

LIS 551, Organization of Information

School of Library and Information Studies
University of Wisconsin-Madison
Fall 2015: M 5:30-8:00

Dorothea Salo (please call me "Dorothea")
Office address: 4261 Helen C. White Hall
Course link page: <http://pinboard.in/u:dsalo/t:551>

salo@wisc.edu
Office Hours: by appointment

Course Description

This course is designed to introduce basic concepts and principles of information structure and organization, and practices that have traditionally been used for this purpose in information organizations, as well as newer developments in the field. Through readings, lectures, discussions and exercises, students will gain an understanding of the theoretical foundations and history of information organization, as well as different methods and tools used for structuring and organizing information.

Phenomena to be examined include (but are not limited to):

- indexes
- bibliographic description
- archival (collection-level) description
- subject analysis
- vocabulary control
- metadata schemas
- relational databases
- markup languages
- linked data and the Semantic Web

Course objectives

Upon completion of the course, you will:

- Understand major standards, systems, and tools used for structuring, describing, and organizing recorded information;
- Acquire the fundamental understanding and skills to effectively use and critically evaluate systems for information organization and retrieval;
- Gain introductory experience in applying standards for creating and encoding metadata;
- Negotiate the social, cultural, and ethical contexts of information organization;
- Have basic knowledge of web standards HTML5, CSS, and XML;
- Appreciate the importance of a user-centered perspective in organizing information.

This course is designed to assess the following SLIS learning outcomes: 2a, 3a, 3c, 3d.

Course Policies

I wish to fully include persons with disabilities in this course. Please let me know within two weeks how I can best meet your needs. I will try to maintain the confidentiality of this information.

Academic Honesty: I follow the academic standards for cheating and plagiarism set forth by the University of Wisconsin.

Attendance in lecture is required. If you miss a class it is your responsibility to a) hand in all assignments due for that day on time, and b) obtain any notes and handouts from other students. Unexcused absences, as well as over two excused absences during the semester, will cost two points apiece from your class-participation total.

Contacting me

READ THE SYLLABUS before asking a question, please; the syllabus may answer it! For any difficulty with the course that is not private or confidential, please bring it up in class or use the Learn@UW help forum; *I will not answer such questions by*

email. Please also do your best to assist your classmates on the forum. I am not available weekends; otherwise, I do my level best to answer forum questions and email within two business days.

Should you see dead links (it does happen, usually with no notice), weird due dates, or other syllabus problems, please post them to the “Syllabus problems” forum on Learn@UW as soon as you see them.

Known schedule disruptions

Expect me to be slower to respond than usual during the following dates:

- October 15-16: Midwest Data Librarians Symposium
- October 22-23: Potomac Technical Processing Librarians Board meeting (I am speaking)
- November 3-6: Wisconsin Library Association conference (I am on a panel)

Textbooks

There are no required textbooks for this course. If you wish, you may refer to Taylor and Joudrey, *Organization of Information*, which is on reserve at the SLIS library. For MARC, I strongly recommend Lois Mai Chan’s *Cataloging and Classification: An Introduction*.

All readings are to be finished by the class meeting time under which they are listed, except for Week 1. Linklists are for enrichment and reference; you are not required to read everything on them.

Unit 1: Building blocks

Week 1: Course introduction. What do information professionals organize; why and how? HTML.

Learning objectives: Problems we solve by organizing information. Working out some information-organization principles by induction. Human/computer asymmetry. “Documents,” records, and statements. HTML, validation, basic HTML tags.

Buckland. “What is a document?” <http://people.ischool.berkeley.edu/~buckland/whatdoc.html>

Buckland. “What is a digital document?” <http://people.ischool.berkeley.edu/~buckland/digdoc.html>

Bush. “As We May Think.” <http://www.theatlantic.com/magazine/archive/1945/07/as-we-may-think/3881/>
(Friendly advice: ignore the materials-engineering jargon; focus on the reading and memory tools that Bush describes, and how he wants them to work.)

Lesk. “The seven ages of information retrieval.” <http://archive.ifla.org/VI/5/op/udtop5/udtop5.htm>

Salo. “MARC, linked data, and human-computer asymmetry.” <http://lj.libraryjournal.com/2015/02/opinion/peer-to-peer-review/marc-linked-data-and-human-computer-asymmetry-peer-to-peer-review/>

Merritt. “The fractal taxonomy of museums.” <http://futureofmuseums.blogspot.com/2012/09/the-fractal-taxonomy-of-museums.html> (mostly for fun!)

(Resources for learning HTML on Learn@UW; use them as you see fit.)

Week 2: Identifiers. Name authority control.

