

Author(s): Paul Conway, Ph.D., 2010

**License:** Unless otherwise noted, this material is made available under the terms of the **Creative Commons Attribution–Noncommercial–Share Alike 3.0 License**: http://creativecommons.org/licenses/by-nc-sa/3.0/

We have reviewed this material in accordance with U.S. Copyright Law and have tried to maximize your ability to use, share, and adapt it. The citation key on the following slide provides information about how you may share and adapt this material.

Copyright holders of content included in this material should contact **open.michigan@umich.edu** with any questions, corrections, or clarification regarding the use of content.

For more information about **how to cite** these materials visit http://open.umich.edu/education/about/terms-of-use.

Any **medical information** in this material is intended to inform and educate and is **not a tool for self-diagnosis** or a replacement for medical evaluation, advice, diagnosis or treatment by a healthcare professional. Please speak to your physician if you have questions about your medical condition.

Viewer discretion is advised: Some medical content is graphic and may not be suitable for all viewers.





#### **Citation Key**

for more information see: http://open.umich.edu/wiki/CitationPolicy

Use + Share + Adapt	
{ Content the copyright holder, author, or law permits you to use, share and adapt. }	
PD-GOV	Public Domain – Government: Works that are produced by the U.S. Government. (17 USC § 105)
© PD-EXP	Public Domain – Expired: Works that are no longer protected due to an expired copyright term.
PD-SELF	Public Domain – Self Dedicated: Works that a copyright holder has dedicated to the public domain.
(cc) ZERO	Creative Commons – Zero Waiver
(cc) BY	Creative Commons – Attribution License
CC) BY-SA	Creative Commons – Attribution Share Alike License
CC BY-NC	Creative Commons – Attribution Noncommercial License
CC) BY-NC-SA	Creative Commons – Attribution Noncommercial Share Alike License
③ GNU-FDL	GNU – Free Documentation License

#### Make Your Own Assessment

{ Content Open.Michigan believes can be used, shared, and adapted because it is ineligible for copyright. }

PD-INEL Public Domain – Ineligible: Works that are ineligible for copyright protection in the U.S. (17 USC § 102(b)) \*laws in your jurisdiction may differ

{ Content Open.Michigan has used under a Fair Use determination. }



**Fair Use**: Use of works that is determined to be Fair consistent with the U.S. Copyright Act. (17 USC § 107) \*laws in your jurisdiction may differ

Our determination **DOES NOT** mean that all uses of this 3rd-party content are Fair Uses and we **DO NOT** guarantee that your use of the content is Fair.

To use this content you should do your own independent analysis to determine whether or not your use will be Fair.

### SI 640 DIGITAL LIBRARIES AND ARCHIVES

2010 Week 2: Definitions – The Archival Dimension

#### THEMES FOR THIS WEEK

Resources on digital libraries
METS at OAC
Archival integrity
Other archival constructs
Research agenda example

"The challenge ahead is to bring our best technical skills to bear on the problem of digital preservation without losing sight of the ultimate human purposes these efforts serve, purposes which cannot be found inside the machines we are busy programming or using."

Levy, "Heroic Measures," 1998.

### **ESOURCES FOR FURTHER STUDY**

- 1. Resources
- 2. METS in OAC
- 3. Integrity
- 4. Concepts
- 5. Research

- Journals
  - See list in CTools
- Organizations
  - DLF <u>http://www.diglib.org/</u>
    - JISC <u>http://www.jisc.ac.uk/</u>
  - DELOS <u>http://www.delos.info/</u>
- Portals

- ACM <u>http://www.acm.org/</u>
- PADI <u>http://www.nla.gov.au/padi/</u>
- Conferences
  - JCDL <u>http://www.jcdl.org/</u>
  - ECDL <u>http://www.ecdl2007.org/index.php</u>
  - ICADL <u>http://www.icadl.org/</u>

### **FINDING METS IN OAC**

- 1. Resources
- 2. METS in OAC
- 3. Integrity
- 4. Concepts
- 5. Research

• Search OAC

- Find Item and record URL
  - <u>http://www.oac.cdlib.org/ark:/</u> <u>13030/hb4b69p0jz/?brand=oac4</u>
- Record descriptive metadata
- Add "mets" to string, per example below
  - <u>http://www.oac.cdlib.org/mets/ark:/</u> <u>13030/hb4b69p0jz/?brand=oac4</u>



