

# INLS 722: Introduction to Metadata Architectures

## Fall 2023

### Basic information

This class is online; most information will be available through Canvas.

### Instructor information

*Instructor:* Megan Winget

*E-mail:* [megan.winget@unc.edu](mailto:megan.winget@unc.edu)

*Office:* Manning Hall

*Office hours:* <https://calendly.com/megan-winget/15min>

### Introduction

An enduring goal of the information professions is to enable precise, reliable, comprehensive retrieval of data and documents. A key means to achieving this goal has been the design and application of systems for structured description—metadata. Consistent, correct, and well-formed metadata, the thinking goes, facilitates access to diverse information resources, making them more discoverable to potential users.

To facilitate retrieval across collections and enable federated information aggregations, such as WorldCat, Europeana, and the Digital Library of America (DPLA), metadata needs to be *interoperable* across systems. Interoperability means that metadata created for one system can work in another system. For example, if one system describes authors and titles of resources in a coordinated statement (“Hamlet by William Shakespeare”) and another separates authors and titles into distinct metadata elements, then author and title information is not easily interoperable between the systems. One system’s records would need to be mapped to the other system’s structure before they could be aggregated.

There are a number of different levels of interoperability. *Syntactic interoperability* has to do with data encodings and formats. For example, are dates written like this: *August 18, 2018* or like this *08/18/2018* or like this *18/08/2018*? Is the metadata record encoded as an XML file or in the MARC format?

*Semantic interoperability* has to do with meaning. What kind of content should be in a metadata element titled Subject? (How specific and exhaustive should a subject term be? Should terms be chosen from a controlled vocabulary? Are metadata creators applying subject terms in the same way, so that they carry the same meaning from record to record?)

*Standards* to specify parameters for what is described, how it is described, and the format and syntax of description are meant to facilitate interoperability. Interoperable metadata via standardization is viewed as an important component of cyberinfrastructure and data integration. Interoperable metadata facilitates the aggregation of research data in all fields, from the sciences to the humanities. Such data integration and availability is often described as an engine of scientific progress, social innovation, and the general advancement of knowledge.

In this course, we will learn fundamental concepts central to structured metadata implementations and survey the many types of standards that attempt to harmonize description and enable interoperable systems. Concurrently, we will put these concepts into practice through a course project. In the project, we will employ standards to create metadata for a particularly complex type of cultural heritage artifact: video games. We will combine our individual metadata records to create a larger aggregated dataset of video game metadata, and we will examine the dataset to understand the interoperability of its records.

Throughout these explorations, we will also seek to understand the limits of metadata standards and to critically examine the inevitable role of interpretive diversity for information systems. All metadata, even the most “objectively” determined measurement, has a subjective component, in that a particular property has been identified as significant for a human purpose. Strategies to understand and work with interpretive diversity across time, across communities, and across cultures are accordingly necessary for effective design, application, and use of metadata. Our course project will engage the challenge of implementing standards for interoperable metadata while acknowledging the messy reality of interpretive diversity.

### **INLS 722 and INLS 720: What’s the difference?**

INLS 722 is a requirement for the online PSM degree in data curation; it is a 1.5-unit class.

INLS 720 is a 3-unit class that is an elective for the MSLS and MSIS degrees and a requirement for the online post-master’s certificate (PMC) in digital curation.

Both INLS 722 and INLS 720 are taught online.

*Although MLIS and MSIS students may take INLS 722, it is suggested that they take INLS 720 instead.* INLS 720 includes the content from INLS 722 plus 8 additional modules that examine particular metadata contexts; additionally, the semester project includes more stages in INLS 720.

*Students cannot receive credit for both INLS 722 and INLS 720.*

### **Learning Objectives**

At the end of this course, students will be able to:

- Identify, define, and disambiguate fundamental concepts of metadata architectures, such as entities, attributes, and relations.
- Compare and relate different types of standards (for example, content standards, structure standards, and encoding standards) and describe the role of intersecting standards in facilitating metadata interoperability.
- Identify, characterize, and assess the interpretive diversity in a dataset created to align with metadata standards.
- Describe and defend a principled position to inform the design, implementation, and use of metadata standards that accounts for inevitable interpretive diversity in metadata environments.

### **Course Structure**

Course content is organized into two sections:

- Metadata basics.
- Metadata standards.

