Metadata and Digital Collections

Jessica Bright (she/her)

Digital Collections & Metadata Librarian

jbright1@brynmawr.edu

What are we talking about today?

- What is metadata?
  - Standards that make up metadata (structure, value, content, and format standards)
    - Where things like Dublin Core and Library of Congress Subject Headings fit in
- How is an online archival collection organized?
- How do you navigate an online archival collection?
  - What to do if pieces of metadata are missing

What is metadata?

- Metadata is most simply data about data
- To be most useful, metadata is structured information

metadata

n.

1. information about data that promotes discovery, structures data objects, and supports the administration and preservation of records   (View Citations)

Notes

Metadata may be embedded or external. It may be applied at a variety of levels of granularity and during different periods in the life cycle of data. It is typically demarcated and standardized, and it often provides context.

Society of American Archivists, Dictionary of Archives Terminology
https://dictionary.archivists.org/entry/metadata.html

What does metadata look like?

BOOK

Fifty years of association work among young women, 1866-1916 : a history of Young Women's Christian Associations in the United States of America

Wilson, Elizabeth, 1867- author

New York : National Board of the Young Women's Christian Associations of the United States of America, 1916

🔗 **Available Online** ›

TOP
SEND TO
VIEW ONLINE
DETAILS
LINKS

Send to

| " CITATION | 🖨 PRINT | 🔗 PERMALINK | ✉ EMAIL | 📄 EXPORT RIS | 📄 EXPORT BIBTEX |

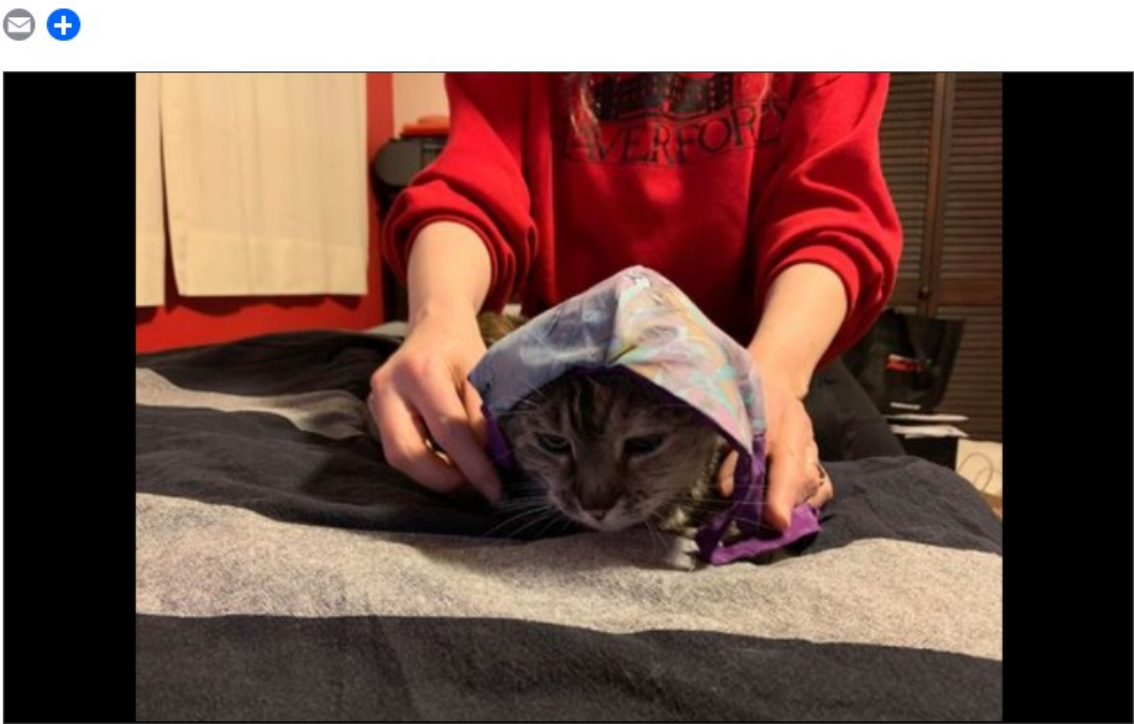## View Online

Nineteenth Century Collections Online ⧉

Nineteenth Century Collections Online ⧉

Details

Title	Fifty years of association work among young women, 1866-1916 : a history of Young Women's Christian Associations in the United States of America
Creator	Wilson, Elizabeth, 1867- author ›
Language	English
Subject	Young Women's Christian Association of the U.S.A -- History ›
Series	Nineteenth Century Collections Online: Women and Transnational Networks ›
	Nineteenth Century Collections Online: Women and Transnational Networks. ›
Publisher	New York : National Board of the Young Women's Christian Associations of the United States of America
Date	1916
Format	1 online resource ([14], 402 pages, [27] leaves of plates) : illustrations, portraits, facsimiles. data file
Notes	Includes index. Reproduction of the original from the Arthur and Elizabeth Schlesinger Library on the History of Women in America. Harvard University.

Tripod library catalog book record
https://tripod.brynmawr.edu/permalink/01TRI_INST/1ijd0uu/alma991019256706004921

Cat Mask

✉ ➕

DOWNLOAD ⬇   Permalink: https://digitalcollections.tricolib.brynmawr.edu/260091

Item Description

Member of	**Documenting COVID-19 at Haverford College**
Description	Description from Adrian Velonis: "April 8, Dobbs Ferry, NY. My cat not understanding how to wear her mask properly. My dear mother, Jeanne Velonis '94, is trying to teach her."
Linked Agent	Photographer (pht): **Velonis, Adrian**
Physical Form	**Still image**
Genre	**photographs** **color photographs**