7

### PART OF A METS RECORD

- <mets schemaLocation="http://www.loc.gov/METS/ http://www.loc.gov/standards/mets/mets.xsd http://cosimo.stanford.edu/sdr/n</p> http://www.loc.gov/standards/mods/v3/mods-3-1.xsd http://www.loc.gov/mix/ http://www.loc.gov/standards/mix/mix.xsd" OBJID=' PROFILE="http://www.loc.gov/standards/mets/profiles/0000002.xml"> -<metsHdr CREATEDATE="2006-04-06T14:25:40" LASTMODDATE="2006-04-06T14:25:40"> -<agent ROLE="CREATOR" TYPE="ORGANIZATION"> -<name> Library Systems Office, University of California, Berkeley </name> </agent> -<altRecordID> http://digitalassets.lib.berkeley.edu/eqf/chs/mets/chs 24 2 00000847.xml </altRecordID> </metsHdr> -<dmdSec ID="DMR1"> <mdRef xlink:href="http://oac.cdlib.org/findaid/ark:/13030/hb8779p2cx" XPTR="xpointer(id('c2000847'))" LOCTYPE="U Collection"/> </dmdSec> -<dmdSec ID="DM1"> -<mdWrap MDTYPE="MODS" LABEL="3. City Hall."> -<vmlData> -<mods:mods> -<mods:titleInfo> <mods:title>3. City Hall.</mods:title> </mods:titleInfo> -<mods:name type="personal"> <mods:namePart>Aydelotte, C. L.</mods:namePart> </mods:name> <mods:typeOfResource>still image</mods:typeOfResource> © FAIR USE Source Undetermined

#### NTEGRITY: WHAT IS A RECORD?

- 1. Resources
- 2. METS in OAC
- 3. Integrity
- 4. Concepts
- 5. Research

- Angelika Menne-Haritz: "what can be read in the texts [of records] is called information... What can be read between the lines, in signs, symbols, or even in the composition of texts ... is evidence."
- ICA Committee on Electronic Records, 1997. (Yeo, pp. 325 and 330)
  - Records comprise "content, context, and structure sufficient to provide evidence" of the activities in which the were produced.
  - Provenance, physical and intellectual form, and incorporation in record-keeping system.

Yeo, Geoffrey, Concepts of Record (2): Prototypes and Boundary Objects. *American Archivist* 71 (Spring-Summer 2008): 118-143.

### NTEGRITY: WHAT IS A RECORD?

- 1. Resources
- 2. METS in OAC
- 3. Integrity
- 4. Concepts
- 5. Research

• Records can be documents or byproducts; activities or probative.

- Yeo: Records are "persistent representations of activities..." p. 337
- "... or other occurrents..."\*\*
- "... created by participants or observers of those occurrents or by their authorized proxies..."
- "...or sets of such representations representing particular occurrents."

p. 136

\*\* "... a collective term for concepts such as function, process, activity, transaction, and event." [Yeo] Often contrasted with "continuant," which is "an existent, or something physical or psychical manifested in space and time, that retains its identity though changing its states or relations or when regarded as having different states or relations." [OED]

Duranti, Reliability and Authenticity, 1995.

#### **INTEGRITY: RELIABILITY**

- 1. Resources
- 2. METS in OAC
- 3. Integrity
- 4. Concepts
- 5. Research

• "... a record is considered reliable when it can be treated as a fact in itself, that is, as the entity of which it is evidence." (p. 6)

- Form (complete) in relation to needs of the socio-juridical system
- Process of creation (body of rules)
- "... degree of completeness and degree of control of the procedure of creation are the only two factors that determine reliability of records."

Luciana Duranti, "Reliability and Authenticity: The Concepts and Their Implications," *Archivaria* 39 (Spring 1995), pp. 5-10.

Duranti, Reliability and Authenticity, 1995.

#### **INTEGRITY: ORIGINAL VERSUS COPY**

- 1. Resources
- 2. METS in OAC
- 3. Integrity
- 4. Concepts
- 5. Research

• By nature original record is more reliable than a copy, because...first complete document with consequence

- primitiveness
- completeness
- effectiveness

• Copies of original records can fully meet criteria for reliability and authenticity.

Duranti, Reliability and Authenticity, 1995.