In the *Metadata Basics* section, we review fundamental concepts of resource description, including entities, attributes of entities, and relations between entities. Although our focus here will be to understand the general concepts that make up metadata architectures, we will make use of several common standards (such as Dublin Core and Functional Requirements for Bibliographic Records, or FRBR) to do so. We will also begin our acquaintance with the video game metadata standard, proposed by Jin Ha Lee and her GAMER group colleagues, that we will use in our course project.

In the *Metadata Standards* section, we will survey the many types of standards employed in research description and understand how these standards relate to each other. We will also examine the creation and implementation of standards in practice, across different communities, and across time.

The course is divided into 7 modules. Modules might include:

- Readings.
- Lectures.
- Activities.
- Discussion.
- Tasks associated with the course project (the course project is described in the next section).

A complete module calendar appears at the end of this syllabus.

### **Course Grading**

In collaboration with the other faculty, we have decided that students will receive either a P (Pass) or F (Fail) grade (There will be no H or L grades). We do this to alleviate stress and allow for a more equitable grading system across sections.

To pass the course, students must:

- Satisfy participation requirements (described below).
- Complete a metadata project to investigate the process of making metadata and developing an interoperable standard within the class.

Assigned work will receive written feedback aligned with each project's documented success criteria. A successful assignment will be marked as "complete." There will be no scores or grades.

Should any assignment that you submit fail to fulfill the success criteria to a minimum standard, you will be invited to resubmit the assignment.

Extensive project details are provided in the following section. *It is your responsibility to read and understand all the instructions. If you have questions, ask. There will be a Sakai discussion board for this purpose.*

For all graded course components, grading is based entirely upon the criteria in this syllabus. *There will be no surprises.*

### **No busy work principle**

No one wants to do boring things for no reason, including me! If something is required for the course, it has a purpose that requires thought. If something seems like busy work, I probably haven't articulated the purpose well. First reread the instructions, and then be sure to ask me about it, so I can better explain why I am asking you to perform that task.

### **Participation**

Because this class is delivered online, consistent and energetic participation is more vital for the success of the course than with an in-person class. The goal of participation requirements is to ensure a level of dialogue with each other. Remember: no busy work.

Participation assessment involves two components:

Syllabus developed by Melanie Feinberg

- Frequency.
- Content.

## Frequency

Module 1 participation is not graded because it's introductions.

For all other modules, you must contribute *at least two substantive* posts to discussion forums (about 200 words each) *per module*. These are the *minimum* posting requirements. These can be posts that you've initiated (i.e., you start a thread) or comments on other students' posts.

I strongly encourage posting on at least two different days, so that you can contribute more effectively to our continued conversation.

Posting is formative, not summative. You do *not* need to complete the readings or activities in a module before contributing to discussion forums. For example, you can post about a lecture on Tuesday, a reading on Thursday, and an activity on Friday; you don't need to have finished all the lectures, readings, and activities before making any posts. Additionally, you may continue discussions from a previous module; for instance, you may post about entities (a topic from Module 2) at 10 a.m. on Monday on the Module 3 discussion board; this posting "counts" toward your minimum requirements for Module 3.

Our goal is to maintain an ongoing conversation; think of contributing to our discussions as something that you do *during* a module to help you learn, and *not* after a module, to demonstrate that you have done the work. Accordingly, try not to focus on your forum posts as being "due" on Friday and Monday morning; this is limiting (and may produce unnecessary anxiety). Instead, think of having an entire week to contribute to class discussions based on *your* schedule. For instance, *if you post once every Monday, Wednesday, and Friday afternoon, you will get through all of your posting requirements during the week, without any worry or stress*. This is the kind of approach I recommend. (It is unfortunate when much of the posting occurs after 10 p.m. on Sunday night; this is not conducive to group interaction, nor to a healthy sleep schedule. There is no reason to stay up late on the weekend to fulfill your posting requirements.)

## Content

If there are no module-specific instructions, you can write postings that address any of our course materials: readings, lectures, and so on. (Occasionally, you will be explicitly directed to write one of your required postings about a module activity, or about the course project tasks that will take place during that module.)

Postings that relate multiple materials and refer to previous weeks or to other students' postings are especially encouraged. While I will provide discussion prompts to help you think about a module's readings, *you do not need to respond to one of my prompts*. You are encouraged to create your own conversations within each module's forum area.