Learning objectives: Why identifiers matter; how they are used. Identifiers and data modeling; identifier scope; unique identifiers; persistent identifiers. Characteristics of good and poor identifiers. Direct and indirect identifiers. Compound identifiers. Identifiers vis-a-vis user/discourse communities. Why most strings are poor identifiers. Why names (personal, geographic, organizational) are poor identifiers. Name authority control: preferred/non-preferred variants, internationalization. VIAF, ORCID. URIs, URLs, URNs. DOIs, handles, ARKs, PURLs, and how they work.

Linklist(s): <https://pinboard.in/u:dsalo/t:identifiers>, <https://pinboard.in/u:dsalo/t:authoritycontrol>

Kent/Anderson. “The unsolvable identity problem.” http://figshare.com/articles/wkent2003_The_Unsolvable_Identity_Problem/156471

McKenzie. “Falsehoods programmers believe about names.” <http://www.kalzumeus.com/2010/06/17/falsehoods-programmers-believe-about-names/>

Smith-Yoshimura et al. “Registering researchers in authority files.” <http://www.oclc.org/content/dam/research/publications/library/2014/oclcresearch-registering-researchers-2014.pdf> (Introduction pp. 6-8)

Callaghan. “Musings on data and identifiers...” <http://citingbytes.blogspot.co.uk/2013/04/musings-on-data-and-identifiers.html>

Waters, “A Digital Gazetteer of Places for the Library of Congress and the NYPL.” <https://thinkwhere.wordpress.com/2015/01/19/a-digital-gazetteer-of-places-for-the-library-of-congress-and-the-nypl/> (The “Background, Requirements, and Challenges” section.)

Richards et al. "A beginner's guide to persistent identifiers." <http://www.gbif.org/resource/80575>
Bruce. "Identifiers, part 2." <https://blog.law.cornell.edu/metasausage/2012/05/15/identifiers-part-2/>
Hickey. "VIAF and other IDs." <http://outgoing.typepad.com/outgoing/2011/07/viaf-and-other-ids.html>

Unit 2: Structuring information

Week 3: RDF and linked data. Schema.org microdata.

Learning objectives: Escaping information silos. Search-engine optimization (SEO). schema.org. RDF. Hypertext. Linked data. Triples as statements. URIs as identifiers. The "open world assumption." Reconciliation.

Linklist(s): <https://pinboard.in/u:dsalo/t:linkeddata>, <https://pinboard.in/u:dsalo/t:schemaorg>

Starting from the divine Billie Holiday https://linkedjazz.org/network/?person=Billie_Holiday (note that the URL is alterable, if you have a favorite!) explore linkedjazz.org. Take a look at at least one set of transcripts establishing a connection between two musicians; how would you characterize that connection, and do you see any issues with it?

Voss/OCLC. "Linked data for libraries." <http://lodlam.net/2012/08/09/linked-data-for-libraries-video-from-oclc/>

Tauberer, "What is RDF and what is it good for?" <https://github.com/JoshData/rdfabout/blob/gh-pages/intro-to-rdf.md#> (Sections 1-4; pay special attention to Tauberer's answers to the second question in his title, please.)

Hilton. "Rise of the machines." <http://blog.wellcomelibrary.org/2013/12/rise-of-the-machines/>

Kelley, Michael. "How the W3C has come to love library linked data." http://www.libraryjournal.com/lj/home/891826-264/how_the_w3c_has_come.html.csp

Page. "Using Google Refine..." <http://iphylo.blogspot.com/2012/02/using-google-refine-and-taxonomic.html>
(The how-to is not as important as understanding what he's doing and why; don't bog down in the procedure.)

"The Periodic Table of SEO Success Factors." <http://searchengineland.com/guide/seo/> (Chapters 1, 3, and especially 4)

Search for a currently-in-theatres movie in Google. Staying on the search results page, look at the available information about the movie in the search result from [imdb.com](http://www.imdb.com), and compare it to the available information in the search result from the movie's official page or from Rotten Tomatoes. Schema.org microdata is the difference!

"What is schema.org?" <http://schema.org/>

"Getting started with schema.org using Microdata." <http://schema.org/docs/gs.html> (Sections 1 and 2)

Week 4: XML for documents; XML for records. CSS.

Learning objectives: Well-formed and valid XML. Delimiters. "Mixed content." Parsers and validators. DTDs and XML Schema. XML authoring tools. Structure vs. presentation; CSS.

Maler and El Andaloussi. "Introduction to SGML." <http://www.xmlgrrl.com/publications/DSDTD/ch01.html>
(Sections 1, 1.1, and 1.2 only. Where you see "SGML," think "XML.")

Hawkins, "Introduction to XML for Text." <http://www.ultraslavonic.info/intro-to-xml/>

Ray, *Learning XML*, sections 1.1, 1.2, 2.1-2.4, 3-3.3, 4-4.3. (Available as an ebook through UW-Madison Libraries; look it up in the catalog.)

