

Metadata & Best practices for Media & Digital Works

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

Besser--Intro class 10/2/07

1

Metadata & Best practices for Media & Digital Works-

- ✧ **Importance of Metadata Standards & Philosophies**
 - ✧ Introduction
 - ✧ Discovery Metadata: The Dublin Core
 - ✧ Administrative & Structural Metadata; Digital Object Standards (METS)
 - ✧ Fitting METS in--Content Management
 - ✧ Content Format Standards (Images)
- ✧ **Other Elements**
 - ✧ Actors Metadata
 - ✧ Preserving Difficult Material

Besser--Intro class 10/2/07

2

Importance of Metadata Standards & Philosophies

Besser--Intro class 10/2/07

3

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

- **Catalog-MARC/AACR2** (http://library.nyu.edu/http://gondolin.rutgers.edu/MIC/text/how/unioncat_registry_xmlschema.htm)
- **Finding Aid** (<http://www.oac.edlib.org/>)
- **Spreadsheets or databases**
(http://www.imappreserve.org/cat_proj/sample_records.html)
- **Embedded within the essence file**

Besser--Intro class 10/2/07

4

Metadata vs. Essence

Besser--Intro class 10/2/07

5

Some Characteristics of the New Information Environment

- **Increased Quantity of Information**
 - With the Web, everyone can become a publisher
 - Varying level of quality
- **Digital Libraries Need to Work With New Classes of Information**
 - Web Pages, Museum Artifacts, GIS, Statistical Information, etc.

Besser--Intro class 10/2/07

6

Characteristics of the New Information Environment (Cont.)

- **Information is Decentralized**
 - Distributed repositories
- **Information is in Proprietary Formats**
 - Everyone has their own method of creating a digital book, journal, manuscript, Etc.

How Do We Cope????

Defining Digital Libraries in the NIE

- **A Series of Collaborating Services & Systems that Allow for the Discovery, Display, Maintenance and Preservation of Complex Digital Objects**

- The Traditional ILS
 - Created to manage physical materials
 - Almost all metadata is descriptive (e.g., MARC)
- Digital Libraries
 - Created to manage complex *digital objects*
 - New types of metadata (administrative, structural, etc.)
 - New Services (content management, digital preservation)

Complex Digital Objects

- **Scrapbook Example**
 - Digitized pages with text entries
 - Photos and newspaper clippings attached to the pages
 - Envelopes glued to the pages that hold
 - Letters & cards
 - DVDs
- **The Scrapbook has**
 - Multiple material types (text, image, audio)
 - Structure (e.g., like a table of contents)
 - Internal Relationships
 - The DVD on page 5 is linked to the file that is the DVD content and to its descriptive metadata

“A Series of Collaborating Services”

- **Content Management Systems (CMS)**
 - Create & maintain complex digital objects
- **Preservation Repositories**
 - Long-term retention of digital objects
- **Access Systems & Integration**
 - Global Access Portals
 - Subject Access Portals
 - Material Type Portals

How Can These Systems Collaborate?

- **Via “Standardized Digital Objects”**
 - A means to “wrap-up” a digital object and send it to another system or repository
 - Same idea as MARC, but for entire digital objects
 - E.g., A CMS sending a digital object to a Preservation Repository
- **The METS Digital Object Standard**
 - Metadata Encoding and Transmission Standard

For Interoperability, Repositories of both analog & digital material 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**

Besser--Intro class 10/2/07

13

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**

Besser--Intro class 10/2/07

14

The Research Process and Functional Categories of Metadata

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

These are different, and each carries its own standards & best practices

- **Data Fields (buckets)**
- **Data Values**

Besser--Intro class 10/2/07

16

Discovery Metadata

- **OPACs**
- **AACR2, MARC, AMIM**
- **Finding Aids**
- **Dublin Core - NISO Z39.85 (3/95)-**
- **CBIR (ongoing)**

