- "Data about Data"
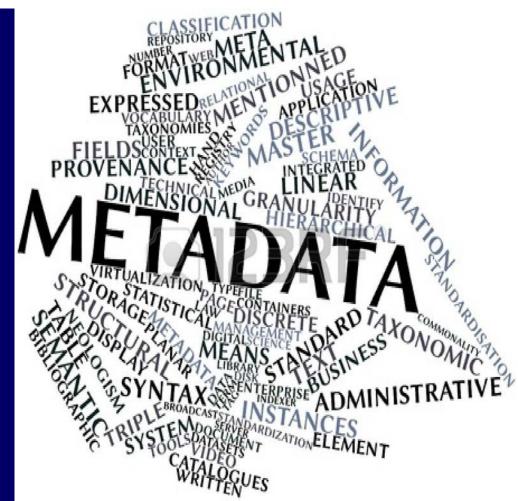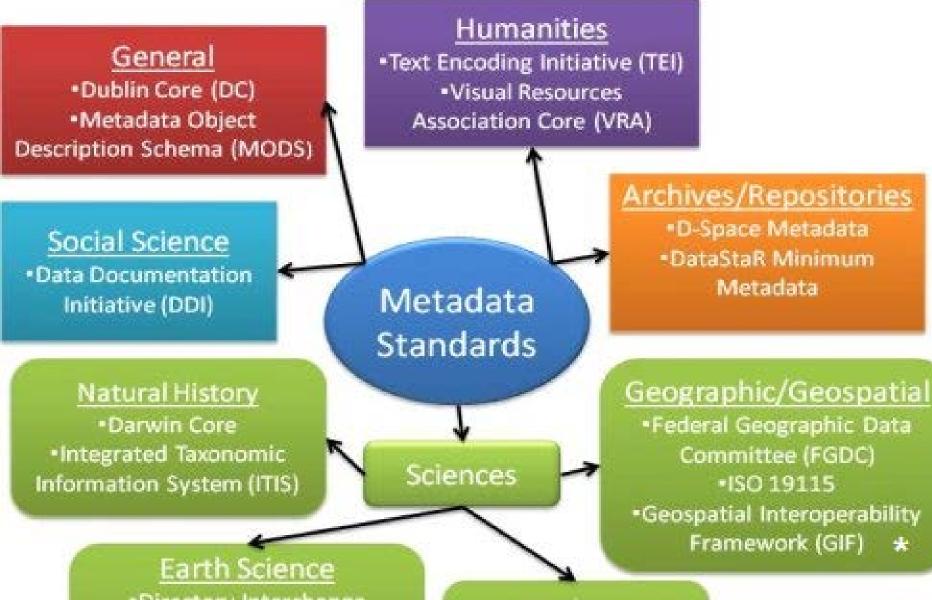 - Notes about Information Sources
- Metadata Standards
  - Involve Cataloging Resources
    - Differ by Subject
    - Differ by Document Type

General
- Dublin Core (DC)
- Metadata Object Description Schema (MODS)

Humanities
- Text Encoding Initiative (TEI)
- Visual Resources Association Core (VRA)

Social Science
- Data Documentation Initiative (DDI)

Archives/Repositories
- D-Space Metadata
- DataStaR Minimum Metadata

Metadata Standards

Natural History
- Darwin Core
- Integrated Taxonomic Information System (ITIS)

Sciences

Geographic/Geospatial
- Federal Geographic Data Committee (FGDC)
- ISO 19115
- Geospatial Interoperability Framework (GIF) *

Earth Science
- Directory Interchange Format (DIF)
- Standard for the Exchange of Earthquake Data (SEED)