Raggett. "Adding a touch of style." <http://www.w3.org/MarkUp/Guide/Style>

Meyer. *CSS: The Definitive Guide*, 3rd ed. Chapter 1, sections 2.1, 2.2, 5.1-5.3, 7.1-7.3, 9.1-9.3. (Available as an ebook through UW-Madison Libraries; look it up in the catalog.)

Week 5: Relational databases

Learning objectives: The problems with spreadsheets as records tools. Tables (entities), columns (attributes), rows (instances). Data typing. Atomicity. Entity-relationship diagrams. Primary and foreign keys. Relationship cardinality. The FRBR Group 1 (WEMI) model.

Veltman. "SQL: The Prequel (Excel vs. Databases)" <http://schoolofdata.org/2013/11/07/sql-databases-vs-excel/>

Allardice. "Foundations of programming: databases." <http://www.lynda.com/Programming-tutorials/Foundations-Programming-Databases/112585-2.html?org=wisc.edu> (Watch/read "Welcome," "What are databases?" "The features of a relational database" I don't care whether you watch the video or read the transcript.)

Chapple, Mike. "Database keys." <http://databases.about.com/od/specificproducts/a/keys.htm>

Huscher, Brent. "Database design and modeling fundamentals." <http://www.sqlteam.com/article/database-design-and-modeling-fundamentals>

Tillett. "What is FRBR?" <http://www.loc.gov/cds/downloads/FRBR.PDF>

Coyle. "FRBR, Twenty Years On." <http://kcoyle.net/FRBR20.pdf> (Stop when you reach "Variations on the FRBR model.")

Week 6: Search indexes, search engines, and search results

Learning objectives: Concordance. Back-of-book index (See reference, See Also reference, first-level entries and subentries). Pointers. Print literature indexes. Search-engine index (inverted index), spider/crawler. TF/IDF, precision vs. recall. Page Rank (and ways to game/spam it). Search-results management: relevance ranking, deduplicating, and faceted limiting. Leveraging index structures to search for keywords, full text, citations. Search-results personalization and filtering.

Mulvany. "Introduction to book indexing." *Indexing Books* (2nd ed). On reserve.

Meyers. "Why an ebook still needs an index." <http://radar.oreilly.com/2011/09/ebook-index-search-discovery.html>

Franklin. "How Internet Search Engines Work." <http://computer.howstuffworks.com/search-engine.htm> (Parts 1-4)

"Learn how Google works." <http://ppcblog.com/how-google-works/>

Brin and Page. "The anatomy of a large-scale hypertextual web search engine." <http://infolab.stanford.edu/~backrub/google.html> (Brin and Page founded Google based on the ideas in this paper.)

Das and Jain. "Indexing the World Wide Web: the journey so far." <http://research.google.com/pubs/archive/37043.pdf> (pp. 1-7)

Catone. "How to use Google search more effectively." <http://mashable.com/2011/11/24/google-search-infographic/>

DuckDuckGo. "Escape your search engine's filter bubble!" <http://dontbubble.us/>

Heller, "Removing the truthiness from Google." <http://acrl.ala.org/techconnect/?p=5011>

Webber. "Should search algorithms be moral?" <http://qz.com/451051/should-search-algorithms-be-moral-a-conversation-with-googles-in-house-philosopher/>

Rochkind, Jonathan. "Information retrieval and relevance ranking for librarians." <http://bibwild.wordpress.com/2011/03/28/information-retrieval-and-relevance-ranking-for-librarians/>

Burton-West. "Practical relevance ranking for 11 million books." <https://www.hathitrust.org/blogs/large-scale-search/practical-relevance-ranking-11-million-books-part-1> and <https://www.hathitrust.org/blogs/large-scale-search/practical-relevance-ranking-11-million-books-part-2-document-length-and-rel>

Unit 3: Vocabularies

Week 7: Controlled vocabularies.

Learning objectives: When and why full-text searching fails. When and why free-keywording fails. Language problems that controlled vocabularies solve. Language problems that controlled vocabularies create. Thesaurus (BT, NT, USE, UF; scope notes; thesaurus use in search engines). Taxonomy. Ontology.

Linklist(s): <https://pinboard.in/u:dsalo/t:controlledvocab>, <https://pinboard.in/u:dsalo/t:ontology>

Leise, Fast, and Steckel. "What is a controlled vocabulary?" <http://boxesandarrows.com/what-is-a-controlled-vocabulary/>

Leise, Fast, and Steckel. "Creating a controlled vocabulary." <http://boxesandarrows.com/creating-a-controlled-vocabulary/>

Rayburn. "Taxonomies and thesauri." <http://www.ischool.utexas.edu/~i385e/readings/Warner-aTaxonomyPrimer.html>

Noy and McGuinness. "Ontology development 101." <http://www.ksl.stanford.edu/people/dlm/papers/ontology-tutorial-noy-mcguinness-abstract.html>

"Implementing a Government-wide semantic solution to thesauri." <http://www.pdfpower.com/XML2005Proceedings/ship/71/sall-71.HTML>

Ackland. "Metadata, controlled vocabulary, and your DAM." <http://blog.databasics.com.au/2014/08/digital-asset-management-and-controlled-vocabulary.html>

Week 8: User-created vocabularies. Tagging systems; folksonomies. Recommender systems.

