1: METADATA STANDARDS AND CREATION

Mark Sullivan
Digital Library of the Caribbean
Metadata for Digital Resources

- What is metadata?
- What is your digital resource?
- Standards
- Standards Comparison
- Tools
- Questions
What is metadata?
What is metadata?
What is metadata?
Library and Museum Metadata
Library and Museum Metadata

Greek Vase
470 – 460 BC
35 centimeters
Odysseus and Eumaios the Swineherd
from Homer’s story of the Odyssey
Format: Greek Vase
Date: 470 – 460 BC
Height: 35 centimeters
Title: [Greek Vase of Odysseus and Eumaiosthe Swineherd]
Notes: from Homer’s story of the Odyssey
Library and Museum Metadata
Title: L' Isle St. Domingue
Alternate Title: Atlas portatif Universel..
Physical Desc: 1 Map. : col; 17 x 22 cm on sheet 23 x 32 cm.
Language: French
Creator: Robert de Vaugondy, Gilles, 1688-1766
Date: 1749
Place of Publ: Paris 1749
Subjects: Maps --Early works to 1800 --West Indies (lcsh)
Maps --Early works to 1800 --Hispaniola (lcsh)
Genre: single map (marcgt)
Maps (lcsh)
Early works to 1800 (lcsh)
Spatial Coverage: Haiti, Dominican Republic
Note: Outline color.
Library and Museum Metadata
Avanzamos o retrocedemos? : reflexiones sobre…

165 p.
Spanish
Mejia, Radhames
Santo Domingo : PUCMM/CIEDHUMANO
2010
Educacion-Republica Dominicana-Ensayos, conferencias
Este libro recoge una serie de artículos sobre temas educativos publicados en los años 2006 y 2007 en el periódico El Caribe.
RD 370.97293 M516a
isbn-9789945415339
Library Metadata

Economic theory, econometrics, and mathematical economics.

Ichilishi, Tatsuro.

Game theory for economic analysis /
Tatsuro Ichilishi. -- New York :
x, 164 p. : ill. ; 24 cm. --
(Economic theory, econometrics, and mathematical economics)

Bibliography: p. 151-156.
Includes indexes.
ISBN 0-12-370180-5

1. Game theory. 2. Equilibrium (Economics) I. Title. II. Series
## Library Metadata - MARC

<table>
<thead>
<tr>
<th>FMT</th>
<th>BK</th>
</tr>
</thead>
<tbody>
<tr>
<td>LDR</td>
<td>008820706s19838nyua00010eng</td>
</tr>
<tr>
<td>010</td>
<td>a82011554</td>
</tr>
<tr>
<td>020</td>
<td>a0123701805</td>
</tr>
<tr>
<td>035</td>
<td>a(OCoLC)08627997</td>
</tr>
<tr>
<td>035</td>
<td>aACB4513 bUF</td>
</tr>
<tr>
<td>040</td>
<td>aDLC cDLC dFUG</td>
</tr>
<tr>
<td>049</td>
<td>aFLGA</td>
</tr>
<tr>
<td>050</td>
<td>aHB144 b128 1983</td>
</tr>
<tr>
<td>082</td>
<td>a330/.01/5193 2 19</td>
</tr>
<tr>
<td>100</td>
<td>aIchiiishi, Tatsuro.</td>
</tr>
<tr>
<td>245</td>
<td>aGame theory for economic analysis / cTatsuro Ichiiishi.</td>
</tr>
<tr>
<td>300</td>
<td>ax, 164 p. : bill. ; c24 cm.</td>
</tr>
<tr>
<td>490</td>
<td>aEconomic theory, econometrics, and mathematical economics</td>
</tr>
<tr>
<td>504</td>
<td>aBibliography : p. 151-156.</td>
</tr>
<tr>
<td>500</td>
<td>aIncludes indexes.</td>
</tr>
<tr>
<td>599</td>
<td>amgh</td>
</tr>
<tr>
<td>650</td>
<td>aGame theory.</td>
</tr>
<tr>
<td>650</td>
<td>aEquilibrium (Economics)</td>
</tr>
</tbody>
</table>
What is your digital resource?
What is your digital resource?