Ecology
- Ecological Metadata Language (EML)
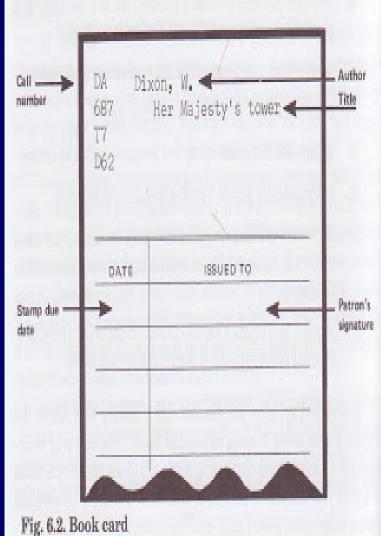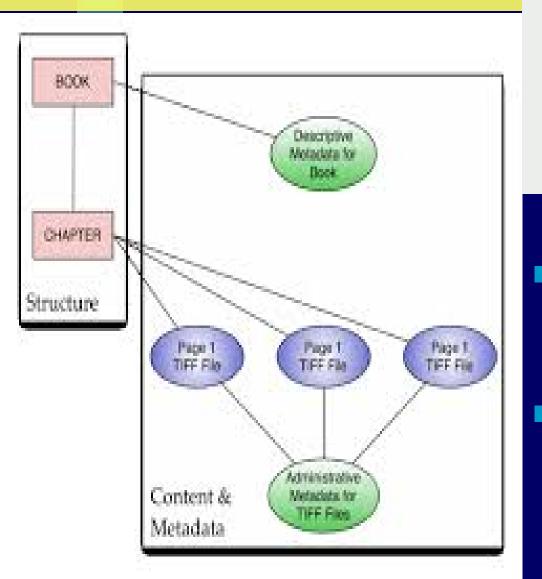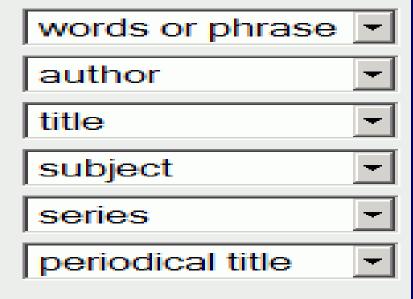
Administrative Metadata

- **Storage and Usage Data**
  - Where are Books Kept?
  - Who Borrowed Books?
- **Collection Development**
  - Patron-Driven Acquisition (PDA)
  - Weeding (Books Not Checked Out)
- **Rights and Access**
  - Who Can Access Information?
- **Meta-Metadata**
  - Information about Metadata
- **Technical Metadata**
  - How Digital Objects were Created

Call number

DA
687
T7
D62

Author

Dixon, W.

Title

Her Majesty's tower

DATE   ISSUED TO

Stamp due date

Patron's signature

Fig. 6.2. Book card

Structural Metadata

words or phrase ▼

author ▼

title ▼

subject ▼

series ▼

periodical title ▼

- ■ Documenting Structure
  - ■ Pages Structured into Chapters in Book
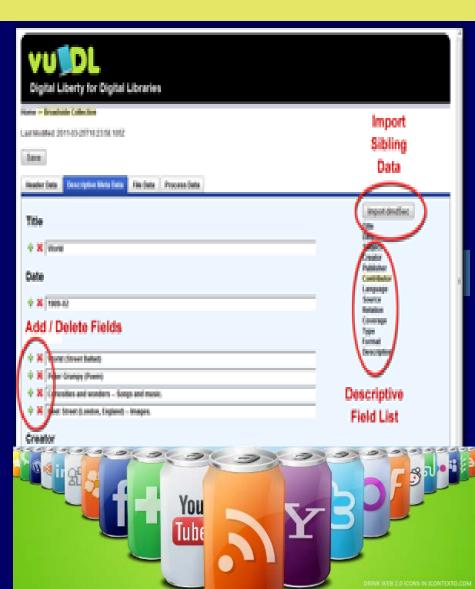- ■ Authority Control Elements
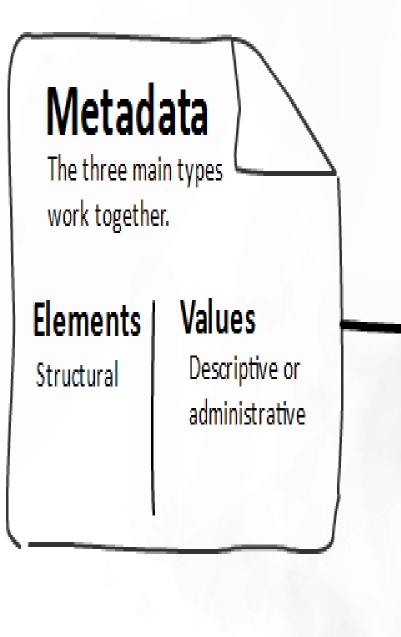
Descriptive Metadata

- **Structural Metadata**
  - Authority Control Elements
- **Descriptive Metadata**
  - Specific Information about Resources
    - Last Name, First Name
    - *Book Title*
    - Synopsis of Material
- **Web 2.0**
  - User Interaction with Web
    - Social Media
  - Users May Add Descriptive Information
    - Reviews

Metadata

The three main types work together.

