# Preliminary INF 384C Syllabus Spring 2015

### Basic information

*Instructor:* Melanie Feinberg

*Class location:* UTA 1.212 (“medium classroom”)

*Date and time:* Wednesdays, 12:00 p.m. to 3:00 p.m.

### Instructor information

*E-mail:* feinberg@ischool.utexas.edu

*Office:* UTA 5.446

*Office phone:* 512-471-8487

*Office hours:* TBA (and always by appointment)

### Introduction

The course catalog description of this class is as follows:

*Introduction to general principles and features of organizing and providing access to information, including varieties and numbers of information-bearing objects, different traditions of practice, user concerns, metadata and metadata formats, document representation and description, subject access, and information system features and evaluation.*

This course provides a general introduction to the organization of information, concentrating on the core operations of describing, grouping, arranging, and relating objects. While the course will focus on the organization of documents, or bibliographic information, the objects most commonly organized in libraries and archives, we will not be unduly concerned with particular implementations for any specific institution. In other words, you will not learn traditional library cataloging or archival description in this class. You will, however, learn the principles that form the basis for all such systems, applicable to any sort of environment, from digital asset management to Web site information architecture.

Accordingly, the assignments for the course **emphasize the application of organization principles in desigining mechanisms for organizing information.** Note, however, that the course concentrates on conceptual aspects of design, not technical ones; you will not learn implementation encodings (such as MARC for cataloging or RDF/XML for the Semantic Web) in this class.

## Policies

### Instructor communication

E-mail is the official mode of communication for the university and the most reliable means of contact for me. I do my best to answer e-mail within a two days of receipt. If you do not receive a response after a few days, please follow up. It is always helpful if your e-mail includes a targeted subject line that begins with “INF 384C.” *Do not use the messaging facilities in Canvas; these messages do not arrive in my e-mail in-box.*

### Late assignments

*Late assignments are not acceptable.* For each day that an assignment is late, ten percent of the possible points will be deducted from the score, rounded up (that is, if the descriptive schema, worth 30 points, is one day late, the maximum number of points for the late assignment is 27, or 30-3). *Students who anticipate difficulties with completing assignments on time should consult with the instructor as soon as possible so that alternate solutions can be discussed.* When negotiated in advance, arrangements can often be made.

### Attendance

Attendance is not taken. You do not need to inform me of absences, nor do you need to “make up” anything if you are absent. While participation is an important part of your grade, and attendance is important, there are no specific requirements for mandatory attendance.

### Students with disabilities

Students with disabilities may request appropriate academic accommodations from the Division of Diversity and Community Engagement, Services for Students with Disabilities, 512-471-6259. (Or see their Web site.)

### Academic integrity

I follow University of Texas standards for academic integrity, as documented on the Dean of Students’ Web site and in associated materials. It is very important that academic work is conducted according to accepted ethical standards. If you uncertain whether an action is in keeping with academic integrity standards, please consult with me before undertaking it.

### Assignments

There are three assignments:

* Designing a metadata schema.
* Designing a subject taxonomy.
* Generating and analyzing a set of aggregated metadata records.

*Assignment Due Date*

Descriptive schema Feb. 25, at the end of class.

Taxonomy April 8, at the end of class.

Metadata generation and aggregation analysis May 6, at the end of class.

All written assignments should be printed and brought to class. They are due at 3 p.m. (the end of class). There is no need to send me an electronic copy.

### Descriptive schema

In this assignment, you will define a set of entities, articulate a motivating purpose for describing them, and then outline a structure of attributes and associated values to systematically represent your entities as metadata. You will use your schema to create metadata for five varied instances of entities in your set.

*Assignment details*

First, you will define a group of entities to describe. This could be anything. Some examples have been:

* Austin-area swimming holes.
* Spoons.
* Camps for the visually impaired.
* Jerky.
* Representations of “the Aztec” in film.
* Knitting patterns.
* Web comics.
* Yoga poses.
* Intellectual movements in literary criticism.

Via e-mail, submit your idea for a group of entities by *5 p.m. on Wednesday, February 4,* for instructor approval.

When defining your class of entities, give a sense of the range of possible instances that would fit into your set. For example, is a photo in the airport an acceptable instance of Iceland vacation photo? Is an establishment that serves burritos but no tacos an acceptable instance of a taqueria? Thinking about border cases will help you create attributes that apply equally to all members of your set of entities.

Next, you will articulate a purpose and associated target audience to motivate your description. For example, you might want to share your Iceland vacation photos with your photography club, or you might want to advance the cause of Icelandic hegemony throughout the world. You might want to help novice knitters find patterns that make nice gifts, or you might want to interrogate stereotypes latent in “Aztec” imagery. You can see how each situation might suggest a different set of attributes for the same entity set.

