Creating Shareable Metadata

Pre-Conference at WebWise 2006:
Inspiring Discovery: Unlocking Collections
Los Angeles, CA   February 15, 2006

Jenn Riley, Indiana University

Sarah Shreeves, University of Illinois at Urbana-Champaign

Liz Milewicz, Emory University

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

Dublin Core record retrieved via the OAI Protocol

Orig. No. 460   Sta. No. ....

Locality Michigan: Trib. to sixteen a., Sec. 26-27,
T. 21M. R. 10 W., Wexford Co.
Trib. Pine River → Manistee

Water: mostly spring-fed; clear; spring floods ~~not so~~ rather bad.
Vegetation: ~~not much~~ absolutely nothing
Bottom: sand; no mud, rather deep   Temp. 47° air 62°
Shore: ~~mostly cleared~~ → farms & some brush   Current: ~~swift~~
Distance from shore: av. width 4 ft   Tide: ....
Depth of capture: ....   Depth of water: ~~~~ 1 ft.
Method of capture: ....
Collected by Metzelaar   Date: V : 18 : 1926
Orig. preserv. ....   Time: ....

Animal life subnormal

And this one?

- [Sample Record](#) (adapted from an Illinois database)

Creating Shareable Metadata:
WebWise 2006 Preconference
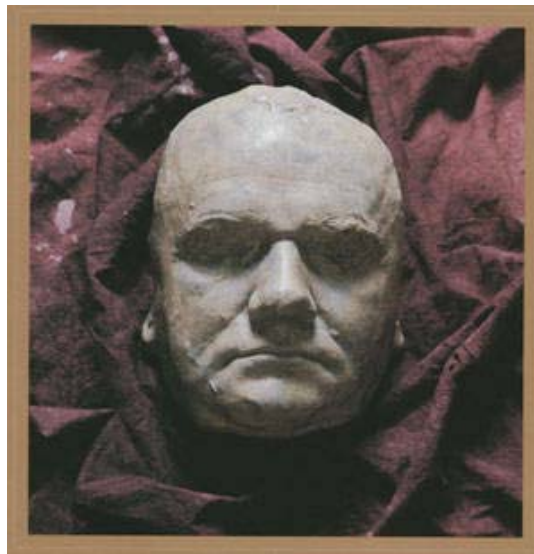
Agenda

1. Introduction
2. Why share metadata anyway?
3. What is shareable metadata?
4. How do you create shareable metadata?
5. Communication and documentation
6. Before you share…

Time for questions between each section

Packets

Break around 2:45

Creating Shareable Metadata:
WebWise 2006 Preconference

We **hope that you** will leave today and…

- Understand the **need** for interoperable or shareable metadata

- Understand the **impact** your metadata has on larger aggregations

- Understand the **key components** of shareable metadata

- **Think critically** about the shareability of their own metadata.

Creating Shareable Metadata:
WebWise 2006 Preconference

What we aren't covering…

- How to share metadata (whether via the OAI Protocol, SRW/U, FTP, CD-ROM, Cooperative Cataloging, Excel Spreadsheet!)
- Details of technical aspects of shareable metadata such as XML schema & namespaces

- How to use specific metadata formats

One Size Does **Not** Fit All!

Creating Shareable Metadata:
WebWise 2006 Preconference

Questions?

Creating Shareable Metadata:
WebWise 2006 Preconference

Why share metadata anyway?

