CDATA, documents, image, pages, minutes, copies	REQUIRED
xml:lang	CDATA	Optional

Specific Usage Rules for Collections

The following elements, attributes, and values may be used with these elements in these specific collections.

Prange Digital Children's Book Collection

- `<physDesc><extent units="documents" label="pcbcensor"> Optional`

Example:

```
<descMeta xml:lang="en">
[... ]
<physDesc>
<extent units="pages">12</extent>
</physDesc>
<physDesc>
<size units="papersize">B5</size>
```

```

</physDesc>
<physDesc>
<size units="cm">26 x 19</size>
</physDesc>
<physDesc type="format">Book</physDesc>
<physDesc type="format" xml:lang="ja-Hani">本</physDesc>
[...]
</descMeta>

```

<extPtr>

Definition:

Extended pointer. An empty linking element which connects the model to an external electronic object.

Must Contain:

EMPTY

May Occur Within:

- [<corpName>](#)
- [<description>](#)
- [<extRef>](#)
- [<linkGrp>](#)
- [<other>](#) *Use of this element is restricted. See the Usage Rules for Specific Collections for this element.*
- [<persname>](#)
- [<relation>](#)

Input Standard(s):

Use the attributes entityref to identify the external object.

Attributes:

behavior	CDATA	Optional
entityref	ENTITY	Optional
href	CDATA	Optional
id	ID	Optional
idref	IDREF	Optional
inline	"true"	
rel	"version" or "part"	Optional
role	CDATA	Optional
show	"embed", "new", "replace", "other", "none"	Optional

targettype	CDATA	Optional
title	CDATA	Optional
type	"simple"	Fixed
xlink	CDATA	Optional
xpointer	CDATA	Optional
xml:lang	CDATA	Optional

Example:

```

<descMeta>
[... ]
<agent type="creator">
<corpName>London Weekend Television, ltd <extPtr type="simple" entityref="lwt"
title="Logo for LWT" show="embed" /></corpName>
</agent>
<agent type="creator">
<corpName>Reiner Moritz Associates</corpName>
</agent>
[... ]
</descMeta>

```

<extRef>

Definition:

Extended reference. A linking element that can include text and subelements as part of a reference to an external electronic object.

May Contain:**Finite:**

- PCDATA

Repeatable:

- [<extPtr>](#)

May Occur Within:

- [<corpName>](#)
- [<description>](#)
- [<linkGrp>](#)
- [<other>](#) *Use of this element is restricted. See the Usage Rules for Specific Collections for this element.*
- [<persName>](#)
- [<relation>](#)

Input Standard(s):

Use the attributes entityref or href to identify the external object.

Attributes:

behavior	CDATA	Optional
----------	-------	----------

entityref	ENTITY	Optional
href	CDATA	Optional
id	ID	Optional
idref	IDREF	Optional
inline	"true"	
rel	"version" or "part"	Optional
role	CDATA	Optional
show	"embed", "new", "replace", "other", "none"	Optional
targettype	CDATA	Optional
title	CDATA	Optional
type	"simple"	Fixed
xlink	CDATA	Optional
xml:lang	CDATA	Optional
xpointer	CDATA	Optional

Example:

```

<descMeta>
[... ]
<agent type="creator">
<corpName>
<extRef type="simple" href="http://www.londonweekend.com" title="ITv.com"
show="new">London Weekend Television, ltd </extRef>
</corpName>
</agent>
<agent type="creator">
<corpName>Reiner Moritz Associates</corpName>
</agent>
[... ]
</descMeta>

```

<form>**Definition:**

The form of the contents of the parent element, according to a non-Dublin Core scheme, e.g. poetry, prose, fiction, painting, sculpture, etc. The form element functions as a qualifier for the [<mediaType>](#) element.

Must Contain:**Finite:**

- PCDATA

May Occur Within:

- [<mediaType>](#)

Content Standard(s):

- <mediaType> = **Collection**:
 - Collection
- <mediaType> = **Dataset**:
 - *To be determined*
- <mediaType> = **Image**:
 - Architecture
 - Decorative art
 - Drawing. Design. Illustration.
 - Painting
 - Photograph, artistic
 - Photograph, documentary
 - Print
 - Sculpture
- <mediaType> = **interactiveResource**:
 - Hypertext systems
 - Interactive video
- <mediaType> = **movingImage**:
 - *To be determined*
- <mediaType> = **software**:
 - *To be determined*
- <mediaType> = **sound**:
 - instrumental music
 - vocal music
 - instrumental and vocal music
 - electronic music
 - spoken word
 - nature or natural sounds
- <mediaType> = **text**:
 - Poetry
 - Drama

[Home](#) | [General Info](#) | [Services](#) | [Research & Development](#) | [Metadata](#) | [Reports](#) | [Questions?](#) | [Virgo](#)

UNIVERSITY of VIRGINIA LIBRARY



Digital Initiatives: Metadata

UVa Metadata Descriptive Elements (UVa DescMeta)

[Metadata Home](#) > UVa Metadata Descriptive Elements

UVa DescMeta is currently under development by the [Metadata Steering Group](#) at the University of Virginia Library. The element set, the minimal requirements, and the DTDs are still considered in-progress. The DTD will be released as UVa DescMeta version 2.0 when development and prototyping are complete.

[DTD](http://text.lib.virginia.edu/dtd/descmeta/descmeta.dtd) at: <http://text.lib.virginia.edu/dtd/descmeta/descmeta.dtd>

[Documentation](http://dl.lib.virginia.edu/html/descmeta/) at: <http://dl.lib.virginia.edu/html/descmeta/>

(note: the documentation is NOT up to date with the current DTD. Please use with caution.)

[Mappings](http://www.lib.virginia.edu/digital/metadata/mappings.html) at: <http://www.lib.virginia.edu/digital/metadata/mappings.html>

[Descriptive Element Set](#)

[Minimal Requirements for Ingest into the UVa Digital Library](#)

Descriptive Element Set

<agent>
<authority>
<covplace>
<covtime>
<culture>
<description>
<identifier>
<language>
<mediatype>
<mimetype>
<physdesc>
<pid>
<place>
<relationships>
<rights>
<style>
<subject>
<surrogate>
<time>
<title>

Minimal Requirements for Ingest into the UVa Digital Library

<pid>

- Surrogate identifier
- The Fedora PID
- “Unknown” is not an option

<title type="main" >

- Original title
- One and only one
- “Unknown” is not an option

<mediatype>

- Mediatype
- One and only one (not repeatable)
- “Unknown” is not an option

<rights type="access" >

- Access rights
- “Unknown” is not an option

<surrogate><time>

- The date the resource was put on the first server
- If unknown, may be populated with date of ingest into Fedora

<time>

- An original date of any date type
- In extreme situations, an appeal can be made to the MSG for populating the field with "Unknown"; the MSG will decide if this is acceptable on a case-by-case basis.

Not required but highly recommended:

<agent>

- At least agent of any type
- Consider using "Unknown", if unavailable

Digital Initiatives
University of Virginia
PO Box 400112
Charlottesville, VA 22904-4112

[Digital Initiatives Home](#) • [UVA Library Home](#)
[Search the Library Site](#) • [UVA Home](#)
Maintained by: dl@virginia.edu
Last Modified: Monday, June 02, 2008
© The Rector and Visitors of the [University of Virginia](#)

*THIS PAGE INTENTIONALLY LEFT BLANK
FOR DOUBLE SIDED COPY*

Texas Digital Library

Application Profile for Electronic Theses and Dissertations

Introduction

This MODS application profile for electronic theses and dissertations (ETDs) describes the best practices for descriptive metadata for members of the Texas Digital Library (TDL). This document defines the mandatory minimum elements for ETDs. Besides these elements, other valid MODS elements may be included in ETD records. Optional elements, subelements, and attributes are described throughout the document.

MODS elements for ETDs:

Title Information (mandatory)
Name of Author (mandatory)
Name of Thesis Advisor (mandatory)
Name of Committee Member (optional)
Name of Degree Grantor (mandatory)
Type of Resource (mandatory)
Genre (mandatory)
Origin Information (mandatory)
Language (mandatory)
Physical Description (mandatory)
Abstract (mandatory)
Subject (mandatory)
Identifier (mandatory)
Location (mandatory)
Degree Information (mandatory)
Record Information (mandatory)

Instructions for formatting and encoding:

Title Information

Mandatory practice: Encode the title information in a <mods:titleInfo> wrapper element. Encode the title proper in a <mods:title> subelement. Encode the subtitle in a <mods:subTitle> subelement.

Optional practice: Other valid subelements or attributes within the <mods:titleInfo> element may be used.

Example:

```
<mods:titleInfo>  
  <mods:title>Critical processes and performance measures for patient safety  
    systems in healthcare institutions  
</mods:title>
```

Texas Digital Library

Application Profile for Electronic Theses and Dissertations

```
<mods:subTitle>a Delphi study</mods:subTitle>
</mods:titleInfo>
```

Name of Author

Mandatory practice: Encode information about the name of the author in the <mods:name> wrapper element with the type attribute set to “personal.” Encode the MARC relator term “Author” in the <mods:roleTerm> subelement under the <mods:role> subelement. Encode the various parts of the name in the <mods:namePart> subelement. Include the type attribute in each <mods:namePart> subelement. The “given” and “family” name types are mandatory.

Optional practice: Encode the birthdate in a <mods:namePart> subelement with type set to “date”. Other valid subelements or attributes within the <mods:name> element may be used.

Example:

```
<mods:name type="personal">
  <mods:namePart type="given">Ralitsa B.</mods:namePart>
  <mods:namePart type="family">Akins</mods:namePart>
  <mods:namePart type="date">1967-</mods:namePart>
  <mods:role>
    <mods:roleTerm authority="marcrelator"
      type="text">Author</mods:roleTerm>
  </mods:role>
</mods:name>
```

Name of Thesis Advisor

Mandatory practice: Encode information about the thesis advisor in the <mods:name> wrapper element with the type attribute set to “personal.” Encode the MARC relator term “Thesis advisor” in the <mods:roleTerm> subelement under the <mods:role> subelement. Encode the various parts of the name in the <mods:namePart> subelement. Include the type attribute in each <mods:namePart> subelement. The “given” and “family” name types are mandatory. The element <mods:name> is repeatable for thesis advisors.

Optional practice: Encode the birthdate in a <mods:namePart> subelement with type set to “date”. Other valid subelements or attributes within the <mods:name> element may be used.

Example:

```
<mods:name type="personal">
  <mods:namePart type="given">Bryan R.</mods:namePart>
  <mods:namePart type="family">Cole</mods:namePart>
  <mods:role>
    <mods:roleTerm authority="marcrelator" type="text">Thesis
      advisor</mods:roleTerm>
  </mods:role>
</mods:name>
```

Texas Digital Library

Application Profile for Electronic Theses and Dissertations

Name of Committee Member

Optional practice: Encode information about committee members in the <mods:name> wrapper element with the type attribute set to “personal.” Encode the term “Committee member” in the <mods:roleTerm> subelement under the <mods:role> subelement. Encode the various parts of the name in the <mods:namePart> subelement. Include the type attribute in each <mods:namePart> subelement. The “given” and “family” name types are mandatory. The element <mods:name> is repeatable for committee members. Encode the birthdate in a <mods:namePart> subelement with type set to “date”. Other valid subelements or attributes within the <mods:name> element may be used.

Example:

```
<mods:name type="personal">
  <mods:namePart type="given">Jane R.</mods:namePart>
  <mods:namePart type="family">Smith</mods:namePart>
  <mods:role>
    <mods:roleTerm type="text">Committee member</mods:roleTerm>
  </mods:role>
</mods:name>
```

Name of Degree Grantor

Mandatory practice: Encode information about the degree grantor in the <mods:name> wrapper element with the type attribute set to “corporate.” Encode the name of the degree granting institution in a <mods:namePart> subelement. Use the form of the name authorized by the Library of Congress Name Authority File. Encode the name of the department that granted the degree in a <mods:namePart> subelement. Encode the MARC relator term “Degree grantor” in the <mods:roleTerm> subelement under the <mods:role> subelement.

Optional practice: Other valid subelements or attributes within the <mods:name> element may be used.

Example:

```
<mods:name type="corporate" authority="lcnaf">
  <mods:namePart>Texas A & M University</mods:namePart>
  <mods:namePart>Philosophy</mods:namePart>
  <mods:role>
    <mods:roleTerm authority="marcrelator" type="text">
      Degree grantor
    </mods:roleTerm>
  </mods:role>
</mods:name>
```

Texas Digital Library

Application Profile for Electronic Theses and Dissertations

Type of Resource

Mandatory practice: Encode the type of resource in the <mods:typeOfResource> element. The element <mods:typeOfResource> is repeatable for ETDs with multiple files.

Example:

```
<mods:typeOfResource>
  text
</mods:typeOfResource>
```

Genre

Mandatory practice: Encode the MARC genre term “theses” in the <mods:genre> element. Set the authority attribute to “marcgt.”

Optional practice: Other valid attributes within the <mods:genre> element may be used.

Example:

```
<mods:genre authority="marcgt">
  theses
</mods:genre>
```

Origin Information

Mandatory practice: Encode relevant dates for the ETD in the <mods:originInfo> wrapper element. The creation date is defined as the date the student graduates or the date the degree is conferred. The publication date is defined as the date the ETD is released to the public.

Encode the month and year of the creation date, according to ISO 8601, in the <mods:dateCreated> subelement. Set the encoding attribute to “iso8601.”

Encode the month and year of the publication date, according to ISO 8601, in the <mods:dateIssued> subelement. Set the encoding attribute to “iso8601.”

Optional practice: The day of the month may be included date encodings. Other valid subelements or attributes within the <mods:originInfo> element may be used.

Examples:

```
<mods:originInfo>
  <mods:dateCreated encoding="iso8601">200408</mods:dateCreated>
  <mods:dateIssued encoding="iso8601">200412</mods:dateIssued>
</mods:originInfo>
```

Texas Digital Library

Application Profile for Electronic Theses and Dissertations

Language

Mandatory practice: Encode information about the language of the ETD in the <mods:language> wrapper element. Encode the language, according to ISO 639-2b, in the <mods:languageTerm> subelement. Set the type attribute to “code” and the authority attribute to “iso639-2b.” The <mods:languageTerm> subelement is repeatable.

Optional practice: Other valid subelements or attributes within the <mods:language> element may be used.

Example:

```
<mods:language>
  <mods:languageTerm type="code" authority="iso639-2b">
    eng
  </mods:languageTerm>
  <mods:languageTerm type="code" authority="iso639-2b">
    spa
  </mods:languageTerm>
</mods:language>
```

Physical Description

Mandatory practice: Encode the physical description in the <mods:physicalDescription> wrapper element. Encode the MARC format term “electronic” in the <mods:form> subelement. Set the authority attribute to “marcform.” Encode the MIME type in the <mods:internetMediaType> subelement. Encode the digital origin (“born digital” or “reformatted digital”) in the <mods:digitalOrigin> subelement.

Optional practice: Other valid attributes within the <mods:physicalDescription> element may be used.

Example:

```
<mods:physicalDescription>
  <mods:form authority="marcform">electronic</mods:form>
  <mods:internetMediaType>application/pdf</mods:internetMediaType>
  <mods:digitalOrigin>born digital</mods:digitalOrigin>
</mods:physicalDescription>
```

Abstract

Mandatory practice: Encode the abstract in the <mods:abstract> element. Include the language attribute encoded in ISO 639-2b.

Optional practice: Valid attributes within the <mods:abstract> element may be used.

Texas Digital Library

Application Profile for Electronic Theses and Dissertations

Example:

```
<mods:abstract lang="eng">This dissertation study presents a conceptual framework for implementing and assessing patient safety systems in healthcare institutions. The conceptual framework consists of critical processes and performance measures identified in the context of the 2003 Malcolm Baldrige National Quality Award (MBNQA) Health Care Criteria for Performance Excellence...</mods:abstract>
```

Subject

Mandatory practice: Encode topical subject terms in the <mods:subject> wrapper element. Encode individual terms or phrases in the <mods:topic> subelement. The <mods:subject> element is repeatable.

Optional practice: Controlled subject headings may be included by using the authority attribute of the <mods:topic> subelement. Other valid subelements or attributes within the <mods:subject> element may be used.

Example:

```
<mods:subject>
  <mods:topic>healthcare</mods:topic>
</mods:subject>

<mods:subject>
  <mods:topic>patient safety</mods:topic>
</mods:subject>

<mods:subject authority="lcsh">
  <mods:topic>Medical care</mods:topic>
  <mods:topic>Quality control</mods:topic>
  <mods:geographic>United States</mods:geographic>
  <mods:temporal>20th century</mods:temporal>
</mods:subject>
```

Identifier

Mandatory practice: Encode the unique identifier in the <mods:identifier> element. The <mods:identifier> element is repeatable.

Optional practice: The type attribute may be used in the <mods:identifier> element. Other valid attributes within the <mods:identifier> element may be used.

Example:

```
<mods:identifier type="hdl">
```

Texas Digital Library

Application Profile for Electronic Theses and Dissertations

```
    http://handle.tamu.edu/1969.1/1042
</mods:identifier>
```

Location

Mandatory practice: Encode the location in the <mods:location> wrapper element. Encode the uniform resource locator (URL) in the <mods:url> subelement.