You will then articulate a set of 10-15 attributes to define your entities in support of the purpose. You will label and describe each attribute in sufficient detail so that an “outside indexer” (that is, someone who is merely given your schema and its documentation) could assign values for entities of the type that you have described. For each attribute, you will set parameters for acceptable values and provide guidelines that show how values should be expressed.

Once you have sufficiently defined your attributes, use the structure that you have developed to preliminarily describe five instances to represent both central and border cases of your entity set. If there are cases where you are unable to satisfactorily describe an instance, use this as an opportunity to revise the schema and clarify your attribute definitions. (You might even need to clarify the boundaries of your group of entities and sharpen its description.) Then use your revised schema to create five final descriptions for your entity instances.

Finally, write a brief critical reflection on your design process and resulting product. You might discuss questions such as the following:

* As you designed your initial attribute set, how did you ensure that your attributes contributed to your defined purpose for description? If this process was difficult, what made it so?
* How did your schema change, if at all, following the initial test? What necessitated the changes, and why weren't you able to anticipate these issues in advance?
* Did your perception of your group of entities change after attempting to describe the five instances? How did your conception of this set change, and what caused your perspective to shift?
* Do you feel like your final schema represents the group of entities well? Why or why not?
* What was difficult about this design project?
* What might you keep in mind for subsequent metadata design projects?

Note that the point of this reflective essay is not to justify why your schema is awesome. Clearly, it is awesome, and you don’t need to persuade me of that. Instead, the goal of this essay is to explore how the practical experience of designing a schema provokes insight onto the conceptual foundations of information organization.

Note: These are *examples* of questions that you *might* discuss. To create a concise yet cohesive paper, you will need to concentrate on a few design issues of particular relevance to your project. Do not merely answer the questions here.

*Deliverables*

Your final assignment should include:

* A set of instructions for an outside indexer to use in adding descriptions of new instances to the set of records that you have already created. These instructions should include the following two elements:
  + A few paragraphs to describe your group of entities, your purpose for describing them, and how your defined attributes work to facilitate the purpose.
  + Your attribute descriptions, value parameters, and associated guidelines for using the schema to describe the entities. The description for each attribute should follow a consistent format. (You can use something similar to the NISO standard for Dublin Core metadata elements or devise your own format. You may use tables if you wish.)
* Your descriptions of five instances. Use a consistent format for each record (perhaps a table for each instance).
* Your critical reflection. This should be written in narrative form, as a cohesive paper of 750-1000 words.

*Grading criteria*

A successful assignment will exhibit these characteristics:

* The reader understands what constitutes a member of the defined set, who makes up the audience, and the goals associated with the purpose.
* The defined attributes effectively represent the selected entities in the context of the described purpose, and the value space effectively represents the extent of the attributes. For example, when describing yoga poses for students, an attribute that indicates level of difficulty might be appropriate. However, such an attribute might seem less appropriate if describing yoga poses in relation to the history of Hindu thought and culture. In addition, the values described for the potential level of difficulty attribute for yoga poses should encompass the full range of possibilities at an appropriate level of detail for the audience and purpose.
* The attribute descriptions, value parameters, and associated guidelines can be easily understood by outside indexers and applied to describe actual entities accurately and comprehensively within the context of the selected purpose. It is clear how to apply the descriptive schema to both standard and border cases.
* The descriptions of entity instances follow the created definitions and guidelines accurately and represent a range of potential entities, from typical, standard examples to less common ones that stretch the potential definition of the entity set.
* The critical reflection thoughtfully considers the design process, product, or both, using the experience of creating the descriptive schema to productively engage larger issues of theory and practice (that is, the reflection does not merely summarize or justify the design process or product; it interrogates it).
* The assignment follows a logical document structure, is clearly written, and uses correct grammar and punctuation.

In grading the assignment, the total number of points are equally distributed amongst these criteria (so 5 points possible per item).

### Subject taxonomy

You will define, label, and relate a set of 30-40 concepts to serve as a means of organizing documents. Your taxonomy will elucidate the conceptual landscape for a single subject area, which you will select.

*Assignment details*

First, you will select a subject area that you will represent in your taxonomy. Because your set of categories will be small, this should be something quite specific, such as:

* Veganism.
* Javanese gamelan musical tradition.
* Sustainable urban planning.
* Costume design.
* Nuclear power.
* Scuba diving.

Via e-mail, submit your idea for a subject area by *5 p.m. on Wednesday, March 4,* for instructor approval.

