

Digital Records Curation Programme

Week 4: Metadata

Week 3 Recap

What did you learn?

- Class on Digital and Hybrid Records Management
- Workshop on File Profiling

Learning Outcomes:

At the end of this class you will be able to:

- explain what metadata is and why it is important for record-keeping
- understand the different types of metadata
- know the main metadata standards used in archiving, records management and digital preservation;
- crosswalk metadata element sets

What is Metadata?

Metadata

- Metadata is data about data.
- All digital objects have some intrinsic metadata.

Can you give some examples of metadata?

Metadata Elements

- Each metadata 'element' is a field such as 'creator' or 'date'. The metadata that populates this field is called a 'metadata value'.
- Groups of metadata elements are called 'metadata schemas' or 'metadata element sets'.
- A 'metadata record' is a set of metadata (elements with values) for one object or resource.

Why is Metadata Important?

Metadata is an Important Aspect of Digital Curation

- Records must be capable of being related to each other
- Records must be reliable, authentic and accessible in order to have integrity
- Records comprise content (what it says), context (where comes from, why created, when, by whom) and structure (format, components)
- Content, context and structure must be assembled and maintained to enable them to serve their purpose, we need to preserve all 3 over time or the record loses integrity

Metadata is the glue that binds records together and to each other

Types of Metadata

- *Descriptive metadata* – data elements or properties that identify a digital resource and are used to find and interpret the resource.
- *Structural metadata* – data elements or properties that document the structural relationships between or within digital resources, for example the file structure within which a digital resource resides, or the linkage between pages in a website. Structural metadata supports proper display and use of complex objects.
- *Administrative metadata* – data elements or properties used to manage the resource. Administrative metadata may include:
 - *Technical metadata* – data elements or properties that provide information about the technical context of the resource
 - *Rights metadata* – data elements or properties describing rights and obligations adhering to the digital resource such as ownership, copyright or other intellectual property rights, usage and security restrictions
 - *Preservation metadata* – data elements or properties describing requirements for preserving the resource over time and across technological change.

Where is the metadata?

- Embedded within a digital file
- Within a database (digital asset management system)
- In a separate XML-encoded file
- In an associated analog record
- A combination of these methods

Building Blocks Exercise

In groups:

1. Use the building blocks to construct a duck
2. Now use the blank index card to describe your duck.
3. Use the random index card you have been given to identify the duck it describes

Metadata Standards

Standards for Descriptive Metadata

Some examples:

- Dublin Core (ISO) 15836:2009
<https://www.dublincore.org/specifications/dublin-core/dcmi-terms/>
- MARC (a library standard for bibliographic information)
- MODS (like MARC, but simpler)
- ISAD(G)
- EAD Tag Library (<https://www.loc.gov/ead/EAD3taglib/>)
- Many of the functional requirements standards include metadata schemas (NoARK, MoReq2010, DoD 5015.02-STD)

A Standard for Preservation Metadata

PREMIS (Preservation Metadata: Implementation Strategies)

- Data Dictionary (element set) and XML schema
- PREMIS Data Model:

[INSERT IMAGE OF PREMIS DATA MODEL]

PREMIS Data Dictionary

[INSERT IMAGE OF ENTRY IN THE DATA DICTIONARY]

PREMIS

Preservation metadata is usually created automatically, rather than manually.

Library of Congress points to PREMIS implementation tools:

- Tools for extracting technical metadata from objects
- Tools for converting extracted metadata into the PREMIS XML schema elements
- Tools for generating a METS object with appropriate slots for PREMIS metadata
- Tools for converting JHOVE output to PREMIS elements
- Tools for recording events and outcomes (e.g. format validation, fixity check, etc.)

Type Specific Metadata Standards

- ANSI/NISO Z39.87 – 2006 (for still images)
- MXF (for audiovisual)
- DDI (Data Documentation Initiative) (for datasets)
- SPECTRUM (for museum documentation)
- CDWA (for works of art)

Interoperability

- Semantics (defined meaning) and syntax (content rules)
- Crosswalking (metadata or field mapping) metadata
- A one way process
- Absolute and relative crosswalking

Crosswalking Exercise

Using Excel you are going to crosswalk the Dublin Core (DCMI) elements to the EAD Tag Library elements:

1. Copy the elements from Section 3: Properties in the /elements/1.1/ namespace (found here: <https://www.dublincore.org/specifications/dublin-core/dcmi-terms/#section-3>) into an excel spreadsheet.
2. Add in the appropriate EAD tags for each element from here: <https://www.loc.gov/ead/EAD3taglib/>

Encoding, Containers and Enrichment

Encoding Metadata

- You can record metadata in a CSV or Excel file if needed, using tools like Google Sheets or MS Office.
- Metadata can also be encoded using markup languages such as XML (eXtensible Markup Language).
- RDF - Resource Description Framework - uses data to form graphs (subject, property, object – a triple) using Uniform Resource Identifiers. Foundational for the semantic web.

Container Formats and Managing Metadata

- Metadata can be embedded (.txt, .xls, .xml, .zip) or external (relational databases)
- When storing or transferring digital objects, it is important to maintain the relationships between the object and its metadata
 - Co-locating objects and their metadata (containers)
 - Using persistent identifiers (ie ISBNs, DOIs,) to maintain a link
- Metadata container format
 - METS (Metadata Encoding and Transmission Standard)
 - expressed in XML documents that include names and locations of files comprising the digital object, their hierarchical structure (structural map) and metadata
 - can be used as a SIP, AIP or DIP

Getting more metadata

- Metadata harvesting
 - Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH)
- Metadata enrichment

Implementation Advice

Application Profiles

- Application profiles are documents or sets of documents that provide:
 - A metadata element set (direct from a standard or partial or modified or developed locally)
 - Rules for applying the element set in a specific context
- DCMI identifies four components:
 - objectives and scope of the application (target group and context of use)
 - functional requirements of the application (what do users need to do? “find”, “select” etc.)
 - data model of the entities described by the application (entities and relationships)
 - a description set profile detailing the classes and properties used in an application, together with constraints on their usage (the element set and its rules)

Application Profiles

When implementing metadata in systems you do not implement one standard in isolation but multiple standards (like EAD + PREMIS + METS), and:

you have the possibility to define your own specific rules on top of the standards (like rules for titles of records, date formats, lists of roles/actions, etc...)

Note that using an international standard does not mean you have to ONLY use this standard – add your own metadata if needed

Descriptive metadata is your main concern because someone usually inputs it manually somewhere. There are tools for creating PREMIS, METS, technical metadata (such as Jhove, DROID, Premis tools, etc).

Groups for Metadata Seminar

- Before the seminar, work in groups to prepare presentations on:
 - Group 1 - ANSI/NISO Z39.87 – 2006 (for still images)
 - Group 2 - MXF (for audiovisual)
 - Group 3 - DDI (Data Documentation Initiative) (for datasets)
 - Group 4 - SPECTRUM (for museum documentation)
 - Group 5 - CDWA (for works of art)

Conclusion

- Metadata is crucial for managing and preserving digital records
- There are many metadata schemas and internationally recognised standards that help to ensure metadata is created, managed and can be shared and transferred as necessary

Any questions?



“Digital records Curation Programme” copyright International Council on Archives, 2021, is licensed under Creative Commons License Attribution-Noncommercial 4.0.