

## Metadata & Other Cataloging

**Howard Besser**  
**NYU Archiving and Preservation Program**  
**and Library Senior Scientist**  
<http://www.tisch.nyu.edu/preservation>  
<http://besser.tsoa.nyu.edu/howard/>

Whether your collection is analog or digital, your metadata will still be digital

- **Catalog**
- **Finding Aid**
- **Spreadsheets or databases**
- **Embedded within the essence file**

## Metadata vs. Essence

## Concepts & Terminology

- **Data Fields (buckets)**
- **Data Values**
- **Controlled Vocabulary**
- **Semantics/Syntax/Structure-**

## Semantics/Syntax/Structure

- **Semantics**
  - meaning, as defined by a community to meet their particular needs (DC)
- **Syntax**
  - a systematic arrangement of data elements for machine processing
  - facilitates the exchange and use of metadata among various applications (HTML, XML, RDF)
- **Structure**
  - a formal arrangement of the syntax with the goal of consistent representation of the semantics (rules defining field contents like 1/11/99)

## Cataloging Principles

- [http://mic.imtc.gatech.edu/catalogers\\_portal/cat\\_principles.htm](http://mic.imtc.gatech.edu/catalogers_portal/cat_principles.htm)

## Cataloging/Metadata approaches

- **Libraries**
    - Item level records
    - Marc/AACR2
  - **Archives**
    - Collection level records
    - Finding Aids, fonds, boxes, folders
    - <http://www.oac.cdlib.org/findaid/ark:/13030/t9779q7sn> ,
    - <http://www.oac.cdlib.org/findaid/ark:/13030/tf796nb4fk> ,
    - <http://oac.cdlib.org/findaid/ark:/13030/tf238n986k>
  - **Museums**
    - Art-item level records
    - Science-collection or item level records
  - **Smaller Institutions**
    - Databases, Spreadsheets
    - IMAP Template
- Besser--SOIMA Metadata 18/11/09

7

## MIC Website

- <http://mic.imtc.gatech.edu/>

Besser--SOIMA Metadata 18/11/09

8

## Moving Image Cataloging

- **Johnson, Jane, and Kniesner, Dan.** "Glossary of Cataloging and General Terms." MIC.  
[http://gondolin.rutgers.edu/MIC/text/how/catalog\\_glossary.htm](http://gondolin.rutgers.edu/MIC/text/how/catalog_glossary.htm)
- **Archival Moving Image Materials: a Cataloging Manual. 2nd ed.**  
<http://www.tlcdelivers.com/tlc/crs/arch0332.htm>
- **Library of Congress Authorities, 2002.** <http://authorities.loc.gov/>
- **Moving Image Genre-Form Guide., 1998.**  
[www.loc.gov/rr/mopic/migintro.html](http://www.loc.gov/rr/mopic/migintro.html)
- **Radio Form/Genre Terms Guide, 2000.**  
<http://www.loc.gov/rr/record/frmggen.html>
- **Independent Media Arts Preservation. "Cataloging Project."**  
IMAP, 2002. [http://www.imappreserve.org/cat\\_proj/index.html](http://www.imappreserve.org/cat_proj/index.html)

Besser--SOIMA Metadata 18/11/09

9

## Cataloging : Data categories

- **Identification**
    - Unique identifiers
    - Titles
    - Location (physical and digital files)
  - **Descriptive**
    - Names of creatives
    - Subject and genre access
    - Dates
    - Summaries
  - **Technical**
    - Physical description (formats)
    - Digital file information
  - **Administrative**
    - Acquisition
    - Access restrictions
  - **Legal**
    - Who owns the copyright (intellectual property)
    - Restrictions on preservation, display, and use
  - **Preservation**
    - Condition
    - Preservation actions
- Besser--SOIMA Metadata 18/11/09

10

## Brief look at Cataloging Records

- **IMAP Cataloging Template**
  - [http://www.imappreserve.org/cat\\_proj/sample\\_records.html](http://www.imappreserve.org/cat_proj/sample_records.html)
- **Library of Congress (AACR2/MARC)**
- **CNN (hybrid News)**
- **MIC**

Besser--SOIMA Metadata 18/11/09

11

## What's challenging about Moving Image Collections?

- **Often important to collect various "unpublished" parts in different formats (A/B rolls, mag-stripe film, separated soundtracks, unedited footage, rushes, screen-tests, ...)**
- **Film Archivists seldom organize films or videos in standard archival arrangements (no fonds, boxes, folders, or Finding Aids)**
- **Ancillary materials (pressbooks, stills, scripts, storyboards, set designs, reviews ...) are often necessary to understand a work**

Besser--SOIMA Metadata 18/11/09

12

## Importance of Metadata Standards & Philosophies

## For Interoperability, Repositories Need Standards (as well as Sustainability & Access)

- ✿ **Descriptive Metadata for consistent description**
- ✿ **Discovery Metadata for finding**
- ✿ **Administrative Metadata for viewing and maintaining**
- ✿ **Structural Metadata for navigation**
- ✿ **... Terms & Conditions Metadata for controlling access...**