- ## Benefits to users
  - Single search of a variety of digital resources
  - Aggregation of subject-specific resources
  - *Higher quality* resources

- ## Benefits to institutions
  - Increased user access to collection by allowing metadata to appear in other places
  - Exposure to broader audience, new users
  - Surfacing rare, unknown, or scattered collections

Creating Shareable Metadata:
WebWise 2006 Preconference

Different shapes and sizes…

- Range of different aggregations:
  - focused subject area v. comprehensive
  - specialized audience v. general audience

- Range of different displays

Creating Shareable Metadata:
WebWise 2006 Preconference

Metadata aggregators

- ## CIC Metadata Portal
  ### http://cicharvest.grainger.uiuc.edu/

  - Records and digital resources shared by consortium of institutions, *provided for*
  - Educators, researchers, and general public

<u>Benefits</u>:

Single comprehensive search of multiple collections and a variety of disciplines

CIC Metadata Portal

February 15, 2006

CIC Metadata Portal

Search CIC Library Resources

Click here to install the Firefox plug-in

Add Firefox search plugin

| Simple search | Bibliographic search | Browse collections | Project Details | Help |

kw: "reading rooms"

[Search]

196 records.

Refine search ☐

Results by [Collection ▼]

[Sort by] [weighted hit frequency ▼]

[Group results by collection]

Next 10 Records

Jump to Records: **1** | 11 | 21 | 31 | 41 | 51 | 61 | 71 | 81 | 91 | 101 | 111 | 121 | 131 | 141 | 151 | 161 | 171 | 181 | 191

Michigan State University:
The Making of Modern Michigan
3 records

Ohio State University:
The OhioLINK Digital Media Collection
1 record

University of Chicago:
Archival Photofiles
186 records

University of Minnesota:
War porters
1 record

University of Wisconsin-Madison:
Portal to Asian Internet Resources
1 record

Record 1 of 196

Title: American Library Association Archives: Association for Library Service to Children (ALSC): Publications: Children's Library Yearbooks

Subject: Booklists; Children, Committee on Library Work With; Children's Libraries; Children's Services; Foreign Born, Library Work With The; Salaries

Type: Text

Collection: American Library association finding-aid

Record 2 of 196

Title: American Library Association Archives: International Relations Office (IRO): American Library School in Paris: American Legion Conference Subject File

Subject: American Legion; Paris, American Library in

Type: Archival Material

Collection: American Library association finding-aid

Creating Shareable Metadata:
WebWise 2006 Preconference

CIC Metadata Portal

Metadata aggregators

- ## National Science Digital Library
  http://nsdl.org

  - Online resources and records pertaining to science & math education and research, vetted for inclusion, *provided for*
  - Educators, researchers, policy makers, and the general public

Benefits:

Single portal serving a range of resources on a specialized topic to a diverse audience

Creating Shareable Metadata:
WebWise 2006 Preconference

NSDL

NSDL

NSDL

Metadata aggregators

- **Collaborative Digitization Program**
  <http://www.cdpheritage.org/>
  - Records and digital resources from archives, museums, and libraries, pertaining to cultural heritage of the American West
    provided for
  - Educators, researchers, and the general public

<u>Benefits</u>:

Single portal serving a range of resources on a specialized topic to a diverse audience

Creating Shareable Metadata:
WebWise 2006 Preconference

Collaborative Digitization Program

Collaborative Digitization Program

Creating Shareable Metadata:
WebWise 2006 Preconference

Collaborative Digitization Program

Creating Shareable Metadata:
WebWise 2006 Preconference

Not only stand alone portals…

- http://www.a9.com/

Creating Shareable Metadata:
WebWise 2006 Preconference

Search: illinois cushman [GO]

Home | Prefs | Toolbar | Sign Out

Advanced Web Search

Hello Sarah L. Shreeves. Click here if this is not you

▼ Hide Column Choices

☑ Web   ☐ Books   ☑ Images   ☐ Yellow Pages   ☐ Reference   ☐ Movies
☐ People   ☐ Blog Search   ☐ Wikipedia   ☐ Your History   ☐ Your Diary   ☐ Your Bookmarks
☑ DLF MODS Portal ☐ [More Choices... ▼]

Join the A9 Instant Reward program to get 1.57% off almost everything on Amazon.com

Web Results [full] [close]

Showing 1 - 10 of about 39,400

CUSHMAN PAGE

Karen **Cushman** was born in Chicago,
Illinois. As a child, she learned about the
public library very early. She would go
alphabetically down the shelves and ...
http://www.scils.rutgers.edu/~kvander/
cushman.html - 15k Cached [Site Info]