Optional practice: Other valid attributes within the <mods:identifier> element may be used.

Example:

```
<mods:location>
  <mods:url>
    http://handle.tamu.edu/1969.1/1042
  </mods:url>
</mods:location>
```

Degree Information

Note: The MODS standard does not have elements specifically for theses and dissertations. In order to encode degree information in MODS, the <mods:extension> element is used to reference the ETD-MS XML schema.

Mandatory practice: Encode information about the conferred degree in the <etd:degree> wrapper element. Encode the degree name in the <etd:name> subelement. Use the fully spelled out form of the degree name. Encode the degree level, from the TDL vocabulary, in the <etd:level> subelement. Encode the degree discipline, from the TDL vocabulary, in the <etd:discipline> subelement.

Example:

```
<mods:extension>
  <etd:degree>
    <etd:name>Doctor of Philosophy</etd:name>
    <etd:level>Doctoral</etd:level>
    <etd:discipline>Educational Administration</etd:discipline>
  </etd:degree>
</mods:extension>
```

Record Information

Mandatory practice: Encode information about the MODS record in the <mods:recordInfo> wrapper element. Encode the name of the agency that created the MODS record in the <mods:recordContentSource> subelement, with the authority attribute set to "marcorg." Encode the month, year, and day of the creation date of the record, according to ISO 8601, in the

Texas Digital Library Application Profile for Electronic Theses and Dissertations

<mods:recordCreationDate> subelement. Set the encoding attribute to “iso8601.” Encode the month, year, and day of the change date, according to ISO 8601, in the <mods:recordChangeDate> subelement. Set the encoding attribute to “iso8601.” Encode the unique record identifier in the <mods:recordIdentifier> subelement.

Optional practice: Other valid attributes within the <mods:recordInfo> element may be used.

Example:

```
<mods:recordInfo>
  <mods:recordContentSource authority="marcorg">
    TxCM
  </mods:recordContentSource>
  <mods:recordCreationDate encoding="iso8601">
    20050826
  </mods:recordCreationDate>
  <mods:recordChangeDate encoding="iso8601">
    20050826
  </mods:recordChangeDate>
  <mods:recordIdentifier>12345678</mods:recordIdentifier>
</mods:recordInfo>
```

Texas Digital Library

Application Profile for Electronic Theses and Dissertations

Full example of MODS record for an ETD

```
<?xml version="1.0" encoding="UTF-8"?>
<mods:mods

xmlns:mods="http://www.loc.gov/mods/v3"
xmlns:etd="http://www.ndltd.org/standards/metadata/etdms/1.0/"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"

xsi:schemaLocation=
"http://www.loc.gov/mods/v3
http://www.loc.gov/standards/mods/v3/mods-3-1.xsd

http://www.ndltd.org/standards/metadata/etdms/1.0/
http://www.ndltd.org/standards/metadata/etdms/1.0/etdms.xsd">

<mods:titleInfo lang="eng">
  <mods:title>Critical processes and performance measures for patient safety
    systems in healthcare institutions</mods:title>
  <mods:subTitle>a Delphi study</mods:subTitle>
</mods:titleInfo>

<mods:name type="personal" authority="lcnaf">
  <mods:namePart>Akins, Ralitsa B., 1967-</mods:namePart>
  <mods:namePart type="given">Ralitsa B.</mods:namePart>
  <mods:namePart type="family">Akins</mods:namePart>
  <mods:namePart type="date">1967-</mods:namePart>
  <mods:role>
    <mods:roleTerm authority="marcrelator"
      type="text">Author</mods:roleTerm>
  </mods:role>
</mods:name>

<mods:name type="personal">
  <mods:namePart type="given">Bryan R.</mods:namePart>
  <mods:namePart type="family">Cole</mods:namePart>
  <mods:role>
    <mods:roleTerm authority="marcrelator" type="text">Thesis
      advisor</mods:roleTerm>
  </mods:role>
</mods:name>

<mods:name type="corporate" authority="lcnaf">
  <mods:namePart>Texas A & M University</mods:namePart>
  <mods:role>
    <mods:roleTerm authority="marcrelator" type="text">Degree
      grantor</mods:roleTerm>
  </mods:role>
</mods:name>

<mods:typeOfResource>
  text
</mods:typeOfResource>

<mods:genre authority="marcgt">
  theses
</mods:genre>
```

Texas Digital Library

Application Profile for Electronic Theses and Dissertations

```
<mods:originInfo>
  <mods:dateCreated encoding="iso8601">200408</mods:dateCreated>
  <mods:dateIssued encoding="iso8601">200411</mods:dateIssued>
</mods:originInfo>

<mods:language>
  <mods:languageTerm type="code" authority="iso639-2b">
    eng
  </mods:languageTerm>
</mods:language>

<mods:physicalDescription>
  <mods:form authority="marcform">electronic</mods:form>
  <mods:internetMediaType>application/pdf</mods:internetMediaType>
  <mods:digitalOrigin>born digital</mods:digitalOrigin>
</mods:physicalDescription>

<mods:abstract lang="eng">This dissertation study presents a conceptual
framework for implementing and assessing patient safety systems in
healthcare institutions. The conceptual framework consists of critical
processes and performance measures identified in the context of the 2003
Malcolm Baldrige National Quality Award (MBNQA) Health Care Criteria for
Performance Excellence. Methodology: The Delphi technique for gaining
consensus from a group of experts and forecasting significant issues in
the field of the Delphi panel expertise was used. Data collection included
a series of questionnaires where the first round questionnaire was based
on literature review and the MBNQA criteria for excellence in healthcare,
and tested by an instrument review panel of experts. Twenty-three experts
(MBNQA healthcare reviewers and senior healthcare administrators from
quality award winning institutions) representing 18 states participated in
the survey rounds. The study answered three research questions: (1) What
are the critical processes that should be included in healthcare patient
safety systems? (2) What are the performance measures that can serve as
indicators of quality for the processes critical for ensuring patient
safety? (3) What processes will be critical for patient safety in the
future? The identified patient safety framework was further transformed
into a patient safety tool with three levels: basic, intermediate, and
advanced. Additionally, the panel of experts identified the major barriers
to the implementation of patient safety systems in healthcare
institutions. The identified "top seven" barriers were directly related to
critical processes and performance measures identified as "important" or
"very important" for patient safety systems in the present and in the
future. This dissertation study is significant because the results are
expected to assist healthcare institutions seeking to develop high quality
patient safety programs, processes and services. The identified critical
processes and performance measures can serve as a means of evaluating
existing patient safety initiatives and guiding the strategic planning of
new safety processes. The framework for patient safety systems utilizes a
systems approach and will support healthcare senior administrators in
achieving and sustaining improvement results. The identified patient
safety framework will also assist healthcare institutions in using the
MBNQA Health Care Criteria for Performance Excellence for self-assessment
and quality improvement.</mods:abstract>

<mods:subject>
  <mods:topic>healthcare</mods:topic>
</mods:subject>
```

Texas Digital Library

Application Profile for Electronic Theses and Dissertations

```
</mods:subject>

<mods:subject>
  <mods:topic>patient safety</mods:topic>
</mods:subject>

<mods:subject authority="lcs" >
  <mods:topic>Medical care</mods:topic>
  <mods:topic>Quality control</mods:topic>
  <mods:geographic>United States</mods:geographic>
  <mods:temporal>20th century</mods:temporal>
</mods:subject>

<mods:identifier type="hdl">
  http://handle.tamu.edu/1969.1/1042
</mods:identifier>

<mods:location>
  http://handle.tamu.edu/1969.1/1042
</mods:location>

<mods:extension>
  <etd:degree>
    <etd:name>Ph. D.</etd:name>
    <etd:level>Doctoral</etd:level>
    <etd:discipline>Educational Administration</etd:discipline>
  </etd:degree>
</mods:extension>

<mods:recordInfo>
  <mods:recordContentSource authority="marcorg">
    TxCM
  </mods:recordContentSource>
  <mods:recordCreationDate encoding="iso8601">
    20050826
  </mods:recordCreationDate>
  <mods:recordChangeDate encoding="iso8601">
    20050826
  </mods:recordChangeDate>
  <mods:recordIdentifier>12345678</mods:recordIdentifier>
</mods:recordInfo>

</mods:mods>
```

*THIS PAGE INTENTIONALLY LEFT BLANK
FOR DOUBLE SIDED COPY*



6. Controlled vocabularies

Metadata Standards and Applications Workshop



Goals of Session

- Understand how different controlled vocabularies are used in metadata
- Learn about relationships between terms in thesauri
- Understand methods of encoding vocabularies
- Learn about how registries are used to document vocabularies



Controlled vocabularies

- Values that occur in metadata
- Often documented and published
- Goal to reduce ambiguity
- Control of synonyms
- Establishment of formal relationships among terms (where appropriate)
- Testing and validation of terms
- Metadata registries may be established



Why bother?

- To improve retrieval, i.e., to get an optimum balance of precision and recall
 - Precision – How many of the retrieved records are relevant?
 - Recall – How many of the relevant records did you retrieve?



Improving Recall and Precision

- Controlled Vocabularies improve recall by addressing synonyms [attire vs. dress vs. clothing]
- Controlled Vocabularies improve precision by addressing homographs [bridge (game) vs. bridge (structure) vs. bridge (dental device)]



Types of Controlled Vocabularies

- Lists of enumerated values
- Synonym rings
- Taxonomy
- Thesaurus
- Classification Schemes
- Ontology



Lists

A list is a simple group of terms

Example:

Alabama
Alaska
Arkansas
California
Colorado
...

Frequently used in Web site pick lists and pull down menus



Synonym Rings

- Synonym rings are used to expand queries for content objects
 - If a user enters any one of these terms as a query to the system, all items are retrieved that contain any of the terms in the cluster
- Synonym rings are often used in systems where the underlying content objects are left in their unstructured natural language format
 - the control is achieved through the interface by drawing together similar terms into these clusters
- Synonym rings are used in conjunction with search engines and provide a minimal amount of control of the diversity of the language found in the texts of the underlying documents



Taxonomies

A *taxonomy* is a set of preferred terms, all connected by a hierarchy or polyhierarchy

Example:

Chemistry

Organic chemistry

Polymer chemistry

Nylon

Frequently used in web navigation systems



Thesauri

A *thesaurus* is a controlled vocabulary with multiple types of relationships

Example:

Rice

UF paddy

BT Cereals

BT Plant products

NT Brown rice

RT Rice straw



Ontology

- One definition: “An arrangement of concepts and relations based on an underlying model of reality.”
 - Ex.: Organs, symptoms, and diseases in medicine
- No real agreement on definition—every community uses the term in a slightly different way



Thesaural Relationships

Relationship types:

- Use/Used For – indicates preferred term
- Hierarchy – indicates broader and narrower terms
- Associative – almost unlimited types of relationships may be used

It is the most complex format for controlled vocabularies and widely used.

ANSI/NISO Z39.19-2005

ISBN: 1-880124-65-3

Guidelines for the Construction, Format, and Management of Monolingual Controlled Vocabularies

Abstract: This Standard presents guidelines and conventions for the contents, display, construction, testing, maintenance, and management of monolingual controlled vocabularies. This Standard focuses on controlled vocabularies that are used for the representation of content objects in knowledge organization systems including lists, synonym rings, taxonomies, and thesauri. This Standard should be regarded as a set of recommendations based on preferred techniques and procedures. Optional procedures are, however, sometimes described, e.g., for the display of terms in a controlled vocabulary. The primary purpose of vocabulary control is to achieve consistency in the description of content objects and to facilitate retrieval. Vocabulary control is accomplished by three principal methods: defining the scope, or meaning, of terms; using the equivalence relationship to link synonymous and nearly synonymous terms; and distinguishing among homographs.



Z39.19 Types of Concepts

- Things and their physical parts
- Materials
- Activities or processes
- Events or occurrences
- Properties or states of persons, things, materials or actions
- Disciplines or subject fields
- Units of measurement
- Unique entities



Examples

- Birds (things)
- Ornithology (discipline)
- Feathers (materials)
- Flying (activity or process)
- Bird counts (event)
- Barn Owl (unique entity)



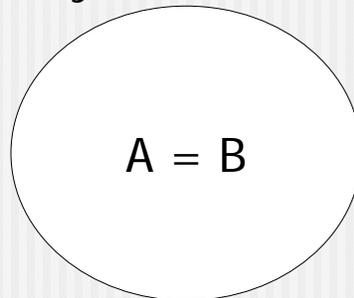
Relationships

- Equivalence
- Hierarchical
- Associative



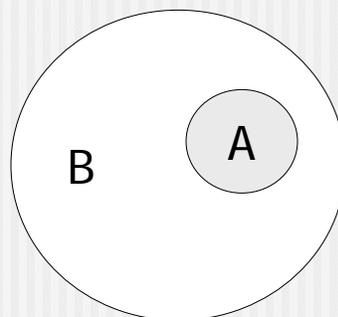
Equivalence Relationships

Term A and Term B overlap completely



Hierarchical Relationships

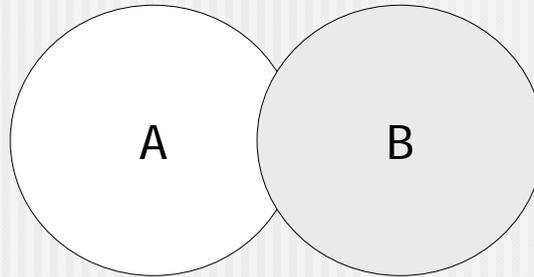
■ Term A is included in Term B





Associative Relationships

- Semantics of terms A and B overlap



Expressing Relationship

Relationship	Rel. Indicator	Abbreviation
Equivalence (synonymy)	Use Used for	None or U UF
Hierarchy	Broader term Narrower term	BT NT
Association	Related term	RT



Vocabulary Management

- The degree of control over a vocabulary is (mostly) independent of its type
 - **Uncontrolled** – Anybody can add anything at any time and no effort is made to keep things consistent
 - **Managed** – Software makes sure there is a list that is consistent (no duplicates, no orphan nodes) at any one time. Almost anybody can add anything, subject to consistency rules
 - **Controlled** – A documented process is followed for the update of the vocabulary. Few people have authority to change the list. Software may help, but emphasis is on human processes and custodianship



Informal Vocabularies

- New movement towards ‘bottom up’ classification goes by many names:
 - Tagging
 - Social bookmarking
 - Folksonomies
- Some in this movement, seeing problems of scale, are moving towards more formalization—others are reframing the vocabulary issue



Encoding controlled vocabularies

- MARC 21
 - Authority Format used for names, subjects, series
 - Classification Format used for formal classification schemes
- MADS (a derivative of MARC)
- Simple Knowledge Organization System (SKOS)
 - Intended for concepts



SKOS

- “SKOS Core provides a model for expressing the basic structure and content of concept schemes such as thesauri, classification schemes, subject heading lists, taxonomies, 'folksonomies', other types of controlled vocabulary, and also concept schemes embedded in glossaries and terminologies.”

--SKOS Core Guide

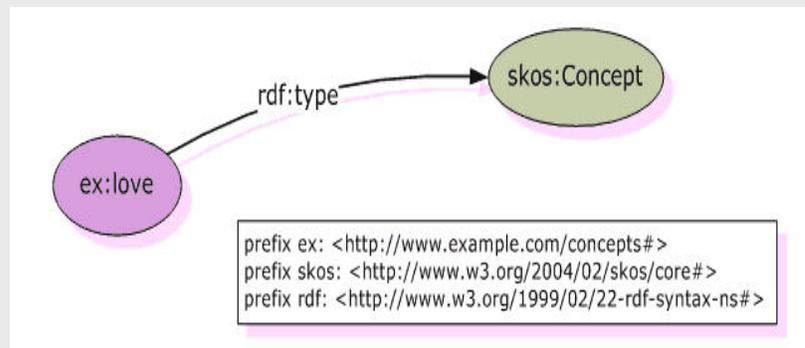


SKOS & RDF

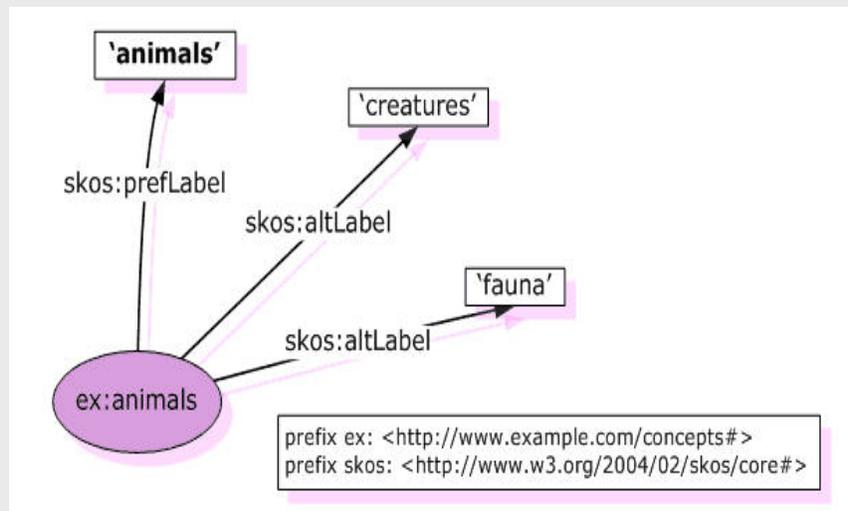
- SKOS Core Vocabulary
- Applies Resource Description Framework (RDF)
- Express a concept scheme as an RDF graph
- Data linked to and/or merged with other data
- Data sources distributed across the web
- Data meaningfully composed and integrated

The [skos:Concept](#) class allows you to assert that a resource is a *conceptual resource*.