Besser--Intro class 10/2/07

17

Dublin Core (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**

Besser--Intro class 10/2/07

18

Dublin Core

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

Dublin Core--further work

- **Warwick Framework**
 - metadata packages for extensible functions
 - layed groundwork for RDF
- **Canberra Qualifiers**
 - refining the semantics of the element set to provide more precise info
 - SUBELEMENT, SCHEME, LANG
- **Granularity**
 - no hierarchical relationships w/i a given DC record; only one record per discrete object (collection or item-level), and relationship field plus qualifier links them

PB Core 1.1

<http://www.pbcore.org>

- **PBCore Intellectual Content**
 - 9 containers; 16 elements
 - metadata elements describing the actual intellectual content of a media asset or resource
- **PBCoreIntellectualProperty**
 - 4 containers; 7 elements
 - metadata elements related to the creation, creators, usage, permissions, constraints, and use obligations associated with a media asset or resource
- **PBCoreInstantiation**
 - 1 container; 3 sub-containers; 28 elements
 - metadata elements that identify the nature of the media asset as it exists in some form or format in the physical world or digitally
- **PBCoreExtensions**
 - 1 container; 2 elements
 - additional descriptions that have been crafted by organizations outside of the PBCore Project

Metadata Mapping

- ✦ Crosswalks
- ✦ Resource Description Framework (RDF)
- ✦ Open Archives & metadata harvesting

Crosswalks

- ✦ mapping btwn differing metadata structures
- ✦ eliminate the need for monolithic, universally adopted standards
- ✦ focus on flexibility and interoperability
- ✦ RDF-based metadata registries

CDWA	Object	CIMI	EDA	VRA Core	SMARC	DUBLIN CORE
OBJECTWORK ORK (core)			Document Classification -Catalog Level (core)			
Object/Work-Type (core)	Type of Object	objectNAME	Document Classification - Document Type (core)	WT Work Type	655 Genre-Form	Type
Object/Work-Components		quantity	Document Classification -Extent		300a Physical Descriptions -Extent	
ORIENTATION/ARRANGEMENT/TITLES OR NAMES (core)	Title	objectTitle bibliographic Title	Group/Item Identification-Repository Title	W2: Title	24Xa Title and Title-Related Information	Title

MPEG to DC Crosswalk (1/2)



Besser-Intro class 10/2/07

25

MPEG to DC Crosswalk (2/2)



Besser-Intro class 10/2/07

26

Resource Description Framework

(RDF, spec released 2/99)

- W3C Metadata activity
- designed to move the Web beyond simple links to semantically-rich relationships btwn resources
- metadata application using XML as a common syntax for exchange and processing
- flexible architecture for managing diverse application-specific metadata packets that can be processed by machines
- associates resources, property types, and corresponding values
- <http://www.w3.org/RDF/>

Besser-Intro class 10/2/07

27

RDF

- **Resources (character strings, names, digital objects)**
- **Property (“is the author of”)**
- **Value**

- **resources+properties=relationships**
- **many different relationships can be reflected**

Besser-Intro class 10/2/07

28

XML-encoded RDF

- `<?xml:namespace ns=http://www.w3.org/RDF/RDF prefix="RDF" ?>`
- `<?xml:namespace ns=http://purl.oclc.org/DC/ prefix="DC" ?>`

- `<RDF:RDF>`
- `<DC:Creator>Howard Besser</DC:Creator>`
- `</RDF:Description>`
- `</RDF:RDF>`

Besser-Intro class 10/2/07

29

Philosophical Metadata Decisions-

- **Warwick vs MARC**
- **Where to put the metadata**

Besser-Intro class 10/2/07

30

Containers and Packages of Metadata

Warwick, not MARC

- **modular**
- **overlapping**
- **extensible**
- **community-based**
- **designed for a networked world to aid commonality btwn communities while still providing full functionality within each community**

Besser--Intro class 10/2/07

31

Some different schemes where Metadata is kept

- **embedded within the object (TIFF headers)**
- **encapsulated with image (MOA2/METS)**
- **in a separate related DB maintained by same organization (OPAC, Finding Aid)**
- **in a separate DB maintained by a separate organization (IMDB, AFI Catalog, ratings systems)**