Learning objectives: Tagging systems. Folksonomies, and how they interact with controlled vocabularies. How recommender systems work. Privacy ethics of recommender systems and other kinds of profiling.

Linklist(s): <https://pinboard.in/u:dsalo/t:recommendersystems>, <https://pinboard.in/u:dsalo/t:folksonomy>, <https://pinboard.in/u:dsalo/t:dataethics>

Shirky. "Ontology is overrated: categories, links, and tags." http://www.shirky.com/writings/ontology_outrated.html

Manzo et al. "By the People, For the People: assessing the value of crowdsourced, user-generated metadata." <http://www.digitalhumanities.org/dhq/vol/9/1/000204/000204.html>

Spalding. "When tags work and when they don't: Amazon and LibraryThing." <http://blog.librarything.com/thingology/2007/02/when-tags-work-and-when-they-dont-amazon-and-librarything/>

Konstan and Rieder. "Deconstructing recommender systems: how Amazon and Netflix predict your preferences and prod you to purchase." <http://spectrum.ieee.org/computing/software/deconstructing-recommender-systems>

Madrigal. "How Netflix re-engineered Hollywood." <http://www.theatlantic.com/technology/archive/2014/01/how-netflix-reverse-engineered-hollywood/282679/>

Singel. "Netflix cancels recommendation contest after privacy lawsuit." <http://www.wired.com/2010/03/netflix-cancels-contest/>

Hill. "How Target figured out a teen girl was pregnant before her father did." <http://www.forbes.com/sites/kashmirh11/2012/02/16/how-target-figured-out-a-teen-girl-was-pregnant-before-her-father-did/>

Week 9: Social justice and controlled vocabularies

Learning objectives: Why naming is powerful. How information-organization activities are culturally-bound. How controlled vocabularies and other information-organization practices can marginalize, insult, and exclude. How algorithms can create or exacerbate this class of problem. Conflicts of social justice with literary warrant. Sandy Berman, Hope Olson, and vocabulary reform.

Linklist(s): <https://pinboard.in/u:dsalo/t:551/t:socialjustice>

For reference: Schomberg. "Cataloging and social justice." <http://catassessmentresearch.blogspot.com/2014/02/cataloging-and-social-justice.html>

Grieve-Smith. "Challenges for radical categorization." <http://grieve-smith.com/blog/2015/07/challenges-for-radical-categorization/>

Olson. "Mapping beyond Dewey's boundaries: constructing classificatory space for marginalized knowledge domains." *Library Trends* 47:2. <http://hdl.handle.net/2142/8210>

"Berman scorecard." <http://www.scribd.com/doc/19082065/Bermanlcshtscorecard0409>

Ockerbloom. "Understanding concept-oriented catalogs." <http://everybodyslibraries.com/2009/12/04/understanding-concept-oriented-catalogs/>

Weinberger. "I bet your ontology never thought of this one!" <http://www.hyperorg.com/blogger/2012/12/18/misc-i-bet-your-ontology-never-thought-of-this-one/>

Hern. "Flickr faces complaints over 'offensive' auto-tagging for photos." <http://www.theguardian.com/technology/2015/may/20/flickr-complaints-offensive-auto-tagging-photos>

Flood. "What happens when Google doesn't think you're a human." <http://www.buzzfeed.com/joeflood/what-happens-when-google-doesnt-think-youre-a-human#.qgrGGMJAV>

Baldrige. "Machine learning and human bias: an uneasy pair." <http://techcrunch.com/2015/08/02/machine-learning-and-human-bias-an-uneasy-pair/>

Reagle and Rhue. "Gender bias in Wikipedia and Britannica." *International Journal of Communication* 5. <http://ijoc.org/ojs/index.php/ijoc/article/view/777>

Week 10: Subject analysis. Classification.

Learning objectives: Subject authority control. Subject analysis and "aboutness." Classification schemes: DDC, LCC. Classification schedules; how to use them. Subject headings: LCSH. Pre- vs. post-coordinated headings. FAST. Faceted classification.

Linklist(s): <https://pinboard.in/u:dsalo/t:classification>, <https://pinboard.in/u:dsalo/t:subjectanalysis>, <https://pinboard.in/u:dsalo/t:fast>

"#WikipediaProblems: How to classify everything." <http://www.theatlantic.com/technology/archive/2013/10/wikipediaproblems-how-do-you-classify-everything/280178/> (How would you do this?)