Elements	Values
Structural	Descriptive or administrative

Metadata example

Structural	Descriptive
Title: Description:	'Mad Men' Season 5 plot predictions The two-hour premiere quickly answered lingering questions from season four.
	Administrative
Creator:	Lucas Shaw
Date:	26/03/2012
Type:	Article

Metadata Schemas

MARC	Machine Readable Cataloguing
BIBFRAME	Bibliographic Framework
DC	Dublin Core
METS	Metadata Encoding Transmission Standards

WHAT IS MARC?

LIBRARY OF CONGRESS

- Stands for Machine Readable Cataloging
- Invented in the 1960s by Henriette Avram and the Library of Congress.
- It is a metadata schema used to create Machine Readable Bibliographic records for printed materials.
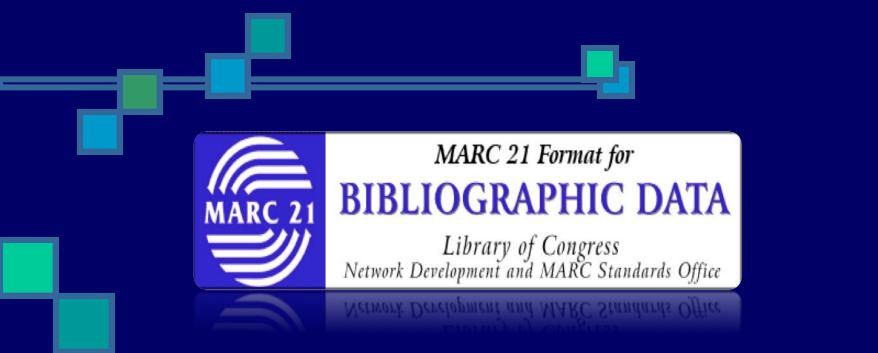- MARC is used primarily by the library community.

LDR			01671cam a2200385 a 4500
001			12901170
005			20041228203623.0
008			020820s2003 nyub 000 1 eng
010			‡a 2002030595
020			‡a 0743233034 (alk. paper)
040			‡a DLC ‡c DLC ‡d DLC
043			‡a n-usp--
050	0	0	‡a PS3563.A319 ‡b W36 2003
082	0	0	‡a 813/.54 ‡2 21
100	1		‡a McMurtry, Larry.
245	1	4	‡a The wandering hill : ‡b a novel / ‡c Larry McMurtry.
260			‡a New York : ‡b Simon & Schuster, ‡c c2003.
300			‡a xiii, 302 p. : ‡b maps ; ‡c 25 cm.
490	1		‡a Berrybender narratives ; ‡v bk. 2
500			‡a Maps on endpapers.
650		0	‡a British ‡z West (U.S.) ‡v Fiction.
650		0	‡a Eccentrics and eccentricities ‡v Fiction.
651		0	‡a Yellowstone River ‡v Fiction.
650		0	‡a Women immigrants ‡v Fiction.
650		0	‡a Young women ‡v Fiction.
650		0	‡a Berrybender family (Fictitious characters) ‡v Fiction.
800	1		‡a McMurtry, Larry. ‡t Berrybender narratives ; ‡v bk. 2.

- When MARC was initially invented, computers were not as powerful as they are today, nor did they have as much storage space. MARC was created around these limitations.
- MARC utilizes a simple three digit numeric code to identify its fields. These three digit codes, referred to as "tags," aid the computer in interpreting the bibliographic data. The numbers for each field range from 001 to 999.
- Within each bibliographic entry, there are fields, such as author and title, which can then be divided into sub-fields.

- MARC is the most commonly used metadata schema used in libraries today.
- There are many different versions of the schema including country specific and lite versions.
- In 1991, the U.S. and Canada merged their national MARC formats to create a new system for the 21st Century, MARC 21.

MARC 21 Format for
BIBLIOGRAPHIC DATA
Library of Congress
Network Development and MARC Standards Office

■ MARC 21 is now the most commonly used version worldwide.