That is, the resource *is itself* a concept.



Preferred and Alternative Lexical Labels



The RDF/XML Encoded Version

```

<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:skos="http://www.w3.org/2004/02/skos/core#">

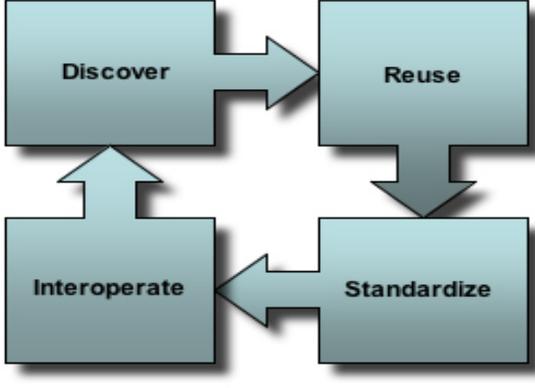
  <skos:Concept rdf:about="http://www.example.com/concepts#animals">
    <skos:prefLabel>animals</skos:prefLabel>
    <skos:altLabel>creatures</skos:altLabel>
    <skos:altLabel>fauna</skos:altLabel>
  </skos:Concept>

</rdf:RDF>

```



Registries: the Big Picture



```
graph TD; Discover --> Reuse; Reuse --> Standardize; Standardize --> Interoperate; Interoperate --> Discover;
```

(Adapted from Wagner & Weibel, "The Dublin Core Metadata Registry: Requirements, Implementation, and Experience" JoDI, 2005)



Why Registries?

- Support interoperability
 - Discovery of available schemes and schemas for description of resources
 - Promote reuse of extant schemes and schemas
 - Access to machine-readable and human-readable services
 - Support for crosswalking and translation
- Coping with different metadata schemes



Declaration, documentation, publication

- To identify the source of a vocabulary, e.g., a term comes from LCSH, as identified in my metadata by a URI
- To clarify a term and its definition
- To publish controlled vocabularies and have access to information about each term



Some uses for registries

- Metadata Schemas
 - Crosswalks between metadata schemas
- Controlled Vocabularies
 - Mappings between vocabularies
- Application Profiles
 - Schema and vocabulary information in combination



Metadata registries

- Some are formal, others are informal lists
- Some formal registries:
 - Dublin Core registry of DC terms
 - NSDL registry of vocabularies used
 - LC is establishing registries
 - MARC and ISO code lists
 - Enumerated value lists
 - LCSH in SKOS (example: <http://lcsch.info/sh85118553>)

Example of ISO 639-2 language code in SKOS

```

<rdf:Description rdf:about=
"http://www.loc.gov/standards/registry/vocabulary/iso639-2/por">
<rdf:type rdf:resource="http://www.w3.org/2008/05/skos #Concept"/>
<skos:prefLabel xml:lang="x-notation">por</skos:prefLabel>
<skos:altLabel xml:lang="en-latn">Portuguese</skos:altLabel>
<skos:altLabel xml:lang="fr-latn">portugais</skos:altLabel>
<skos:notation rdf:datatype="xs:string">por</skos:notation>
<skos:definition xml:lang="en-latn">This Concept has not yet been
defined.</skos:definition>
<skos:inScheme rdf:resource=
"http://www.loc.gov/standards/registry/vocabulary/iso639-2"/>
<vs:term_status>stable</vs:term_status>
<skos:historyNote rdf:datatype="xs:dateTime">2006-07-
19T08:41:54.000-05:00</skos:historyNote>
<skos:exactMatch rdf:resource=
"http://www.loc.gov/standards/registry/vocabulary/iso 639-1/pt"/>
<skos:changeNote rdf:datatype="xs:dateTime">2008-07-
09T13:49:05.321-04:00</skos:changeNote>
</rdf:Description>
  
```

*THIS PAGE INTENTIONALLY LEFT BLANK
FOR DOUBLE SIDED COPY*



7. Approaches to Models of Metadata Creation, Storage and Retrieval

Metadata Standards and Applications



Goals of Session

- Understand the differences between traditional vs. digital library Metadata Creation
- Storage, and
- Retrieval/Discovery



Creating Metadata Records

- The “Library Model”
 - Trained catalogers, one-at-a-time metadata records
- The “Submission Model”
 - Authors create metadata when submitting resources
- The “Automated Model”
 - Automated tools create metadata for resources
- “Combination Approaches”



The Library Model

- Records created “by hand,” one at a time
- Shared documentation and content standards (AACR2, etc.)
- Efficiencies achieved by sharing information on commonly held resources
- Not easily extended past the granularity assumptions in current practice



The Submission Model

- Based on author or user generated metadata
- Can be wildly inconsistent
 - Submitters generally untrained
 - May be expert in one area, clueless in others
- Often requires editing support for usability
- Inexpensive, may not be satisfactory as an only option



The Automated Model

- Based largely on text analysis; doesn't usually extend well to non-text or low-text
- Requires development of appropriate evaluation and editing processes
- Still largely research; few large, successful production examples, yet
- Can be done in batch
- Also works for technical as well as descriptive metadata



Combination Approaches

- Combination machine and human created metadata
 - Ex.: LC Web Archives (<http://www.loc.gov/minerva>)
- Combination metadata and content indexing
 - Ex.: NSDL (<http://nsdl.org>)



Content “Storage” Models

- ‘Storage’ in this context relates to the relationships between metadata and content (not the systems that purport to ‘store’ content for various uses)
- These relationships affect how access to the information is accomplished, and how the metadata either helps or hinders the process (or is irrelevant to it)



Common 'Storage' Models

- Content with metadata
- Metadata only
- Service only



Content with metadata

- Examples:
 - HTML pages with embedded 'meta' tags
 - Most content management systems (though they may store only technical or structural metadata)
 - Text Encoding Initiative (TEI)
- Often proves difficult to update



Metadata only

- Library catalogs
 - Web-based catalogs often provide some services for digital content
- Electronic Resource Management Systems (ERMS)
 - Provide metadata records for title level only
- Metadata aggregations
 - Using API or OAI-PMH for harvest and re-distribution



Service only

- Often supported partially or fully by metadata
 - Google, Yahoo (and others)
 - Sometimes provide both search services and distributed search software
 - Electronic journals (article level)
 - Linked using 'link resolvers' or available independently from Websites
 - Have metadata behind their services but don't generally distribute it separately



Common Retrieval Models

- Library catalogs
- Web-based (“Amazoogole”)
- Portals and federations



Library Catalogs

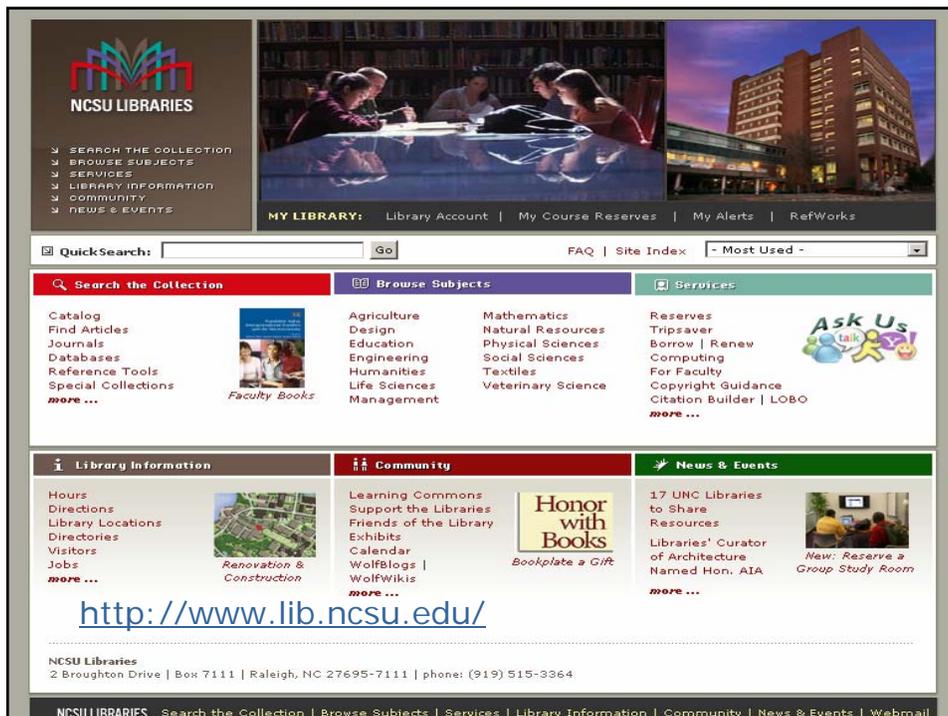
- Based on a consensus that granular metadata is useful
- Expectations of uniformity of information content and presentation
- Designed to optimize recall and precision
- Addition of relevance ranking and keyword searching of limited value (only ‘text’ used is the metadata itself)
- Retrieval options limited by LMS vendor decisions



Endeca

- Information Access Platform (software)
 - optimizes huge volumes of data and can be used by everyday people

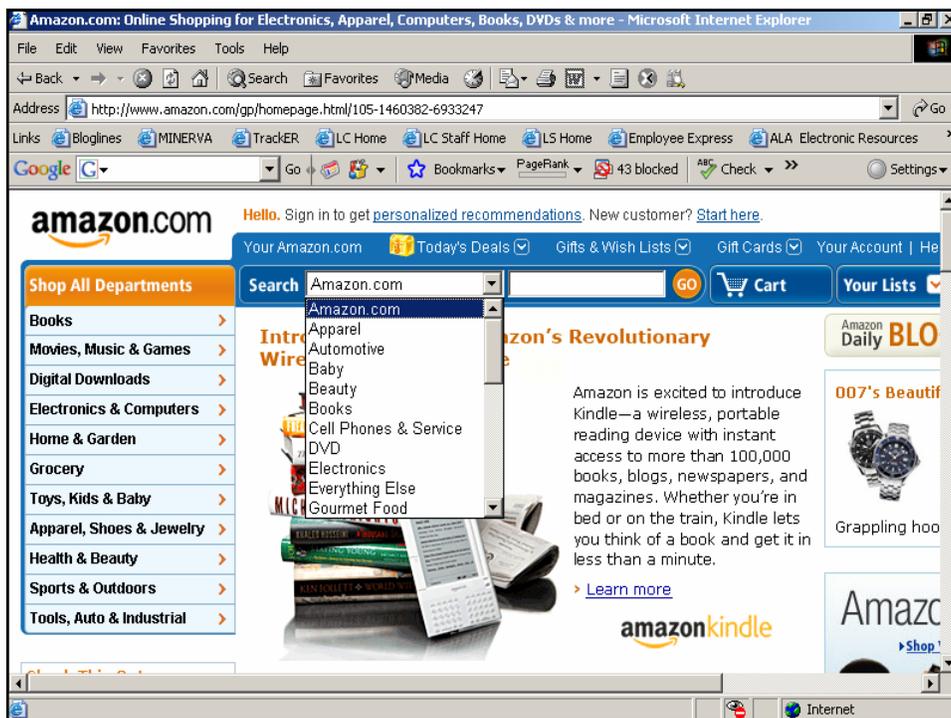
- North Carolina State University Libraries
 - 2006, first library to use this technology
 - Fast, flexible & as simple as browsing the Web, but unmatched by Web search engines due to rich [metadata] content & cutting-edge capabilities.





Web-based

- The “Amazoogole” model:
 - Lorcan Dempsey: “Amazon, Google, eBay: massive computational and data platforms which exercise strong gravitational Web attraction.”
 - Based primarily on full-text searching and link- or usage-based relevance ranking (lots of recall, little precision)
 - Some efforts to combine catalog and Amazoogole searches (ex.: collaborations with WorldCat)
 - Google is using metadata





Portals and Federations

- **Portals:** defined content boundaries
 - Some content also available elsewhere
 - ex.: Specific library portals, subject portals like Portals to the World (<http://www.loc.gov/rr/international/portals.html>)
- **Federations:** protected content and services
 - Often specialized services based on specifically purposed metadata
 - ex.: BEN-<http://www.biosciencednet.org/portal/>)



XML based digital library application

- Similar to a portal application
- May use a database for record creation and maintenance
- Often uses open source tools
- Files are indexed for searching and presented on the Web using an XML based publishing framework
- Combines some of the other metadata creation, storage and retrieval approaches
- <http://www.loc.gov/performingarts/>



Information Discovery and Retrieval

- Z39.50
- SRU
- Federated search (Metasearch)



Z39.50

- An international (ISO 23950) standard defining a protocol for computer-to-computer information retrieval.
- Makes it possible for a user in one system to search and retrieve information from other computer systems (that have also implemented Z39.50)
- Originally approved by the National Information Standards Organization (NISO) in 1988

The Library of Congress >> Switch to [Library of Congress Authorities](#)



LIBRARY OF CONGRESS ONLINE CATALOG

[About Displaying and Searching Using Non-Roman Characters](#)
[Frequently Asked Questions](#) - [Help Table of Contents](#) - [What's New](#)

Basic Search

Using a fill-in box, search by:

- Title or Author/Creator
- Subject
- Call number
- LCCN, ISSN, or ISBN
- Keywords

Note: Search limits are available only for title and keyword searches.

Guided Search

Using a series of forms and menus:

- Construct keyword searches
- Restrict all or part of the search to a [particular index](#)
- Combine search words or phrases with [Boolean operators](#)

Note: Search limits are available for all searches.

Other Online Catalogs:

Prints and Photographs Online Catalog (PPOC)
[About - Start Searching](#)

Sound Online Inventory & Catalog (SONIC)
[About - Start Searching](#)

Alternative Interface to the LC Online Catalog (Z39.50)
[About - Start Searching](#)

[Copyright Records: Registrations and Documents](#)

[Other Libraries' Catalogs](#)

Information about the images: Two pendentive paintings by Edward J. Holslag are displayed from the Librarian's Room (Librarian's Ceremonial Office) located in the Thomas Jefferson Building of the Library of Congress. On the left, "Efficiunt clarum studio" (Study, the watchword of fame), on the right, "Dulce ante omnia musae" (The Muses, above all things, delightful).

The Library of Congress
July 25, 2007

[Contact Us](#)



SRU Search/Retrieval via URL

- **SRU** is the successor to Z39.50
- **SRU** is a standard XML-focused search protocol for Internet search queries, utilizing CQL (Contextual Query Language), a standard syntax for representing queries
- To learn more about it see:
<http://www.loc.gov/standards/sru/index.html>



Federated search

- Some institutions are using federated search (meta-search) to search multiple data sources
- LC has a new limited version available:
<http://www.loc.gov/search/new/>



Can You Tell?

- Can you tell what's going on behind these sites?
- How are they organized?
- What creation and storage models are used?
 - *Plant and Insect Parasitic Nematodes:*
<http://nematode.unl.edu/>
 - *Public Radio Market:*
<http://www.prms.org/>
 - *Brown University Library Center for Digital Initiatives: Alcohol, Temperance & Prohibition*
[\(http://dl.lib.brown.edu/temperance/\)](http://dl.lib.brown.edu/temperance/)
 - *Country walkers*
<http://www.countrywalkers.com/>

*THIS PAGE INTENTIONALLY LEFT BLANK
FOR DOUBLE SIDED COPY*



8. Metadata Interoperability and Quality Issues

Metadata Standards and Applications Workshop



Goals of Session

- Understand interoperability protocols (OpenURL for reference, OAI-PMH)
- Understand crosswalking and mapping as it relates to interoperability
- Investigate issues concerning metadata quality



Tools For Sharing Metadata/Interoperability

- Protocols
 - OpenURL for reference linking
 - OAI-PMH for harvesting
- Good practices and documentation
- Crosswalking



What's the Point of Interoperability?

- For users, it's about resource discovery (user tasks)
 - What's out there?
 - Is it what I need for my task?
 - Can I use it?
- For resource creators, it's about distribution and marketing
 - How can I increase the number of people who find my resources easily?
 - How can I justify the funding required to make these resources available?



What's an OpenURL?