Karen Cushman

Karen **Cushman** was born in Chicago,
Illinois and lives now on Vashon Island west
of Seattle, Washington. She received an MA
in human behavior and one in ...
http://www.houghtonmifflinbooks.com/
catalog/authordetail.cfm?authorID=1451 -

Image Results [full] [close]

Showing 1 - 17 of 17

Digital Library Federation MODS Portal [full] [close]

1 - 10 of about 2,488

Cairo, Illinois . Court House. Cairo, Ill.

Title: Cairo, Illinois. Court House. Cairo, Ill.
Alternative Title: County [...] House. **Author/Creator:**
Cushman, Charles Weever, 1896-1972, photographer,
pht **Type:** [...] Fire hydrants **Subject: Hierarchical:**
United States; Illinois; Alexander; Cairo Junction
Identifier: 1952.14; P06185 [...] **URL:**
http://purl.dlib.indiana.edu/iudl/archives/cushman/P061
Rights: Copyright and reproduction rights for all [...]
W. Cushman photographs are held by Indiana
University and administered by the University [...]
Indiana University Digital Library Program: Charles W.
Cushman Photograph Collection [...]

Cairo, Ill. Honey locust ? St. Mary's Park. Cairo, Ill.

[...] tree ? (with pods) in St. Mary's Pk.
Author/Creator: Cushman, Charles Weever,
1896-1972, photographer, pht **Type:** [...] (Botany);
Acacia **Subject: Hierarchical:** United States; Illinois;
Alexander; Cairo Junction **Identifier:** 1952.13; P06184
[...] **URL:**
http://purl.dlib.indiana.edu/iudl/archives/cushman/P061
Rights: Copyright and reproduction rights for all [...]
W. Cushman photographs are held by Indiana
University and administered by the University [...]
Indiana University Digital Library Program: Charles W.

Common Problems with Metadata in Aggregation

- Consistency
- Sufficiency
- Compatibility

Consistency problems

- Appearance of data
- Application of format
- Granularity of records
- Vocabulary usage

<u>Result:</u>

Service Provider must normalize data
(if can determine what "normal" is)

Creating Shareable Metadata:
WebWise 2006 Preconference

Sufficiency problems

- Too little info for understanding *what resource is,* especially outside of local context

Result:

Users don't know whether a resource is relevant or not

Creating Shareable Metadata:
WebWise 2006 Preconference

Compatibility problems

Information in records is

– Erroneous

– Unnecessary

– Incompatible

<u>Result:</u>

Interferes with harvesting and indexing

Questions?

Creating Shareable Metadata:
WebWise 2006 Preconference

What is Shareable Metadata?

Creating Shareable Metadata:
WebWise 2006 Preconference

Shareable Metadata…

- Promotes search interoperability

> "the ability to perform a search over diverse sets of metadata records and obtain meaningful results." (Priscilla Caplan)

- Is (at the least) human understandable outside of its local context (the field note!)

- Is useful outside of its local context
   (Can we build something off of it?)

- Preferably is machine processable!

The Four C's (and lots of S's) of Shareable Metadata

Consistency
Coherence
Context
Conformance

Metadata standards
(and not just DC)

Vocabulary and encoding standards

Descriptive content standards
(AACR2, CCO, DACS)

Technical standards
(XML, Character encoding, etc)

Questions?

Creating Shareable Metadata:
WebWise 2006 Preconference

How do you create shareable metadata?