■ MARC 21 has five different formats for different types of data, which include bibliographic, authority, community information, holdings, and classification data formats.

- Has been in use and in existence for over 5 decades.

- Has become very outdated and is unnecessarily complicated.

- Data is locked into library catalogs and cannot be shared outside of libraries or linked to other data on the web.

- Will be replaced by BIBFRAME, which is currently being created by The Library of Congress.

bf:
BIBFRAME

Bibliographic Framework Initiative

"the future of bibliographic description"

- Library of Congress launched the BIBFRAME Initiative in May 2011

- Uses the data elements used in MARC records to create a linkable information resource

- Maintains a controlled vocabulary

Collaboration among Libraries

- BIBFRAME allows the bibliographic record of an item to be shared or "linked" by many libraries to avoid duplication

- Even the smallest library will have the ability to access this shared information and help users search for information

Linked Data and BIBFRAME
Library catalogs meet the Internet

- URI or Universal Resource Identifier
  - <http://domain.com/collection/1>
- RDF or Resource Description Framework
  - triples made up of subject, predicate and object

    <http://dbpedia.org/resource/Divine_Comedy>

    <http://bibframe.org/vocab/creator>

    <http://dbpedia.org/resource/Dante_Alighieri

Core Classes

- Defined set of core classes identifies a type of BIBFRAME resource
  - BIBFRAME Work
  - BIBFRAME Instance
  - BIBFRAME Authority
  - BIBFRAME Annotation
- Core classes may have subclasses, which have their own properties

BIBFRAME vs. MARC

```
100 1  $aHarvey, D. R.$q(Douglas Ross),$d1951-
245 10 $aDigital curation :$ba how-to-do-it manual /$cRoss Harvey.
260    $aNew York :$bNeal-Schuman Publishers,$cc2010.
300    $axxii, 225 p. ;$c28 cm.
490 1  $aHow-to-do-it manuals ;$vno. 170
504    $aIncludes bibliographical references and index.
505 0  $aThe changing landscape -- Conceptual models -- Defining data -- Curation and cu
rators -- Description and representation information -- Preservation planning and policy
-- Sharing knowledge and collaborating -- Designing data -- Creating data -- Deciding w
hat data to keep -- Ingesting data -- Preserving data -- Storing data -- Using and reusi
ng data.
650 0  $aDigital libraries.
650 0  $aDigital preservation.
650 0  $aDigital libraries$xManagement.
830 0  $aHow-to-do-it manuals for libraries ;$vno. 170.
```

MARC records describe print materials only

MARC records lock data into specific numbered fields

Use of special characters makes it impossible to use this information outside of a library environment

There is no flexibility in the data structure to allow for creating relationships between other elements or linking different numbered fields

BIBFRAME vs. MARC

```
<http://id.loc.gov//resources/bibs/16246164> a bf:Text,
      bf:Work ;
  bf:authorizedAccessPoint "Harvey, D. R. (Douglas Ross), 1951- Digital curation : a h
ow-to-do-it manual / Ross Harvey.Digital curation :a how-to-do-it manual",
      "harveydrdouglasross1951digitalcurationahowtodoitmanualengworktext"@x-bf-hash ;
  bf:classification [ a bf:Classification ;
        bf:classificationEdition "22",
          "full" ;
        bf:classificationNumber "025.00285" ;
        bf:classificationScheme <http://id.loc.gov/authorities/classSchemes/ddc> ;
        bf:label "025.00285" ] ;
  bf:classificationLcc <http://id.loc.gov/authorities/classification/ZA4080> ;
  bf:creator [ a bf:Person ;
        bf:authorizedAccessPoint "Harvey, D. R. (Douglas Ross), 1951-" ;
        bf:hasAuthority [ a madsrdf:Authority ;
              madsrdf:authoritativeLabel "Harvey, D. R. (Douglas Ross), 1951-" ] ;
        bf:label "Harvey, D. R. (Douglas Ross), 1951-" ] ;
```
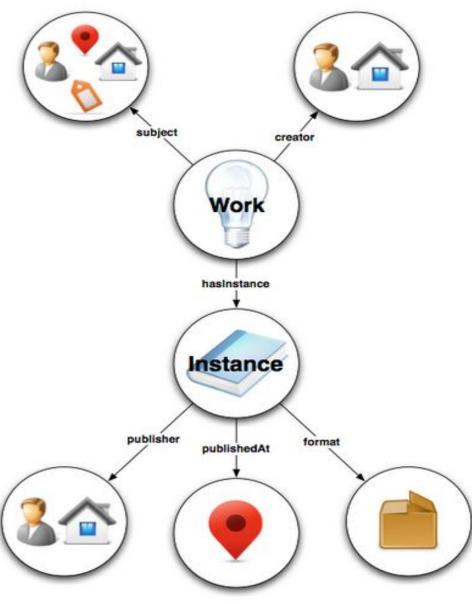
- **BIBFRAME uses web-compatible language**

