LIS 551, Organization of Information

School of Library and Information Studies
University of Wisconsin-Madison
Spring 2013: T 1:30-4:00

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, 608-265-4733
Office Hours: by appointment
Skype: dorotheasalo

Course Description

This course is designed to introduce the basic concepts and principles of information organization, and the practices that have traditionally been used for this purpose in libraries and cultural-heritage 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 organizing information.

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

- back-of-book indexing,
- bibliographic description,
- subject analysis,
- vocabulary control,
- metadata schemas,
- database and search-engine index structure,
- markup languages,
- linked data and the Semantic Web.

Course objectives

Upon completion of the course, you will:

- Have an understanding of 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;
- Gain introductory experience in applying standards for creating and encoding metadata;
- Appreciate the importance of a user–centered perspective in organizing information.

This course is designed to assess the following SLIS learning outcomes: 2a, 3a, 3b, 4a, 4b.

Course Policies

I wish to fully include persons with disabilities in this course. Please let me know within two weeks if you require accommodation. 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

For any difficulty with the course that is not private or confidential, including group-project issues, please 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 Fridays or weekends; otherwise, I do my level best to answer forum questions and email within two business days. If you need to speak with me, please make an appointment with me *directly on WiscCal*, which will email me the appointment information and help ensure I'm not double-booked.

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.

Textbooks

There are no required textbooks for this course. If you wish, you may refer to Taylor and Joudry, *Organization of Information*, which is on reserve at the SLIS library.

All readings are to be finished by the class meeting time under which they are listed.

Week 1: Course introduction

Learning objectives: Problems we solve by organizing information. Working out information-organization principles by induction. Evaluating an information-organization system. Course themes: the affordances of organized things, the constraints of organization systems, identifiers.

Week 2: What do librarians and archivists organize; why and how?

Learning objectives: Content vs. carrier; informational vs. evidential value. "Document." Bibliographic objectives. User tasks and their evolution over time. Known-item searching. Subject and author searching. Collocation and browsing. Access points. Item-level vs. collection-level descriptions. Metadata. Machine-readable description; schema.org.

Svenonius. "Information organization" and "Bibliographic objectives" (through page 24). *The intellectual foundation of information organization.* Boston: MIT, 2000.

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`

Coyle, "Content and carrier." http://kcoyle.blogspot.com/2012/04/content-and-carrier.html

archivists.org message thread, "Are finding aids obsolete?" `http://forums.archivists.org/read/messages?id=42591#42591` (As you read, try to come up with a definition of "finding aid" that you're happy with.)

"What is schema.org?" `http://schema.org/`

Wallis, "Forming consensus on schema.org for libraries and more." `http://dataliberate.com/2012/12/forming-consensus-on-schema-org-for-libraries-and-more/`

Week 3: Document- and collection-internal organization

Learning objectives: Concordance. Back-of-book index (See reference, See Also reference, first-level entries and subentries). Hypertext. Indexing ebooks. Finding aid. DACS. Google sitemap. Navigation. Information architecture. User vs. author language; why string searches can fail.

Svenonius, Elaine. "Bibliographic objectives" (pp 24-chapter end). *The intellectual foundation of information organization.* Boston: MIT, 2000.

Mulvany, Nancy. "Introduction to book indexing." *Indexing Books* (3rd ed).

Bush, Vannevar. "As We May Think." `http://www.theatlantic.com/magazine/archive/1945/07/as-we-may-think/3881/` (What of Bush's vision has been fulfilled by the modern Internet? Modern libraries? What hasn't?)

ISAD(G). 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.")

SAA. "Describing Archives: A Content Standard." `http://www.archivists.org/governance/standards/dacs.asp`

Sitemaps.org. "Protocol." `http://www.sitemaps.org/protocol.php` (Ask yourself: what information does this protocol capture? For what unit of measure? At what scale?)

