

Metadata: Describing Digital Objects

Jay L. Colbert

April 1, 2022

Who am I?

Jay L. Colbert, jay.colbert@unh.edu
Metadata & Discovery Strategy Librarian @ UNH
Metadata Librarian @ NHDL

What is metadata?

“data about data”

Boring, vague, not helpful!

Instead...

Perhaps a more useful, “big picture” way of thinking about metadata is as the sum total of what one can say at a given moment about any *information object* at any level of aggregation.

- Metadata is what we can say about an object
- Metadata schemas and standards tell us how to do it

Why is it important?

- Discovery
- Access
- Preservation
- Administration

How is it used?

The Good

- Describing library collections
- Files on your computer
- Your driver's license

The Bad

- NSA collecting metadata about phone calls
- Police using location metadata (from photos or other sources) to arrest protestors and activists
- Companies harvesting information about our internet habits to sell to advertisers

NHDL & Omeka

Required & strongly recommended elements

<https://nhs1.libguides.com/digitization>

However, some of the metadata recommendations for NHDL/DPLA differ from what is listed here.

NHDL/DPLA is still in its infancy, and documentation should be finished by the end of the month.

If we have time at the end, I'll get into some of the fields just for DPLA. If you partner with us, I will work with you individually to create a plan.

Required

- Identifier
- Language (if available/applicable)
- Title
- Type
- Data Provider*

- Is Shown At*
- Rights

Asterisks are just for DPLA.

Strongly recommended

- Collection*
- Creator
- Date
- Description
- Format
- Place
- Publisher*
- Subject

Asterisks are just for DPLA/NHDL

Interpretation of elements

Required

These elements are required. Language is the exception. It is only required when available/applicable.

- Identifier
 - An unambiguous reference to the resource
 - Might be generated by whatever software you're using
 - State library recommends that each library use their HSA code as a way of identifying objects from their library and avoiding duplication across the state. For instance, NH State Library's HSA code is NHSL, therefore, each digital object would have an identifier beginning with NHSLxxxx, with the x's being a number.
- Language

- Language(s) of described resource. Strongly recommended for text materials.
- Use controlled vocabularies or other established standards, such as ISO 639-3
- eng (for English)
- Title
 - Primary name given to the described resource
 - Transcribe the title exactly how it is presented on the resource
 - If the resource has no title, try to avoid “Untitled”
 - * However, titles do not have to be unique (that’s why we have identifiers!)
 - * and you should avoid interpreting the content to create a title
 - You should also avoid redundancy when possible
- Type
 - What the *original* resource is, *not* the file format.
 - Assign the resource a DCMI type, based on the original resource’s carrier.
 - Scanned items are not automatically images. If the scanned image is a book, its type is Text, not StillImage
- Rights
 - Information about rights held in and over the described resource.
 - Typically, rights information includes a statement about various property rights associated with the described resource, including intellectual property rights.
 - For inclusion in DPLA/NHDL, you will need to use a URI from <https://rightsstatements.org/>
 - However, you may include additional rights/access information in an additional rights field after the element containing the URI.
 - * See: <https://bit.ly/dpla-rights-guidelines>

Recommended

These are not required.

However, they enrich your record and are strongly recommended.

- Creator
 - Entity primarily responsible for making the described resource. May be a person(s), corporate body, or family.
 - If possible, place multiple names in repeated instances of the element; otherwise, separate consistently (e.g. with a semicolon or pipe)
 - Use controlled vocabularies when possible. If a controlled form does not exist, create one following the syntax of an existing vocabulary
 - *Editors, translators, and illustrators go in the Contributors field.*
- Date
 - Date of creation of the original resource
 - Prefer use of YYYY-MM-DD format
 - For uncertain dates, use “circa” or “?”
 - For approximate dates, use “~”
 - Avoid placeholder values like “Unknown” or “n.d.”
 - See: <https://bit.ly/dpla-geo-styleguide>
- Description
 - A free-text account of the resource that describes what the item is about.
 - Could include a summary, an abstract, a table of contents, etc.
 - Description should apply to the object being described, not to a collection to which it belongs
- Format
 - This is the *file format* of the digital manifestation of the item
 - You should use the Internet Media Types (MIME) for the appropriate terms
 - FYI: PDF is under the “application” registry

* application/pdf

- Place

- Spatial characteristics of described resource, such as a country, city, region, address or other geographical term.
- *Covers aboutness. This is not where an item was published!*
- If the name of the place has changed over time, list the one used at the time/by the resource first.
- If the place is on colonized or otherwise stolen, unceded land, list the indigenous territory name(s) first. The Indigenous New Hampshire Collaborative Collective has good resources for finding the appropriate place name: <https://indigenousoh.com>

- Subject

- Topic of the item
- What is it *about*
- Use controlled vocabularies like LCSH, AAT, Homosaurus, and others as appropriate.
- Separate subjects with repeated fields or with a consistent delimiter (such as a semicolon or pipe)

The most important thing to remember

Be consistent!

Questions?

Demonstration!

Breakout 1

Breakout 2

Wrap-up

DPLA Metadata Application Profile & Introduction
Inclusive Metadata & Conscious Editing Resources
Introduction to Metadata, Third Edition