- **Physical archival containers**
  - Archival box?
  - Archival folder?
  - Item-level?

- **Scrapbooks**
  - Pages
  - Photographs

- **Newspapers**
  - Issues
  - Article
Metadata Standards
Why metadata standards?
Standards

“The nice thing about standards is there are so many to choose from.”

- Andrew Tannenbaum (1988)
Types of Metadata Standards

- Bibliographic Description
- Metadata “Wrappers” and Transport
- Collection Descriptions
- Other Standards
  - Authority Standards
  - Proprietary Standards
  - etc…
Types of Metadata Standards

- Bibliographic Description
- Metadata “Wrappers” and Transport
- Collection Descriptions
- Other Standards
  - Authority Standards
  - Proprietary Standards
  - etc…
Bibliographic Description Standards

- MARC
- Dublin Core
- MODS
- VRA Core
Bibliographic Description : MARC

- **MARC**
  - Originally developed in 1960s by Library of Congress
  - Most embraced metadata standard in libraries

- **MARC 21**
  - Combination of USMARC and CAN/MARC
  - Metadata standard for 21st century

- **MarcXML**
  - Same MARC format, except encoded in XML
Bibliographic Description: MARC

- **Positives:**
  - Can encode a very granular amount of data
  - Very well adopted and works well with machine readers

- **Negatives:**
  - Not very human-readable
  - High learning curve
  - Mixes data with display (as commonly implemented)
Bibliographic Description : DC

- **Dublin Core**
- Originally developed between 1994 and 1995 in Dublin Ohio by OCLC
- Used widely on web pages to assist search engines
- **Simplified Dublin Core**

<table>
<thead>
<tr>
<th>Title</th>
<th>Creator</th>
<th>Subject</th>
</tr>
</thead>
<tbody>
<tr>
<td>Description</td>
<td>Publisher</td>
<td>Contributor</td>
</tr>
<tr>
<td>Date</td>
<td>Type</td>
<td>Format</td>
</tr>
<tr>
<td>Identifier</td>
<td>Source</td>
<td>Language</td>
</tr>
<tr>
<td>Relation</td>
<td>Coverage</td>
<td>Rights</td>
</tr>
</tbody>
</table>
Bibliographic Description : DC

- Qualified Dublin Core
  - Added three elements (Audience, Provenance, RightsHolder)
  - More importantly, added some basic refinements

- Positives
  - Widely accepted and easy to read and encode

- Negatives
  - Even with qualified dublin core, difficult to encode complex data
  - Lack of data refinement leads to loss of UI options
Bibliographic Description : MODS

- **Metadata Object Description Scheme**
- Developed in 2002 by Library of Congress
- Beginning to be the de facto standard for digital libraries (MODS/METS)

**Positives**
- Can handle a (large) subset of MARC tags
- Handles complex objects and can easily be extended

**Negatives**
- Hard to do round trip portability from MARC → MODS → MARC
Bibliographic Description : VRA Core

- Visual Resource Association
- Developed in 1996 by Visual Resource Association
- Used for describing visual/cultural materials
- Includes all the standard tags
  - Date, Title, Description, Rights, Subject, etc..
- Also more unique tags
  - Cultural Context
  - Style
  - Technique
Bibliographic Description Standards

- MARC
- Dublin Core
- MODS
- VRA Core
Types of Metadata Standards

- Bibliographic Description
- Metadata “Wrappers” and Transport
- Collection Descriptions
- Other Standards
  - Authority Standards
  - Proprietary Standards
  - etc…
Metadata “Wrappers” and Transport