Ockerbloom. "Understanding concept-oriented catalogs." <http://everybodyslibraries.com/2009/12/04/understanding-concept-oriented-catalogs/>

Ockerbloom. "Some concepts and their catalogs." <http://everybodyslibraries.com/2009/12/10/some-concepts-and-their-catalogs/>

Rochkind, Jonathan. "Broad categories from class numbers." <http://bibwild.wordpress.com/2011/04/04/broad-categories-from-class-numbers/>

Rondeau. "The life and times of aboutness: a review of the library and information science literature." <http://ejournals.library.ualberta.ca/index.php/EBLIP/article/view/19091> (Come to class with a list of "aboutness" theories and the problems/edge cases they collided with. It's fine to build this list collaboratively!)

Betts. "Safari aggregations and content categorization." <https://www.safaribooksonline.com/blog/2014/09/03/safari-aggregations/>

Wood and Kovalchik. "Deciphering the Dewey Decimal." <https://web.archive.org/web/20110426212923/http://www.mentalfloss.com/quiz/quiz.php?q=1194> (Memorize these. Just do it, okay?)

Library of Congress. "Library of Congress Classification Outline." <http://www.loc.gov/catdir/cpso/lcco/> (Memorize these too. You don't have to drill down, though I recommend doing so in your area of greatest expertise; just memorize the top level.)

McGrath. "Facet-based search and navigation with LCSH: problems and opportunities." <http://journal.code4lib.org/articles/23>

Unit 4: Organization systems

Week 11: User needs and usability. Culture, user needs, and usability. Evaluating information-organization systems. Evaluating data quality. The database industry; scholarly communication.

Markets, sources, and coverage (i.e. what's in a "library database" or "in Google" and why.) User tasks and their evolution over time (bibliographic objectives, FRBR FISO). Known-item searching vs. subject and author searching. Collocation, browsing, "serendipity." Access points. Item-level vs. collection-level descriptions. Dewey vs. BISAC and other alternatives.

Svenonius. "Information organization" and "Bibliographic objectives" *The intellectual foundation of information organization*. Boston: MIT, 2000. (Available as an ebook through UW-Madison Libraries; check the catalog.)

Watters. "We aren't the world." <http://www.psmag.com/books-and-culture/joe-henrich-weird-ultimatum-game-shaking-up-psychology-economics-53135>

Coyle. "Works, expressions, and the bibliographic universe." <http://kcoyle.blogspot.com/2014/04/works-expressions-and-bibliographic.html> (Come prepared to discuss "user needs." Warning: slippery concept!)

Carr. "Serendipity in the stacks: libraries, information architecture, and the problems of accidental discovery." <http://dx.doi.org/10.5860/crl.76.6.831>

Shreeves et al. "Moving towards shareable metadata." <http://www.firstmonday.org/ojs/index.php/fm/article/view/1386/1304>

Fister. "The Dewey dilemma." <http://lj.libraryjournal.com/2010/05/public-services/the-dewey-dilemma/>

Coyle. "Library signage." <http://kcoyle.blogspot.com/2012/09/library-signage.html>

Harris. "Summer Project: Kill Dewey." *School Library Journal*. <http://www.thedigitalshift.com/2012/08/k-12/summer-project-kill-dewey/>

Week 12: Content standards for description

Learning objectives: AACR2; what goes into a Level 1 record. RDA; major differences from AACR2. DACS (required elements). CCO. Dublin Core. VRA Core. Access point. Main entry. Uniform titles. "Shareable metadata," interoperability, and problems of metadata-content (in)consistency. Descriptive library cataloging. Cataloging and metadata workflows.

Linklist(s): <https://pinboard.in/u:dsalo/t:rda>, <https://pinboard.in/u:dsalo/t:cataloguing> (in my world this word contains a "u!")

As you read, ask yourself *what problem(s) is this standard trying to solve?*

OCLC. "Online cataloging." <http://www.loc.gov/catdir/cpso/app-lvls.html> ("Types of cataloging" and "Levels of cataloging" only.)

"Cataloging Cultural Objects. Part One: General Guidelines." http://cco.vrafoundation.org/downloads/PartOne_GeneralGuidelines.pdf

SAA. "Describing Archives: A Content Standard." http://files.archivists.org/pubs/DACS2E-2013_v0315.pdf (pp. xxi-xxiv, Chapter 1. Sections 2.1-2.6, 3.1, 4.1, 4.5 and be sure you understand why I assigned these and not others!)