Appropriate representation of the resource through metadata

- Metadata as a *view* of the resource
- Standards promote interoperability
- Appropriate formats
- Appropriate content
- Appropriate context

Creating Shareable Metadata:
WebWise 2006 Preconference

Metadata as a *view* of the resource

- Metadata for the same thing is different depending on *use* and *audience*
- Affected by format, content, and context
- Harry Potter as represented by...
  - a <u>public library</u>
  - an <u>online bookstore</u>
  - a <u>fan site</u>

Creating Shareable Metadata:
WebWise 2006 Preconference
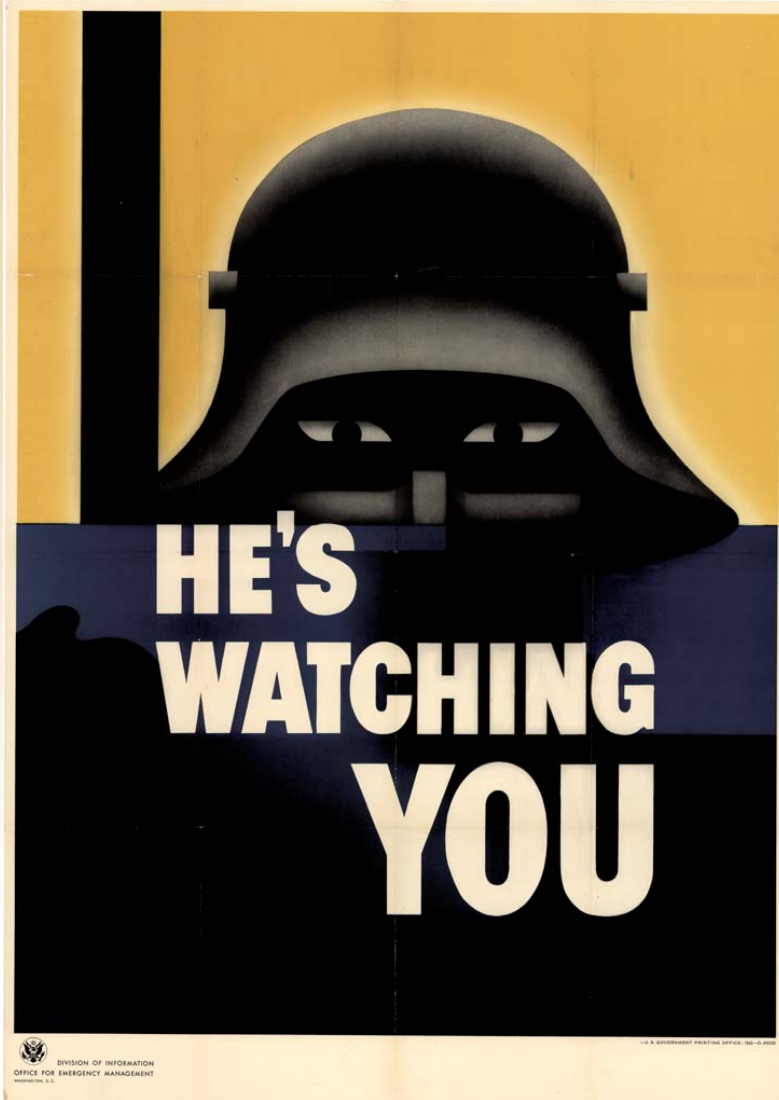
**Same Object,
Different Descriptions,
Different Metadata**

Illinois Digital Archives (**alt**)

University of Minnesota Libraries
(**alt**)

Northwestern Poster Collection
(**alt**)

Teaching with Digital Content (**alt**)

World War II Poster

Creating Shareable Metadata:
WebWise 2006 Preconference

Choice of vocabularies as a *view*

- Names
  - LCNAF: Michelangelo Buonarroti, 1475-1564
  - ULAN: Buonarroti, Michelangelo
- Places
  - LCSH: Jakarta (Indonesia)
  - TGN: Jakarta
- Subjects
  - LCSH: Neo-impressionism (Art)
  - AAT: Pointillism

Standards promote interoperability

Types of standards to consider

- Data structure standards
  - "Buckets" of information (fields)
  - Both label and scope important
  - e.g., MARC, MODS, Dublin Core

- Data content standards
  - Selection, structure and formatting of value within a field
  - e.g., AACR2, DACS, CCO