- The OpenURL provides a standardized format for transporting bibliographic metadata about objects between information services
- Provides a basis for building services via the notion of an *extended service-link*, which moves beyond the classic notion of a *reference link* (a link from metadata to the full-content described by the metadata)



Additional Open URL Services

- Link from a record in an abstracting and indexing database (A&I) to the full-text described by the record
- Link from a record describing a book in a library catalogue to a description of the same book in an Internet book shop
- Link from a reference in a journal article to a record matching that reference in an A&I database
- Link from a citation in a journal article to a record in a library catalogue that shows the library holdings of the cited journal



OpenURL Examples & Demo

- <http://sfxserver.uni.edu/sfxmenu?issn=1234-5678&date=1998&volume=12&issue=2&page=134>
- An OpenURL demo:
 - <http://www.ukoln.ac.uk/distributed-systems/openurl/>



OAI-PMH

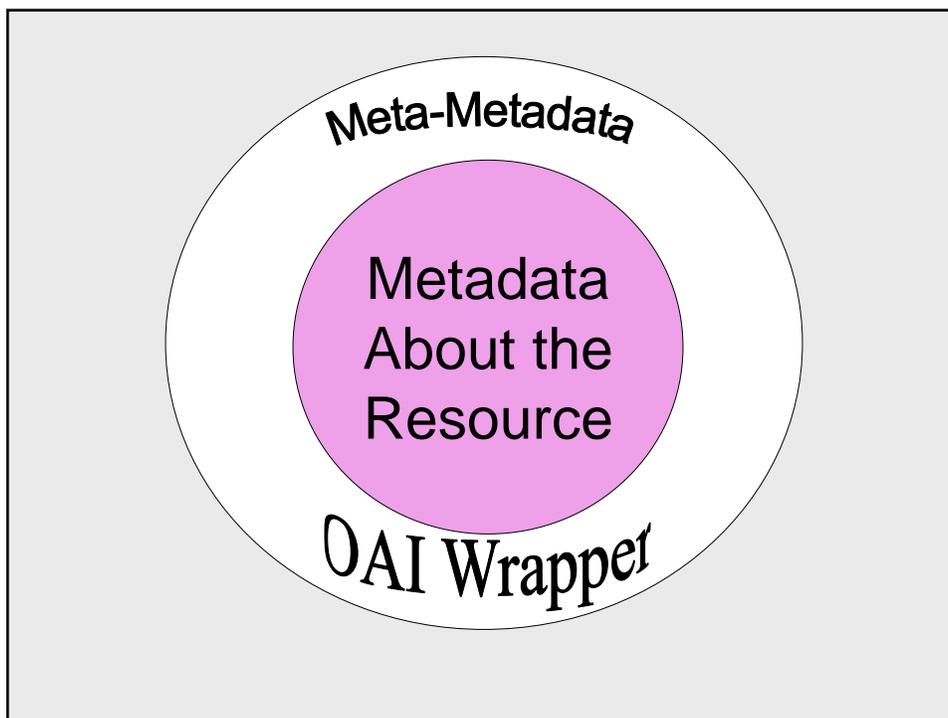
- Open Archives Initiative Protocol for Metadata Harvesting (<http://www.openarchives.org/>)
- Roots in the ePrint community, although applicability is much broader
- Mission: “The Open Archives Initiative develops and promotes interoperability standards that aim to facilitate the efficient dissemination of content.”
- Content in this context is actually “metadata about content”



OAI-PMH in a Nutshell

- Essentially provides a simple protocol for “harvest” and “exposure” of metadata records
- Specifies a simple “wrapper” around metadata records, providing metadata about the record itself
- OAI-PMH is about the **metadata**, not about the **resources**

ARTstor cdwa-Lite experiment <http://www.artstor.org/index.shtml>





What was OAI-PMH designed for?

- Way to distribute records to other libraries
- Low barrier to entry for record providers
- Based on
 - Records must be in XML
 - OAI-PMH supports any metadata format encoded in XML—Simple Dublin Core is the minimal format specified
- Not Z39.50
 - Not a way to support federated search
 - No “on-the-fly” sets.
- More like CDS service, but it’s free,
 - users “pull” records when they want, at intervals that are convenient for them (every day, every hour, on any schedule, or ad hoc)



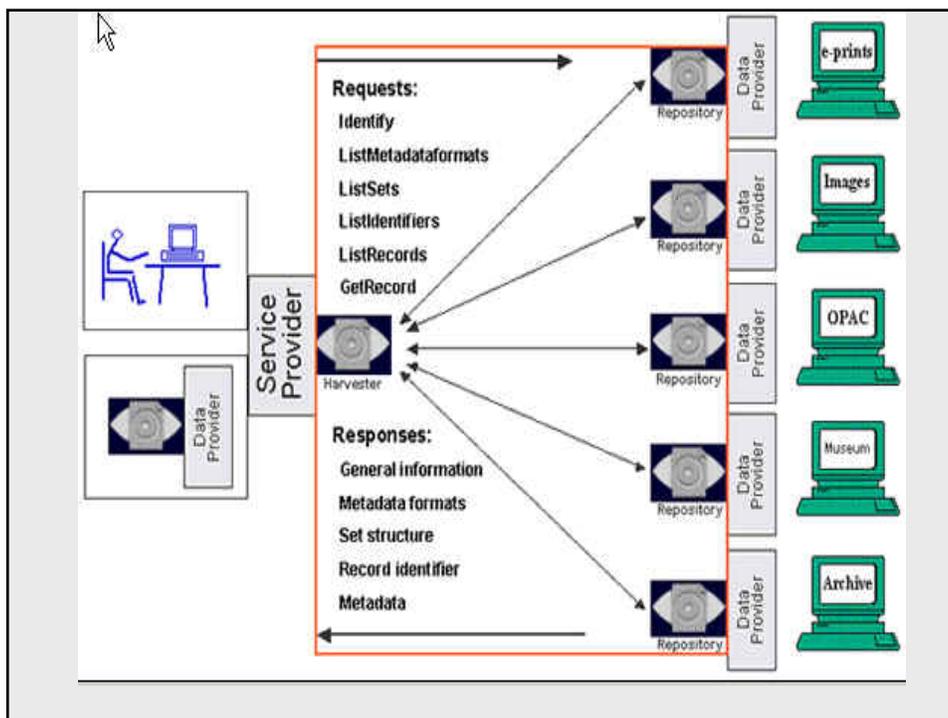
OAI-PMH: Data Provider

- Has records to share
- Runs system that responds to requests
 - following protocol
- Advertises base URL from which records are harvestable
- Just leaves system running
 - No human intervention needed to service requests
 - Can control level of activity to protect performance for primary users



OAI-PMH: Service Provider

- Assumed to be providing “union catalog” service
 - OAIster { <http://www.oaister.org/> }
- or a specialist, value-added service
 - Sheet Music Consortium { <http://digital.library.ucla.edu/sheetmusic/> }
- Harvester records, with ability to select limited to
 - Records updated in a certain timespan
 - Predetermined sets of records (like CDS)
 - Known records by identifiers (OAI identifiers, not LCCNs)



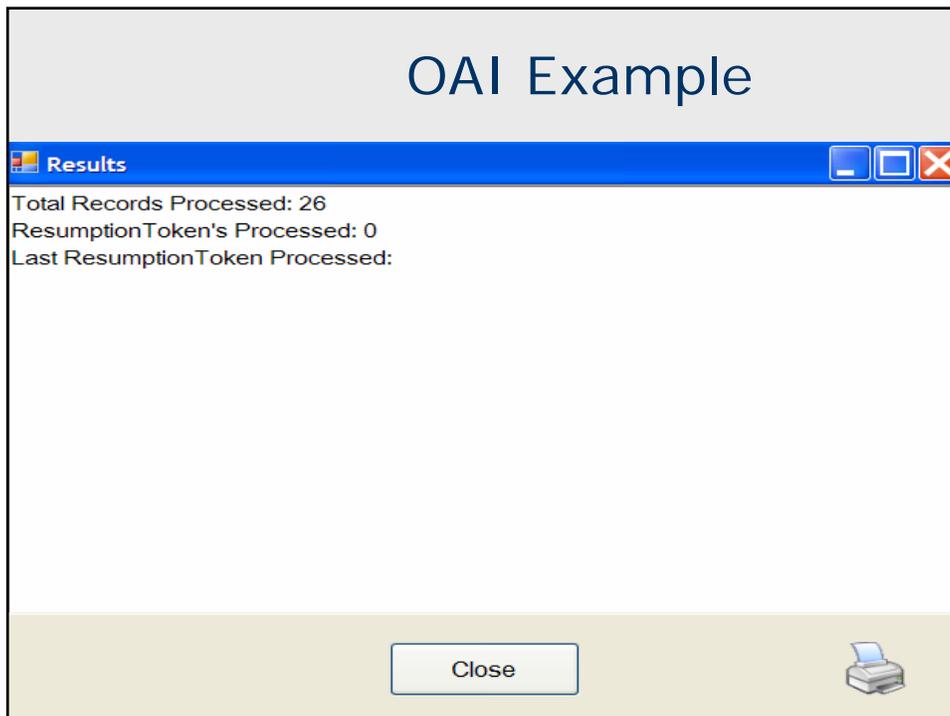
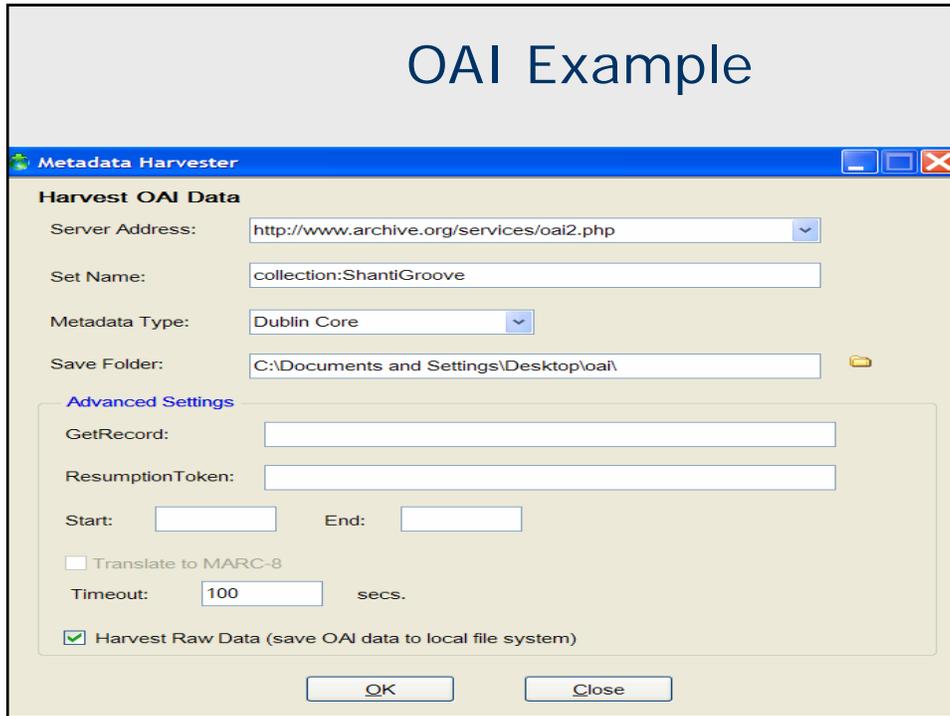


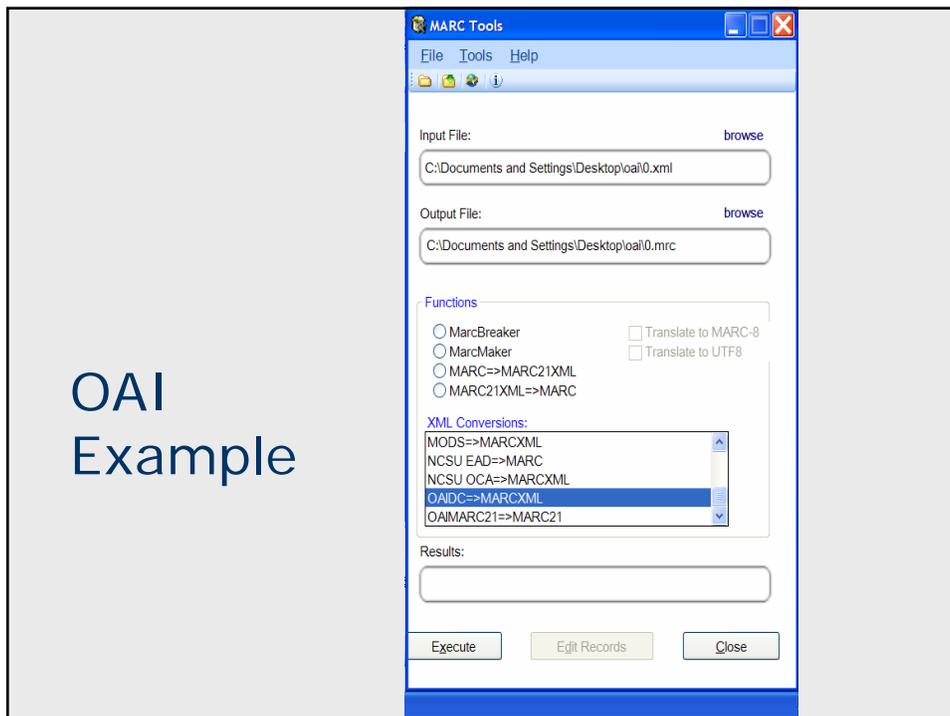
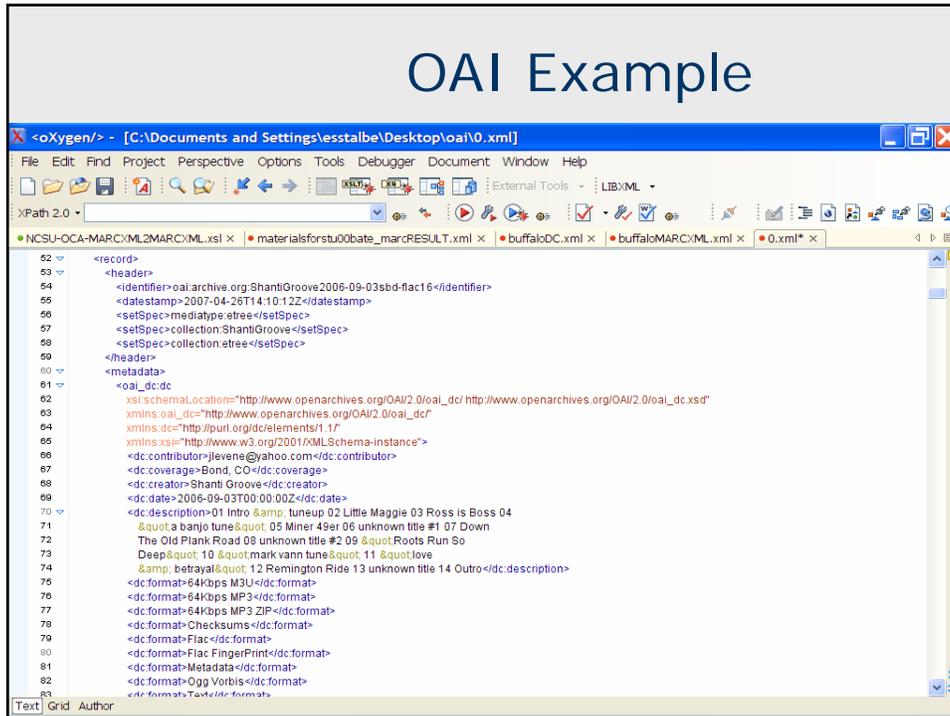
OAI Best Practices Activities

- Sponsored by Digital Library Federation (DLF)
- Guidelines for data providers and service providers
 - <http://oai-best.comm.nsdsl.org/cgi-bin/wiki.pl>
 - Not just DLF, also NSDL
 - Best Practices for Shareable Metadata
 - <http://oai-best.comm.nsdsl.org/cgi-bin/wiki.pl?PublicTOC>
- Workshops to encourage DLF members to make records for their digitized content harvestable
 - Also sponsored by IMLS

OAI Example

The screenshot shows the MarcEdit software interface. The title bar reads "MarcEdit By Terry Reese". The menu bar includes "File", "Tools", "Add-ins", "Plug-ins", and "Help". A yellow navigation bar contains "FAQs: StartPage | MARC Tools | Add-ins | Tutorials | About the Author". The main content area is titled "StartPage" and lists "Frequently Used Tools:" with the following items: MarcBreaker, MarcMaker, MarcEditor, Harvest OAI Records (circled in red), MarcEdit Z39.50/SRU Client, and MarcEdit Help. To the right of the tool list is a colorful illustration of a bucket and tools. At the bottom, it says "What's New in MarcEdit" with a small icon.







OAIster

- **A union catalog of digital resources.** Provides access to digital resources by "harvesting" their descriptive metadata (records) using OAI-PMH.
- Currently provides access to 14,900,092 records from 939 contributors.
- <http://www.oaister.org/>



<http://www.oaister.org/>

Home Search Help About Using OAIster

About OAIster

OAIster currently provides access to 14,900,092 records from 939 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 providers and harvesters.