- **A description can be created for print works, non-print media and cultural materials**

- **By using controlled identifiers, relationships can be established between resources, resulting in linked data**

- **Relationships can also be made by linking to materials already available on the Internet**

BIBFRAME is a
web
of Linked Data
and
resource
relationships

Dublin Core

- In 1995 the National Center for Supercomputing Applications (NCSA) and Online Computer Library Center (OCLC) met in Dublin, Ohio to create "an international consensus on a simple resource description format (Miller, p. 49, 2011)."

- At this meeting the Dublin Core Metadata Element Set was created.

ORIGIN OF DUBLIN CORE

- The DC was named after the city the meeting was held in, Dublin, Ohio.

- As for the "core" portion of the Dublin Core, the members created 15 elements that would allow for a wide range of resources to be applied.

Dublin Core Metadata Initiative (DCMI)

After the creation of the Dublin Core a community quickly built itself around this

THE 15 ELEMENTS OF THE DC

- Contributor
- Coverage
- Creator
- Date
- Description
- Format
- Identifier
- Language
- Publisher
- Relation

- Rights
- Source
- Subject
- Title
- Type

DC QUALIFIERS

3 Extra Elements:

- Audience
- Provenance
- Rights Holder

Two Large Categories:

- Element Refinement
- Encoding Scheme

DC ENCODING EXAMPLE

```html
<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="Tosbourn">
<meta name="DC.Creator" content="Toby Osbourn">
<meta name="DC.Subject" content="Web Development">
<meta name="DC.Subject" content="PHP">
<meta name="DC.Subject" content="MySQL">
<meta name="DC.Subject" content="Linux">
<meta name="DC.Subject" content="Apache">
<meta name="DC.Description" content="A blog about all aspects of web development (plus some general tech articles thrown in for good measure)">
<meta name="DC.Publisher" content="Toby Osbourn">
<meta name="DC.Contributor" content="Toby Osbourn">
<meta name="DC.Type" scheme="DCTERMS.DCMIType" content="Text">
<meta name="DC.Type" scheme="DCTERMS.DCMIType" content="Tech Articles">
<meta name="DC.Language" content="en">
<meta name="DC.Coverage" content="Web Development">
<meta name="DC.Audience" content="Web Developers">
<meta name="DCTERMS.EducationLevel" content="Intermediate to Higher">
<meta name="DCTERMS.RightsHolder" content="Toby Osbourn">
<meta name="DCTERMS.InstructionalMethod" content="Learning by doing">
<meta name="DCTERMS.AccrualPolicy" content="Active">
```

METS

- Metadata Encoding and Transmission Standard
- Completed by the Digital Library Federation (DLF)
- "an XML schema for encoding structurally complex digital objects into a single document" (Taylor & Joudrey, 2009, p. 101)

Background

- Replaced Electronic Binding DTD schema (Ebind) and Making of America Project II (MOA2)
  - "proposed a standard encoding for digital objects" (Taylor & Joudrey, 2009, p. 19)
  - Provided an encoding format for metadata in regards to textual and image-based works
- The DLF built upon the digitization project
- The page-turner model is another successful example of structural metadata

Why METS?

- It standardizes the structural metadata of digital objects
  - **Specific** enough to increase interoperability
  - **Flexible** enough to be used with other schemas
  - **Simple** enough for digital object creators to implement

METS Sections

- METS consists of seven sections:
    - METS Header (metsHdr)
    - Descriptive Metadata (dmdSec)
    - Administrative Metadata (amdSec)
    - File Group (fileSec)
    - Structural Map (structMap)
    - Structural Map Linking (structLink)
    - Behavior (behaviorSec)

    (Linda, 2005, p. 239)

METS Sections (cont.)