Your postings can respond to someone else's post or initiate a new topic.

Postings must be germane to the course, but they can be related to content from either the current module *or to a previous one*. In other words, it is perfectly acceptable to contribute a post about a Module 2 topic during Module 3.

## Other Considerations: Quality

Postings are meant to encourage conversation and help us all to learn. You do not need to have an answer

when you post something. Indeed, postings that describe your difficulties in understanding are especially appreciated. If you're mystified, odds are that someone else is also mystified, and sharing your experience will help everyone.

However, postings should not merely ask endless series of questions but should constitute an initial attempt to work through something you don't understand, or to relate observations, insights, and comments that might extend the learning of others. For example, if a concept is confusing, your posting can explain in detail what you find confusing and describe how you've tried to resolve that confusion, even if those attempts did not succeed and you are more confused than ever.

Postings that reflect on or make use of your professional experiences in light of the course topics are also encouraged.

### Other Considerations: Reciprocity

We embark on learning as a group because we learn from interacting with each other, as peers, in dialogue. To that end, it's important to read the posts that others make, and to set your own contributions in response to others when appropriate.

As you develop your own ideas, put them into conversation with the posts that are already there. If someone else has written about something that you wanted to comment upon, continue the discussion rather than creating a new thread. You can respond to others' contributions in a variety of ways, such as:

- Agree with the initial post by providing confirmatory evidence from other sources (readings, activities, experiences).
- Extend the initial post with complementary observations, reflections, or examples.
- Question the initial post by relating contradictory evidence (from readings, activities, experiences).
- Take the initial post in a new direction by providing a different emphasis or perspective.

Presenting alternative or opposing viewpoints is an important aspect of scholarly discourse, and everyone should feel empowered to disagree with the perspectives presented in our readings, in my lectures, and in other discussion posts. Take care to be respectful and well-reasoned in disagreement; show that you understand the argument being put forth even as you suggest a new approach.

### Discussion Grading criteria

I will post discussion "grades" at the beginning of each module. Each discussion forum is worth 2 points total.

Points will be distributed amongst the two participation components as follows:

- Did you post twice per module?
- Was your post thoughtful and relevant to the discussion?

## ***Project: Local Implementation and Global Aggregation***

In this course project, you will create metadata records using a complex schema for video game metadata. You will also use controlled vocabularies created to accompany the schema.

Then, you will examine and interpret the dataset created by the class to identify, characterize, and assess its areas of interpretive diversity, and to form your own position on interpretive diversity in metadata and on the role of standards in forming the character of a dataset. Is interpretive diversity bad—or can it be useful? How can we more perceptively understand interpretive diversity in metadata, as we attempt to create standards and implement large-scale aggregated datasets?

Our domain for this project will be video game metadata. Video games are complex information resources that have been insufficiently described with standard schemas for cultural heritage. Jin Ha Lee and colleagues at the University of Washington (the GAMe METadata Research, or GAMER group) have been developing a schema for video game description that they propose as a content and structure standard for this domain. We will read several articles over the course of the semester that describe the design, evaluation, and revision of the GAMER video game metadata schema. (We will use the 2.1 version of the schema for our project, which was released in February 2015.)

### *Project components*

- Part 1: 10 metadata records (7 for games of your choice and 3 for games selected for the entire class).
- Part 2: Essay that analyzes the aggregated dataset of metadata records created by the class as a means to formulating your own position on interpretive diversity in resource description.

### Project due dates

- September 1, 11:59pm:** Identify 7 video games to describe (can be any 7 games of your choice)
- September 15, 11:59pm:** Component #1, metadata generation (excel / sheets spreadsheet) and reflective essay.
- October 13, 11:59pm:** Component #2, a position and plan for interpretive diversity

The following sections describe each project component in more detail.

## **Project Component 1: Generating Metadata Records from a Standard Schema**

In this phase of the project, you will create 10 metadata records.