#### **INTEGRITY: AUTHENTICITY**

- 1. Resources
- 2. METS in OAC
- 3. Integrity
- 4. Concepts
- 5. Research

- "...a record is authentic when it is the document that it claims to be."
- "Authentic, unreliable records are of no use to present and future users." (p. 8)
- "...electronic records might be authentic (because of a fairly high attention to security and transmission processes) but they are certainly not reliable."
- ...we should be more concerned with the creation of reliable records....

• O'Toole, Uniqueness, 1993.

#### IDEA OF UNIQUENESS IN ARCHIVES

- 1. Resources
- 2. METS in OAC
- 3. Integrity
- 4. Concepts
- 5. Research

• Records (artifactual)

- Information (duplication/publication)
- Processes (functional view)
- Aggregations (distinct assemblages)

James M. O'Toole, "On the Idea of Uniqueness," American Archivist 57 (Fall1994), pp.632-658.

# PRESERVING DIGITAL NFORMATION

1. Resources

- 2. METS in OAC
- 3. Integrity
- 4. Concepts
- 5. Research



Don Waters, Mellon Foundation

• The Task Force defines digital archives strictly in functional terms as repositories of digital information that are collectively responsible for ensuring, through the exercise of various migration strategies, the integrity and longterm accessibility of the nation's social, economic, cultural and intellectual heritage instantiated in digital form."

# PRESERVING DIGITAL NFORMATION

- 1. Resources
- 2. METS in OAC
- 3. Integrity
- 4. Concepts
- 5. Research



 Don Waters, Mellon Foundation

 PD-GOV

 Library of Congress

• Goal: preserve information integrity: "that is, to define and preserve those features of an information object that distinguish it as a whole and singular work."

- Content (bits, structure/format, ideas)
- Fixity (no change or record of change)
- Reference (stable location)
- Provenance (origin and chain of custody)
- Context (technical, linkages, communication, social)

### A MEASURED CRITIQUE

- 1. Resources
- 2. METS in OAC
- 3. Integrity
- 4. Concepts
- 5. Research



David Levy, U Washington

- Document (talking things) as point of departure.
- The goal must be to preserve its essential features, where what is essential can only be determined with respect to use. (p. 156)
- *all* documents, regardless of medium, are fixed and fluid. All documents hold their form and or content fixed for periods of time and undergo change (fluidity) at others.
- Critique: internalist, technical dominance, users last and fuzzy

### ARCHIVES BY NATURE"

- 1. Resources
- 2. METS in OAC
- 3. Integrity
- 4. Concepts
- 5. Research



image removed

- "But if we think more carefully about digital libraries we easily observe that they may be libraries by name, but they are archives by nature." (p. 8).
  - Hold "unique" exemplars
  - Trust and verification of authenticity
  - Knowledge of context of creation
  - Evidence of provenance

17

Seamus Ross, U. Toronto Please see original image of Seamus Ross at http:// www.research.utoronto.ca/edge/december2008/1.html

# RELEVANCE OF ARCHIVAL PRINCIPLES

- Evidence (think records)
- Provenance
- Life Cycle of Records
- Organic Nature of Records

- Integrity
- Metadata
- Knowledge Management
- Risk Management
- Knowledge Preservation



image removed

Anne Gilliland, UCLA Please see original image of Anne Gilliland at http:// gseis.ucla.edu/about/people/gilliland

#### RESEARCH AGENDA

- 1. Resources
- 2. METS in OAC
- 3. Integrity
- 4. Concepts
- 5. Research

• Technical architectures for archival repositories

- Attributes of archival collections
- Digital archiving tools and technologies
- Organizational, economic, and policy issues

Reality Check: significant progress can be seen in all four areas over the past 7 years.

# Thank you!

#### **Paul Conway**

Associate Professor School of Information University of Michigan www.si.umich.edu

#### **Additional Source Information**

for more information see: http://open.umich.edu/wiki/CitationPolicy

- Slide 6: Online Archive of California, http://www.oac.cdlib.org/
- Slide 7: Source Undetermined
- Slide 14: Library of Congress, http://www.loc.gov/preservation/outreach/tops/waters\_don/index.html
- Slide 15: Library of Congress, http://www.loc.gov/preservation/outreach/tops/waters\_don/index.html
- Slide 16: JenWaller, Flickr.com, http://www.flickr.com/photos/jenwaller/2648545130/
- Slide 17: Please see original image of Seamus Ross at http://www.research.utoronto.ca/edge/december2008/1.html
- Slide 18: Please see original image of Anne Gilliland at http://gseis.ucla.edu/about/people/gilliland