"What are sitemaps?" `http://www.sitemaps.org/index.html`

Google. "Creating sitemaps." `http://www.google.com/support/webmasters/bin/answer.py?answer=183668`

Week 4: Vocabulary control

Learning objectives: Why controlled vocabularies are useful. Thesaurus. BT, NT, USE, UF. Scope notes. Thesaurus use in search engines. Synonym ring. Taxonomy. Ontology.

Leise, Fast, and Steckel. "What is a controlled vocabulary?" `http://www.boxesandarrows.com/view/what_is_a_controlled_vocabulary_`

Leise, Fast, and Steckel. "Creating a controlled vocabulary." `http://www.boxesandarrows.com/view/creating_a_controlled_vocabulary`

Rayburn. "Taxonomies and thesauri." `http://www.ischool.utexas.edu/~i385e/readings/Warner-aTaxonomyPrimer.html` (download and read the presentation PDF. Pay special attention to thesaurus entry syntax.)

Noy and McGuinness. "Ontology development 101." `http://protege.stanford.edu/publications/ontology_development/ontology101-noy-mcguinn`

Week 5: The basics of cataloging: AACR2 and MARC

Learning objectives: Descriptive cataloging. MARC. Henriette Avram. AACR2. ISBD(G). Access point. Main entry. Transcribed vs. controlled fields. Uniform titles. Copy cataloging. Cataloging as user-interface tool. Name authority control. Name-strings vs. identifiers vs. URLs/URIs. VIAF, ORCID.

N.B.: I am assigning the second edition of Chan from 1994, not the current one from 2007! (I'll discuss why in class.)

"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.)

Chan. "Bibliographic control and library catalogs." *Cataloging and classification: an introduction.* 2nd ed. (1994), pp. 3-28.

Chan. "Descriptive cataloging." *Cataloging and classification: an introduction.* 2nd ed. (1994), pp. 38-47.

Chan. "Description." *Cataloging and classification: an introduction.* 2nd ed. (1994), pp. 49-52. Skim remainder of chapter; use it as reference for Weeks 5 and 6 assignments.

Chan. "Choice of access points." *Cataloging and classification: an introduction.* 2nd ed. (1994), pp. 107-112. Skim remainder of chapter; use it as reference for Weeks 5 and 6 assignments.

Chan. "Name authority control" and "Uniform titles." *Cataloging and classification: an introduction.* 2nd ed. (1994), pp. 123-4, 141-4. Skim remainder of chapter; use it as reference for Weeks 5 and 6 assignments.

"About VIAF." http://www.oclc.org/research/activities/viaf/

ORCID. http://orcid.org/

Week 6: Subject cataloging. Classification. More on MARC.

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. Commonest MARC tags and subfields.

Chan. "Subject cataloging." *Cataloging and classification: an introduction.* 2nd ed. (1994), pp. 155-69.

Chan. "Library of Congress Subject Headings." *Cataloging and classification: an introduction.* 2nd ed. (1994), pp. 171-5, 179-82 ("Headings for named entities"). Skim remainder of chapter; use it as reference for Weeks 5 and 6 assignments.

Chan. "General principles of classification." *Cataloging and classification: an introduction.* 2nd ed. (1994), pp. 259-67.

Chan. "Dewey Decimal Classification." *Cataloging and classification: an introduction.* 2nd ed. (1994), pp. 275-84.

Chan. "Library of Congress Classification." *Cataloging and classification: an introduction.* 2nd ed. (1994), pp. 327-49. Skim remainder of chapter; use it as reference.

Week 7: Subject and classification vocabularies: social-justice and usability issues

Learning objectives: Why naming is powerful. How library activities are culturally-bound. How controlled vocabularies and other library practices can marginalize, insult, and exclude. Sandy Berman, Hope Olson, and vocabulary reform. Dewey vs. BISAC and other alternatives.