Next, even if you are an expert in the subject area that you selected, do some basic research. Look at a variety of documents in the subject area and see what they include and how they differ. This should give you a better sense of your subject area and its extent. Based on this research, you might need to define your subject area more narrowly (for example, concentrating on ecofeminism as one type of feminist theory). You should also at this point articulate a purpose and associated user group for your taxonomy. You might create a taxonomy to organize documents on veganism in order to help persuade current meat eaters to become vegans, for example, or you might alternately create a veganism taxonomy for current vegans that focuses on making the most out of the vegan lifestyle. In other words, your audience and purpose will help you interpret the subject in a specific way, which you will then represent in a system of related categories. (Or you can say that your selection of an audience and purpose, in combination with your research into the subject literature, will help form the “semantic warrant” of your classification; we will read about warrant in Beghtol’s article in Week 7.) Write a few paragraphs to describe your intended audience, associated purpose, and resulting interpretation of the subject. This will help you as you develop and shape the taxonomy.

Now, create a system of 30-40 related categories to represent the subject. You might begin generating potential classes by looking through documents in your subject area to “harvest” concepts as “source material.” (This is a bottom-up approach.) Or you might begin by defining some top-level categories in advance and working downwards (a top-down approach).

Relate your categories in one or more hierarchies. (You might choose a truly synthetic, faceted approach, as discussed in Week 8, but this is not a requirement.) You will need to manipulate your categories, aggregating, splitting, rearranging, and relabeling them, in order to generate a cohesive structure. Pay attention to the order within sibling categories at any particular level; you should have a reason for each arrangement. You may choose to relate categories associatively (across hierarchies) if you like.

You can represent the taxonomy in any way that makes sense to you: indented text is fine as long as the levels are clear. You can also use diagrams.

We will have some time in class to share work on your taxonomies with others and to have individual consultations with the instructor.

After you have the classified structure established, you will create an alphabetical list of the categories in your taxonomy. For each category, you will provide a brief definition of the category in the context of your taxonomy. For example, in a veganism taxonomy to persuade meat eaters to adopt a vegan diet, the concept “animal fat” might be significant primarily as a contributor to cardiac disease, and thus as a contributing factor in the health risks of a meat-eating diet. Where warranted, you should also include a usage note that clarifies what kinds of resources should be assigned to the category. For example, a veganism taxonomy for current vegans, designed to increase community, might note that resources that discuss health risks associated with veganism should be classed under “establishment propaganda.” As you proceed, you may wish to test your structure and associated documentation by attempting to classify a variety of documents in the subject area.

You will also write an introduction to the taxonomy that briefly describes your subject area in general, along with your target audience, intended purpose, and associated interpretation of the subject. After reading this description, an outside indexer should be able to understand the rationale behind the taxonomy and use it to assign documents to categories as appropriate. Your introduction should include brief guidelines for using the taxonomy: can resources be assigned to multiple categories, for example?

Finally, you will write a concise critical reflection on your design process and resulting product. In writing this essay, you might consider questions such as:

* How did you determine when a potential category was a likely candidate to be included in the taxonomy?
* What led you to decide on the structural implementation that you ended up with?
* Does your resulting taxonomy seem complete and consistent? Why or why not?
* What might you do differently in your next classification design project?

As with the descriptive schema assignment, note that the point of this reflective essay is not to justify why your taxonomy is awesome. Instead, the goal of this essay is to explore how the practical experience of designing a taxonomy provokes insight onto the conceptual foundations of information organization.

Also note that these are *examples* of questions that you *might* discuss. To create a concise yet cohesive paper, you will need to concetrate on a few design issues of particular relevance to your project. Do not merely answer the questions here.

*Deliverables*

Your final assignment should include:

* The taxonomy introduction, which describes the subject area in general; the target audience, intended purpose, and associated interpretation of the subject to be represented in this taxonomy; guidelines for assigning resources to categories.
* The classified structure, which shows all the categories that you created and their relationships.
* An alphabetical representation of each category, with an associated definition and, where appropriate, a usage note.
* Your critical reflection. This should be written in narrative form, as a cohesive paper of 750-1000 words.

*Grading criteria*

A successful taxonomy will exhibit the following characteristics:

* The introduction succinctly articulates a clear and forceful case for the selected subject interpretation and summarizes, in a way understandable to outside indexers, general principles by which the interpretation is to be implemented via placing resources within the assembled categories.
* The classified stucture persuasively represents the subject according to the selected interpretation, via a cohesive strategy for including, naming, and arranging categories. Hierarchical and other category relationships either follow standard design principles (such as using a single principle of division at one level of hierarchy, with jointly exhaustive and mutually exclusive sibling classes) or the rationale for violating these standard principles is cogently explained.
* For each category, the alphabetical structure adequately clarifies, through definitions and usage notes, how the subject interpretation is to be supported through the assignment of resources to the category.
* After reading the introduction, classified structure, and alphabetical structure, an outside indexer should be able to accurately assign resources to the established categories, so that the selected interpretation maintained and solidified.
* The critical reflection engages the design process and product as a means to illuminating areas of both theory and practice (that is, it examines and evaluates the process and product, instead of simply summarizing them).
* The assignment follows a logical document structure, is clearly written, and uses correct grammar and punctuation.