2007 Dec 12
Using this toolkit, we data provide (and to harvest) from our [MBooks](#) collection, made digitized by Google in partnership with UM

[More...](#)

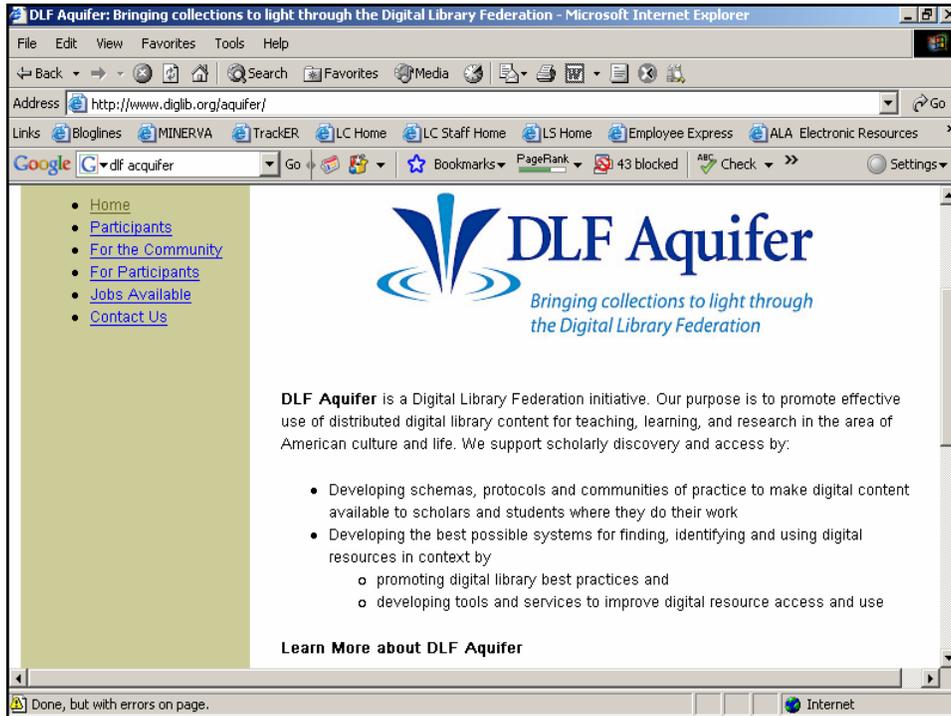
Last updated: 1 Feb 2008
Next update: 29 Feb 2008 (full re-harvest)

New data contributors harvested:

Clark Art Institute Library Repository
Columbia University Libraries Digital Pro Division
Digital Commons@IWL
Lincoln University Research Archive
RWTH Aachen Hochschulbibliothek OPU



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





Crosswalking

“Crosswalks support conversion projects and semantic interoperability to enable searching across heterogeneous distributed databases. Inherently, there are limitations to crosswalks; there is rarely a one-to-one correspondence between the fields or data elements in different information systems.”

-- Mary Woodley, *“Crosswalks: The Path to Universal Access?”*



Crosswalks

- Semantic mapping of elements between source and target metadata standards
- Metadata conversion specification: transformations required to convert metadata record content to another
 - Element to element mapping
 - Hierarchy and object resolution
 - Metadata content conversions
- Stylesheets are created to transform metadata based on crosswalks



Problems With Converted Records

- Differences in granularity (complex vs. simple scheme)
- Some data might be lost
- Differences in semantics
- Differences in use of content standards
- Properties may vary (e.g. repeatability)
- Converting may not always be the solution



Example: Mapping MODS:title to DC:title

- Includes attribute for type of title
 - Abbreviated
 - Translated
 - Alternative
 - Uniform
- Other attributes:
ID, authority, displayLabel, xLink
- Subelements: title, partName, partNumber, nonSort
- Title definition reused by: Subject, Related Item



Mapping MODS:title to DC:title

- DC has one element refinement:
alternative
- DC title has no substructure; MODS allows for subelements for partNumber, partName
- Best practice statement in DC-Lib says include initial article; MODS parses into <nonSort>
- MODS can link to a title in an authority file if desired



Metadata Crosswalks

- Dublin Core-MARC
- Dublin Core-MODS
- ONIX-MARC
- MODS-MARC
- EAD-MARC
- EAD-Dublin Core
- Etc.



Crosswalks

Library of Congress

<http://www.loc.gov/marc/marcdocz.html>

MIT

<http://libraries.mit.edu/guides/subjects/metadata/mappings.html>

Getty

http://www.getty.edu/research/conducting_research/standards/intrometadata/crosswalks.html

MARC to Dublin Core Crosswalk (Unqualified)

- Conventions:
 - 1. "\$" is used to specify the subfield used. If none is specified, use all subfields.
 - 2. DC element is repeated if multiple MARC fields are used.

MARC fields	DC Element	Implementation Notes
100, 110, 111, 700, 710, 711	Contributor	
720		
651, 662	Coverage	
751, 752		
	Creator	Creator element not used.
008/07-10	Date	
260\$c\$g		
500-599, except 506, 530, 540, 546	Description	
340	Format	
856\$q		
020\$a, 022\$a, 024\$a	Identifier	
856\$u		
008/35-37	Language	
041\$a\$b\$d\$e\$f\$g\$h\$j		
546		
260\$a\$b	Publisher	
530, 760-787\$o\$t	Relation	
506, 540	Rights	
534\$t	Source	
786\$o\$t		
050, 060, 080, 082	Subject	
600, 610, 611, 630, 650, 653		
245, 246		
Leader06, Leader07	Type	See Appendix 2 for Leader-Type rules.
655		

MARC to DC Qualified

<http://www.loc.gov/marc/marc2dc.html#qualifiedlist>

III. MARC to Dublin Core Crosswalk (Qualified)

MARC fields	DC Element	DC Qualifier(s)	Implementation Notes
541\$c	Accrual Method		
310\$a	Accrual Periodicity		
521	Audience		
100, 110, 111, 700, 710, 711\$e	Contributor	Value in \$e.	Roles may be used as refinements of Contributor if using qualified DC. See "Roles" note below.
720\$e			
255, 034	Coverage	Spatial	Some 255 information equivalent to DC encoding scheme but different syntax.
522			
630\$a, 651, 662			
751, 752			
043\$c, 044\$c	Coverage	Spatial ISO3166	



NISO's Metadata Principles

- **1:** Good metadata conforms to community standards in a way that is appropriate to the materials in the collection, users of the collection, and current and potential future uses of the collection.
- **2:** Good metadata supports interoperability.
- **3:** Good metadata uses authority control and content standards to describe objects and collocate related objects



NISO's Metadata Principles Continued

- **4:** Good metadata includes a clear statement of the conditions and terms of use for the digital object.
- **5:** Good metadata supports the long-term curation and preservation of objects in collections.
- **6:** Good metadata records are objects themselves and therefore should have the qualities of good objects, including authority, authenticity, archivability, persistence, and unique identification.



Quality issues

- Defining quality
- Criteria for assessing quality
- Levels of quality
- Quality indicators



Determining and Ensuring Quality

- What constitutes quality?
- Techniques for evaluating and enforcing consistency and predictability
- Automated metadata creation: advantages and disadvantages
- Metadata maintenance strategies



Quality Measurement: Criteria

- Completeness
- Accuracy
- Provenance
- Conformance to expectations
- Logical consistency and coherence
- Timeliness (Currency and Lag)
- Accessibility



Basic Quality Levels

- Semantic structure (“format,” “schema” or “element set”)
- Syntactic structure (administrative wrapper and technical encoding)
- Data values or content



Quality Indicators: Tier 1

- Technically valid
 - Defined technical schema; automatic validation
- Appropriate namespace declarations
 - Each element defined within a namespace; not necessarily machine-resolvable
- Administrative wrapper present
 - Basic provenance (unique identifier, source, date)



Quality Indicators: Tier 2

- Controlled vocabularies
 - Linked to publicly available sources of terms by unique tokens
- Elements defined and documented by a specific community
 - Preferably an available application profile
- Full complement of general elements relevant to discovery
- Provenance at a more detailed level
 - Methodology used in creation of metadata?



Quality Indicators: Tier 3

- Expression of metadata intentions based on documented AP endorsed by a specialized community and registered in conformance to a general metadata standard
- Source of data with known history of updating, including updated controlled vocabularies
- Full provenance information (including full source info), referencing practical documentation



Improving Metadata Quality ...

- Documentation
 - Basic standards, best practice guidelines, examples
 - Exposure and maintenance of local and community vocabularies
 - Application Profiles
 - Training materials, tools, methodologies



Exercise

- Evaluate a small set of machine- and human-created metadata

SECTION 8: EXERCISES

Evaluate a small set of human and machine-created metadata

More Information

NSDL OAI identifier	oai:infomine-nsdl.ucr.edu:207689
Title	Ages
Subject Keyword(s)	Forest management -- United States, Forests and forestry -- United States, Branching patterns, Distinguishing characteristics, Needles borne, Terminal buds, Tree identification card, Winter key, Trees in winter, Tree identification, Terminal bud, Identifying trees
Description	Objective: Children will learn several features to observe in order to identify trees in winter. Children will identify several trees common to our area based on careful observation of those features.
Link	http://www.nbtccornell.edu/mainstreetscience/grab_and_go/treeswinterpg.htm
HTML Title	4-H Grab and Go with Science



Dublin Core metadata editor



Results for URL: <http://purl.org/nsdlregistry/> [summary]

```
<link rel="schema.DC" href="http://purl.org/dc/elements/1.1/" />
<link rel="schema.DCTERMS" href="http://purl.org/dc/terms/" />
<meta name="DC.title" content="NSDL Metadata Registry - Portal" />
<link rel="DC.creator" href="http://eg2.ischool.washington.edu/registry/author/rjlaundr" />
<meta name="DC.subject" content="Contact; Home; WCAG; Document Actions; New SOA Registry from Infravio; News; Advanced Search; RDF; 1; Usable in any browser; Goals; More news; SKOS; Links to this content; Powered by Plone, the Open Source Content Management System; National Science Foundation; Log in; ; NSDL Registry Blog; Navigation; Section 508; Portal; Accessibility; Skip to navigation; ; Events; NSDL Metadata Registry; Sections; Take the Survey; Personal tools; NSDL Annual Meeting; Ryan Laundry; National Science Digital Library; Join; Documents; The Registry Blog; Skip to content; Valid CSS; NSDL Registry Public Wiki; 2005-10-03; Valid XHTML; NSDL Registry Grant Awarded; Site Map" />
<meta name="DC.date" scheme="DCTERMS.W3CDTF" content="2005-11-29" />
<meta name="DC.type" scheme="DCTERMS.DCHIType" content="Text" />
<meta name="DC.format" content="text/html" />
<meta name="DC.format" content="4858 bytes" />
<meta name="DC.identifier" scheme="DCTERMS.URI" content="http://purl.org/nsdlregistry/" />
```

[\[how to use the XHTML description\]](#)

Album Title: Brazilliance Vol. 3
Subtitle: Bud Shank, arrangements by Laurindo Almeida
Notes: Bud Shank: Flute
Performer: Almeida, Laurindo
Composer: Almeida, Laurindo
 Monk/Hanigan/Williams
 Shank, Bud
 Webster, H. D. L.
 Gershwin, George

Album back/Liner notes: <http://digital-library.csun.edu/IGRA/images/B1D13b.gif>

Track 1: Side a-1 : Harlem Samba : Laurindo Almeida (composer) : Laurindo Almeida (performer) : 2 Guitars; or Flute, Guitar
Track 2: Side a-2 : North of the Border : Laurindo Almeida (composer) : Laurindo Almeida (performer) : 2 Guitars; or Flute, Guitar
Track 3: Side a-3 : Sunset Baion : Laurindo Almeida (composer) : Laurindo Almeida (performer) : 2 Guitars; or Flute, Guitar
Track 4: Side a-4 : 8Round Midnight : Monk/Hanigan/Williams (composer) : Laurindo Almeida (performer) : 2 Guitars; or Flute, Guitar
Track 5: Side a-5 : Toro Dance : Bud Shank (composer) : Laurindo Almeida (performer) : 2 Guitars; or Flute, Guitar
Track 6: Side b-1 : Serenade for alto : Laurindo Almeida (composer) : Laurindo Almeida (performer) : 2 Guitars; or Flute, Guitar
Track 7: Side b-2 : Xana-Lyn : Bud Shank (composer) : Laurindo Almeida (performer) : 2 Guitars; or Flute, Guitar
Track 8: Side b-3 : Blowing Wild : H. D. L. Webster (composer) : Laurindo Almeida (performer) : 2 Guitars; or Flute, Guitar
Track 9: Side b-4 : Gershwin Prelude : George Gershwin (composer) : Laurindo Almeida (performer) : 2 Guitars; or Flute, Guitar
Track 10: Side b-5 : Frio y color : Bud Shank (composer) : Laurindo Almeida (performer) : 2 Guitars; or Flute, Guitar

Record Company: World Pacific Records
Release Date: 1963
Music Genre: Jazz
Media: 331/3
Collection: John Tanno Collection
Serial number: Rec-46
Box number: 1
Disc number: 13
Identifier: B1D13
Record Label catalog number: WP-1425
Matrix No. A: A-826
Matrix No. B: B-826

NSDL OAI identifier	oai:nsdl.org:nsdl:00126
Title	ICSD Science Zone
Subject Keyword(s)	Science--Study and teaching--New York (State)--Ithaca, General science
Description	Welcome to the Ithaca City School District's Science Zone. This supports science education for students in the Ithaca City School District and the global community. Learn science and let it take you places. Enjoy!
Publisher / Resource Provider	Ithaca City School District
Resource type	text/html, image
Link	http://ithacasciencezone.com
Language	en
Grade Level	Grades Pre-K to 12
HTML Title	ICSD Science Zone

Ages:

8 to 12

HANDOUT: IDENTIFYING TREES IN WINTER

Contributor: Susan Jaquette, Cornell Plantations volunteer

Main idea: Trees have distinguishing characteristics that enable identification even in winter.

Objective: Children will learn several features to observe in order to identify trees in winter. Children will identify several trees common to our area based on careful observation of those features.

Materials:

- ❑ Small branches from several deciduous trees to illustrate the different branching patterns and terminal buds.
- ❑ Small branches from several evergreen trees to illustrate pine needle bundles, needles borne singly, needles borne on a stem or not on a stem, the scale-like foliage of arborvitae or cedar, and differences among cones.
- ❑ Book or chart showing characteristic silhouettes of common trees.
- ❑ *Know Your Trees* (4-H Bulletin 85) has a winter key that identifies hardwoods based on twigs and buds. There is an online version of the winter key available at: <http://cyrus.bh.cornell.edu/tree/winterkey.htm>
- ❑ Other tree handbooks for reference.

Motivator: To get at the idea of identification through different features, ask children how they might recognize someone they know who was far away, for example across a field. (You might elicit posture, shape, size, and silhouette.) Then ask how do you recognize someone who is nearby, for example, someone in the group. (More specific features will be elicited ? blue eyes, freckles, curly hair, etc.)

Explain that trees, too, have features and can be recognized by careful attention to their features. Scientists have organized many of these features into what's called a taxonomic key for identifying trees.

Questions: Before you begin the activity, ask the children what features they might look at if you were trying to distinguish one tree from another in winter? Elicit as many characteristics as possible.

Activity:

1. Beforehand, the instructor should locate 4 to 6 nearby trees that will be easy to identify based on their particular characteristics. Trees such as red pine, white pine, hemlock, spruce, sycamore, oak, maple, and ash lend themselves to using the distinguishing characteristics presented to the students in Step 3. For example:
 - ❑ The two pines have a different number of needles in their clusters.
 - ❑ The hemlock and spruce have different cones, and the hemlock's needles are on short stems while the spruce needles lack stems.
 - ❑ The sycamore lacks a terminal bud and has a zigzag branching pattern while the oak has a terminal bud and an alternate branching pattern.
 - ❑ The maple and the ash are the only common local large trees with opposite buds and branches.
2. Post a number on each of the trees selected. Make a tree identification card for each tree. Each card should have the tree's name, the descriptive characteristics needed to identify that tree (as stated, for example, in Step 1), possibly a silhouette sketch of the tree, and any other prominent characteristics. (For example, the ash has stout twigs or the sycamore has peeling bark high on its trunk.)
3. Have the children examine the silhouette poster, the branches and cones looking for differences among them. Show them or help them discover the different branching patterns. Point out branches with terminal buds and those without. Have them count pine needles in bundles and observe needles borne singly on or without stems. Have them describe how the cones differ. Everyone should have hands-on experience and practice at this time.
4. Divide children into as many groups as trees you have numbered. Give each group a tree identification card and instruct them to examine each numbered tree until they are sure they have found the one on their card.
5. After all groups have found their trees, gather the entire group together and go from tree to tree having each group point out the distinguishing characteristics of their tree.

Learning checks: In addition to the built in check in Step 5, instructor can have the group examine other nearby trees and demonstrate their knowledge of branching patterns, terminal buds, and needle arrangement.

Background: Basic knowledge about tree identification and familiarity with distinguishing characteristics of the 4 to 6 trees chosen for the activity.

Vocabulary:

Terminal bud: The bud found at the end of a twig.

Opposite: Arrangement of leaves or buds so that they connect to branches at points directly across from each other.

Alternate: Arrangement of leaves or buds where they are staggered or not placed directly across from each other on the twig.

Zig-zag branches: Change direction at each bud, creating a zig-zag pattern.

Needle cluster or bundle: A group of needles attached together at one point to a twig.

Taxonomic key: A guide that helps you find the name of an organism, usually by asking sets of paired questions.

Extensions:

- Have groups switch cards and identify other trees (writing down their answers) before moving on to Step 5.

Number additional "distracter" trees so students must consider and observe more trees.

ICSD SCIENCE ZONE

Welcome to the Ithaca City School District's Science Zone. This supports science education for students in the Ithaca City School District and the global community. Learn science and let it take you places. Enjoy!