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

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

"Falsehoods programmers believe about names." http://www.kalzumeus.com/2010/06/17/falsehoods-programmers-believe-about-names/

LaLonde. "Sanford Berman: action librarian." http://www.slideshare.net/shinyinfo/sanford-berman-action-librarian

"Berman scorecard." http://www.scribd.com/doc/19082065/Bermanlcshscorecard0409

Coyle, "Organizing knowledge." http://kcoyle.blogspot.com/2011/10/organizing-knowledge.html

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

"DBPedia." https://en.wikipedia.org/wiki/DBpedia

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

McCoppin, "Who's killing the Dewey decimal system?" Chicago Tribune, 18 February 2011. http://articles.chicagotribune.com/2011-02-18/news/ct-met-drop-dewey-20110218_1_dewey-decimal-system-main-library-newer-books

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 8: Cataloging and classification futures: FRBR and RDA.

Learning objectives: Defining a "record." Issues with MARC/AACR2/ISBD and current discovery tools. FRBR. Work, expression, manifestation, item. RDA. Relationships.

Coyle, "Models of bibliographic data." http://kcoyle.blogspot.com/2011/08/models-of-bibliographic-data.html

Coyle, "From MARC to principled metadata." http://kcoyle.blogspot.com/2011/05/from-marc-to-principled-metadata.html

Coyle, "MARC vs RDA." http://kcoyle.blogspot.com/2011/09/marc-vs-rda.html

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

Ockerbloom, "Content and context in concept-oriented catalogs." http://everybodyslibraries.com/2010/01/07/content-and-context-in-concept-oriented-catalogs/

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

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)

Week 9: Searching, search indexes, search engines, and search results

Learning objectives: Index (inverted index), spider/crawler. TF/IDF, precision vs. recall. Search engine optimization, Page Rank (and ways to game it). Relevance ranking, deduplicating, and faceted browsing.

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

Brin and Page. "The anatomy of a large-scale hypertextual web search engine." http://infolab.stanford.edu/~backrub/google.html

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

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

Meyer, "Searching and the indexes." http://search.library.wisconsin.edu/moving-forward/?p=797

Week 10: Relational databases and SQL

Learning objectives: The problems with spreadsheets as information-organization tools. Tables. Primary and foreign keys. Relationship cardinality.

Grussell, "Introduction: The Database Approach." http://db.grussell.org/section002.html (NOT the rest of the page.)

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

Week 11: XML and metadata

Learning objectives: Well-formed and valid XML. Parsers and validators. DTDs and XML Schema. XML authoring tools. Dublin Core. MODS.

"A Gentle Introduction to XML." http://www.tei-c.org/release/doc/tei-p5-doc/en/html/SG.html (Through "An example schema," but keep going if you like.)

See Learn@UW for my introduction to XML syntax.

Week 12: Metadata standards

Learning objectives: METS. PREMIS. TEI. Library-application-specific metadata formats (DSpace, Fedora Commons).

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

Riley, Jenn. "Seeing Standards." http://www.dlib.indiana.edu/~jenlrile/metadatamap/ (Download the poster and read the legend and definitions carefully. With luck, you will be able to see it in all its glory in my office.)

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/

TEI Consortium. "About these guidelines." http://www.tei-c.org/release/doc/tei-p5-doc/en/html/AB.html

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 13: RDF and linked data

Learning objectives: Escaping information silos. RDF. Linked data. URIs as identifiers. The "open world assumption." SPARQL. LOD-LAM. Open data and data-rights issues.

Voss. "Linked data for libraries." http://lodlam.net/2012/08/09/linked-data-for-libraries-video-from-oclc/
Roberts, "An introduction to RDF." http://learnlinkeddata.com/articles/introduction-to-rdf
Roberts, "Understanding RDF serialisation formats." http://learnlinkeddata.com/articles/rdf-serialisation-formats
Chamberlain, "SPARQL tutorial - Introduction." http://data.lib.cam.ac.uk/sparql.php (Introduction, #1, #2)
Stephens, "What to do with Linked Data?" http://www.meanboyfriend.com/overdue_ideas/2012/08/what-to-do-with-linked-data/
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
Library Linked Data Incubator Group, "Final report." http://www.w3.org/2005/Incubator/lld/XGR-lld/
John, "Linked library data in the wild." http://www.slideshare.net/philjohn/linked-library-data-in-the-wild-8593328

Week 14: Cataloging and classification futures: BIBFRAME and replacing MARC

Learning objectives: Problems with MARC/ISBD/AACR2. BIBFRAME initiative and its challenges.