Zabel and Miller. "Resource Description and Access (RDA): an introduction for reference librarians." <http://blog.rusq.org/2011/04/03/resource-description-and-access-rda-an-introduction-for-reference-librarians/> (both pages, please)

"An introduction to VRA Core." http://www.loc.gov/standards/vracore/VRA_Core4_Intro.pdf

"[Dublin Core] User Guide." http://wiki.dublincore.org/index.php/User_Guide

Scott. "How much description is enough? A brief history on the debate over the Dublin Core Metadata Initiative." http://www.eliotscott.com/documents/dublin_core.pdf

Week 13: Syntax/data-structure standards for description

Learning objectives: MARC. Henriette Avram. Karen Coyle. ISBD(G). ISAD(G). BIBFRAME and other linked-data surrogate-bibliographic-record representations. Transcribed vs. controlled fields. Copy cataloging.

Linklist(s): <https://pinboard.in/u:dsalo/t:marc>, <https://pinboard.in/u:dsalo/t:bibframe>

ISAD(G). [http://www.icacds.org.uk/eng/ISAD\(G\).pdf](http://www.icacds.org.uk/eng/ISAD(G).pdf) (Sections I.1 through I.6. Also skim the glossary, section 0; look particularly at the definition of “finding aid.”)

Coyle. “Models of bibliographic data.” <http://kcoyle.blogspot.com/2011/08/models-of-bibliographic-data.html>

“Superseded ISBDs.” <http://www.ifla.org/isbd-rg/superseded-isbd-s> (Skim this to get a sense of how ISBD divides up the universe of describable things, and what describable things it privileges.)

“Full ISBD Examples.” http://www.ifla.org/files/assets/cataloguing/isbd/isbd-examples_2013.pdf (pp. 7-8 and read all examples in your first language, or English if your first language is not available)

“MARC Records Frequently Asked Questions.” <http://www.ilsa.lib.ia.us/marcfaq.htm>

Enis. “Ending the invisible library.” <http://lj.libraryjournal.com/2015/02/technology/ending-the-invisible-library-linked-data/>

“Bibliographic Framework as a web of data.” <http://www.loc.gov/bibframe/pdf/marclid-report-11-21-2012.pdf> (pp. 1 to the top of 16)

“Henriette D. Avram, Modernizer of Libraries, Dies at 86.” <http://www.nytimes.com/2006/05/03/us/03avram.html>

Laughlin, “Fantastic (and free!) cataloging tools.” http://www.railslibraries.info/sites/default/files/fantastic_free_cat_tools_101112.pdf (For your reference and use.)

Week 14: Metadata standards

Learning objectives: Descriptive, structural, administrative, and technical metadata (with examples of each type: MODS, METS, PREMIS, MIX). Creating, salvaging, cleaning up, and editing metadata.

Linklist(s): <https://pinboard.in/u:dsalo/t:metadata>

As you read, ask yourself *what problem(s) is this standard trying to solve?*

Riley, Jenn. “Seeing Standards.” <http://www.dlib.indiana.edu/~jenlriley/metadatamap/> (Download the poster and read the legend and definitions carefully. A print copy hangs in the SLIS Library classroom, if you prefer.)

Gradman, Stefan. “Cataloguing vs. metadata: old wine in new bottles?” <http://archive.ifla.org/IV/ifla64/007-126e.htm>

“Dublin Core’s dirty little secret.” <http://reprog.wordpress.com/2010/09/03/bibliographic-data-part-2-dublin-cores-dirty-little-secret/>

Cundiff and Trail, “Using METS and MODS...” <http://www.loc.gov/standards/mods/presentations/mets-mods-morgan-ala07/>

MIX. <http://www.loc.gov/standards/mix/>

NISO. “Data dictionary - Technical metadata for digital still images.” http://www.niso.org/kst/reports/standards?step=2&gid=None&project_key=b897b0cf3e2ee526252d9f830207b3cc9f3b6c2c (Download the full report and read sections 1 and 2. This data dictionary is serialized in XML as MIX.)

OCLC/RLG. “Data dictionary for preservation metadata.” <http://www.oclc.org/research/projects/pmwg/premis-final.pdf> pp. vii-x, 1-1 through 1-10.

Week 15: (No Week 15 this fall due to bizarre semester schedule; “Week 15” assignments due Monday of exam week)

Weekly due-date table

	Class day	Due at class time
Week 1	9/14	
Week 2	9/21	
Week 3	9/28	Info Resources website
Week 4	10/5	Info Resources schema.org microdata
Week 5	10/12	Info Resources CSS
Week 6	10/19	
Week 7	10/26	QUIZ 1 in class
Week 8	11/2	Information Resources thesaurus terms
Week 9	11/9	
Week 10	11/16	
Week 11	11/23	Information Resources subject analysis and classification
Week 12	11/30	QUIZ 2 in class
Week 13	12/7	Metadata collection according to content standards
Week 14	12/14	Metadata expression according to MARC standards
Week 15		Due Monday of exam week: Metadata expression in MODS, e-portfolio writeup

ASSIGNMENTS

All assignments are due **at class time** on the dates listed in the table above (the week designations below are mostly for me). One final-grade percentage point will be lost per day or fraction thereof late. If you are comfortable working ahead, feel free.