In grading the assignment, the total number of points are equally distributed amongst these criteria (so 5 points possible per item).

### Metadata generation and aggregation analysis

Design of organizational systems doesn’t end with a schema to describe a set of entities, and a taxonomy to relate them to each other. We could say that design continues as these structures are used to describe and relate different collections of resources, as the concepts represented through schemas, and through elaborately structured taxonomies (or controlled vocabularies of values), are given further shape and meaning through use. In their paper about creating a metadata schema for video games (which we read in Week 3), Jin Ha Lee and colleagues describe some of the semantic diversity encountered in the application of certain attributes. For Lee and colleagues, this diversity constitutes a problem to eventually be rectified. Is interpretive diversity a problem? Is it sometimes a problem? Is it sometimes useful? How can we understand interpretive diversity in metadata generation better, and how can we accommodate this diversity in design of schemas and taxonomies, in application of schemas and taxonomies, and in the use of the resulting metadata? In the last portion of our class, we will look at application of organizational systems, across time, across collections, across different sorts of use situations, and we will consider how we might comprehend and take account of semantic variation of all sorts.

To do this in a concrete way, we will create a course “database” of metadata records using the current version of the complex video game schema devised by Lee and colleagues, along with several taxonomies (or controlled vocabularies) created for certain attributes in the schema. But we won’t create this database communally, discussing our interpretations of the schema and its vocabularies as we go along. We’ll create this database as many collections are created, through aggregating records that have the same conceptual structure. That is, we will each create a set of records independently, and these records will be merged into a single collection. We will then examine, assess, and interpret this aggregate collection to determine both the extent of semantic diversity across the records and the function of this diversity. If different records apply attributes differently, what are the effects?

This assignment has several steps:

1. Create metadata records for 10 video games using the latest version of the video game metadata schema (this is assigned reading for Week 3 of the class).

You will create 7 records for video games that you select yourself. These can be anything that you like, for any sort of device, and of any level of complexity.

You will create 3 records for video games that have been selected for the entire class.

You will submit your metadata in two ways.

The first is with a spreadsheet. You will e-mail the spreadsheet to the instructor.

The second is with an online system, Collective Access. Information for accessing and using Collective Access will be provided in class.

It is probably a good idea to save your data in the spreadsheet first, and then transfer it to Collective Access.

Collective Access should make it easier to search and browse the aggregated records. But because it will be a new setup, the spreadsheets will provide both a backup and an alternate access mechanism.

1. Examine and assess the aggregated set of metadata records, identifying sources of interpretive diversity and considering their effects.

How might you account for these effects in designing a schema and associated vocabularies? In creating records with a schema and associated vocabularies? In planning an aggregation? (Note that you may find these effects to be positive and want to accentuate them, or you might find them negative and think of how to minimize them...it’s up to you and how you read the aggregated collection.)

1. In class, discuss your initial interpretation of the aggregated collection with others.

Do you see the same sources of diversity? Do you characterize the effects in the same way? Would you account for these effects similarly?

1. Write an essay that considers each of these three activities—record creation, your initial assessment of the aggregation, and your assessment after discussion with other record creators—in formulating a position on interpretive diversity in metadata generation and on how we might then perceive and work with metadata aggregations.

*Deliverables*

While the primary outcome of this assignment is the final essay, there are several preparatory parts.

Via e-mail to the instructor, you will indicate your 7 personal video games to describe by *Wednesday, February 25, at 5 p.m*. It does not matter if people select the same games, or what kinds of games you select. You will likely need access to either the game itself or some fairly detailed information about it in order to generate full metadata records; consider the schema and its elements when making your selections.

You will submit the 10 records for the video games that you selected by *Wednesday, March 25, at 5 p.m.* You will input data into Collective Access online (instructions provided in class). You will send your backup spreadsheet to the instructor via e-mail.

**Because others will be relying on this information in order to proceed with this assignment, this task cannot be late.**

If you do not submit your records on time, you will get a ZERO for this assignment. Plan accordingly and do NOT WAIT UNTIL THE LAST MINUTE to attempt doing this (the video game schema is very complicated, and you might need to do some research to complete your records).