You will create 7 records for games of your choice. You will create the remaining 3 records for the following games:

- Final Fantasy 7
- Skyrim
- Journey

*You are not assumed to have any knowledge at all of video games to complete this task.* The creators of metadata often do not have particular subject-matter expertise. You are also not required to own, or have played, or intend to play, any of the games that you describe. The creators of metadata often have little time to become familiar with the resources that they describe. The schema documentation

suggests particular sources of information for creating metadata; however, as with all metadata creation processes, it will be *up to your individual discretion, as a metadata creator, how you follow such suggestions*, given your own personal expertise, available time, and so on. If you don't play any games and use the Internet to determine how to assign appropriate values to your selected resources, that is perfectly fine. There are no metadata police, here or in the real world.

**Creating these records is quite time-consuming; you will have to do some research. Do not underestimate the time that this task will take to complete.**

#### *Deliverables*

- 10 metadata records:
  - 7 for games of your choice.
  - One each for the following games: Skyrim, Journey, and Final Fantasy 7.
  - All 10 metadata records must be implemented in Excel (I will provide a template for students to fill out).
- A reflective essay of about 750-1000 words (3-4 pages) that critically interrogates how the experience of creating metadata informs upon your understanding of metadata concepts, descriptive practices, and standardization.

#### *Grading criteria*

A successful project will exhibit these characteristics:

- The 10 metadata records are complete and represent an honest attempt to fulfill the goals expressed by the schema documentation. (25 points)
- The reflective essay thoughtfully considers the experience of applying the schema documentation to productively engage larger issues of theory and practice (that is, the reflection does not merely summarize or justify the process you followed or the records you created; the essay interrogates those things). (15 points)
- The essay follows a logical document structure, is clearly written, and uses correct grammar and punctuation. (10 points)

*Points will be awarded as specified for each criterion.*

### **Project Component 2: Position Paper on Interpretive Diversity, Standardization, and Metadata**

In the second component of this project, you will analyze the aggregated dataset of video game metadata records created by the students in this course (I will provide this aggregated spreadsheet by Wednesday July 10 in the Resources section of Sakai). You will examine, assess, and interpret our 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 on the aggregated database? This analysis will help you to formulate a *position* on interpretive diversity in metadata generation and aggregation.

In modules 5-7, we will look at some aspects of the aggregated dataset and discuss them in our course forums, as preparation for writing this essay. But you will also want to examine the dataset on your own. There are many approaches you can use to do this. For example, you can compare the metadata generated by individual people. Do certain creators have a particular "style"? How does that style manifest? (Through the amount of information described? The amount of detail for certain elements as opposed to others? The aspects of each game that are described as opposed to other aspects? The choices of values?) Is there a value to style, and what is it?

You can also compare the metadata generated for particular attributes as well, for particular creators, for particular communities (project groups), and across the dataset. Does the meaning of "Mood"

change or remain consistent across indexers, communities, and the entire collection, for example? How is the interpretation of “Mood” different?

And you can compare metadata for the common three games that everyone described. Where are areas of divergence and convergence? Where is it problematic to have differences, and why? Is it ever useful or interesting to have differences?

These are just a few strategies you might adopt in exploring the dataset. Be creative in locating and interpreting dimensions for comparison.

You will use your analysis of the dataset, in conjunction with your experiences creating metadata yourself, as well as your understanding of course readings, lectures, and discussions, to write an essay that *constitutes an argument for how to approach interpretive diversity* in the context of creating and using metadata standards in local and global contexts. Is interpretive diversity a problem that can be eradicated? Is it a pestilence that can never be stopped, only contained (like cockroaches!)? Is it a pervasive necessity, like breathing? Is it sometimes worthwhile or useful? If you determine that interpretive flexibility needs to be constrained, how do you do that?

After you establish and justify your position, you will use it to formulate a course of action or set of guiding principles for metadata creation, use, and aggregation. In other words, if interpretive diversity is a problem to be solved, how do you solve it? Or how do you contain it, minimize it, live with it, enhance it, draw attention to it, or whatever actions the position that you establish might entail? If, for example, your plan is to “write better documentation,” what would make the documentation “better”? As you develop your plan, you might consider the activities of standards developers, standards implementers, metadata creators, data aggregators, and data users. Each of these stakeholders might have different needs.

### *Deliverables*

You will write an essay of about 1,500 words. Your essay should make some reference to the following:

- Your experience creating metadata yourself with the schema documentation.
- Your analysis of the aggregated dataset.
- Concepts from course readings and other activities.