Assignments at-a-glance

	% of final grade	Due class time of:
Quiz 1	15%	In-class week 7
Quiz 2	15%	In-class week 12
Information Resources website	10%	Week 3
Information Resources microdata	5%	Week 4
Information Resources CSS	5%	Week 5
Information Resources thesaurus terms	5%	Week 8
Information Resources subject analysis and classification	5%	Week 11
Metadata collection according to content standards	5%	Week 13
Metadata expression according to MARC standards	10%	Week 14
Metadata expression in MODS	10%	Week 15
E-portfolio writeup	5%	Week 15
Class participation	10%	

No extra credit opportunities are available in this class.

Final grade scale: 100-93.5 A; 93.4-89.5 AB; 89.4-83.5 B; 83.4-79.5 BC; 79.4-73.5 C, 69.5-73.4 D, below 69.5 F

Quizzes

These will be on vocabulary, concepts, and techniques. Each quiz will have a few points above the listed maximum available. Points earned above the listed maximum will not be counted.

Class participation

Students earn points for participation by attending class regularly, contributing to group and full-class discussions, asking questions, and participating usefully in in-class group work. Conduct contravening the Recurse Center's four social rules (see <https://www.recurse.com/manual#sec-environment>) will be penalized. (I am "the faculty" for Rule 4 reporting.)

Assignment evaluation criteria

A full-credit assignment will:

- follow all the instructions specified in the assignment;
- address the problem posed in the exercise;
- be mechanically correct, insofar applicable;
- be handed in on time.

Assignment descriptions

Unless otherwise stated, turn in all homework as a Word or PDF file to the appropriate homework dropbox on Learn@UW.

Information Resources website

Hand-encode (no Dreamweaver, no WordPress, no content-management system! just you, your text editor, and your favorite validator!) a website consisting of five individual HTML5 pages, each of which must link to all the others:

- One page must describe a *book*.
- One page must describe a *person or organization* involved in the production of the book.
- One page must describe/discuss the main *subject matter* of the book.
- One page must describe/discuss a *place* either mentioned in the book OR important to the production of the book (e.g. author's hometown, place of publication; places existing or previously existing on planet Earth only, please)
- One page must describe a *non-book information resource* derived from the book OR from which the book is itself derived (e.g. a movie or TV show, a piece of visual art, a piece of music, a non-book reference work, a scholarly article about the book, a cosplay photo or design — fan wikis and other fan works encouraged!)

The information you capture in your website, and the manner in which you present it, are largely up to you. (You will be adding to the site over the course of the semester, as well as using it as grist for other assignments.) The website as a whole must minimally contain:

- an HTML5 declaration (must be present and correct on all pages)
- `<title>` tag (must be present and correct on all pages)
- two levels of heading tags
- a paragraph
- a block quotation
- a list (ordered or unordered)
- an image
- a table (with tabular data, please; "layout tables" not acceptable)
- an external hyperlink

You may use other HTML5 tags and structures as you see fit. Put your website in a folder, zip the folder, and upload the zip file to the Learn@UW Information Resources Website dropbox.

Add schema.org microdata to your Information Resources website

Encode metadata about your book, person/organization, topic, other information resource, and place as correct schema.org microdata. You must correctly identify the type of item described on each page (with the `itemscope` and `itemtype` attributes) and add at least three different correctly-placed properties (with the `itemprop` attribute) per page. It is fine to add extra text or HTML to accommodate enough schema.org attributes. (Hint: find and work from examples! The schema.org website incorporates many!) You may work through this process with your colleagues and discuss questions and problems freely on Learn@UW, since you are all describing different objects. (N.b. the best you may be able to do with your topic is Thing. Do better if you can.)

Rezip your site folder and upload it to the Learn@UW Microdata dropbox.

Styling your Information Resources website with CSS

Add an external CSS stylesheet to your website that all its pages use. Minimally, this stylesheet must:

- Change the font to sans-serif for all headers
- Change the font for all other tags away from the browser default
- Change the link color; remove link underlining except while the link is being hovered over or clicked on
- Change the background color of the page
- Add margins around the page
- Indent paragraphs a visually-pleasing amount of space

You may add as many additional CSS rules as you like. Rezip your site folder and upload it to the Learn@UW CSS dropbox.

Finding thesaurus terms

Find the best approximation to your Information Resources website *topic* in each of the following thesauri:

- Art and Architecture Thesaurus <http://www.getty.edu/research/tools/vocabularies/aat/>
- ERIC Thesaurus <http://eric.ed.gov/>
- Library, Information Science, and Technology Abstracts thesaurus (go through the UW-Madison Libraries website)

There will not always be a good match, of course! Keep a list of close matches along with your top choices.

In ERIC and LISTA, do a subject search on the thesaurus term you chose, and create a proper APA citation for the *newest* available resource under that term.