## Uses of Metadata

- **Discovery & Retrieval**
- **Identification/Provenance**
- **Rights Management**
- **Viewing**
- **Integrity**
- **Longevity**
- **Content rating**

## Why are Standards and Metadata consensus important?

- ✿ **Managing digital files over time**
- ✿ **Longevity**
- ✿ **Interoperability**
- ✿ **Veracity**
- ✿ **Recording in a consistent manner**
- ✿ **Will give vendors incentive to create applications that support this**

## Semantics/Syntax/Structure

- **Semantics**
  - meaning, as defined by a community to meet their particular needs (DC)
- **Syntax**
  - a systematic arrangement of data elements for machine processing
  - facilitates the exchange and use of metadata among various applications (HTML, XML, RDF)
- **Structure**
  - a formal arrangement of the syntax with the goal of consistent representation of the semantics (rules defining field contents like 1/11/99)

## The Research Process and Functional Categories of Metadata

- **Discovery**
- **Retrieval**
- **Collation**
- **Analysis**
- **Re-presentation**

## Dublin Core (NISO Z39.85, 3/95)

- improve resource discovery
- anticipate precision problems of Web Crawler-based searching tools
- existing metadata could be “dumbed down”
- elements should be simple to understand and use, so that any individual should be able to assign terms him/herself
- software might eventually automatically generate very base-level metadata

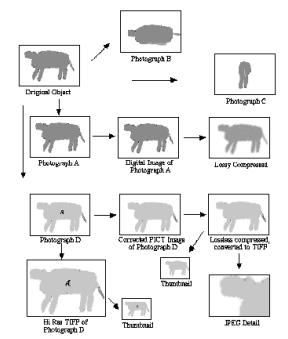
## Dublin Core

- |                |              |
|----------------|--------------|
| ✿ Title        | ✿ Format     |
| ✿ Creator      | ✿ Identifier |
| ✿ Subject      | ✿ Source     |
| ✿ Description  | ✿ Language   |
| ✿ Publisher    | ✿ Relation   |
| ✿ Contributors | ✿ Coverage   |
| ✿ Date         | ✿ Rights     |
| ✿ Type         |              |

## The number of variant forms of a work can be enormous

- ✿ different views of the same object
- ✿ different scans of the same photo
- ✿ different resolutions
- ✿ different compression schemes
- ✿ different compression ratios
- ✿ different file storage formats
- ✿ different details of the same image
- ✿ ...

### Image Families by Howard Besser



## Incorporate parts of Functional Requirements for Bibliographic Records (FRBR)

- work
  - expression
  - manifestation
  - item
- (and push into “change history” section of Technical Image Metadata)

## Various Other Metadata Standards

## NISO/DLF Technical Image Metadata Workshop--4/99

(Z39.87-2002 draft)

- ❖ create metadata needed to manage images in digital repositories over long periods of time (full life-cycle mgmt)
- ❖ document image provenance & history
- ❖ ensure that the images will be rendered accurately on any output device

Besser--SOIMA Metadata 18/11/09

25

<http://www.delos-nsf.actorswg.cdlib.org/>

DELOS/NSF Working Group

## Reference Models for Digital Libraries: Actors and Roles

**DELOS** Network of Excellence on Digital Libraries

Besser--SOIMA Metadata 18/11/09



## NSF/DELOS Actors/Roles Project

- **Classes of Actors, including**
  - Persons
  - Organizations
  - automata
- **Roles & implications**
  - Production
  - Dissemination
  - Management
  - use

Besser--SOIMA Metadata 18/11/09

27

## Multimedia & Collaborative Authorship imply

- **Not only:**
  - Authors
  - Editors
  - Publishers
- **But also creators of**
  - Text
  - Illustrations
  - Composers
  - Musicians...

Besser--SOIMA Metadata 18/11/09

28

## And goes beyond conventional authors

- **Others that are part of digital library process**
  - Users
  - Catalogers
  - Reference librarians
- **Even other groups/entities**
  - Software agents
  - Mediators
  - Special rights holders...

Besser--SOIMA Metadata 18/11/09

29

## Benefits for

- **Linking metadata to authority records**
- **Rights management**
- **Privacy protection**

Besser--SOIMA Metadata 18/11/09

30

## Data Structures: The VRA Core

- ✿ 28 elements specifically for visual resource collections
- ✿ Work Description Categories-
- ✿ Visual Document Description Categories-
- ✿ <http://www.oberlin.edu/~art/vra/dsc.html>

## VRA Core: Work Description Categories

- ✿ Work type
- ✿ Title
- ✿ Measurements
- ✿ Material
- ✿ Technique
- ✿ Creator
- ✿ Role
- ✿ Date
- ✿ Repository name
- ✿ Repository place
- Repository number
- Current site
- Original site
- Style/period/group /movement
- Nationality/culture
- Subject
- Related work
- Relationship type
- Notes