Creating Shareable Metadata:
WebWise 2006 Preconference

Appropriate formats

- Depends upon:
  - nature of materials
  - nature of holding institution
  - depth of description needed
  - community practice
  - need for distinction between versions
  - need for repeating elements
- MARC, MODS, Dublin Core, EAD, and TEI may all be appropriate for a single item

Creating Shareable Metadata:
WebWise 2006 Preconference

Break

Creating Shareable Metadata:
WebWise 2006 Preconference

Appropriate content

- Choose appropriate vocabularies
- Choose appropriate granularity
- Make it obvious what to display
- Exclude unnecessary "filler"
- Make it clear what links point to
- Handout has some guidance for "classes of metadata elements"

Creating Shareable Metadata:
WebWise 2006 Preconference
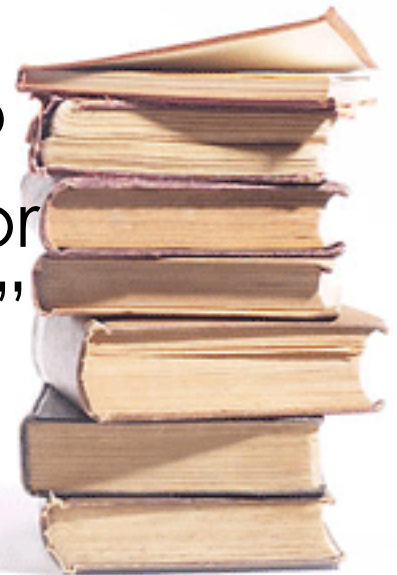
Common content mistakes

- No indication of vocabulary used - [Example](#)

- Shared record for a single page in a book

- Link goes to search interface rather than item being described

- "Unknown" or "N/A" in metadata record

Creating Shareable Metadata:
WebWise 2006 Preconference

Appropriate context

- Thinking about shareability
  - Include information not used locally
  - Exclude information *only* used locally

- Current safe assumptions
  - Users discover material through shared record
  - User then delivered to your environment for full context

- Context driven by intended use

Creating Shareable Metadata:
WebWise 2006 Preconference

Common context mistakes

- Leaving out information that applies to an entire collection ("On a horse")

- Location information lacking parent institution

- Geographic information lacking higher-level jurisdiction

- Inclusion of administrative metadata

Creating Shareable Metadata:
WebWise 2006 Preconference

What are you describing?

Physical object w/ links to the digital? (Digital surrogate approach)

Both digital and physical in the same flat record?

Content but not the carrier?

Both digital and physical in the same record but in a hierarchy?

A record for the analog and the digital item with linkage? (one to one principle)

Creating Shareable Metadata:
WebWise 2006 Preconference

Questions?

Creating Shareable Metadata:
WebWise 2006 Preconference

Communication

- Metadata providers can learn from aggregators

- Aggregators can learn from metadata providers

- Providing supplemental information to make records more intelligible

Creating Shareable Metadata:
WebWise 2006 Preconference

Metadata providers can learn from aggregators

- Crosswalking methods and rules

- Information to include and exclude

- Choice of standards
  - Metadata formats
  - Vocabularies

- Where to spend normalization effort

- Use can drive future priorities

  [Example from the CIC Portal](...)… ([alt](...))

Creating Shareable Metadata:
WebWise 2006 Preconference

Aggregators can learn from metadata providers

- Where to spend normalization effort

- Context, importance, and primary uses of resources shared

- Variety of resource types and descriptive practices encountered

- Local, robust metadata formats

Creating Shareable Metadata:
WebWise 2006 Preconference

Providing supplemental information to make records more intelligible

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

Creating Shareable Metadata:
WebWise 2006 Preconference

Before you share…

- Check your metadata
  - Appropriate view?
  - Consistent?
  - Context provided?
  - Does the aggregator have what they need?
  - Documented?

Can a stranger tell you what the record describes?

Creating Shareable Metadata:
WebWise 2006 Preconference

Questions?