- METS flexibility
  - Only the Structural Map is a mandatory section
  - Creators have many options for the descriptive and administrative sections
    - They can "choose from a number of extension schemas" (Taylor & Joudrey, 2009, p. 102)
  - The metadata record can be stored both within or outside of a METS document

Bibliographic Record

Traditional Music and Spoken Word Catalog
from the American Folklife Center

Print Subscribe   Share/Save

Alabama blues / Booker T Sapps [sound recording]

Title
Alabama blues [sound recording]

Recordist(s)
Lomax, Alan, 1915-2002
Hurston, Zora Neale
Barnicle, Mary Elizabeth, 1891-1978

Performers
Sapps, Booker T.
Matthews, Roger
Flowers, Willy

Type of Material
sound recording

Language
English

Form
sound recording

Collection Description
A. Lomax Z.N. Hurston and Barnicle Expedition

Performance note
Sung by Booker T. Sapps with harmonica, with harmonica by Roger Matthews, and guitar by Willy Flowers.

Instrument
Harmonica (mouth organ)
Harmonica (mouth organ)
Guitar

Subject
United States of America --Florida --Belle Glade

AFC Number
AFC 1935/001

AFS Number
AFS 00368 A
AFS 00368 B

Repository
American Folklife Center, Library of Congress

METS Documentation

```
– <mets:mets OBJID="loc.afc.afc9999005.1153" xsi:schemaLocation="http://www.loc.gov/METS/ http://ww
  – <mets:dmdSec ID="dmd1">
    – <mets:mdWrap MDTYPE="MODS">
      – <mets:xmlData>
        – <mods:mods ID="mods1">
          – <mods:titleInfo>
              <mods:title>Alabama blues</mods:title>
            </mods:titleInfo>
          – <mods:name type="personal">
              <mods:namePart>Lomax, Alan</mods:namePart>
```

```
        – <mods:location>
            <mods:physicalLocation authority="marcorg">DLC</mods:physicalLocation>
          </mods:location>
        – <mods:recordInfo>
            <mods:recordContentSource>IHAS</mods:recordContentSource>
            <mods:recordChangeDate encoding="marc">111123</mods:recordChangeDate>
            <mods:recordIdentifier source="IHAS">loc.afc.afc9999005.1153</mods:recordIdentifier>
          </mods:recordInfo>
        </mods:mods>
      </mets:xmlData>
    </mets:mdWrap>
  </mets:dmdSec>
– <mets:fileSec>
  – <mets:fileGrp USE="MASTER">
    – <mets:file MIMETYPE="image/tiff" GROUPID="G1" ID="f0178m">
        <mets:FLocat LOCTYPE="URL" xlink:href="http://lcweb4.loc.gov/natlib/ihas/warehouse/afc99
      </mets:file>
    </mets:fileGrp>
  – <mets:fileGrp USE="SERVICE">
    – <mets:file MIMETYPE="image/jpeg" GROUPID="G1" ID="f0178s">
        <mets:FLocat LOCTYPE="URL" xlink:href="http://lcweb4.loc.gov/natlib/ihas/service/afc99990
      </mets:file>
    </mets:fileGrp>
  </mets:fileSec>
– <mets:structMap>
  – <mets:div DMDID="mods1" TYPE="bib:bibRecord">
    – <mets:div TYPE="bib:card">
      – <mets:div TYPE="lc:image">
          <mets:fptr FILEID="f0178m"/>
          <mets:fptr FILEID="f0178s"/>
        </mets:div>
      </mets:div>
    </mets:div>
  </mets:structMap>
</mets:mets>
```

XML

- Extensible Markup Language
- Developed by the World Wide Web Consortium (WC3)
- Preferred markup language for "the encoding and exchange of structured data" (Duval, Hodgins, Sutton, and Weibel, 2002, B. Syntax and Semantics para. 5)
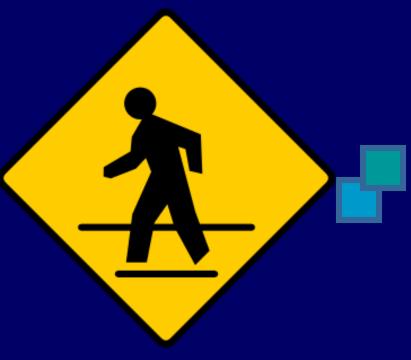
XML (cont.)

