Shareable Metadata in the Museum Community

Jenn Riley
Metadata Librarian
Indiana University Digital Library Program
What does this record describe?

**identifier:** http://name.university.edu/IC-FISH3IC-X0802]1004_112

**publisher:** Museum of Zoology, Fish Field Notes

**format:** jpeg

**rights:** These pages may be freely searched and displayed. Permission must be received for subsequent distribution in print or electronically.

**type:** image

**subject:** 1926-05-18; 1926; 0812; 18; Trib. to Sixteen Cr. Trib. Pine River, Manistee R.; JAM26-460; 05; 1926/05/18; R10W; S26; S27; T21N

**language:** UND

**source:** Michigan 1926 Metzelaar, 1926--1926;

**description:** Flora and Fauna of the Great Lakes Region
Trib Pine River, Mainstem

Water: mostly spring fed; clear, spring floods water bad.
Vegetation: not much, absolutely nothing.
Bottom: sand, no mud, rather deep. Temp. 47° air 62°
Shore: mostly cleared, farma, some brush. Current:

Distance from shore: about 4 ft
Depth of capture: Depth of water:
Method of capture:

Collected by: Metzler
Date: V: 15: 1926

Orig. preserv.

Animal life subnormal
Why share metadata?

- Benefits to users
  - One-stop searching
  - Aggregation of subject-specific resources

- Benefits to institutions
  - Increased exposure for collections
  - Broader user base
  - Bringing together of distributed collections

Don’t expect users will know about your collection and remember to visit it.
Sharing can be hard

- Some initiatives have fizzled out
  - CIMI
  - AMICO

- Some are still going
  - ARTstor
  - RLG Cultural Materials
  - CAMIO and other AMICO derivatives

- Note focus on art museums
But it’s getting easier

- **Open Archives Initiative Protocol for Metadata Harvesting** (OAI-PMH)
  - Popular “low barrier” mechanism
  - Shares metadata, not necessarily content
  - Any metadata format with XML Schema can be shared

- Museum-centric OAI initiatives are emerging
  - CDWA Lite from the Getty
  - RLG Museum Collections Sharing Working Group
  - UC Berkeley Art Museum leading project to develop [MOAC Community Toolbox](#)

- Other sharing mechanisms: Z39.50->SRU, A9/OpenSearch
How OAI works

Diagram from OAI for Beginners - the Open Archives Forum online tutorial at http://www.oaforum.org/tutorial/english/intro.htm
Typical service provider behavior

- “Generic”
  - Collect and normalize metadata
  - Provide basic discovery
  - Send user back to home institution for more information and/or access to content
  - **OAIster** is a good example

- Domain-specific
  - More advanced discovery capabilities
  - Selling branded products
  - ????
Three possible architectures

CMS or DAMS

- Metadata creation module
- Transformation
- OAI data provider module
  - QDC
  - MODS
  - DC
  - CDWA Lite

Stand-alone OAI data provider

- Transformation
- Static Repository Gateway
- XML File
- QDC
- MODS
- DC
- CDWA Lite

OAI Harvester
Shareable metadata defined

- Promotes search interoperability - “the ability to perform a search over diverse sets of metadata records and obtain meaningful results” (Priscilla Caplan)
- Is human understandable outside of its local context
- Is useful outside of its local context
- Preferably is machine processable
Finding the right balance

- Metadata providers know the materials
  - Document encoding schemes and controlled vocabularies
  - Document practices
  - Ensure record validity
- Aggregators have the processing power
  - Format conversion
  - Reconcile known vocabularies
  - Normalize data
  - Batch metadata enhancement
Metadata as a view of the resource

- There is no monolithic, one-size-fits-all metadata record
- Metadata for the same thing is different depending on use and audience
- Affected by format, content, and context
- Descriptive vs. administrative vs. technical, etc. data
Choice of vocabularies as a view

- **Names**
  - LCNAF: Michelangelo Buonarroti, 1475-1564
  - ULAN: Buonarroti, Michelangelo

- **Places**
  - LCSH: Bloomington (Ind.)
  - TGN: Bloomington

- **Subjects**
  - LCSH: Neo-impressionism (Art)
  - AAT: Pointillism
Choice of metadata format(s) as a view

- Many factors affect choice of metadata formats
- Many different formats may all be appropriate for a single item
- High-quality metadata in a format not common in your community of practice is not shareable
- Museum-focused formats still developing
  - CDWA Lite for art museums
  - CIMI had a good start, but no longer maintained
Focus of description as a view

- Link between records for analog and digital
- Hierarchical record with all versions
- Physical with link to digital
- All versions in flat record
- Content but not carrier
6 Cs and lots of Ss of shareable metadata

- Content
- Consistency
- Coherence
- Context
- Communication
- Conformance

Metadata standards
Vocabulary and encoding standards
Descriptive content standards
Technical standards
Content

- Choose appropriate vocabularies
- Choose appropriate granularity
- Make it obvious what to display
- Make it obvious what to index
- Exclude unnecessary “filler”
- Make it clear what links point to
Consistency

- Records in a set should all reflect the same practice
  - Fields used
  - Vocabularies
  - Syntax encoding schemes
- Allows aggregators to apply same enhancement logic to large groups of records
Coherence

- Metadata format chosen makes sense for materials and managing institution
  - Not just Dublin Core!
  - Museums have specific needs: context, interpretation, relationships between objects, provenance, etc.

- Record should be self-explanatory
- Values must appear in appropriate elements
- Repeat fields instead of “packing”
Context

- Include information not used locally
- Exclude information *only* used locally
- Appropriate context driven by intended use
Communication

- Method for creating shared records
- Vocabularies and content standards used
- Record updating practices and schedules
- Accrual practices and schedules
- Existence of analytical or supplementary materials
- Provenance of materials
Conformance to Standards

- Metadata standards, e.g., CDWA Lite
- Vocabulary and encoding standards, e.g., TGN
- Descriptive content standards, e.g., CCO
- Technical standards e.g., Sharing protocol, XML, Character encoding
The reality of sharing metadata

- Creating shareable metadata requires thinking outside of your local box
- Creating shareable metadata will require more work from you and your technical staff
- Creating shareable metadata will require our vendors to support (more) standards
- Creating shareable metadata is no longer an option, it’s a requirement
For more information

- jenlrile@indiana.edu
- DLF/OAI Best Practices for Shareable Metadata
  <http://oai-best.comm.nsd1.org/cgi-bin/wiki.pl?PublicTOC>
- These presentation slides:
  <http://www.dlib.indiana.edu/~jenlrile/presentations/mcn2006/shareableMetadata.ppt>