Contents
- Files and Structural data
- Administrative data

Choices
- METS
- OAI-PMH
- *No wrapper necessary*
Metadata “Wrappers” : METS

- **Metadata Encoding & Transmission Standard**
- Contains
  - Descriptions section(s)
  - Administrative section(s)
  - File section(s)
  - Structure map(s)
- Accepts any XML schema
Metadata “Wrappers” : METS

- Metadata Encoding and Transmission Standard (METS)
  - metsHeader – information about the METS
  - dmdSec – descriptive metadata
  - amdSec – administrative metadata
  - fileSec – list of all files grouped by type
  - structMap – divisions/pages linked to files
Metadata “Wrappers” : OAI-PMH

- Open Archives Initiative - Protocol for Metadata Harvesting (2000-2001)
- Defines a protocol (over HTTP)
- Main Verbs
  - Identify
  - ListMetadataFormats – usually dublin core (99%)
  - ListSets
  - ListRecords
Types of Metadata Standards

- Bibliographic Description
- Metadata “Wrappers” and Transport
- Collection Descriptions
- Other Standards
  - Authority Standards
  - Proprietary Standards
  - etc…
Collection Metadata Standards

- (METS)

- EAD
  - Encoded Archival Description
  - Electronic Finding Guide/Aid
  - Main Sections
    - Descriptions
    - Container List – can link to digital objects
Types of Metadata Standards

- Bibliographic Description
- Metadata “Wrappers” and Transport
- Collection Descriptions

Other Standards
- Authority Standards
- Proprietary Standards
- etc…
Still to come...

- Authority Metadata Standards
  - NACO
  - MADS
  - EAC
- Other Standards
  - KML
  - DarwinCore
  - Z39.50 / ZING
- Standards are like weeds....
Proprietary and Custom Formats

- Proprietary File Formats
  - Greenstone Document File (doc.xml)
  - Fedora Object File (FOXML)

- Custom Schemas
  - SobekCM Schema (UFDC and dLOC)
Back to the beginning?

MODS?

Dublin Core?

KML?

VRA Core?

MADS?

EAC?

MARC?

EAD?

METS?

To wrap?
Evidence-based

- Universities within Florida (10 schools)
- Digital Objects
  - MARC (linked to catalog)
  - Dublin Core (no wrapper, simple objects)
  - METS wrapper (for complex items)
  - MODS
- EAD for Finding Aids / Guides
- OAI-PMH Harvesting
“One standard to rule them all”
“One standard to rule them all”

METS Wrapper
“One standard to rule them all”

METS Wrapper

DC ➔ MODS

Additional custom metadata
“One standard to rule them all”

- METS Wrapper
  - Dublin Core
  - Additional custom metadata
“One standard to rule them all”

METS Wrapper

MODS

Additional custom metadata
“One standard to rule them all”

METS Wrapper

MODS

VRA Core

Darwin Core

Additional custom metadata
“One standard to rule them all”

Archival

METS

MODS

SobekCM

Service

Greenstone
“One standard to rule them all”

Archival

METS
MODS
SobekCM

Service

Fedora, ContentDM, etc…
Standards Comparison

Link to external comparison

http://digital.uflib.ufl.edu/development/standards_comparison.pdf
Lessons / Standards Conclusions

- Try not to tie your metadata to your system
  - Archival vs. Service

- Don’t be afraid of extending for your own needs, but continue to follow the standards

- Don’t cripple your metadata / Prepare for the most information you would want
  - METS / MODS v. simple Dublin Core
Tools
Metadata Tools

- Library of Congress site of METS utilities
  - [http://www.loc.gov/standards/mets/mets-tools.html](http://www.loc.gov/standards/mets/mets-tools.html)

- SobekCM METS Editor (UF)
  - [http://ufdc.ufl.edu/metseditor](http://ufdc.ufl.edu/metseditor)

- MarcEdit Tool
Review of Topics

- What is metadata?
- What is your digital resource?
- Standards
- Standards Comparison
- Tools
Questions?