- An extended form of HTML
  - HTML is primarily descriptive while XML is geared towards the structural aspect of Web page creation
- Able to accommodate multimedia files and identify the formats of encoded elements
- Expected to "[play] an increasingly crucial role in the exchange of a variety of information on the Web" (NISO, 2004, p. 3)

Metadata Management Tools: Overview

- Definition: Tools that facilitate "interoperability"- "the ability of two or more information systems to exchange metadata with minimal loss of information." (Neiswender, 2009)
- Examples:
  - Crosswalks
  - Application Profiles
  - Registries

Crosswalk - Overview

- Definition: An "authoritative mapping" or "table" that denotes "the relationships and equivalencies between two or more metadata schemes" (Caplan, 2003; Dublin Core Metadata Initiative, 2005)

Crosswalk: Example

MARC to Dublin Core Crosswalk (Unqualified)

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	
260cg		
500-599, except 506, 530, 540, 546	Description	
340	Format	
856$q		
020$a, 022$a, 024$a	Identifier	
856$u		
008/35-37	Language	
041abdefghj		
546		
260ab	Publisher	
530, 760-787ot	Relation	
506, 540	Rights	
534$t	Source	
786ot		
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		

Crosswalks: Applications and Drawbacks

- Application: Allows data created in one scheme to be repurposed in another
  - Efficient
  - Saves costs

- Weaknesses:
  - Mappings are one-directional.
  - Conversion back to original scheme unlikely without potential data loss.

Applications Profiles - Overview

- Definition: "A declaration of the metadata terms used in metadata"
  - Elements can be from one or more element sets (DCMI, 2005)
  - Created by synthesizing elements from various metadata schemes, rather than infusing new ones. See, e.g., Taylor & Joudery (2009); Baker, Dekkers, Heery, Patel, & Salokhe (2006)

Applications Profiles: Example

Dublin Core Library Application Profile
(DCMI Libraries Application Profile Task Group, 2010)

Applications Profiles: Application and Drawbacks

- Application: Fulfills scheme needs where choosing a single one would not adequately serve an organization's needs

- Drawbacks:
    - Policy statements are essential for optimal use.
    - Maintenance requires commitment. (Heery & Patel, 2000)

- Definition: Human-readable repositories of authoritative information on available metadata terms and use (Caplan, 2003; Baker, Dekkers, Heery, Patel, & Salokhe, 2006)
- Few exist at this time

References

- Baker, T., Dekkers, M., Heery, R., Patel, M., & Salokhe, G. (2001). What terms does your metadata use? Application profiles as machine-understandable narratives. *Journal of Digital information, 2*(2). Retrieved from http://opus.bath.ac.uk/11353/3/Patel_Jodi_2001.pdf

- Barroso, I., Hartmann, N., & Ribeiro, C. (2015) Metadata crosswalk for a museum collection in a thematic digital library. *Journal of Library Metadata, 15*(1), 36-49. doi:10.1080/19386389.2015.1011025

- Bibliographic Framework Initiative. (n.d.). Retrieved from http://bibframe.org/

- Cantara, L. (2005). METS: The Metadata Encoding and Transmission Standard. *Metadata: A Cataloger's Primer, 40*(3/4), 237-253. doi:10.1300/J104v40n03_11

- Caplan, P. (2003). Metadata Fundamentals for All Librarians. Chicago, IL, USA: ALA Editions. Retrieved from http://www.ebrary.com

References

- Chan, L. M., & Hodges, T. (2007). Cataloging and classification : an introduction (3rd ed.). Lanham, Md.: Scarecrow Press, 2007.

- Coyle, K. (2005). Managing technology: understanding metadata and its purpose. *The Journal of Academic Librarianship, 31*(2), 160-163.