## VRA Core: Visual Document Description Categories

- ✿ Visual document type
- ✿ Visual document format
- ✿ Visual document measurements
- ✿ Visual document date
- ✿ Visual document owner
- ✿ Visual document owner number
- ✿ Visual document view description
- ✿ Visual document subject
- ✿ Visual document source

## Data Value Metadata (vocabularies)

- ✿ LCSH
- ✿ TGM
- ✿ AAT
- ✿ ULAN
- ✿ TGN
- ✿ VRA Core

## LCSH

- ✿ very general

## Thesaurus for Graphic Materials

- ✿ designed for subject indexing of pictorial materials, particularly large general collections of historical images
- ✿ for cataloging and retrieval
- ✿ good for general audiences and broad approaches to the material
- ✿ TGM-I: Subject Terms & TGM-II: Genre and Physical Characteristic Terms
- ✿ <http://lcweb.loc.gov/rr/print/tgm/toc.html>

## AAT

- ✿ 120,000 terms
- ✿ for describing objects, textual materials, images, architecture, and material culture from antiquity to present
- ✿ large and complex
- ✿ <http://www.getty.edu/gri/vocabularies/>

## ULAN

- ✿ name authority
- ✿ <http://www.getty.edu/gri/vocabularies/>

## Thesaurus of Geographic Names

- ✿ over 1 million records
- ✿ hierarchical and global
- ✿ throughout history
- ✿ most records include coordinates and descriptive notes

## Metadata for Digital Commerce

- ✿ DOI
- ✿ <indecs>

## <Indecs>

- ✿ formal structure for describing and uniquely identifying intellectual property itself, the people and businesses involved in its trading, and the agreements which they make about it (primarily for publishing, music, and visual arts)
- ✿ will develop high-level specifications for the services that will be required to implement a global IP trading system based on this <indecs> generic data model
- ✿ focus is on encoding rights at a high level, not on resource discovery
- ✿ likely to involve metadata schema registration and directory to allow interoperation of personal identifiers for rightsholders and users
- ✿ supported by EEC DG-13
- ✿ First meeting July 1999
- ✿ <http://www.indecs.org/>

## Cataloging, Metadata & Best Practices-

### Howard Besser, NYU Archiving & Preservation Program

- [http://www.firstmonday.dk/issues/issue7\\_6/besser/](http://www.firstmonday.dk/issues/issue7_6/besser/)
- Besser, Howard. Introduction to Imaging, Los Angeles: Getty Information Institute, 1995, 2003 ([http://www.getty.edu/research/conducting\\_research/standards/introimages/](http://www.getty.edu/research/conducting_research/standards/introimages/))
- Baca, Murtha (ed). Introduction to Metadata, Los Angeles: Getty Information Institute, 1998 ([http://www.getty.edu/research/conducting\\_research/standards/intrometadata/](http://www.getty.edu/research/conducting_research/standards/intrometadata/))
- <http://www.gseis.ucla.edu/~howard/Metadata/UC-May00/>
- <http://sunsite.berkeley.edu/Metadata/sp2000.html>
- <http://www.niso.org/commitau.html>
- <http://www.ifla.org/II/metadata.htm>
- METS official site: <http://www.loc.gov/standards/mets>
- [http://www.oclc.org/digitalpreservation/presmeta\\_wp.pdf](http://www.oclc.org/digitalpreservation/presmeta_wp.pdf)
- University of California Digital Image Collection Standards & Best Practices (<http://www.cdlib.org/inside/groups/stas/standards/>)

## Online Access

- **Internet Archive/Prelinger (high quality)**
  - <http://www.archive.org/movies>
- **Participant contributions (low quality)**
  - <http://www.youtube.com>

## Access--NARA/Amazon

washingtonpost.com

### Amazon to Copy and Sell Archives' Footage

First DVDs Already Available Under Non-Exclusive Deal

By Michael E. Ruane  
Washington Post Staff Writer  
Tuesday, July 31, 2007; C01

The National Archives and Records Administration announced yesterday that it has reached a non-exclusive agreement with [Amazon.com](http://www.amazon.com) and one of its subsidiaries to reproduce and sell to the public copies of thousands of historic films and videotapes in the Archives' holdings.

The arrangement allows Amazon and a [California](http://www.customfix.com) subsidiary, CustomFix Labs, to make digitized copies of some of history's most famous, and infamous, footage and make them available in DVD form for purchase via the Internet.

Archives and CustomFix officials stressed yesterday that the agreement is non-exclusive, as [Smithsonian Institution](http://www.smithsonian.org) struck recently with the cable television network Showtime.

That deal angered Congress, which funds the Smithsonian, and also angered filmmakers who documentarians using Smithsonian materials to offer their work first to Showtime.

"We would never go into an exclusive deal with anybody," National Archives spokeswoman Archives is to broaden, not restrict, access to its holdings, she said.

"The more you can make [the archives] accessible to the public the better for everyone it is,