### *Grading criteria*

A successful essay, which will receive a grade of “complete” will exhibit the following characteristics:

- Develops a clear, defensible 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 metadata yourself and analysis of the aggregated dataset.
- Appropriately synthesizes material from course readings to support your argument.
- Uses the defined position on interpretive diversity to present a course of action or set of principles to guide the activities of standards creators, standards implementors, metadata creators, data aggregators, and data users.
- Follows a logical document structure, is clearly written, and uses correct grammar and punctuation.



# MODULE CALENDAR

Modules will become available at 9 a.m.

Module	Dates	Key Questions	Readings
1: <i>Metadata basics</i> Introduction	August 21-27	What is metadata and why should we care about it?	Class syllabus Gilliland, 2015 <i>Optional</i> Greenberg, 2009 Metadata Basics (Zeng)
2: <i>Metadata basics</i> Entities and identifiers	August 28 – September 3	What is being described?  How can entities be persistently identified?	Kent (ch. 1) Sangers-Katz, 2016 IFLA (selections about Group 1 entities: 13-14, 17-24, 31-49) Thompson, 2010 <i>Optional</i> Coyle, 2006 Bates, 1986 (selections)
3: <i>Metadata basics</i> Properties of entities (attributes and values)	September 4 - 10	What significant properties distinguish each entity?  What kinds of values best express these properties?	ANSI/NISO Z39.85-2012 (Dublin Core metadata element specification) CDP Metadata Working Group, 2006 Lee, et al, 2013 Lee, et al (GAMER group), 2015, schema version 2.1 <i>Optional</i> Global Terrorism Database (GTD) codebook Armed Conflict Location Event Data project (ACLED) codebook

<p>4: <i>Metadata basics</i></p> <p>Relations between entities (models)</p>	September 11 - 17	How are entities and properties related?	<p>Dublin core abstract model</p> <p>Johnston, 2006</p> <p>Urban, 2014</p> <p>IFLA (again)</p> <p>Jett, et al 2015</p> <p><i>Optional</i></p> <p>Lee, et al (GAMER group) 2015, schema version 3.1</p>
<p>5: <i>Metadata standards</i></p> <p>Types of standards and the work of creating them</p>	September 18 - 24	What is interoperability and how might it be achieved?	<p>Zeng and Chan, 2009</p> <p>Elings and Weibel, 2007</p> <p><i>Optional</i></p> <p>Millerand and Bowker, 2009</p>
<p>6: <i>Metadata standards</i></p> <p>Semantic diversity and change over time</p>	September 25 – October 1	How does our understanding of attributes and their values change over time and across communities of practice?	<p>Buckland, 2012</p> <p>Bowker and Star, 2000</p> <p><i>Optional</i></p> <p>Mai, 2011</p>
<p>7. <i>Metadata standards</i></p> <p>Implementation of standards in practice</p>	October 2 - 8	<p>How is a standard implemented in different situations?</p> <p>What happens when data from different sources is aggregated?</p>	<p>Weagley, Gelches, and Park, 2010</p> <p>Jackson and Barrow, 2015</p> <p>Lee, Clarke, and Perti, 2016</p> <p><i>Optional</i></p> <p>Goodwin, 1994</p>

## **Policies**

### **Instructor communication**

For specific and personal inquiries, e-mail is the most reliable means of contact for me. I do my best to answer e-mail within a day 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 "INLS 722."

For general questions about the class, logistics, and requirements, please post your question to the Logistics and Requirements forum on Sakai. If you were confused about something, it's likely that others are as well, and posting your question enables others to get the information as well.

### **Academic integrity**

The UNC Honor Code states that:

*It shall be the responsibility of every student enrolled at the University of North Carolina to support the principles of academic integrity and to refrain from all forms of academic dishonesty...*

This includes prohibitions against the following:

- Plagiarism.
- Falsification, fabrication, or misrepresentation of data or citations.
- Unauthorized assistance or collaboration.
- Cheating.

All scholarship builds on previous work, and all scholarship is a form of collaboration, even when working independently. Incorporating the work of others, and collaborating with colleagues, is welcomed in academic work. However, the honor code clarifies that you must always acknowledge when you make use of the ideas, words, or assistance of others in your work. This is typically accomplished through practices of reference, quotation, and citation.

*If you are not certain what constitutes proper procedures for acknowledging the work of others, please ask the instructor for assistance.* It is your responsibility to ensure that the honor code is appropriately followed.