- DCMI Libraries Application Profile Task Group. (2010). DCMI Libraries Community. Retrieved June 3, 2015, from http://dublincore.org/groups/libraries/

- DCMI: History of the Dublin Core Metadata Initiative (2015). Retrieved from: http://dublincore.org/about/history/

- DCMI-Libraries Working Group. (2004). Library Application Profile. Retrieved June 3, 2015, from http://dublincore.org/documents/library-application-profile/

- Development and MARC Standards Office, Library of Congress. (2008). MARC to Dublin Core Crosswalk. Retrieved June 3, 2015, from http://www.loc.gov/marc/marc2dc.html

- Dublin Core Metadata Initiative. (2005). DCMI Glossary. Retrieved from http://dublincore.org/documents/usageguide/glossary.shtml

References

- Dublin Core Metadata Initiative. (n.d.) DCMI registry community. Retrieved from http://dublincore.org/groups/registry/

- Duval, E., Hodgins, W., Sutton, S., & Weibel, S. (2002). Metadata principles and practicalities. *D-Lib Magazine*, *8*(4). doi:10.1045/april2002-weibel

- Furrie, B. (2003). Understanding MARC bibliographic: machine-readable cataloging. (7th ed.). Washington, D.C.: Library of Congress, Cataloging Distribution Service, in collaboration with Follett Software Company, 2003.

- Gonzales, B. M. (2014). Linking libraries to the web: Linked Data and the future of the bibliographic record. *Information Technology & Libraries, 33*(4), 10-22.

- Heery, R. & Patel, M. (2000). Application Profiles: Mixing and Matching Metadata Schemas, *Ariadne*, *2000*(25). Retrieved from www.ariadne.ac.uk/issue25/app-profiles/

- Home page for ISO/IEC 11179 information technology -- metadata registries. (2014). Retrieved from http://metadata-standards.org/11179/

References

- Keenan, T. M. (2010). Why purchase when you can repurpose? Using crosswalks to enhance user access. *Code4lib Journal, 11*(2010-09-21). Retrieved from http://journal.code4lib.org/articles/3604

- Library of Congress. (n.d.). Bibliographic framework initiative. Retrieved from http://loc.gov/bibframe/

- Library of Congress. (2011, July 1). METS Example Documents. In Metadata Encoding & Transmission Standard Official Website (Bibliographic Record). Retrieved from http://www.loc.gov/standards/mets/mets-examples.html

- Lopatin, L. (2010). Metadata practices in academic and non-academic libraries for digital projects: a survey. *Cataloguing & Classification Quarterly, 48*(8), 716-742. doi:10.1080/01639374.2010.509029

- Miller, S. J. (2011). Metadata digital collections. New York, NY: Neal-Schuman Publishers, Inc.

- National Information Standards Organization (NISO). (2004). Understanding metadata. Bethesda, MD: NISO Press. Retrieved from http://www.niso.org/publications/press/ UnderstandingMetadata.pdf

References

- Neiswender, C., Montgomery, E. 2009. Metadata interoperability—what is it, and why is it important? *The MMI Guides: Navigating the World of Marine Metadata*. Retrieved from http://marinemetadata.org/guides/mdataintro/mdatainteroperability

- Otto, JJ. (2014). Administrative metadata for long term preservation and management of resources: a survey of current practices in ARL libraries. *LRTS 58*(1), 4-32.

- Rolan, G. (2015) Towards archive 2.0: issues in archival systems interoperability. *Archives and Manuscripts, 43*(1), 42-60. doi:10.1080/01576895.2014.959535

- Spicher, K. M. (1996). The development of the MARC format. *Cataloging & Classification Quarterly, 21*(3/4), 75-90. doi:10.1300/J104v21n03_06

- Taylor, A.G. & Joudrey, D.N. (2009). *The organization of information* (3rd ed.). Library and information science text series. Westport, CT: Libraries Unlimited.

- Tennant, R. (2002). MARC must die. *Library Journal 127*(17), 26-27.

- Tharani, K. (2015). Linked Data in libraries: a case study of harvesting and sharing bibliographic metadata with BIBFRAME. *Information Technology and Libraries, 34*(1), 5-19.

References

Urbano, C., Zhang, Y., Downey, K., & Klingler, T. (2015, May). Library catalog log analysis in e-book patron-driven acquisitions (PDA): a case study. *College & Research Libraries, 76*(4), 412-426. doi:10.5860/crl.76.4.412

WC3. (n.d.). Linked data. Retrieved from http://www.w3.org/standards/semanticweb/data

Zeng, M. L. & Chen, L. M. (2006). Metadata interoperability and standardization: a study of methodology part II. D-*Lib Magazine, 12*(6), 11. doi:10.1045/june2006-zeng