Final essay

The primary element of this assignment is an essay of between 2,000 and 3,000 words. Your final essay should make some reference to the following:

* Your experience of using Lee’s schema to create your own records.
* Your initial reading of the aggregated collection prior to discussing it with others in the class.
* Your reading of the aggregated collection after discussing it with others in the class.
* Concepts from course readings and other activities.

Ultimately, your final essay should *constitute an argument for how to approach intepretive diversity* when creating and using metadata schemas and associated controlled vocabularies. Consider schema designers, catalogers, aggregators, and database users in your discussion. Should everyone do whatever they can all the time in order to diminish interpretive diversity, even though this is ultimately impossible, as metadata practice guidelines often imply? If so, what should be done and how? If not, is semantic diversity nothing to worry about? Ever? Should there never be any standards then? Etc.

This essay is your chance to develop a real, nuanced position on an important area of information practice, a position that can inform many subsequent professional activities.

*Grading criteria*

A successful essay on metadata generation and aggregation will exhibit the following characteristics:

* Develops an actual position on interpretive diversity in aggregated collections that clarifies the effects of such diversity, assesses the impact of those effects, and suggests how to either minimize any negative effects or maximize positive ones.
* Incorporates, as evidence for the argument, reflections on your activities: creating records from Lee’s schema, initial reading of the aggregated collection of records, synthesis of your own interpretation with the interpretations of others.
* Appropriately synthesizes material from course readings to support your argument.
* Anticipates and responds to potential objections to the argument and its associated suggestions for action.
* Follows a logical document structure, is clearly written, and uses correct grammar and punctuation.

In grading the assignment, the total number of points are equally distributed amongst these criteria (so 6 points possible per item).

### Course grading

Descriptive schema: 30 points

Subject taxonomy: 30 points

Metadata generation and aggregation: 30 points

Class participation: 10 points

100 points total.

Class participation is a vital component of the course. I expect everyone to come to class having read the required material, prepared to critically enage the concepts. Participation is not based on attendance, but *excessive* absences will of course be noticed.

For me, student participation is not about having the correct answer, nor about demonstrating that you have done the readings. It is especially not about saying what you think I want to hear. (In fact, I am happiest when students disagree with me copiously.) Participation is about us all being able to engage in honest dialogue and learn from each other. Asking questions and having the class work through something that you find perplexing are *excellent* forms of participation.

I will use the following schedule in calculating final grades:

|  |  |  |
| --- | --- | --- |
| A = 95-100 | A- = 90-94 | B+ = 85-89 |
| B = 79-84 | B- = 74-78 | C+ = 69-74 |
| C = 64-69 | F < 64 |  |
|  |  |  |

### Calendar: Readings and Assignments

*Note:* Optional readings are exactly that. This extra material provides some additional nuance to the week’s topics for those with time and interest.