Date Created	2020-04-08
Subject (Topic)	**COVID-19 Pandemic, 2020-**

Digital Collections image record https://digitalcollections.tricolib.brynmawr.edu/260091

What does metadata look like?

Embedded technical metadata in Adobe Bridge for a digital image

Metadata record from TriArte http://triarte.brynmawr.edu/objects-1/info/153198

Why is metadata important?

- "Certifies the authenticity and degree of completeness of the content"
  ❑Was the entire letter digitized and who is the creator?

- "Establishes and documents the context of the content"
  ❑Provenance information, who/how was it acquired?
  ❑What collection is it from?

- "Identifies and exploits the structural relationships that exist within and between information objects"
  ❑Filenames/identifiers can track the page numbers of a digitized book

- "Provides a range of intellectual access points for an increasingly diverse range of users"
  ❑Different search terms like subject headings, descriptions, creator, and general keywords are available in a robust metadata record

Baca, Murtha, ed. *Introduction to Metadata: Setting the Stage* 3rd ed. Los Angeles: Getty Publications, 2016. http://www.getty.edu/publications/intrometadata

What does this really mean?

- Searching!
- Structured data allows us to query information in a (human *and* machine) predictable way
- This allows us to search more effectively and efficiently
  - Example: Think of how difficult it is to search on Instagram for things that aren't an account or place name and don't fit neatly into a hashtag.

What makes up metadata?

Type	Examples
Data *structure* standards (metadata element sets, schemas). These are "categories" or "containers" of data that make up a record or other information object.	MARC (Machine-Readable Cataloging) Format, Encoded Archival Description (EAD), BIBFRAME (Bibliographic Framework), Dublin Core Metadata Element Set, Categories for the Description of Works of Art, VRA Core
Data *value* standards (controlled vocabularies, thesauri, controlled lists). These are the terms, names, and other values that are used to populate data structure standards or metadata element sets.	Library of Congress Subject Headings, Name Authority File, and Thesaurus for Graphic Materials; Getty Art & Architecture Thesaurus, Union List of Artist Names (ULAN), and Thesaurus of Geographic Names; ICONCLASS; Medical Subject Headings
Data *content* standards (cataloging rules and codes). These are guidelines for the format and syntax of the data values that are used to populate metadata elements.	Anglo-American Cataloguing Rules, Resource Description and Access, International Standard Bibliographic Description, Cataloging Cultural Objects, Describing Archives: A Content Standard

Baca, Murtha, ed. *Introduction to Metadata: Setting the Stage* 3rd ed. Los Angeles: Getty Publications, 2016. http://www.getty.edu/publications/intrometadata

Pause and consider

- Of the different types of standards, which seem the most accessible/inaccessible?
- What kinds of guidelines would you anticipate seeing in a content standard that governs how you format data?
  - Date encoding
  - Name structures (direct order, inverse, etc.)