### **Students with disabilities**

Students with disabilities should request accommodations from the UNC office of Accessibility Resources and Service (<https://accessibility.unc.edu/>).

### **Late work**

I must have your 10 metadata records to move forward with the next assignment. Please turn it in on time. If that is impossible, please send me an email with a plan for turning the work in as soon as you know your work will be late.

### ***Extensions and "make-up" work***

Extensions for project work may be granted under reasonable circumstances, *when negotiated with the instructor in advance*. A request for an extension minutes before a due date will likely be denied. A request for an extension made a month before the due date will be much better received.

Regarding participation, students who encounter extreme circumstances may consult with the instructor regarding possible "make-up" work. Such accommodations will only be granted for significant and unforeseen life events, and any alternate arrangements will likely be more onerous than the baseline module participation.

Students who anticipate difficulties with completing work on time, or who encounter unexpected and severe challenges, should consult with the instructor as soon as possible so that alternate solutions can be discussed.

### Citation policy and paper presentation details

All written work needs to properly acknowledge the ideas of others via in-text references, even when not directly quoting from a source.

In making in-text references or preparing reference lists for outside sources, you may adopt any standard citation style you prefer (such as APA or the Chicago Manual of Style).

You may select whatever font, font size, margin, spacing, and other options that you like, as long as your paper is professionally presented. I will not actually count the words in your paper; directions about length are guidelines only.

### Use of Generative AI

#### Introduction

ChatGPT and other Generative Artificial Intelligence (AI) can produce text, images, and other media. These tools can assist with brainstorming, finding information, and even reading and creating materials; however, they must be used appropriately and ethically, and you must understand their limitations. Regardless of your use of any AI tools, you are responsible for the final product of your work, both academically and in the workforce. Generative AI is extremely useful; however, it has the following limitations:

- It is unclear how the system generates content, as the internal processes used to produce a particular output within the generative AI cannot be determined.
- The output is based on existing data (often scraped from online sources) and may reflect biases that should be acknowledged; it may also be inaccurate or entirely fabricated, even if it appears reliable or factual.
- AI evokes a range of intellectual property concerns; sourcing and ownership of information is unclear, and the status of AI output raises numerous questions—e.g., is output equivalent to a published resource? What citational responsibilities are in place for various AI interactions?

The following sections provide the philosophy and specific guidelines for using these tools and features (increasingly, generative AI capabilities will be integrated with everyday applications). **Unless I provide other guidelines for an assignment or exam, you should follow these guidelines.**

#### Usage Philosophy:

Use of generative AI in your coursework is based on the following principles:

1. **AI should help you think. Not think for you.** Use these tools to give you ideas, perform research (in compliance with point 2 below), and analyze problems. Do not use them to do your work for you, e.g., do not enter an assignment question into ChatGPT and copy & paste the response as your answer.
2. **Engage with AI Responsibly and Ethically:** Engage with AI technologies responsibly, critically evaluating AI-generated outputs and considering potential biases, limitations, and ethical implications in your analysis and discussions. Utilize AI technologies ethically, respecting privacy, confidentiality, and intellectual property rights. Ensure that the data used for AI applications is obtained and shared responsibly and in compliance with relevant regulations.

3. **You are 100% responsible for your final product.** You are the user. If the AI makes a mistake, and you use it, it's your mistake. If you don't know whether a statement about *any item in the output* is true, then your responsibility is to research it. If you cannot verify it as factual, you should delete it. You hold full responsibility for AI-generated content as if you had produced the materials yourself. This means ideas must be attributed, facts are true, and sources must be verified.
4. **The use of AI must be open and documented.** The use of any AI in the creation of your work must be declared in your submission and explained. Details on how to source your AI usage are explained below.
5. **These guidelines are in effect unless I give you specific guidelines for an assignment or exam.** It is your responsibility to ensure you are following the correct guidelines.
6. **Data that are confidential or personal should not be entered into generative AI tools.** Putting confidential or personal data (e.g., your One Card details) into these tools exposes you and others to the loss of important information. Therefore, do not do so.

## Bibliography of readings

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

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

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

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

Coyle, Karen. 2006. Identifiers: unique, persistent, global. *The Journal of Academic Librarianship* 32(4): 428-431.