| **Date** | **Assignments** | **Themes** | **Readings** |
| --- | --- | --- | --- |
| January 21  Week 1 |  | Introduction to the course and the topic.  What is information organization?  What are systems for organizing information? | Borges, 1962  Borges, 1964  Glushko, introduction |
| January 28  Week 2 |  | Entities (objects, resources)  What kinds of things are we organizing? What is in and what is out? | Buckland  Wilson (ch. 1)  Kent (ch. 1)  IFLA (Read about Group 1 entities only: pages 13-14, 17-24, 31-49)  Bates (selected portions only) |
| February 4  Week 3 | *Via e-mail, submit your idea for a set of entities to describe with your schema by 5 p.m.* | Schemas: sets of attributes and values    What are our dimensions for description? | Wilson (ch. 2)  ANSI/NISO Z39.85-2012 (Dublin Core metadata element specification)  CDP Metadata Working Group, 2006  Jin Ha Lee, et al, 2013  Jin Ha Lee, et al (GAMER group), latest schema (TBA)  Gilliland |
| February 11  Week 4 |  | Control of values  The equivalence relationship: when are two values functionally the same? | Furnas  Berlin, et al  Rader, 2010  Taylor and Joudrey  Rosenfeld and Morville |
| February 18  Week 5 | *Peer work groups: Discuss work on descriptive schemas in the second half of class; bring your rough drafts.* | Rules and standards; introduction to interoperability | Greenberg, 2009  Zeng and Chan, 2009  Elings and Weibel, 2007  Chan and Zeng, 2006 |
| February 25  Week 6 | *Descriptive schema due in class at 3 p.m.*  *Via e-mail, submit the names of 7 video games to describe with Lee’s schema for the metadata generation and aggregation assignment by 5 p.m.* | Subjects: a type of value of particular interest in information science  Indexing (or classifying): specifying a document’s subject | Wilson (ch. 5)  Taylor and Joudrey  Fidel  ISO 5963-1985  *Optional*  Hjorland (1992) |
| March 4  Week 7 | *Via e-mail, submit your idea for a subject area for your taxonomy by 5 p.m.* | Subject languages (a set of values to control subject assignment)  How do we go about identifying the concepts to include in a taxonomy? And what do we do when we find them?  Relationships between subject categories: hierarchical and associative | Ranganathan (ch. 1-5)  Kwasnik  Beghtol  Hjorland and Albrechtsen  Broughton |
| March 11  Week 8 | *Discussion class: come prepared to talk about the readings in a seminar-style discussion, and with at least one substantive question prepared to generate conversation and debate* | Subject languages 2: alternate foundations | Mai, 2011  Hur-Li Lee  Mayor and Robinson, 2014b  *Optional*  Mayor and Robinson, 2014a |
| March 18 |  | Spring break |  |
| March 25  Week 9 | *Submit metadata records for your video games by 5 p.m. E-mail the spreadsheet to the instructor, and input the same data into Collective Access online (instructions provided in class)* | Faceted classification  Creating a complex category from simple parts | Hunter (ch. 1-5)  Vickery (ch. A-M only)  Yee, Swearingen, and Hearst  Broughton |
| April 1  Week 10 | *Discussion class: come prepared to talk about the readings in a seminar-style discussion, and with at least one substantive question prepared to generate conversation and debate.*  *Peer work groups and instructor check: Discuss subject taxonomy project in second half of class; bring your classified structure. .* | Applying classification schemes and other metadata: in practice | Wild, 2009  Nakasone and Sheffield, 2013  Pollack, 2013  Jin Ha Lee, Clarke, and Perti, in press |
| April 8  Week 11 | *Subject taxonomy due in class at 3 p.m.*  *Discussion class: come prepared to talk about the readings in a seminar-style discussion, and with at least one substantive question prepared to generate conversation and debate.* | Applying classification schemes and other metadata: across time | Tennis, 2012  Buckland, 2012  Carlyle, 2015  Bowker and Star, 2000 |
| April 15  Week 12 | *Discussion class: come prepared to talk about the readings in a seminar-style discussion, and with at least one substantive question prepared to generate conversation and debate.* | Applying classification schemes and other metadata: across institutions and collections | Waigley, Gaches, Park, 2010  Windnagel, 2014  Kansa, Kansa, and Arbuckle 2014  *Optional*  Arbuckle, et al 2014 |
| April 22  Week 13 | *Discussion class: come prepared to talk about the readings in a seminar-style discussion, and with at least one substantive question prepared to generate conversation and debate.*  *Peer work groups: Be prepared to discuss the aggregated video game metadata database in the second part of class.* | Applying classification schemes and other metadata: without rules (social media) | Furner, 2009  Chan, 2007  Klavans, LaPlante, and Golbeck 2014  Srinivasan et al, 2009  *Optional*  Good, Tennis, and Wilkinson, 2009 |
| April 29  Week 14 | *Discussion class: come prepared to talk about the readings in a seminar-style discussion, and with at least one substantive question prepared to generate conversation and debate.*  *Peer work groups: Be prepared to discuss the aggregated video game metadata database in the second part of class.* | Applying classification schemes and other metadata: values, ethics, fairness | Mai, 2013  Olson, 2007  Tennis, 2013 |
| May 6  Week 15 | *Metadata generation and aggregation essay due at 3 p.m.*  *Discussion class: come prepared to talk about the readings in a seminar-style discussion, and with at least one substantive question prepared to generate conversation and debate.* |  | Doty |

### Readings

Course readings are available via Canvas.

Although it is not assigned, for reference purposes, you might be interested in:

Glushko, Robert J. (ed). *The Discipline of Organizing.* MIT Press, 2013.

The book is available from Amazon and other sources.

### Complete bibliography

Arbuckle, B., et al. (2014) Data sharing reveals complexity in the westward spread of domestic animals across neolithic Turkey. *PLOS One* 9(6): e99845. doi:10.1371/journal.pone.0099845.

Bates, Marcia J. What is a reference book: a theoretical and empirical analysis. *RQ* 26 (Fall 1986): 37-57. (Selected excerpts only.)

Beghtol, Clare (1986). Semantic validity: concepts of warrant in bibliographic classification systems. *Library Resources & Technical Services* 30:109–125.

Berlin, Lucy M., Robin Jeffries, Vicki O’Day, Andreas Paepcke, and Cathleen Wharton. (1993) Where did you put it? Issues in the design and use of a group memory. *Proceedings ACM CHI 1993,* 23-30.