Coyle, "Is MARC dead?" http://www.kcoyle.net/marcdead.html
Code4Lib. "MARC Problems." http://wiki.code4lib.org/index.php/MARC_Problems
Dueber, Bill. "Why programmers hate free text in MARC records." http://robotlibrarian.billdueber.com/why-programmers-hate-free-text-in-marc-records/
Dueber, Bill. "Bad MARC data #1" http://robotlibrarian.billdueber.com/isbn-parenthetical-notes-bad-marc-data-1/
Murray, Peter. "MARC isn't dead, but it is a dead end." http://dltj.org/article/marc-as-dead-end/ (Read the comments too. Strongly recommended: clicking through to Tennant, "MARC Must Die.")
Rochkind, Jonathan. "Broad categories from class numbers." http://bibwild.wordpress.com/2011/04/04/broad-categories-from-class-numbers/
Ford, "LC's Bibliographic Framework Initiative: an update." http://3windmills.com/kefo-swib12-bfi/kefo-swib12.pdf
Miller, "Bibliographic Framework Initiative update: MARC linked data model." http://de.slideshare.net/zepheiraorg/bibliographic-14207718

Week 15: Final-project presentations

ASSIGNMENTS

All assignments are due ***at class time*** on the dates listed. 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

Assignments	Percentage	Due date
Metadata system review	30%	7 May
Choice of system		5 February
Quiz 1	10%	5 March
Quiz 2	10%	9 April
Create two simple resource descriptions	5%	5 February
Turn your descriptions into schema.org microdata	5%	19 February
Find the best WorldCat record (copy cataloging exercise)	5%	26 February
Turn your descriptions into MARC records	5%	5 March
Apply subject headings and call numbers to your records	5%	12 March
Analyze your records for FRBR entities and relationships	5%	19 March
Crosswalk your records to Dublin Core	5%	16 April
Re-encode your records as valid MODS	5%	23 April
Class participation	10%	

No extra credit opportunities are available in this class.

Grading scale: 94-100 A; 88-93 AB; 82-87 B; 77-81 BC; 72-76 C

Quizzes

These will be on vocabulary and concepts, and will consist of labeling and short-answer questions. Each quiz will be at least 12 points of work; full credit is 10 points. Points above 10 will not be counted.

Class participation

Students earn points for participation by attending class regularly, contributing to class discussions, asking questions, and participating usefully in in-class group work.

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

Create two simple resource descriptions

Choose two information objects that you will describe and re-describe throughout the semester. One MUST be a print book; the other must NOT be print (but can be music, film, unpublished manuscript, photography, realia, digital object, etc). Provide a description of the object that you believe suffices to help someone discover it through browsing or searching. Post the result in the designated Learn@UW forum.

Turn your descriptions into schema.org microdata

Using appropriate item types from `http://schema.org/docs.schemas.html`, turn your descriptions into microdata on a valid HTML page. (Hint: find and work from examples!) 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. Upload the resulting page to your MyWebSpace; post its URL to the designated Learn@UW forum.

Find the best WorldCat record

For the book given you in class, find as many records in WorldCat as you can. Compare them: why are there so many, and what are the differences among them? Make a list of the distinctions you find. Find the record you think is best for the exact book mentioned in class. Post your list and the permalink for the "best record" you chose to the designated (closed) Learn@UW forum.

Turn your descriptions into MARC records

Turn your descriptions into MARC records, capturing as much of the information you have already amassed as you can. You will be marked down for MARC and ISBD syntax/punctuation errors. You are expected to perform name authority control (though you obviously will not be marked down for nonexistent authority records). You DO NOT (yet) need to add subject headings or classification numbers. Post your records in the designated (closed) Learn@UW forum.