Dublin Core

- Originally created to describe web resources
- Designed in 1998 in Dublin, Ohio
- It is a basic element set with 15 terms, each are optional and repeatable
- Qualified Dublin Core expanded these 15 terms

Dublin Core Metadata Element Set

Dublin Core Element	Use	Possible Data Value Standards
Title	A name given to the resource.	
Subject	The topic of the resource.	Library of Congress Subject Headings (LCSH) Faceted Application of Subject Terminology (FAST) Art & Architecture Thesaurus (AAT)
Description	An account of the resource.	
Creator	An entity primarily responsible for making the resource.	Library of Congress Name Authority File (LCNAF) Faceted Application of Subject Terminology (FAST)
Publisher	An entity responsible for making the resource available.	
Contributor	An entity responsible for making contributions to the resource.	Library of Congress Name Authority File (LCNAF) Faceted Application of Subject Terminology (FAST)
Date	A point or period of time associated with an event in the lifecycle of the resource.	W3CDTF
Type	The nature or genre of the resource.	DCMI Type Vocabulary
Format	The file format, physical medium, or dimensions of the resource.	Internet Media Types (MIME) Library of Congress Genre/Form Terms (LCGFT) Thesaurus for Graphical Materials (TGM)
Identifier	An unambiguous reference to the resource within a given context.	
Source	A related resource from which the described resource is derived.	
Language	A language of the resource.	ISO 639
Relation	A related resource.	
Coverage	The spatial or temporal topic of the resource, the spatial applicability of the resource, or the jurisdiction under which the resource is relevant.	Thesaurus of Geographic Names (TGN) Geonames
Rights	Information about rights held in and over the resource.	RightsStatements.org

Qualified Dublin Core Elements

Element	Qualifier
Title	Alternative
Description	Table Of Contents
"	Abstract
DateCreated	Valid
"	Available
"	Issued
"	Modified
FormatExtent	Medium
Relation	Is Version Of
"	Has Version
"	Is Replaced By
"	Replaces
"	Is Required By
"	Requires
"	Is Part Of
"	Has Part
"	Is Referenced By
"	References
"	Is Format Of
"	Has Format
"	Is Version Of
"	Has Version
"	Is Replaced By
"	Replaces
"	Is Required By
"	Requires
"	Is Part Of
"	Has Part
"	Is Referenced By
"	References
"	Is Format Of
"	Has Format
Coverage	Spatial
"	Temporal

Example of a DC type record from DPLA

- This isn't exactly DC, but it is based on DC
- The beauty of standardization is that you can aggregate mass amounts of data for searching
- Many libraries digitize their materials and DPLA will harvest their metadata records for greater visibility
- Libraries can use different metadata schemas and then map the elements to Dublin Core, which is more basic and flexible
- End users then search on DPLA and find more materials than just searching their local institution

Cat

View Full Item

COPYRIGHT NOT EVALUATED

Created Date	1940-1959
Description	Photo shows a cat, probably a pet of Charles and Harriette Kelly
Creator	Kelly, Charles, 1889-1971
Partner	Mountain West Digital Library
Contributing Institution	University of Utah - J. Willard Marriott Library
Collection	University of Utah Photograph Archives
Subjects	Utah Capitol Reef National Park Cats Cats--Photographs
Type	image
URL	https://utah-primoprod.hosted.exlibrisgroup.com/primo-explore/fulldisplay?docid=digcoll_uuu_11uum_map/963153&context=L&vid=MWDL
Standardized Rights Statement	**Copyright Not Evaluated:** The copyright and related rights status of this Item has not been evaluated. Please refer to the organization that has made the Item available for more information. You are free to use this Item in any way that is permitted by the copyright and related rights legislation that applies to your use.
Rights	http://rightsstatements.org/vocab/CNE/1.0/

https://dp.la/item/3f1ac38f0baeaa66d3c57e5dbb633b8f?q=cat&page=2

LCSH and LCNAF

✓LCSH = Library of Congress Subject Headings
✓LCNAF= Library of Congress Name Authority File