Borges, Jorge Luis. (1962) The library of Babel. In *Labyrinths.* Edited by Donald A. Yates and James E. Irby. New York: New Directions.

Borges, Jorge Luis. (1964) The analytical language of John Wilkins. In *Other Inquisitions 1937-1952.* Translated by Ruth Simms. Austin: University of Texas Press.

Bowker, Geoffrey, and Susan Leigh Star. (2000*) Sorting Things Out: Classification and Its Consequences.* Cambridge, MA: MIT Press. Chapter 3, 106-133.

Broughton, Vanda. 2006. *Essential Thesaurus Construction.* London: Facet Publishing.

Buckland, Michael. (1997) What is a “document”? *Journal for the American Society of Information Science* 48 (9): 804–809.

Buckland, Michael. (2012) Obsolescence in subject description. *Journal of Documentation* 68(2): 154-161.

Carlyle, Allyson. (2015) The policeman’s beard was what? Representation and reality in knowledge organization and description. *Proceedings of the iConference 2015.*

Chan, Lois Mai, and Marcia Lei Zeng. (2006). Metadata interoperability and standardization: a study of methodology, part 1. *D-Lib* 12(6). Available at: http://www.dlib.org/dlib/june06/chan/06chan.html

Collaborative Digitization Program (CDP) Metadata Working Group. (2006). Dublin Core Metadata Best Practices version 2.1.1.

Chan, Seb. 2007 http://www.archimuse.com/mw2007/papers/chan/chan.html

Elings, Mary, and Gunter Weibel. (2007) Metadata for all: descriptive standards and metadata sharing across libraries, archives, and museums. *First Monday* 12(3). Available at: http://firstmonday.org/article/view/1628/1543

Fidel, Raya. (1994) User-centered indexing. *Journal of the American Society for Information Science* 45 (8): 572–576.

Furnas, G.W., Thomas K. Landauer, L. M. Gomez, and S. T. Dumais. (1987) The vocabulary problemin human-system communication. *Communications of the ACM* 30(11): 964–971.

Furner, Jonathan. 2009. Folksonomy. In Bates, Marcia, and Mary Niles Maack (eds). *Encyclopedia of Library and Information Sciences.* 3rd ed. CRC Press.

Greenberg, Jane. (2009) Metadata and digital information. In Bates, Marcia, and Mary Niles Maack (eds). *Encyclopedia of Library and Information Sciences.* 3rd ed. CRC Press.

Gilliland, Anne. Setting the stage. In *Introduction to Metadata.* 3rd ed (online edition). Edited by Murtha Baca. Available at: http://www.getty.edu/research/conducting\_research/standards/intrometadata/setting.html

GAMER Group. (Lee, J.H., and colleagues.) (2014). Video game metadata schema version 2.0.

Glushko, Robert J. (ed). *The Discipline of Organizing.* MIT Press, 2013. Introduction, 1-19.

Good, Benjamin, Tennis, Joseph, and Wilkinson, Mark. (2009). Social tagging in the life sciences: characterizing a new metadata resource for bioinformatics. *BMC Bioinformatics* 10. Available at: http://www.biomedcentral.com/1471-2105/10/313

Hjorland, Birger. (1992) The concept of “subject” in information science. *Journal of Documentation* 48(2): 172–200.

Hjorland, Birger, and Albrechtsen, Hanne. (1995) Toward a new horizon in information science: domain-analysis. *Journal for the American Society of Information Science* 46 (6): 400–425.

Hunter, Eric J. (2002) *Classification made simple.* 2nd ed. Aldershot, England: Ashgate. (Ch. 1-5)

IFLA. Functional Requirements for Bibliographic Records final report. Available at: http://www.ifla.org/VII/s13/frbr/frbr.pdf (Read about Group 1 entities, sections 3.1-3.11, and sections 4.1-4.4 only.)

International Organization for Standardization (ISO) (1985) *Documentation. Methods for Examining Documents, Determining their Subjects and Selecting Indexing Terms*. International Organization for Standardization. ISO 5963-1985.

Kansa, E., Kansa, S., and Arbuckle, B. (2014) Publishing and pushing: mixing models for communicating research data in archeology. *International Journal of Digital Curation* 9(1): 57-70.

Kent, William. (1978) *Data and reality: basic assumptions in data processing reconsidered.* Amsterdam: North Holland Press. (Ch. 1, Entities.)

Klavans, Judith, LaPlante, Rebecca, and Golbeck, Jennifer. (2014) Subject matter categorization of tags applied to digital images from art museums. *Journal of the American Society for Information Science and Technology* 65(1): 3-12.