 **ICSD Elementary Science Zone** 

ICSD Middle School Science Zone



Earth Science Zone



Biology Zone

The ICSD EnviroZone

ICSD Chemistry Zone



The Physics Zone

Science Teacher Zone

Search All Science Zones for:

Search the entire internet using Google:



*THIS PAGE INTENTIONALLY LEFT BLANK
FOR DOUBLE SIDED COPY*

Dublin Core to MARC Crosswalk

Network Development and MARC Standards Office
Library of Congress

Date issued: 2008-04-23

Previous version: http://www.loc.gov/marc/dccross_20010312.html

I. Introduction

The following is a crosswalk between the metadata terms in the [Dublin Core Element Set](#) and [MARC 21](#) bibliographic data elements. The crosswalk may be used for conversion of Dublin Core metadata into MARC, for instance as a tool for developing XSLT transformations. For conversion of MARC 21 into Dublin Core, the [MARC to Dublin Core Crosswalk](#) should be used, since in this conversion many fields are mapped into a single Dublin Core element.

In the Dublin Core to MARC mapping, two mappings are provided, one for unqualified Dublin Core elements (i.e. the main fifteen elements as in [Dublin Core Metadata Element Set, version 1.1](#)) and the other for qualified (those in addition that appear only in [DCMI Metadata Terms](#)). The latter includes both refinements of the original fifteen as well as syntax and vocabulary encoding schemes.

MARC 21 fields are listed with field number, then two indicator values with field name/subfield name in parentheses. If both the field and subfield have the same name, the subfield name is not included. A blank (H'20') is indicated in this document by "#". The label is a shortened form of the element name.

Definitions are taken from [Dublin Core Metadata Element Set, Version 1.1: Reference Description](#). For further information about Dublin Core elements, including application notes (given in Comment), refer to that document. All Dublin Core elements are optional and repeatable. In this document elements are listed in alphabetical order by Dublin Core label.

II. Dublin Core to MARC Crosswalk (15 elements in version 1.1 and refinements)

Contributor -- An entity responsible for making contributions to the resource.

Unqualified:

720 ##\$a (Added Entry--Uncontrolled Name/Name)

Qualified:

Element refinements may be used from the MARC relator list for those specified roles that refine dc:contributor. If DC metadata includes a role refinement (only used for Contributor), use the term in 720\$t or the code in 720\$t. See: [MARC Relator Terms and Dublin Core](#)

Note: there is no way to specify whether the Contributor is a person or organization because it is not in the Dublin Core data. If it can reasonably be determined that the contributor is a

person or organization, fields 700 1#\$a (Added Entry--Personal Name) or 710 2#\$a (Added Entry--Corporate Name) may be used.

Coverage -- The spatial or temporal topic of the resource, the spatial applicability of the resource, or the jurisdiction under which the resource is relevant.

Unqualified:

500\$a (General note)

Qualified:

Spatial: 522 ##\$a (Geographic Coverage Note)

Temporal: 513 ##\$b (Type of Report and Period Covered Note/Period covered)

Syntax encoding schemes:

Box: 507\$a (Scale Note for Graphic Material)

ISO3166: 043\$c with \$2 iso3166 (Geographic Area Code/ISO code)

Point: 507\$a (Scale Note for Graphic Material)

Vocabulary encoding scheme under spatial:

TGN: 651 #7 \$a (Subject Added Entry--Geographic Name) with \$2=tgn

Period: 045 0# \$b (Time Period of Content/Formatted 9999 B.C. through C.E. time period)

Creator -- An entity primarily responsible for making the resource.

Unqualified:

720 ##\$a (Added Entry--Uncontrolled Name/Name) with \$e=author

Note: there is no way to specify whether the Creator is a person or organization because it is not in the Dublin Core data. If it can reasonably be determined that the Creator is a person or organization, fields 100 1#\$a (Main Entry--Personal Name) or 110 2#\$a (Main Entry--Corporate Name) may be used for the first Creator, and 700 1#\$a (Added Entry--Personal Name) or 710 2#\$a (Added Entry--Corporate Name) for others may be used.

Date -- A point or period of time associated with an event in the lifecycle of the resource.

Unqualified:

260 ##\$c (Date of publication, distribution, etc.)

Qualified:

Available: 307 ##\$a (Hours, Etc.)

Created: 046\$k (Special Coded Dates/Date created)

Date Accepted: 502##\$a (Dissertation Note) with initial label "Date accepted"

Date Copyrighted: 260##\$c (Date of publication, distribution, etc.) or 542 \$g (Information Related to Copyright Status/Copyright date) *Note: this field was defined in MARC in January 2008. Prefer 542 if field is available.*

Date Submitted: 502##\$a (Dissertation Note) with initial label "Date submitted"

Issued: 260 ##\$c (Date of publication, distribution, etc.)

Modified: 046##\$j (Special Coded Dates/Date modified)

Valid: 046##\$m (Special Coded Dates/Date valid)

Syntax encoding schemes:

Period: 045##\$c (Special Coded Dates/Date 1, C.E. date)

W3CDTF: 260 ##\$c (Date of publication, distribution, etc.); date may also be generated in 008/07-10; see below under Notes. Remove hyphens in 008.

Description -- An account of the resource.

Unqualified:

520 ##\$a (Summary, etc. note)

Qualified:

Abstract: 520 ##\$a (Summary, etc. note)

TableofContents: 505 0#\$a (Formatted Contents Note)

Format -- The file format, physical medium, or dimensions of the resource.

Unqualified:

856 ##\$q (Electronic Location and Access/Electronic format type)

Qualified:

Extent: 300 ##\$a (Physical Description)

Medium: 340 ##\$a (Physical Medium)

Syntax encoding scheme:

IMT: 856 ##\$q (Electronic Location and Access/Electronic Format Type)

Identifier -- An unambiguous reference to the resource within a given context.

Unqualified:

If string begins http://: 856 40 \$u (Electronic Location and Access/URI)

Otherwise: 024 8#\$a (Other Standard Identifier/Standard number or code)

Qualified:

Bibliographic Citation: 500 ##\$a (General note) with initial label "Bibliographic citation"

Syntax encoding scheme:

URI: 856 40\$u (Electronic Location and Access/Uniform Resource Locator)

Note: other types of identifiers (e.g. ISSN, ISBN) may be expressed in the form of URIs in dc:identifier.

Language -- A language of the resource.

Unqualified:

546 ##\$a (Language note)

Qualified:

Syntax encoding schemes:

ISO 639-2: 0410#\$a (Language code)

RFC 1766: 041 07\$a (Language code) with \$2=rfc1766

RFC 3066: 041 07\$a (Language code) with \$2=rfc3066

RFC 4646: 041 07\$a (Language code) with \$2=rfc4646

Publisher -- An entity responsible for making the resource available.**Unqualified:**

260 ##\$b (Publication, Distribution, etc. (Imprint)/Name of publisher, distributor, etc.)

Relation -- A related resource.**Unqualified:**

787 0#\$n (Nonspecific Relationship Entry/Note)

Qualified:

URI (Encoding scheme): 787 0#\$o (Nonspecific Relationship Entry/Other identifier)

Conforms To: 787 0#\$n (Nonspecific Relationship Entry/Other identifier) with \$i=Conforms to

HasFormat: 776 0#\$n (Additional Physical Form Entry/Note)

HasFormat (Scheme=URI): 776 0#\$o (Additional Physical Form Entry/Other identifier)

IsFormatOf: 776 0#\$n (Additional Physical Form Entry/Note)

IsFormatOf (Scheme=URI): 776 0#\$o (Additional Physical Form Entry/Other identifier)

IsPartOf: 773 0#\$n (Host Item Entry/Note)

IsPartOf (Scheme=URI): 773 0#\$o (Host Item Entry/Other identifier)

HasPart: 774 0#\$n (Constituent Unit Entry/Note)

HasPart (Scheme=URI): 774 0#\$o (Constituent Unit Entry/Other identifier)

IsVersionOf: 775 0#\$n (Other Edition Entry/Note)

IsVersionOf (Scheme=URI): 775 0#\$o (Other Edition Entry/Other identifier)

HasVersion: 775 0#\$n (Other Edition Entry/Note)

HasVersion (Scheme=URI): 775 0#\$o (Other Edition Entry/Other identifier)

Is Based On: 786 0#\$n (Data Source Entry/Note)

Is Based On (Scheme=URI): 786 0#\$o (Data Source Entry/Other identifier)

Is Referenced By: 510 0#\$a (Citation/References Note/Name of source)

Requires: 538 ##\$a (System Details Note)

Is Required By: 787 0#\$n (Nonspecific Relationship Entry/Other identifier) with \$i=Is required by

Replaces: 780 00\$n (Preceding entry)

Replaces (Scheme=URI): 780 00\$o (Preceding entry)

Is Replaced By: 785 00\$n (Succeeding entry)

Is Replaced By (Scheme=URI): 785 00\$o (Succeeding entry)

References: 787 0#\$n (Nonspecific Relationship Entry/Other identifier) with \$i=References

References (Scheme=URI): 787 0#\$o (Nonspecific Relationship Entry/Other identifier) with \$i=References

Rights -- Information about rights held in and over the resource.

Unqualified:

540 ##\$a (Terms Governing Use and Reproduction Note)

Qualified:

Access Rights: 506##\$a (Restrictions on Access Note)

License: 540##\$a (Terms Governing Use and Reproduction Note)

License (if value is a URI): 540##\$u

Source -- A related resource from which the described resource is derived.

Unqualified:

786 0#\$n (Data Source Entry/Note)

Qualified:

Syntax encoding scheme

URI : 786 0#\$o (Data Source Entry/Other identifier)

Subject -- The topic of the resource.

Unqualified:

653 ##\$a (Index Term--Uncontrolled)

Qualified:

Vocabulary encoding schemes:

DDC: 082 ##\$a (Dewey Decimal Call Number/Classification number)

LCC: 050 ##\$a (Library of Congress Call Number/Classification number)

LCSH: 650 #0\$a (Subject added entry--Topical term)

MeSH: 650 #2\$a (Subject added entry--Topical term)

UDC: 080 ##\$a (Universal Decimal Classification Number)

Title -- A name given to the resource.

Unqualified:

245 00\$a (Title Statement/Title proper)

If repeated, all titles after the first: 246 33\$a (Varying Form of Title/Title proper)

Qualified:

Alternative: 246 33\$a (Varying Form of Title/Title proper)

Type -- The nature or genre of the resource.

Unqualified:

655 #7\$a (Index Term--Genre/Form) with \$2=local

Qualified:

DCMI Type vocabulary (Encoding scheme): 655 #7\$a (Index Term--Genre/Form) with \$2=dc

See Section III for use to determine Leader/06 (Type of Record) values.

III. Additional Dublin Core Elements

Accrual Method: 541##\$c (Source of Acquisition Note)

Accrual Periodicity: 310##\$a (Current Publication Frequency)

Audience: 521##\$a (Target Audience Note)

Education Level: 521##\$a (Target Audience Note)

Provenance: 561##\$a (Ownership and Custodial History)

IV. Notes

In addition to the variable length fields listed in the mapping, a MARC 21 record will also include a Leader and field 008 (Fixed-Length Data Elements). Certain character positions in each of these fixed length fields of a USMARC record will need to be coded, although most will generate default values.

Leader: a fixed field comprising the first 24 character positions (00-23) of each record that provides information for the processing of the record. The following positions should be generated:

Character Position 06: Type of record

Leader/06 value should be set according to value in Type as follows (these values are from Dublin Core List of Resource Types (DC Type Vocabulary)):

Type value	Leader/06 value
collection	p
dataset	m
event	r
image	k
interactive resource	m
service	m
software	m
sound	i

text	a
------	---

If no type is indicated, use value "a". If two type values are indicated, and one of these is "collection" use the other value for setting Leader/06. If more than two, use "m".

Character Position 07: Bibliographic level

If Type value is collection, use value "c" (Collection)

All others, use value "m" (Monograph).

Character Position 08: Type of control

Use value "#" (blank: no specific type of control).

Character Position 09: Character coding scheme

Use value "#" (blank: MARC-8).

Character Position 17: Encoding level

Use value "3" (Abbreviated level) or other value as appropriate to application

Character Position 18: Descriptive cataloging form

Use value "u" (Unknown) to indicate that the descriptive cataloging form is unknown.

008 Fixed Length Data Elements: Forty character positions (00-39) containing positionally-defined data elements that provide coded information about the record as a whole or about special bibliographic aspects of the item being cataloged. For records originating as Dublin Core, the following character positions are used:

Character positions 00-05: Date the MARC 21 record was created or converted (generate by date record entered system; formatted as YYMMDD)

Character positions 07-10: Date of Publication (YYYY portion from Date if present).
Qualified DC: Date.Issued in ISO 8601 (only YYYY portion).

Character positions 35-37: Language. May be generated from data in Language if scheme=ISO 639-2.

Other character positions can default to fill characters (ASCII 7C)

042\$a Authentication Code: Use "dc" (identifies that MARC 21 record is derived from Dublin Core style record).

IV. Uses for mapping Dublin Core to MARC

A mapping between the elements in the Dublin Core and MARC 21 fields is necessary so that conversions between various syntaxes can occur accurately. Once Dublin Core style metadata is widely provided, it might interact with MARC records in various ways such as the following:

Enhancement of simple resource description record. A cataloging agency may wish to extract the metadata provided in Dublin Core style and convert the data elements to MARC 21 fields, resulting in a skeletal record. That record might then be enhanced as needed to add additional information generally provided in the particular catalog. Some projects convert data and use as basic record for reporting to national bibliography.

Searching across syntaxes and databases. Libraries have large systems with valuable information in metadata records in MARC format. It will be important for systems to be able to search metadata in different syntaxes and databases and have commonality in the definition and use of elements. A primary use for this mapping is for harvesting metadata records from different sources, where there may be a need to translate Dublin Core records to MARC or vice versa.

Go to:

[MARC Home Page](#)

[Library of Congress Home Page](#)



Library of Congress

[Library of Congress Help Desk](#) (04/25/2008)

MARC to Dublin Core Crosswalk

Development and MARC Standards Office
Library of Congress

Date issued: 2008/04/24

Previous version: <http://www.loc.gov/marc/marc2dc-2001.html>

Table of Contents

- [I. Introduction](#)
- [II. MARC to Dublin Core \(Unqualified\)](#)
- [III. MARC to Dublin Core \(Qualified\)](#)
- [IV. Appendix 1--Notes](#)
- [V. Appendix 2--Conversion rules for Leader/06 to dc:Type mapping](#)

I. Introduction

The following is a crosswalk between core [MARC 21](#) bibliographic data elements and elements in the [Dublin Core Element Set](#). It may be used in conversion of metadata from MARC into Dublin Core. Since MARC is richer in data than Dublin Core, it differs from the [Dublin Core to MARC Crosswalk](#) in that multiple MARC fields are mapped to a Dublin Core element. The Dublin Core to MARC crosswalk maps a Dublin Core element to a single MARC field. In both crosswalks there are different mappings for Dublin Core simple or qualified. Not all possible MARC fields are included in this mapping, but only those considered useful for broad cross-domain resource discovery. Applications may wish to include other MARC elements that are prevalent in their data but are not listed here, or they may not include all that are listed.

MARC 21 fields are listed by field number with specific subfields if applicable. In many cases specific subfields are not provided, since applications may differ in subfields used. Applications may not need control subfields such as \$2, \$5. Notes concerning implementation are given. Further information about the mapping is given at the end of the document. Definitions of MARC elements may be found at the [MARC Bibliographic format](#) site and definitions of Dublin Core elements from the [Dublin Core Metadata Element Set Reference Description, Version 1.1](#) and [DCMI Metadata Terms](#).

Note that it is not expected that round-trip mapping is possible using this crosswalk. Once MARC data is converted to Dublin Core, not enough information is retained to allow for mapping back to MARC accurately. This is inevitable when mapping a complex set of data elements to a simpler set. Where a tag is used in mapping to more than one DC element, a decision may need to be made which Dublin Core element to map it to.

II. MARC to Dublin Core Crosswalk (Unqualified)

Conventions:

1. "\$" is used to specify the subfield used. If none is specified, use all subfields.