Apply subject headings and call numbers to your records

Apply three appropriate Library of Congress Subject Headings to each of your records. Assign call numbers to your objects in Library of Congress Classification and Dewey Decimal Classification. Do not forget to add Cutter numbers! (Honor system: don't cheat by copying from a library catalog, please. This assignment is about process as much as product.) Add these to your MARC records; post them to the designated (closed) Learn@UW forum.

Analyze your records for FRBR entities and relationships

List all Group 1 and Group 2 entities important to your records, with appropriate identifiers (authority strings, URLs) when you can find them. (Not "Manifestation," but "Manifestation with the ISBN…") List all relevant Work(s) and Manifestation(s); look for additional Expression(s) and list any you find. Post your results to the designated (closed) Learn@UW forum.

Crosswalk your records to Dublin Core

Capture as much information as you can about your items in the DSpace batch-import XML metadata format, which is based on Dublin Core. Put the resulting XML files in the designated Learn@UW dropbox.

Re-encode the data as MODS

Crosswalk your MARC records to valid MODS, recapturing as much information as possible. Put the resulting records in the designated Learn@UW dropbox.

Review of a Metadata System (group project)

For this project, you will work in a small group - maximum number of members is 5. Your group will select an information retrieval system used in a particular institution (e.g., library catalogs, archival finding aids, museum registers, digital libraries). You will describe, analyze and evaluate it, and produce a report on your findings. Once you have selected a specific system, please submit your choice to the instructor for approval.

Describe and analyze the metadata standards, elements and contents, available through the system. Evaluate the system from the users' perspective. Your group will submit a written report with the following sections:

1. Metadata standards: What metadata schema is in use?
2. Metadata elements: What are the elements included for each item described? List and describe each element.
3. Metadata contents: What is in each element? How is it presented? Any content standards, rules are used for the creation of metadata? What kind of searchable data is there?
4. Metadata encoding: How is the metadata encoded? Any standards are used?
5. Usability: Try searching and review how a metadata record is displayed. Give your opinions on how well the system works for the untrained user. For example,
    a. Is it possible to tell what search terms to use on first encountering the system?
    b. Is it easy to understand what is on a record?
    c. Does a record provide enough metadata that users would want?
    d. If it is not easy to use the system or to understand the metadata, what can be done to improve it?
6. Prepare a searching guide for the system, which would instruct a novice user on the most effective way to search the collection.

Groups will submit and present their reports on the last class meeting of the semester.

At the end of the semester, everyone will post to a locked Learn@UW forum a short "360 evaluation" of the other members of their group: a brief description of the contributions of all other group members. I will use this information to assign individual project-participation grades as I see fit; only I will see the posts.

On group projects: The idea that group projects are uniquely designed to torture library school students is a snare and a delusion. Librarianship includes immense amounts of collaborative work, from local committees and task forces to involvement in national professional organizations and everything in between. None of the obstacles to working in groups – scheduling, free riders, personality conflicts – disappears when you receive your degree. If you are not good at working in a team, now is the time to learn!

SLIS Goals	551 Objectives	551 Measurable Outcomes
2a. Students evaluate and debate information policy and ethics applicable in local, national, or global contexts.	1. Have an understanding of the major standards, systems, and tools used for organizing and cataloging recorded information 2. Acquire the fundamental understanding and skills to effectively use and critically evaluate systems for information organization and retrieval	Week 7 discussion and class exercise on locating and redressing potentially problematic vocabularies: reflected in participation grade.
3a. Students organize and describe print and digital information resources.	1. 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.
3b. Students search, select, and evaluate print and digital information resources.	2. Acquire the fundamental understanding and skills to effectively use and critically evaluate systems for information organization and retrieval 3. Gain introductory experience in applying standards for creating and encoding metadata	Final project.
3c. Students analyze information needs of diverse individuals and communities.	2. Acquire the fundamental understanding and skills to effectively use and critically evaluate systems for information organization and retrieval	Week 7 discussion and class exercise on locating and redressing potentially problematic vocabularies: reflected in participation grade.
4a. Students participate effectively as team members to solve problems.		Final project: 360 evaluation.
4b. Students demonstrate good oral and written communication skills.		Final project presentation and report.