Kwasnik, Barbara. (1999) The role of classification in knowledge representation and discovery. *Library Trends 4*8(1) 22-47.

Lee, Hur-Li. (2012) Epistemic foundation of bibliographic classification in early China: A ru classicist perspective. *Journal of Documentation* 68(3): 378-401.

Lee, J. H., Cho, H., Fox, V., and Perti, A. (2013). User-centered approach in creating a metadata schema for video games and interactive media. *Proceedings of the 13th ACM/IEEE-CS Joint Conference on Digital Libraries (JCDL),* 229-238.

Lee, J.H., Clarke, R., and Perti, A. (In press). Empirical evaluation of metadata for video games and interactive media. *Journal for the Association of Information Science and Technology (JASIST).*

Mai, J-E. (2011) The modernity of classification. *Journal of Documentation* 67(4) 710-730.

Mai, J-E. (2013) Ethics, values, and morality in contemporary library classifications. *Knowledge Organization* 40(4): 242-253.

Mayor, Charlie, and Lyn Robinson. (2014a). Ontological realism, concepts, and classification in molecular biology. *Journal of Documentation* 70(1): 173-193.

Mayor, Charlie, and Lyn Robinson. (2014b) Ontological realism and classification: structures and concepts in the gene ontology. *Journal of the American Society for Information Science and Technology,* in press.

Nakasone, Sonoe, and Carolyn Sheffield. (2013). Descriptive metadata for field books: methods and practices of the field book project. D*-Lib* 19(11/12). Available at: http://www.dlib.org/dlib/november13/nakasone/11nakasone.html

National Information Standards Organization (NISO). (2013) ANSI/NISO Z39.85 2012. Dublin Core Metadata Element Set.

Olson, Hope. (2007). How we construct subjects: a feminist analysis. *Library Trends* 56(2): 509-541.

Pollack, Andrew. (2013) Roughed up by an orca? There’s a code for that. *New York Times,* December 29, 2013.

Rader, Emilee. (2010) The effect of audience design on labeling, organizing, and finding shared files. *Proceedings ACM CHI 2010,* 777-786.

Ranganathan, S. R. (1959) *Elements of Library Classification.* London: The Association of Assistant Librarians. (Chapters 1-5)

Rosenfeld, Louis, and Peter Morville. (2006) Thesauri, controlled vocabularies, and metadata. In *Information architecture for the World Wide Web.* 3rd ed. Sebastopol, CA: O’Reilly. (Chapters 5 and 9.)

Srinivasan, Ramesh, Robin Boast, Katherine Becvar, and Jonathan Furner. (2009). Blobgects: digital museum catalogs and diverse user communities. Journal for the American Society of Information Science and Technology 60(4): 666–678.

Taylor, Arlene and Daniel Joudrey. (2009) *The Organization of Information.* 3rd ed. (Chapters 8 and 9.)

Tennis, Joseph T. (2012) The strange case of eugenics: a subject’s ontogeny in a long-lived classification scheme and the question of collocative integrity. *Journal of the American Society for Information Science and Technology* 63(7): 1350-1359.

Tennis, Joseph T. (2013) Ethos and ideology of knowledge organization: toward precepts for an engaged knowledge organization. *Knowledge Organization* 40(1): 43-49.

Vickery, Brian. (1960) *Faceted classification: a guide to construction and use of special schemes.* London: Aslib. (Chapters A-M only.)

Weagley, Julie, Gelches, Ellen, and Park, Jung-Ran. (2010) Interoperability and metadata quality in digital video repositories: a study of Dublin Core. *Journal of Library Metadata* 10(1): 37-57.

Wild, Peter, Giess, Matt, and McMahon, Chris. (2009) Describing engineering documents with faceted approaches: observations and reflections. *Journal of Documentation* 65(3): 420-445.

Wilson, Patrick. (1968) *Two kinds of power: an essay on bibliographic control.* Berkeley and Los Angeles: University of California Press. (Chapters 1, 2, 5)

Windnagel, A (2014). The use of simple Dublin Core metadata in digital math and science repositories. *Journal of Library Metadata* 14(2): 77-102.

Yee, K-P., Swearingen, K., Li, K., and Hearst, M. (2003) Faceted metadata for image search and browsing. *CHI Proceedings 2003.* (April 5–10, 2003, Ft. Lauderdale, Florida, USA.)

Zeng, Marcia Lei, and Lois Mai Chan. (2009) Semantic interoperability. In Bates, Marcia, and Mary Niles Maack (eds). *Encyclopedia of Library and Information Sciences.* 3rd ed. CRC Press.