Now, browse (do not search!) the Thesaurus of Geographic Names <http://www.getty.edu/research/tools/vocabularies/tgn/> until you find your Information Resources website *place*, keeping a list of all its broader terms and their place types. (If your place is not in TGN, get as close as you can to it.) Turn in your terms, citations, and place list to the Learn@UW Thesaurus Terms dropbox.

Classification of and subject analysis for your information resources

- Apply three appropriate Library of Congress Subject Headings to your Information Resources website's book and non-book information resource (n.b. this may be challenging!).
- Find the closest LCSH for your chosen topic and your chosen place.
- Assign call numbers to your chosen book and non-book information resource in Library of Congress Classification and Dewey Decimal Classification. Do not forget to add Cutter numbers; use the LC cutter tables at <http://www.itsmarc.com/crs/mergedProjects/cutter/cutter/contents.htm>.
- Find the closest Cutter number for your chosen place from the LC Geographic Cutter tables linked to from <http://www.itsmarc.com/crs/mergedProjects/cutter/cutter/contents.htm>.

When you are done, look up your book in WorldCat and copy any subject headings applied. Now go to the catalog for each of the three closest-to-you libraries holding the book (note: go further down the list if two libraries share a single catalog) and copy their subject headings and call numbers (LCC or DDC; note which is which for each number). Where their decisions differ from yours, briefly explain which decision you think better and why, with specific reference to your understanding of user behavior. Turn in your work to the Learn@UW Classification dropbox.

Metadata collection according to content standards

- For your book and your non-book information resource, collect the necessary information to create a Level 1 record in AACR2. Do not worry about punctuation, capitalization, or serialization yet; just collect the information!
- Treating your entire Information Resources website as a collection, collect all required elements in DACS for it. Again, do not bother about mechanics. You may keep narrative elements brief (1-2 sentences).
- Make the most extensive Simple (unqualified) Dublin Core records possible for your book and your non-book information resource.

Turn in your work to the Learn@UW Content Standards dropbox.

Metadata expression according to MARC syntax

Create a MARC/AACR2 or MARC/RDA record (your choice) for your book and another for your non-book information resource. Each record should contain the following fields and subfields:

- 020 subfield a if available

- 050 (from earlier assignment)
- 082 (from earlier assignment)
- 100/110 (find the authorized heading! you know how to do this!)
- 245 (subfields a, b if necessary, c)
- 250 (subfields a, b if necessary)
- 260 (subfields a, b, c)
- 300 (subfields a, b if necessary, c)
- for each of the subject headings you provided in the subject-analysis assignment, a 600/610/630/650/651 field as appropriate
- any appropriate 700/710 fields (find the authorized heading!)
- 856, if appropriate.

You will lose points if your ISBD punctuation is incorrect! Work from (at least check yourself against) examples! Turn in your work to the Learn@UW MARC dropbox.

Metadata expression in MODS

Create a valid MODS record for your book, and another for your non-book information resource, based on the MARC records you created in your prior assignment. Capture as much of the MARC information as you can; for anything in the MARC record you cannot capture, leave an XML comment in your file with the MARC field/subfield and a brief explanation. Turn in your work to the Learn@UW MODS dropbox.

E-portfolio writeup

Treating all the work you have done with your three information resources as a single assignment, write a justification statement for your e-portfolio aimed at SLIS learning outcome 3a “Students organize and describe print and digital information resources.” You may also write toward other outcomes (see table below!) if you wish. Your writeup should make clear to me (pretend I am a prospective boss!) your improved understanding of information-organization tools and techniques, your sense of situations in which certain techniques are and are not appropriate, and your ability to implement organization standards (as well as similar standards you have not yet encountered) in a real-world work situation. Turn this in to the Learn@UW E-portfolio dropbox, though of course you may use it in your e-portfolio as well.

Learning outcomes table

SLIS Goals	551 Objectives	551 Measurable Outcomes
2a. Students evaluate and debate information policy and ethics applicable in local, national, or global contexts.	Have an understanding of the major standards, systems, and tools used for organizing and cataloging recorded information Acquire the fundamental understanding and skills to effectively use and critically evaluate systems for information organization and retrieval	Discussion and in-class exercise on locating and redressing potentially problematic vocabularies: reflected in participation grade.
3a. Students organize and describe print and digital information resources.	Have an understanding of the major standards, systems, and tools used for organizing and cataloging recorded information	All graded course assignments concern organization and description of information resources.
3c. Students analyze information needs of diverse individuals and communities.	Negotiate the social, cultural, and ethical contexts of information organization	Discussion and in-class exercise on locating and redressing potentially problematic vocabularies: reflected in participation grade.
3d. Students understand and use appropriate information technologies.	Have basic knowledge of web standards HTML5, XML, and CSS	HTML, CSS, XML, microdata assignments.