2. DC element is repeated if multiple MARC fields are used.

MARC fields	DC Element	Implementation Notes
100, 110, 111, 700, 710, 711	Contributor	
720		
651, 662	Coverage	
751, 752		
	Creator	Creator element not used.
008/07-10	Date	
260\$c\$g		
500-599, except 506, 530, 540, 546	Description	
340	Format	
856\$q		
020\$a, 022\$a, 024\$a	Identifier	
856\$u		
008/35-37	Language	
041\$a\$b\$d\$e\$f\$g\$h\$j		
546		
260\$a\$b	Publisher	
530, 760-787\$o\$t	Relation	
506, 540	Rights	
534\$t	Source	
786\$o\$t		
050, 060, 080, 082	Subject	
600, 610, 611, 630, 650, 653		
245, 246	Title	Repeat dc:title for each. Some applications may wish to include 210, 222, 240, 242, 243, and 247.
Leader06, Leader07	Type	See Appendix 2 for Leader-Type rules.
655		

III. MARC to Dublin Core Crosswalk (Qualified)

MARC fields	DC Element	DC Qualifier(s)	Implementation Notes
541\$c	Accrual Method		
310\$a	Accrual Periodicity		
521	Audience		
100, 110, 111, 700, 710, 711\$e	Contributor	Value in \$e.	Roles may be used as refinements of Contributor if using qualified DC. See "Roles" note below.
720\$e			
255, 034	Coverage	Spatial	Some 255 information equivalent to DC encoding scheme but different syntax.
522			
650\$z, 651, 662			
751, 752			
043\$c,044\$c	Coverage	Spatial ISO3166	
651 if ind2=7 & \$2=tgn		Spatial TGN	
033\$a	Coverage	Temporal	
533\$b			
260\$c if precedes date	Date	Copyrighted	This field was defined in MARC in January 2008.
542\$g			
260\$c\$g	Date	Created	
533\$d			
008/07-10	Date	Issued	
260\$c			
046\$j	Date	Modified	
046\$m\$n	Date	Valid	
520 if ind1=# or 3	Description	Abstract	
505	Description	TableofContents	

300\$a	Format	Extent	
533\$e			
856\$q			IMT
340\$a			Medium
856\$u	Identifier	URI	
008/35-37	Language	ISO369-2	
041 with no \$2		ISO639-2	
041 with \$2=iso639-3		ISO639-3	
041 with \$2=rfc1766		RFC1766	
041 with \$2=rfc3066		RFC3066	
041 with \$2=rfc4646		RFC4646	
561		Provenance	
530	Relation	HasFormat	
776\$n\$t			
530\$u	Relation	HasFormat URI	
776\$o			
774\$n\$t	Relation	HasPart	
774\$o	Relation	HasPart URI	
775\$n\$t	Relation	HasVersion	
775\$o	Relation	HasVersion URI	
530	Relation	IsFormatOf	
776\$n\$t			
530\$u	Relation	IsFormatOf URI	
776\$o			
440, 490, 800, 810, 811, 830	Relation	IsPartOf	
760, 773\$n\$t			
760, 773\$o	Relation	IsPartOf	

		URI	
510	Relation	IsReferencedBy	
785\$n\$t	Relation	IsReplacedBy	
785\$o	Relation	IsReplacedBy URI	
775,786\$n\$t	Relation	IsVersionOf	
775, 786\$o	Relation	IsVersionOf URI	
780\$n\$t	Relation	Replaces	
780\$o	Relation	Replaces URI	
538	Relation	Requires	
506\$a\$d	Rights	AccessRights	
540\$a\$d			
542\$d	Rights	RightsHolder	This field was defined in MARC in January 2008.
786\$o	Source	URI	
082	Subject	DDC	
050	Subject	LCC	
600, 610, 611, 630, 650 if ind2=0	Subject	LCSH	
600, 610, 611, 630, 650 if ind2=2	Subject	MeSH	
060	Subject	NLM	
655 if ind2=7 & \$2=tgn	Subject	TGN	
080	Subject	UDC	
245	Title		
130, 210, 240, 242, 246, 730, 740	Title	Alternative	
Leader06, Leader07	Type	DCMI Type Vocabulary	See Appendix 2 for Leader-Type rules.
655 if \$2=dct			

Appendix 1--Notes

DC Refinements and encoding schemes. Some DC refinements or encoding schemes have not been included because they are not generally found or can not be specifically identified in MARC data. Examples include DCMI Point, DCMI Box and Mediator.

Roles. Those roles in the MARC Relators list that indicate that they refine "Contributor" may be used in qualified DC. See: [Relator Terms and Dublin Core Elements](#).

Appendix 2 - Conversion rules for Leader06 - dc:Type mapping

Multiple Type fields may be used; conversions below may result in 2 or 3 Type fields. There are several additional sources of type information in the MARC record; only coded values in Leader/06 and Leader/07 are detailed in this chart. Field 655 may also be used for more specific type information.

Leader/06 value	Leader/07 value	Type value
a,c,d,t		text
e,f,g,k		image
i,j		sound
m,o,p,r		no type provided
p	c,s	collection

Institutions may want to consider generating additional type values, such as "map" or "cartographic" for codes e or g; "musical notation" for codes c or d, etc.

Go to:

[Dublin Core to MARC Crosswalk](#)
[MARC Home Page](#)
[Library of Congress Home Page](#)



Library of Congress

[Library of Congress Help Desk](#) (05/02/2008)

Terminology

Metadata: In general, data about data; functionally, structured data about data. Metadata includes data associated with either an information system or an information object for purposes of description, administration, legal requirements, technical functionality, use and usage, and preservation. Traditional library cataloging is a form of metadata. (Dublin Core Metadata Initiative glossary)

Schema: Defines the vocabulary of a particular set of metadata (i.e., element names and formatting rules). A schema is usually defined by some authority to describe data in a standard way so that it may be accessed by other users or applications. (Tom Sheldon's Linktionary.com)

Metadata Schema: Sets of metadata elements designed for a specific purpose, such as describing a particular type of information resource. The definition or meaning of the elements themselves is known as the semantics of the scheme. The value given to the metadata elements are the content. (Understanding Metadata)

Dublin Core Metadata Element Set: Includes terms intended to facilitate the discovery of resources. The metadata is associated with the intellectual content, intellectual property, and/or instantiation characteristics of an information resource. (Dublin Core Metadata Initiative glossary)

Standard Generalized Markup Language (SGML): A non-proprietary language/enabling technology for describing information. Information in SGML is structured like a database, supporting rendering in and conversion between different formats. Both XML and later versions of HTML are instances of SGML. (Dublin Core Metadata Initiative glossary)

Hypertext Markup Language (HTML): The standard text-formatting language for documents on the World Wide Web. HTML text files contain content that is rendered on a computer screen and markup, or tags, that can be used to tell the computer how to format that content. HTML tags can also be used to encode metadata and to tell the computer how to respond to certain user actions, such as a mouse click. (Dublin Core Metadata Initiative glossary)

eXtensible Markup Language (XML): A subset of SGML. XML is designed to bring the power and flexibility of generic SGML to the World Wide Web, while maintaining interoperability with full SGML and HTML. (Dublin Core Metadata Initiative glossary)

XML Schemas: Express shared vocabularies and allow machines to carry out rules made by people. They provide a means for defining the structure, content and semantics of XML documents. (World Wide Web Consortium Web page)

Namespaces An effort to allow markup from different XML applications to be used in the same document without conflict (even if element names used in each namespace are the same).

XSL Extensible Stylesheet Language, an XML application for transforming XML documents into a form that could be viewed in a web browser. Eventually split to form: XSLT and XSL-FO.

XSL-FO An XML application for describing the layout of both printed pages and webpages, sometimes compared to PostScript.

CSS Cascading Style Sheets, used originally for HTML, and, when XML was invented, it was used for that, too.

Xlink Extensible Linking Language, used to define more powerful linking constructs to connect XML documents, going beyond the “a” tag used in HTML.

Xpointer Addresses individual parts of an XML document.

XPath Extracted from the addressing parts of XLink and XPointer.

Metadata Encoding and Transmission Standard (METS): XML schema for encoding descriptive, administrative, and structural metadata regarding objects within a digital library. (METS Web page)

Metadata Object Description Schema (MODS): XML schema for a bibliographic element set that may be used for a variety of purposes, and particularly for library applications. MODS is intended to be able to carry selected data from existing MARC 21 records as well as to enable the creation of original resource description records. It includes a subset of MARC fields and uses language-based tags rather than numeric ones, in some cases regrouping elements from the MARC 21 bibliographic format. (MODS Web page)

Metadata Authority Description Schema (MADS): XML schema for an authority element set that may be used to provide metadata about agents (people, organizations), events, and terms (topics, geographics, genres, etc.). MADS was created to serve as a companion to the Metadata Object Description Schema (MODS). As such, MADS has a relationship to the MARC 21 Authority format, as MODS has to MARC 21 Bibliographic. Both carry selected data from MARC 21. (MADS Web page)

Web Style Sheets: Describe how documents are presented on screens, in print, or perhaps how they are pronounced. By attaching style sheets to structured documents on the Web (e.g. HTML), authors and readers can influence the presentation of documents without sacrificing device-independence or adding new HTML tags. (World Wide Web Consortium Web page)

XSLT: Language for transforming XML documents into other XML documents. A transformation expressed in XSLT is called a style sheet. (World Wide Web Consortium Web page)

MARC XML: Framework for working with MARC data in an XML environment. This framework is intended to be flexible and extensible to allow users to work with MARC data in ways specific to their needs. The framework will contain many components such as schemas, style sheets, and software tools. (MARC XML Web page)

Validation Checking a document against a schema or DTD.

Well-formed A document is considered ‘well-formed’ if it satisfies XML grammatical rules, including where tags are placed (especially beginning and ending), what the tags look like, what element names are legal, etc.

Parser A program that divides the XML document up into individual elements, attributes, etc., to determine if well-formed.

Validating parser Checks for both well-formed XML and conformance with a schema.

DTD Document Type Definition, used first with SGML. Also used with XML but gradually being phased out in favor of schemas.

*THIS PAGE INTENTIONALLY LEFT BLANK
FOR DOUBLE SIDED COPY*

References Handout: Course 2

ANSI/NISO Z39.19 - 2005 Guidelines for the Construction, Format, and Management of Monolingual Controlled Vocabularies <http://www.niso.org/standards/resources/Z39-19.html>

Baker, Thomas and Makx Dekkers, Identifying metadata elements with URIs: the CORES Resolution. D-Lib magazine, July/August 2003.
<http://www.dlib.org/dlib/july03/baker/07baker.html>

Bibliographic Control of Web Resources: A Library of Congress Action Plan
<http://www.loc.gov/catdir/bibcontrol/actionplan.pdf>

CanCore Guidelines (an AP of IEEE-LOM, best guidelines available for this schema). Available at: <http://www.cancore.ca/en/>

DCMI Abstract Model. <http://dublincore.org/documents/abstract-model/>

DCMI Initiative. Available at: <http://dublincore.org/>

DLESE Metadata page <http://www.dlese.org/Metadata/>

DLF MODS Implementation Guidelines For Cultural Heritage Materials
http://www.diglib.org/aquifer/DLF_MODS_ImpGuidelines_ver4.pdf

DLF OAI Best Practices Guidelines <http://oai-best.comm.nsdlib.org/cgi-bin/wiki.pl?PublicTOC>

Dublin Core/MARC/GILS Crosswalk <http://www.loc.gov/marc/dccross.html>

Dushay, Naomi and Diane I. Hillmann. Analyzing Metadata for Effective Use and Re-Use, by Available at: <http://dc2003.ischool.washington.edu/Archive-03/03dushay.pdf>

Godby, Carol Jean, Devon Smith, and Eric Childress, *Two Paths to Interoperable Metadata*, a paper presented at the 2003 Dublin Core Conference
<http://www.oclc.org/research/publications/archive/2003/godby-dc2003.pdf>

Guidelines for Assigning Identifiers to Metadata Terms
<http://www.ukoln.ac.uk/metadata/dcmi/term-identifier-guidelines/>

IEEE 1484 Learning Objects Metadata (IEEE LOM). Official standard available at: http://ltsc.ieee.org/wg12/files/LOM_1484_12_1_v1_Final_Draft.pdf
Simpler chart and DC mapping available at:
<http://www.ischool.washington.edu/sasutton/IEEE1484.html>

IFLA Study Group on the Functional Requirements for Bibliographic Records.

Functional Requirements for Bibliographic Records: Final report. Available at:
<http://www.ifla.org/VII/s13/frbr/frbr.pdf>

MARC 21. Available at: <http://lcweb.loc.gov/marc/>

MARC 21 to Dublin Core Crosswalk <http://www.loc.gov/marc/marc2dc.html>

Metadata Encoding & Transmission Standard (METS). Available at:
<http://www.loc.gov/standards/mets/>

Metadata Encoding and Transmission Standard: Primer and Reference Manual
<http://www.loc.gov/standards/mets/METS%20Documentation%20final%20070930%20msw.pdf>

Metadata for Images in XML Schema (MIX): <http://www.loc.gov/mix>

Metadata Mappings (Crosswalks)
<http://libraries.mit.edu/guides/subjects/metadata/mappings.html>

Metadata: Mapping Between Metadata Formats
<http://www.ukoln.ac.uk/metadata/interoperability/>

Metadata Object Description Schema (MODS). Available at:
<http://www.loc.gov/standards/mods/>
MODS User Guidelines, vers. 3 <http://www.loc.gov/standards/mods/v3/mods-userguide.html>

Morville, Peter, Ambient Findability: What We Find Changes Who We Become.
Cambridge, Mass.: O'Reilly, 2005.

National Science Digital Library Metadata Guidelines
<http://metamanagement.comm.nsl.org/outline.html>

National Information Standards Organization. Understanding Metadata.
<http://www.niso.org/standards/resources/UnderstandingMetadata.pdf>

ONIX for Books. Available at: <http://www.editeur.org/onix.html>

Open Archives Initiative <http://openarchives.org>

The OpenURL Framework for Context-Sensitive Services
http://www.niso.org/committees/committee_ax.html

PREMIS: <http://www.loc.gov/standards/premis/>

Readings in metadata and cataloging education:
<http://www.loc.gov/catworkshop/readings.html>

Resource Description and Access (RDA):

<http://www.collectionscanada.gc.ca/jsc/index.html>

Rosenfeld, Louis and Peter Morville, *Information Architecture for the World Wide Web: Designing Large-Scale Web Sites*. Cambridge, Mass.: O'Reilly, 2002.

Rust, Godfrey and Mark Bide, *The <INDECS> Metadata Framework: Principles, Model and Data Dictionary*. Available at:

http://www.doi.org/topics/indecs/indecs_framework_2000.pdf

SKOS Core Guide <http://www.w3.org/TR/swbp-skos-core-guide/>

SKOS Quick Guide to Publishing a Vocabulary on the Semantic Web

<http://www.w3.org/TR/swbp-thesaurus-pubguide/>

St. Pierre, Margaret and William P. La Plant, *Issues in crosswalking content metadata standards*. (NISO White Paper). Bethesda, Md.: National Information Standards Organisation, 15 October 1998.

Available from: <http://www.niso.org/press/whitepapers/crswalk.html>

Texas Digital Library Metadata Working Group. *MODS Application Profile for Electronic Theses and Dissertations*. Version 1. December 2005

<http://metallogger.files.wordpress.com/2007/06/tld-etd-mods-profile.pdf>

Tillett, Barbara B. *A taxonomy of bibliographic relationships*. *Library Resources & Technical Services*, 35(2):150--158, 1991.

University of Maryland Descriptive Metadata Tag Library

<http://www.lib.umd.edu/dcr/publications/taglibrary/umdm.html>

UVa Metadata

<http://lib.virginia.edu/digital/metadata/>

Using Dublin Core <http://dublincore.org/documents/usageguide/>

Using Metadata Standards in Digital Libraries: Implementing METS, MODS, PREMIS and MIX: <http://www.loc.gov/standards/mods/presentations/litaprogram-an2007.html>

Working Group on the Future of Bibliographic Control:

<http://www.loc.gov/bibliographic-future/>

World Digital Library

<http://www.worlddigitallibrary.org/project/english/index.html>

XHTML: Extensible Hypertext Markup Language. Available at:
<http://www.w3.org/TR/xhtml1/>

XML: Extensible Markup Language. Available at: <http://www.w3.org/XML/>

XML Tutorial. Available at: <http://www.w3schools.com/xml/default.asp>

Metadata Standards and Applications

Handout: Monitoring and Participating in Metadata Developments

Recommended e-journals, current awareness resources and discussion lists

1. D-Lib Magazine (<http://www.dlib.org/>)
2. Ariadne (<http://www.ariadne.ac.uk/>)
3. Current Cites (<http://lists.webjunction.org/currentcites/>) Also available as mailing list or RSS feed
4. Journal of Digital Information (<http://jodi.tamu.edu/>)
5. NISO Newslines (<http://www.niso.org/news/newslines/>)
6. Metadata Librarians listserv (<http://metadatalibrarians.monarchos.com/>)

Blogs

Sign on with a blog aggregator (Bloglines is a good basic one) and start reading. Some recommendations:

1. Lorcan Dempsey's weblog: (<http://orweblog.oclc.org/>)
2. Inquiring Librarian: (<http://inquiringlibrarian.blogspot.com/>)
3. Weibel Lines: (<http://weibel-lines.typepad.com/weibelines/>)
4. The Shifted Librarian (<http://www.theshiftedlibrarian.com/>)

Recommended Format Specific Discussion Lists

1. DC-General (consider a Working Group list as well) (<http://dublincore.org/about/contact/#dcgeneral>)
2. MODS list (<http://listserv.loc.gov/listarch/mods.html>)
3. VRA Core (<http://vraweb.org/vra-l/index.html>)

Recommended Conferences

1. ALA and the specialized library divisions (especially LITA and ALCTS)
2. Dublin Core (next is DC 2008: <http://dc2008.de/>)
3. Digital Library Federation forums (<http://www.diglib.org/forums.htm>)

Organize a local forum where you and your peers can have regular discussions about the work you're doing (e.g., Cornell Metadata Working Group <http://metadata-wg.mannlib.cornell.edu/>)

*THIS PAGE INTENTIONALLY LEFT BLANK
FOR DOUBLE SIDED COPY*

Handout: XML Definitions and Genealogy

Definitions:

Validation	<i>Checking a document against a schema or DTD.</i>
Well-formed	<i>A document is considered 'well-formed' if it satisfies XML grammatical rules, including where tags are placed (especially beginning and ending), what the tags look like, what element names are legal, etc.</i>
Parser	<i>A program that divides the XML document up into individual elements, attributes, etc., to determine if well-formed.</i>
Validating parser	<i>Checks for both well-formed XML and conformance with a schema.</i>
DTD	<i>Document Type Definition, used first with SGML. Also used with XML but gradually being phased out in favor of schemas.</i>

XML Schema Genealogy:

SGML	<i>Standard Generalized Markup Language, invented in the 1970's, became ISO 8879 in 1986. Very powerful, used extensively in the military and government, as well as for markup for narrative structures (see Text Encoding Initiative for more detail).</i>
HTML	<i>Hypertext Markup Language, an SGML application. Restricted set of tags, designed to describe web pages for purposes of presentation. Current version is HTML 4.0.</i>
XML	<i>Extensible Markup Language, originally intended as a 'lite' version of SGML, first version unveiled in 1998.</i>
Namespaces	<i>An effort to allow markup from different XML applications to be used in the same document without conflict (even if element names used in each namespace are the same).</i>
XSL	<i>Extensible Stylesheet Language, an XML application for transforming XML documents into a form that could be viewed in a web browser. Eventually split to form: XSLT and XSL-FO.</i>
XSLT	<i>Extensible Stylesheet Language Transformations, a general purpose language for transforming one XML document into another, sometimes (not always) for web page display.</i>

- XSL-FO** *An XML application for describing the layout of both printed pages and webpages, sometimes compared to PostScript.*
- CSS** *Cascading Style Sheets, used originally for HTML, and, when XML was invented, it was used for that, too.*
- XLink** *Extensible Linking Language, used to define more powerful linking constructs to connect XML documents, going beyond the “a” tag used in HTML.*
- XPointer** *Addresses individual parts of an XML document.*
- XPath** *Extracted from the addressing parts of XLink and XPointer.*

Exercise for Session 2: Descriptive Metadata
Instructors' information and Answer sheet

- Each group will be given a printout of a digital object
- Create a brief metadata record based on the standard assigned to your group (MODS or DC)
- Take notes about the issues and decisions made
- Appoint a spokesperson to present the metadata record created & the issues involved (2-5 minutes)

Provide participants with printout of different digital objects (selected by trainer).
Two examples (provided):

Example 1:

America's Pinch Hit March (sheet music from Baseball Sheet Music collection, LC)
<http://lcweb2.loc.gov/diglib/ihhas/loc.natlib.ihhas.200033287/default.html>

Example 2:

52nd Street, New York, N.Y.
<http://lcweb2.loc.gov/diglib/ihhas/loc.natlib.gottlieb.02771/default.html>

Participants will provide descriptive metadata in the standard assigned with the following elements (element names may vary according to scheme):

Title/subtitle
Creator/name (with role defined if possible)
Type of resource
Publication/origin information with place, publisher, date
Physical description
Subject
URL

Answers are given in the Metadata Template below:

MODS metadata: example 1 (sheet music)

Data Element (Field)	Data Value (Content)	Controlled Vocabulary (Yes / No) (please specify)
titleInfo title	America's pinch hit march	
titleInfo subTitle	the hit that ended the worlds greatest war	
name namePart	Dempsey, Bertha Stanfield Type="personal" Authority="lcnaf" (optional)	
name role	composer	Yes: MARC relators
typeOfResource	notated music	Yes (MODS)
originInfo place	Joplin, Missouri	
originInfo publisher	The McMillan Music Co.	
originInfo dateIssued	1919	

Data Element (Field)	Data Value (Content)	Controlled Vocabulary (Yes / No) (please specify)
PhysicalDescription	3 p.	
subject topic	Baseball—Songs and music	Yes: LCSH
location/url	http://lcweb2.loc.gov/diglib/has/loc.natlib.ihas.200033287/default.html	
Also useful: RelatedItem type="host" Title	Baseball sheet music collection	

Dublin Core metadata: example 1

Data Element (Field)	Data Value (Content)	Controlled Vocabulary (Yes / No) (please specify)
Title	America's pinch hit march : the hit that ended the worlds greatest war	
Creator	Bertha Stanfield Dempsey	
Resource Type	Image	Yes: DCMI type
Publisher	The McMillan Music Co.	
Date	1919	
Format	3 p.	
Subject	Baseball—Songs and music	Yes: LCSH
Identifier	http://lcweb2.loc.gov/diglib/has/loc.natlib.ihas.200033287/default.html	
Also useful: Relation	Baseball sheet music collection	

MODS metadata: example 2 (image)

Data Element (Field)	Data Value (Content)	Controlled Vocabulary (Yes / No) (please specify)
titleInfo title	52 nd Street, New York, N.Y., ca. 1948	
titleInfo subtitle		
Name type="personal" NamePart	Gottlieb, William P.	
name namePart type="date"	1917-	
name role	photographer	Yes: MARC relator
typeOfResource	still image	Yes (MODS)
originInfo place	New York	

Data Element (Field)	Data Value (Content)	Controlled Vocabulary (Yes / No) (please specify)
originInfo publisher		
originInfo dateCreated	1948	
physicalDescription form	graphic	Yes (gmd)
physicalDescription extent	1 negative : color ; 4 x 5 in.	
subject geographic	New York (N.Y.)	Yes (lcsh)
subject topic temporal		
Location/url	http://lcweb2.loc.gov/diglib/ihas/loc.natlib.gottlieb.02771/default.html	
Also useful: accessCondition	Original negative not served.	

Dublin Core metadata: example 2

Data Element (Field)	Data Value (Content)	Controlled Vocabulary (Yes / No) (please specify)
Title	52 nd Street, New York, N.Y., ca. 1948	
Creator	Gottlieb, William P.	
Resource Type	image	Yes: DCMI type
Publisher		
Date	1948	
Format	image/jpeg	Yes: (MIME)
Subject	New York (N.Y.)	
Identifier	http://lcweb2.loc.gov/diglib/ihas/loc.natlib.gottlieb.02771/default.html	
Also useful: Rights	Original negative not served.	

Session 3: Administrative metadata (PREMIS)

Example 1: America's pinch hit march

We will only fill in a template for one of the files. A complete example would include a record for each file, showing the relationship with the relationship element. For this example we would have 3 records, one for each file (representing each page).

Data Element (Field)	Data Value (Content)	Controlled Vocabulary (Yes / No) (please specify)
objectIdentifier objectIdentifierType	URI	yes
objectIdentifier objectIdentifierValue	http://lcweb2.loc.gov/natlib/ihas/service/encyclopedia/200033287/0001v.jpg	
objectCharacteristics size	629507	
objectCharacteristics format formatDesignation formatName	image/jpeg	yes
objectCharacteristics format formatDesignation formatVersion		
creatingApplication creatingApplicationName	Adobe photoshop	
creatingApplication creatingApplicationVersion	version CS3	
creatingApplication datecreatedByApplication	2008-01-02	
Environment Software SwName	Macromedia Fireworks MX	maybe
Environment software swVersion	6.0	
relationship RelationshipType	structural	yes
relationship relationshipSubType	hasSibling	
relatedObjectIdentification relatedObjectIdentifierType	URI	
relatedObjectIdentification relatedObjectIdentifierValue	http://lcweb2.loc.gov/natlib/ihas/service/encyclopedia/200033287/0002v.jpg	
relatedObjectIdentification relatedObjectIdentifierSequence	2	

Example 2: 52nd Street, New York, N.Y., ca. 1948

We will only fill in a template for one of the files. A complete example would include a record for each file, showing the relationship with the relationship element. There is a relationship to the JPEG file, which is a derivative of this master file.

Data Element (Field)	Data Value (Content)	Controlled Vocabulary (Yes / No) (please specify)
objectIdentifier objectIdentifierType	URI	yes
objectIdentifier objectIdentifierValue	http://lcweb2.loc.gov/natlib/ihas/warehouse/gottlieb/02771/ver01/0001.tif	
objectCharacteristics size	60158210	
objectCharacteristics format formatDesignation formatName	image/tiff	yes
objectCharacteristics format formatDesignation formatVersion		
creatingApplication creatingApplicationName	Image Alchemy PS	
creatingApplication creatingApplicationVersion	v1.11	
creatingApplication datecreatedByApplication	20030215	
Environment Software swName	Firefox	maybe
Environment software swVersion	5.0	
relationship RelationshipType	derivation	yes
relationship relationshipSubType	source of	
relatedObjectIdentification relatedObjectIdentifierType	URI	
relatedObjectIdentification relatedObjectIdentifierValue	http://lcweb2.loc.gov/natlib/ihas/service/gottlieb/02771/ver01/0001v.jpg	
RelatedObjectIdentification relatedObjectIdentifierSequence		

Exercise for Session 4: XML syntax
Answer sheet

Example 1: America's pinch hit march (MODS)

```
<?xml version="1.0" encoding="UTF-8"?>
<mods:mods version="3.2" ID="MODS1"
xsi:schemaLocation="http://www.loc.gov/mods/v3
http://www.loc.gov/standards/mods/mods.xsd">
  <mods:titleInfo>
    <mods:title>America's pinch hit march</mods:title>
    <mods:subTitle>the hit that ended the world's greatest war</mods:subTitle>
  </mods:titleInfo>
  <mods:name type="personal">
    <mods:namePart>Dempsey, Bertha Stanfield</mods:namePart>
    <mods:role>
      <mods:roleTerm authority="marcrelator"
type="text">Composer</mods:roleTerm>
    </mods:role>
  </mods:name>
  <mods:typeOfResource>notated music</mods:typeOfResource>
  <mods:originInfo>
    <mods:place>
      <mods:placeTerm>Joplin, Missouri</mods:placeTerm>
    </mods:place>
    <mods:publisher>McMillan Music Co.</mods:publisher>
    <mods:dateIssued>1919</mods:dateIssued>
  </mods:originInfo>
  <mods:physicalDescription>
    <mods:form authority="ihas">sheet music</mods:form>
    <mods:extent>3 p. </mods:extent>
  </mods:physicalDescription>
  <mods:subject authority="lcsht">
    <mods:topic>Baseball</mods:topic>
    <mods:genre>Songs and music</mods:genre>
  </mods:subject>
  <mods:relatedItem type="host">
    <mods:titleInfo>
      <mods:title>Baseball sheet music collection</mods:title>
    </mods:titleInfo>
  </mods:relatedItem>
  <mods:location>
    <mods:physicalLocation authority="marcorg">DLC</mods:physicalLocation>

  <mods:url>http://lcweb2.loc.gov/diglib/ihas/loc.natlib.ihas.200033287/default.html
  </mods:url>
  </mods:location>
  <mods:recordInfo>
    <mods:recordContentSource>IHAS</mods:recordContentSource>
    <mods:recordChangeDate
encoding="marc">060412</mods:recordChangeDate>
    <mods:recordIdentifier
source="IHAS">loc.natlib.ihas.200033287</mods:recordIdentifier>
```

```

</mods:recordInfo>
</mods:mods>

```

Example 1: America's pinch hit march (DC)

```

<?xml version="1.0" encoding="UTF-8"?>
<metadata
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  http://dublincore.org/schemas/xmls/qdc/2003/04/02/dcterms.xsd
  xmlns:dc="http://purl.org/dc/elements/1.1/"
  xmlns:dcterms="http://purl.org/dc/terms/"
  xmlns:dcmitype="http://purl.org/dc/dcmitype/">

  <dc:title xml:lang="en "> America's pinch hit march : the hit that ended the
worlds greatest war </dc:title>
  <dc:creator> Bertha Stanfield Dempsey</dc:creator>
  <dc:type> http://purl.org/dc/dcmitype/Image</dc:type>
  <dc:publisher>The McMillan Music Co.</dc:publisher>
  <dc:date>1919</dc:date>
  <dc:format>3 p.</dc:format>
  <dc:subject xsi:type="dcterms:LCSH"> Baseball—Songs and music </dc:subject>
  <dc:relation> Baseball sheet music collection</dc:relation>

</metadata>

```

Example 2: 52nd Street, New York, N.Y., ca. 1948 (MODS)

```

<?xml version="1.0" encoding="UTF-8"?>
<mods:mods version="3.2" ID="MODS1"
  xsi:schemaLocation="http://www.loc.gov/mods/v3
http://www.loc.gov/standards/mods/mods.xsd">
  <mods:titleInfo>
    <mods:title>52nd Street, New York, N.Y., ca. 1948</mods:title>
  </mods:titleInfo>
  <mods:name>
    <mods:namePart> Gottlieb, William P. </mods:namePart>
    <mods:namePart type="date">1917-</mods:namePart>
    <mods:role>
      <mods:roleTerm type="text"
authority="marcrelator">photographer</mods:roleTerm>
    </mods:role>
  </mods:name>
  <mods:typeOfResource>still image</mods:typeOfResource>
  <mods:originInfo>
    <mods:place>
      <mods:placeTerm>New York</placeTerm>
    </mods:place>
    <mods:dateCreated>1948</mods:dateCreated>
  <mods:physicalDescription>
    <mods:form authority="gmd">graphic</mods:form>
  </mods:physicalDescription>
  <mods:accessCondition> Original negative not served.</mods:accessCondition>
  <mods:subject authority="lcsch">

```

```

    <mods:geographic>
New York (N.Y.)</mods:geographic>
    </mods:subject>
    <mods:location>
      <mods:physicalLocation authority="marcorg">DLC</mods:physicalLocation>
    </mods:location>
    <mods:recordInfo>
      <mods:recordContentSource>DLC</mods:recordContentSource>
      <mods:recordChangeDate
encoding="marc">2007030616</mods:recordChangeDate>
      <mods:recordIdentifier source="DLC">got99000277</mods:recordIdentifier>
    </mods:recordInfo>
  </mods:mods>

```

Example 2: 52nd Street, New York, N.Y., ca. 1948 (DC)

```

<?xml version="1.0" encoding="UTF-8"?>
<metadata
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  http://dublincore.org/schemas/xmls/qdc/2003/04/02/dcterms.xsd
  xmlns:dc="http://purl.org/dc/elements/1.1/"
  xmlns:dcterms="http://purl.org/dc/terms/"
  xmlns:dcmitype="http://purl.org/dc/dcmitype/" >

  <dc:title xml:lang="en "> 52nd Street, New York, N.Y., ca. 1948</dc:title>
  <dc:creator> Gottlieb, William P., 1917-</dc:creator>
  <dc:type> http://purl.org/dc/dcmitype/Image </dc:type>
  <dc:date>1948</dc:date>
  <dc:format>image/jpeg</dc:format>
  <dc:subject> </dc:subject>
<dc:coverage>New York (N.Y.) </dc:coverage>
  <dc:rights Original negative not served.</dc:rights>
</metadata>

```

Note that dc:format could use image/tif or image/jpeg (or nothing) depending upon whether the object of cataloging is considered the intellectual object or the file itself.

Evaluation Form

Metadata Standards and Applications

Your evaluation of this workshop is very important to the future development of this course and other similar courses. Your honest, candid answers to the following questions will assist us in providing quality programs.

Please rate the following aspects of today's workshop by checking the box that best reflects your evaluation:

1. The overall content of the workshop:

	5	4	3	2	1	
a. was extremely valuable						was of little value
b. provided enough detail						was too general
c. was current & relevant						was outdated
d. was cohesive & logical						was fragmented/difficult to follow
e. was appropriate to my needs						was not at all appropriate
f. met its stated objectives						did not meet objectives

2. Presenter 1:

	5	4	3	2	1	
a. was knowledgeable						was unsure of the material
b. had good presentation skills						had poor presentation skills
c. encouraged participation						discouraged participation
d. addressed my level of understanding						did not consider my level
e. answered questions directly						did not answer questions
f. was prepared						was not prepared
g. understood the audience dynamics						ignored audience dynamics

3. Presenter 2:

	5	4	3	2	1	
a. was knowledgeable						was unsure of the material
b. had good presentation skills						had poor presentation skills
c. encouraged participation						discouraged participation
d. addressed my level of understanding						did not consider my level
e. answered questions directly						did not answer questions
f. was prepared						was not prepared
g. understood the audience dynamics						ignored audience dynamics

4. The handouts:

	5	4	3	2	1	
a. are excellent						are poor
b. followed course content						are disjointed/out of sequence
c. are valuable for future reference						are of no value

5. The PowerPoint slides:

- a. were clear and easy to read
- b. were well organized
- c. illustrated concepts clearly
- d. covered an appropriate amount of information
- e. were visually effective
- f. were enhanced by and supported the presenter's remarks

5	4	3	2	1

- were hard to read
- were poorly organized
- were confusing
- contained too much or not enough information
- were not effective
- were poorly related to the presenter's remarks

Please give the following information about yourself:

- 6. Your level of knowledge in the subject of this workshop before today: expert 5 4 3 2 1 novice
- 7. Your level of experience in the subject of this workshop before today: very experienced 5 4 3 2 1 beginner
- 8. Other comments:

Comments on specific sessions:

THANK YOU!

Note to workshop organizer: Please send copies of all completed evaluation forms to:

CLW Editorial Team Chair, The Library of Congress, Cooperative Cataloging Team, Washington, DC 20540-4373 USA