Besser--Intro class 10/2/07

32

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)

Besser--Intro class 10/2/07

33

Open Archives & metadata harvesting

Besser--Intro class 10/2/07

34

Metadata Concepts (from MOA II/METS)

- **Administrative Metadata**
 - for enhancing resource management
- **Structural Metadata**
 - for reflecting internal hierarchies and relationships btwn parts
- **Raw/Seared/Cooked**

Besser--Intro class 10/2/07

35

Administrative Metadata

to uniquely identify a resource and manage it over time

- **Information about where the various pieces/versions of the object reside**
- **Information to view the object (equipment, encoding)**
- **Information about the capture process**

Besser--Intro class 10/2/07

36

Structural Metadata:

that which is relevant to presentation of the digital object to the user

- metadata defining the "object": a book, a diary, a photo album
- metadata defining the "sub-objects": pages (physical) or chapters and subheads (intellectual)

Standardized Digital Objects

METS

Metadata Encoding & Transfer Syntax

(example slides courtesy of Bernie Hurley, UCB Library Chief Scientist)

Structural & Administrative Metadata

- **Not enough to merely capture still images (book example)**
- **Must capture Behaviors**

What is a "Digital Object?"

- **Combined Digital Content & Metadata**
 - Digital Content
 - Digitized materials -- photographs, page images from a book, maps, digitized audio or video...
 - Born Digital -- GIS maps, digitally captured audio or video, numeric datasets (census files, scientific dataset), Web sites...
 - Metadata
 - Descriptive
 - Administrative
 - Structural
 - Behavior

What is METS?

- **An XML Schema that is used to Encode all the Content and Metadata for a Digital Object**
 - The relationships between content and metadata are also captured
- **METS Object -- METS Document**
- **A METS Document can be**
 - A single file with all content & metadata
 - A "hub document" that points to content and metadata
 - A combination of the above

Uses of METS

- **Transfer Syntax**
 - Standard for transmitting/ exchanging digital objects.
 - SIP (Open Archival Information Systems Reference Model)
- **Functional Syntax**
 - basis for providing end users with the ability to view and navigate digital content and its associated metadata
 - DIP
- **Archiving Syntax**
 - standard for archiving digital objects.
 - AIP

Why Is METS Important?

- **Interoperability**
 - Share objects between digital library systems
 - Allow a DL to work with objects from other repositories
- **Scalability**
 - Same software can be used to index, navigate and display different content types
 - E.g., book, diary, scrapbook, music score, etc.
- **Preservation**
 - Aids Migration Strategies

Besser--Intro class 10/2/07

43

History of METS

- **Originates in Making of America II Initiative**
 - Making of America II (MOA2) was a NEH funded Digital Library Federation initiative started in 1997. Participants included UC Berkeley (lead), Stanford, Penn State, Cornell, and NYPL.
 - **GOAL:** to create a digital object standard for encoding structural, descriptive and administrative metadata along with primary content
 - **RESULT:** MOA2.DTD (an XML DTD)
 - Adopted by UC Libraries

Besser--Intro class 10/2/07

44

History of METS (cont'd)

- **Concerned Parties Meet at NYU in February, 2001 to Discuss Future of MOA2**
 - **Additional needs emerge**
 - Support for time-based content
 - More flexibility in Descriptive and Administrative metadata
 - **Outcome**
 - MOA2 revised & renamed to METS
 - Outcome: mets.xsd is endorsed by DLF
- **METS Governance Structure**
 - **Editorial Board, Jerry McDonough is Chair**
 - RLG coordinates editorial board activities
 - Library of Congress is the Maintenance Agency for METS