- Controlled vocabularies of subject headings and names maintained by the US Library of Congress for use in bibliographic records
- Since they're so robust, many libraries use them for cataloging books, describing digital collection materials, and more
  - They're controlled by LC, so they are *not* perfect, but they help with standardization and searching

College teachers

URI(s)
- http://id.loc.gov/authorities/subjects/sh85028378 ⧉
- http://id.loc.gov/authorities/sh85028378#concept

Variants
- Academicians
- Academics (Persons)
- College instructors
- College lecturers
- College professors
- College science teachers
- Lectors (Higher education)
- Lecturers, College
- Lecturers, University
- Professors
- Universities and colleges--Teachers
- University academics
- University instructors
- University lecturers
- University professors
- University teachers

Use For
- http://id.loc.gov/authorities/subjects/sh2005000056

Broader Terms
- Teachers

Narrower Terms
- Agricultural colleges--Faculty
- Christian college teachers
- Communist college teachers
- Community college teachers
- Counselor educators
- Gay college teachers
- Graduate teaching assistants
- Immigrant college teachers
- Indian college teachers
- Jewish college teachers
- Lesbian college teachers
- Library schools--Faculty

LCSH ⟵

LCNAF ⟶

Prince

URI(s)
- http://id.loc.gov/authorities/names/n84079379 ⧉

Variants
- Artist Formerly Known as Prince
- Nelson, Prince Rogers
- TAFKAP

Identifies LC/NAF RWO
- http://id.loc.gov/rwo/agents/n84079379 ⧉

Identifies RWO
- https://isni.org/isni/0000000120964892 ↗
- https://d-nb.info/gnd/118866060 ↗

- **Birth Date**
  1958-06-07

- **Death Date**
  2016-04-21

- **Descriptor**
  African American musicians

- **Birth Place**
  Minneapolis (Minn.)

- **Associated Language**
  English

- **Field of Activity**
  Popular music
  Funk (Music)
  Rhythm and blues music
  Rock music

- **Occupation**
  Composers
  Instrumentalists
  Sound recording executives and producers
  Singers

https://id.loc.gov/authorities/subjects/sh85028378.html

https://id.loc.gov/authorities/names/n84079379.html

How collections are organized physically/digitally

- Collections can be organized in a number of ways, but generally you will always see a link between the physical and digital
  - Generally, this refers to the filename/identifier linking back to where the item came from within the collection
    - E.g. filename: BMC-M124_b2f02
      - BMC, Manuscript Collection 124, box 2, folder 2
- When you have a digital collection, there is no one perfect way to organize things because you don't know how people are coming in and searching your site
- They can sort it and organize it themselves in any number of ways

How to navigate a collection

- https://digitalcollections.tricolib.brynmawr.edu/collections/college-women
- https://digitalcollections.tricolib.brynmawr.edu/collections/her-own-right
- https://digitalcollections.tricolib.brynmawr.edu/collections/1918-pandemic-bryn-mawr-college?items_per_page=50
- https://digitalcollections.tricolib.brynmawr.edu/collections/summer-school-women-workers-industry

- Facets/filters are your friend
- Visual review
- Keyword search is most helpful

What to do if metadata is missing?

- Oftentimes there are missing pieces of metadata. You might not know:
  - Who is in the photo (description) or who took it (creator)
  - The year a digitized photographic print was originally taken (date)
- This doesn't necessarily mean that you can't use the material! (You do need to have the required permissions to use it).
  - Note in the metadata any unknown information (in a standard way, so you can retrieve it via search)
  - Crowdsourcing, especially with community-based projects, can be a powerful way to solve some of these unknowns.

Wrap-up

- Metadata can seem confusing, but we all use and interact with metadata more than you might think
- Standardization and clarity is key!
- There are lots of helpful resources out there to help create metadata

Any questions?

Please feel free to reach out!

Jessica Bright (she/her)

Digital Collections and Metadata Librarian

Special Collections, Canaday Library 2nd Floor

jbright1@brynmawr.edu