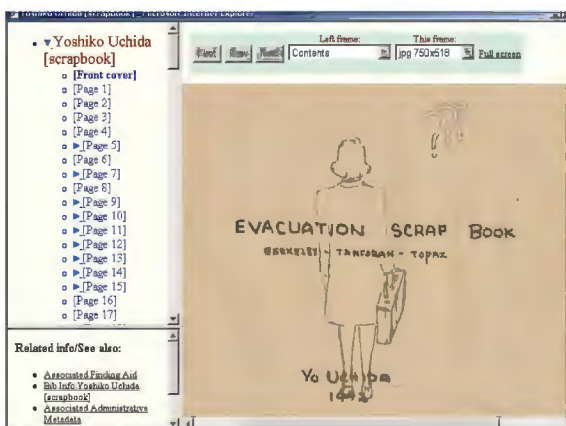
Besser--Intro class 10/2/07

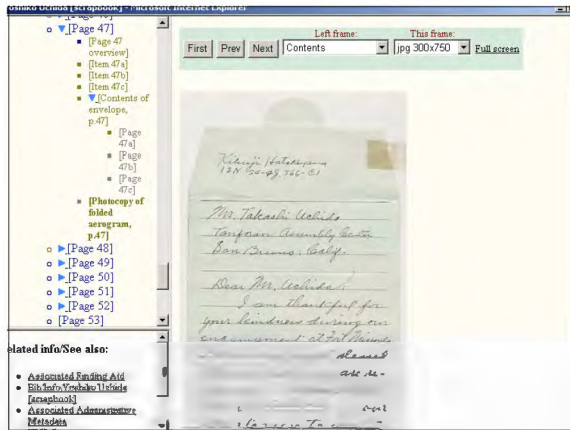
45

Display of METS Objects

Besser--Intro class 10/2/07

46





How Does METS Work?

- METS uses XML to
 - 1) Identify the digital pieces (files) that together comprise a digital object
 - Scrapbook: Digitized pages, photographs, newspaper clippings, digital audio, etc.
 - 2) Specify the location of these pieces
 - Are we pointing to these files?
 - Are they embedded in the METS document?
 - A combination of the above?

Metadata Object Description Schema (MODS)

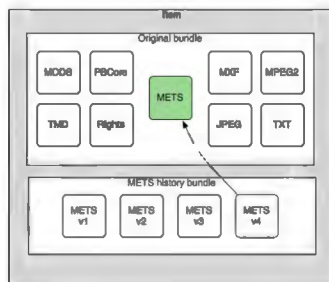
<http://www.loc.gov/standards/mods/>

- For encoding METS

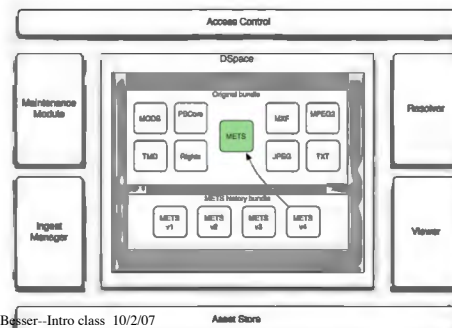
PTV Metadata and Content

- PBCore
- MXF wrapped 50Mbps MPEG2
- AS-PBS MXF wrapped 8Mbps broadcast
- Derivatives - iTunes, Real, YouTube, ...
- Transcripts, closed captions
- Promotional material
- ...

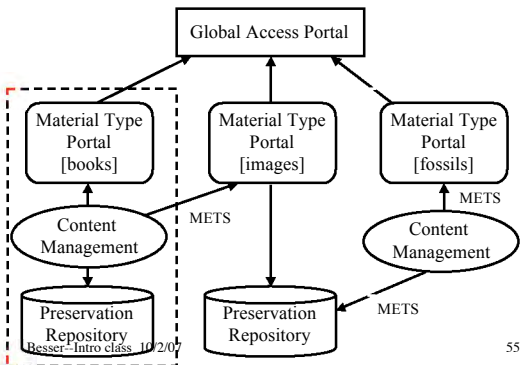
Item versioning



Components



Illustrative Digital Library Services Diagram



55

Content Management Systems

Besser-Intro class 10/2/07

56

Content Management Systems

- **Used to...**
 - Create and edit digital objects
 - Import & export digital objects
 - Manage objects (acquire, inventory, validate)
- **Content Management Systems will Vary Depending on the Materials they Support**
 - Metadata schemes will vary
 - Descriptive Metadata
 - MARC/MODS/Dublin Core for Books
 - Code books for numeric datasets
 - Administrative Metadata
 - Images, audio, test, etc.

Besser-Intro class 10/2/07

57

AES X98B Schema

- **Administrative & Structural Metadata for Audio Objects**
- **Specs for each file, not for each digital object**
- **Still under discussion within AES SC-03-07**

Besser-Intro class 10/2/07

58

OCLC/RLG Efforts

Working Group I: Preservation Metadata Framework

- ...to define the concept of preservation metadata, describe its importance in context of the overall digital preservation process, examine the "state-of-the-art" in the use of metadata in support of digital preservation, and evaluate the prospects for a community-wide, consensus-building activity in the area of preservation metadata (Preservation Metadata for Digital Objects: A Review of the State of the Art http://www.oclc.org/research/pmwg/presmeta_vp.pdf)
- ...to develop a framework outlining the types of information—i.e., metadata—that should be associated with an archived digital object. (A Metadata Framework to Support the Preservation of Digital Objects http://www.oclc.org/research/pmwg/pm_framework.pdf)
 - an expanded conceptual structure for the Open Archival Information System (OAIS) information model, and
 - a set of metadata elements, mapped to the conceptual structure and reflecting the information concepts and requirements articulated in the OAIS model.

Besser-Intro class 10/2/07

59

OCLC/RLG Efforts

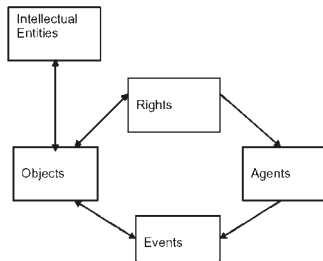
Working Group II: PREservation Implementation Strategies (PREMIS)

- develop a core set of implementable preservation metadata elements, with broad applicability within the digital preservation community
- develop a data dictionary to support the preservation metadata element set
- examine and evaluate alternative strategies for the encoding, storage, and management of preservation metadata within a digital preservation system, as well as for the exchange of preservation metadata between systems
- develop a pilot program for testing the group's recommendations and best practices in a variety of systems settings
- explore opportunities for the cooperative creation and sharing of preservation metadata

Besser-Intro class 10/2/07

60

OCLC/RLG Efforts
PREMIS Data Model



OCLC/RLG Efforts
PREMIS Data Dictionary Example

Semantic unit	fixity
Semantic components	messageDigestAlgorithm, messageDigest, messageDigestOriginator
Definition	Information used to verify whether an object has been altered in an unauthorized or unauthorized way
Data constraint	Container
Object category	Representation File Bitstream
Applicability	Not applicable (see message notes) Applicable Applicable (see message notes)
Repeatability	Repeatable Repeatable
Obligation	Optional Optional
Creation/ Maintenance notes	Automatically calculated and recorded by repository
Usage notes	For perform a fixity check, a message digest calculated at some earlier time is compared with a message digest calculated at a later time. If the digests are the same, the object was not altered in the interim. Recommended practice is to use two or more message digests calculated by different algorithms. The act of performing a fixity check and the date it occurred would be recorded as an Event. The result of the check would be recorded as the event outcome. Therefore, only the messageDigestAlgorithm and messageDigest need to be recorded as object characteristics for future comparison.

File Formats

- **AVI** **Microsoft format**
- **MOV** **Quicktime format**
- **MPEG** **compressed standardized format**
- **RM** **Real Media compressed format**
- **ASF** **Microsoft streaming format**
- **DV25** **25 Mbits/sec, fixed compression, 4:1:1 color encoding, 3.6 MB per sec, 60 min. mini-DV cassette has 13 GB of storage**
- **DV50** **50 Mbits/sec, 4:2:2 or 3:3:1 color encoding**
- **DV100** **100 Mbits/sec, to be used for HDTV production**

From Steve Puglia from Adobe's "A Digital Video Primer" and <http://www.csl.cornell.edu/atc/materials/dig/videoformats.shtml>

MPEG

- **MPEG 2** -- currently most commonly used for distribution, but difficult to use for production
- **MPEG 4** -- used with QuickTime and Windows Media files
- **MPEG 7 & 21** -- metadata standards, not used for compression

Ideal digital moving image file format

(Jerome McDonough)

- **Non-proprietary file format**
- **supports 10-bit/pixel**
- **no compression or lossless compression using non-proprietary CODEC**
- **supports multiple frame rates/frame sizes**
- **supports time code data in file**
- **supports audio (multichannel) and video in single file**

Limitations of present file formats

- **MPEG seems to be only non-proprietary format**
- **AVI and Quicktime with extensions incorporate most features, but are proprietary**

Structural Metadata Standards for Encoding Multimedia-

(no time for details)

- **SMIL**
- **MPEG 4, 7, 20**

Wrappers for time-based formats

- **METS**
- **AAF**
- **MXF**

Content Format Standards (Images)

Images-

- **Content Format & Best Practices**
- **Identification/Provenance**
- **Technical Imaging metadata**
- **Special discovery & descriptive metadata**

Best practices

- **Use/Users/Collection:**
- **Benchmarking**
- **Masters vs. Derivatives**
- **Scanning-**
- **Administrative Metadata-**
- **Structural Metadata-**

Scanning Best Practices

- Think about users (and potential users), uses, and type of material/collection
- Scan at the highest quality that does not exceed the likely potential users/uses/material
- Do not let today's delivery limitations influence your scanning file sizes; understand the difference between digital masters and derivative files used for delivery
- Many documents which appear to be bitonal actually are better represented with greyscale scans
- Include color bar and ruler in the scan
- Use objective measurements to determine scanner settings (do NOT attempt to make the image good on your particular monitor or use image processing to color correct)
- Don't use lossy compression
- Store in a common (standardized) file format
- Capture as much metadata as is reasonably possible (including metadata about the scanning process itself)

Besser--Intro class 10/2/07

73

Why Scale is important



Besser--Intro class 10/2/07

74

Identification/Provenance (Images)-

- ✧ The number of variant forms of a work can be enormous
- ✧ Image Families
- ✧ A digital image frequently has many layers of parentage
- ✧ Information about the parentage that can indicate the quality and veracity of the image (Dublin Core "Source" and "Relation")
- ✧ how to deal with different versions derived from the same scan or different encoding schemes
- ✧ Vocabulary Standards to express this

Besser--Intro class 10/2/07

75

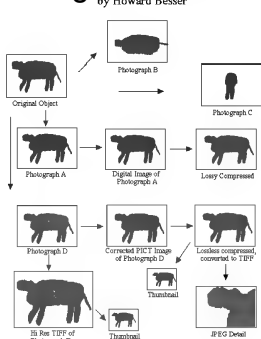
The number of variant forms of a work can be enormous

- ✧ 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
- ✧ ...

Besser--Intro class 10/2/07

76

Image Families



Identification/Provenance

- ✧ how to deal with different versions (browse, hi-res, medium res) derived from the same scan or different encoding schemes (TIFF, PICT, JFIF)
- ✧ Vocabulary Standards to express this
 - VRA Surrogate Categories
 - CIMI's "Image Elements"

Besser--Intro class 10/2/07

78

Incorporate parts of Functional Requirements for Bibliographic Records (FRBR)

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

Besser--Intro class 10/2/07

79

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--Intro class 10/2/07

80

Technical Image Metadata

Focus on Metadata that may prove helpful for

- ✿ management
- ✿ use
- ✿ preservation
- ✿ ...

Besser--Intro class 10/2/07

81

Technical Image Metadata

In Scope

- ✿ still, bit-mapped pictorial images
- ✿ scanned/reformatted images (+ born digital)

Besser--Intro class 10/2/07

82

Technical Image Metadata

Out of Scope

- ✿ vector images
- ✿ moving images
- ✿ images of OCR-able text
- ✿ structural and hierarchical relationships between images
- ✿ rights management, terms of use
- ✿ (authenticity/security)

Besser--Intro class 10/2/07

83

Technical Image Metadata

Technical Image Metadata^{Z39.87}

- ✿ Image parameters (MIME type, compression, colorspace & profile, ...)
- ✿ Image Creation (source, capture info, etc.)
- ✿ Image performance assessment (sampling, colormap, whitepoint, target data, etc.)
- ✿ Change history (source, processing, etc.)

Besser--Intro class 10/2/07

84

Technical Image Metadata
Technical Image Metadata^{Z39.87}

✿ **additional XML implementation schema (MIX)**

Besser--Intro class 10/2/07 85

Other Metadata

- **Description of depiction/surrogate (What VRA calls its "Surrogate Categories")**
- **Description of original object**

- **Rights and Reproduction Information**
- **Location Information**
- **VRA Core, LCSH, TGM, AAT, ULAN, TGN, DOI, <indec>, ...**

Besser--Intro class 10/2/07 86

Other Elements-

- **Actors Metadata**
- **Other Metadata**
- **Preserving Electronic Art**

Besser--Intro class 10/2/07 87


<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--Intro class 10/2/07



88

NSF/DELOS Actors/Roles Project

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

Besser--Intro class 10/2/07 89

**Multimedia & Collaborative Authorship
imply**

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

Besser--Intro class 10/2/07 90

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--Intro class 10/2/07

91

Benefits for

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

Besser--Intro class 10/2/07

92

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>**

Besser--Intro class 10/2/07

93

VRA Core: Work Description Categories

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

Besser--Intro class 10/2/07

94

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**

Besser--Intro class 10/2/07

95

Data Value Metadata (vocabularies)

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

Besser--Intro class 10/2/07

96

LCSH

- ✿ **very general**

Besser--Intro class 10/2/07

97

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>**

Besser--Intro class 10/2/07

98

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/>**

Besser--Intro class 10/2/07

99

ULAN

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

Besser--Intro class 10/2/07

100

Thesaurus of Geographic Names

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

Besser--Intro class 10/2/07

101

Metadata for Digital Commerce

- ✿ **DOI**
- ✿ **<indecs>**

Besser--Intro class 10/2/07

102

<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/>

Besser--Intro class 10/2/07

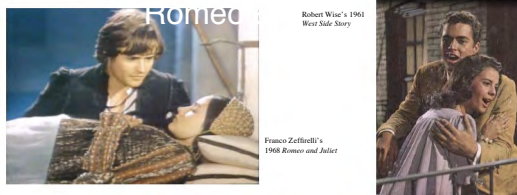
103

Functional Requirements for Bibliographic Records (FRBR)

- work
- expression
- manifestation
- item

Besser--Intro class 10/2/07

104



Besser--Intro class 10/2/07

105

Digital Repository Traditions & Services require

- ✦ Sustainability
- ✦ Interoperability
- ✦ Access

- ✦ And all of these require Standards and Metadata

Besser--Intro class 10/2/07

106

Metadata & Best practices for Media & Digital Works

Howard Besser, NYU Archiving & Preservation Program

- <http://www.pbcore.org>
- METS official site: <http://www.loc.gov/standards/mets>
- <http://www.loc.gov/standards/mods/>
- <http://www.niso.org/commitau.html>
- <http://www.ifla.org/II/metadata.htm>
- http://www.oclc.org/digitalpreservation/presmeta_wp.pdf
- Besser, Howard. Introduction to Imaging, Los Angeles: Getty Information Institute, 1995, 2003 (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.gseis.ucla.edu/~howard/Metadata/UC-May00/>
- <http://sunsite.berkeley.edu/Metadata/sp2000.html>
- University of California Digital Image Collection Standards & Best Practices (<http://www.cdlib.org/inside/groups/stas/standards/>)