Lorraine Daston. 2015. Cloud physiognomy: describing the indescribable. *Representations* 135, Summer 2015, 45-71.

Duval, Eric, Wayne Hodgins, Stuart Sutton, and Stu Weibel. (2002) Metadata principles and practicalities. *D-Lib*. Available at: <http://dlib.org/dlib/april02/weibel/04weibel.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>

GAMER Group. (Lee, J.H., and colleagues.) (2014). Video game metadata schema version 2.1 and 3.1 (optional).

Gilliland, Anne. 2016. Setting the stage. In *Introduction to Metadata*. 3rd ed (online edition). Edited by Murtha Baca. Available at: <http://www.getty.edu/publications/intrometadata/setting-the-stage/>

Gill, Tony. (2004) Building semantic bridges between museums, libraries, and archives: the CIDOC conceptual reference model. *First Monday* 9(5). Available at: <http://firstmonday.org/ojs/index.php/fm/article/view/1145/1065>

Goodwin, Charles. (1994) Professional vision. *American Anthropologist* 96(3): 606-633.

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.

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.)

Jackson, Steven, and Sarah Barbrow. 2015. Standards and/as innovation: protocols, creativity, and interactive systems development in ecology. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (CHI '15)

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

Jett, J., S. Sacchi, J.H. Lee, and R. Clarke. (2015) A conceptual model for video games and interactive media. *Journal of the Association of Information Science and Technology (JASIST)*, doi: 10.1002/asi.23409

Johnston, P. (2006) Why an abstract model for Dublin Core metadata? eFoundations blog. Available at: [http://efoundations.typepad.com/efoundations/2006/11/why\\_an\\_abstract.html](http://efoundations.typepad.com/efoundations/2006/11/why_an_abstract.html)

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. (2015). 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.

Millerand, F., and Bowker, G. (2009) Metadata standards: trajectories and enactment in the life of an ontology. In S. L. Star & M. Lampland (Eds.), *Formalizing Practices: Reckoning with Standards, Numbers and Models in Science and Everyday Life*.

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

Oldman, Dominic. (2013) The costs of cultural heritage data services: CIDOC CRM or aggregator formats? Available at: <http://www.oldman.me.uk/blog/costsofculturalheritage/>

Sanger-Katz, Margot. 2016. Is terrorism getting worse? In the West, yes. In the world, no. *New York Times*, August 16, 2016.

Thompson, H. (2010) What is a URI and why does it matter? Available at: <http://www.ltg.ed.ac.uk/~ht/WhatAreURIs/>

Urban, Richard. (2014) The 1:1 principle in the age of linked data. *Proceedings of the International Conference on Dublin Core and Metadata Applications 2014*, 119-128.

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

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.

### **Resources and references**

These are not assigned, but you may find them helpful.

#### *Metadata fundamentals*

Marcia Lei Zeng. (2016) Metadata Basics tutorial. Available at: <http://metadataetc.org/metadatabasics/>

Marcia Lei Zeng and Jian Qin. (2016) *Metadata*. 2nd ed. New York: Neal-Schuman.

Web site to accompany the book is available here:

<http://metadataetc.org/book-website2nd/>

NISO (National Information Standards Organization). (2004) *Understanding metadata*. Available at: <http://www.niso.org/publications/press/UnderstandingMetadata.pdf>

Amy Brand, Frank Daly, and Barbara Myers. (2003) *Metadata demystified: a guide for publishers*. Available at: [http://www.niso.org/publications/press/Metadata\\_Demystified.pdf](http://www.niso.org/publications/press/Metadata_Demystified.pdf)

Paul Miller. (1996) Metadata for the masses. *Ariadne* (5) Available at: <http://www.ariadne.ac.uk/issue5/metadata-masses/>

#### *Metadata standards*

List of (primarily structural) standards from Metadata book Web site by Zeng and Qin: <http://www.metadataetc.org/book-website/readings/appendixaschemas.htm>

#### *Linked data, Semantic Web, RDF*

Linked Data Tools. Semantic Web Primer. Available at: <http://www.linkeddatatools.com/semantic-web-basics>

Tom Heath and Christian Bizer. (2011) *Linked data: evolving the Web into a global data space*. Available at: <http://linkeddatabook.com/>

World Wide Web Consortium (W3C). (2014) *RDF Primer 1.1*. Available at: <http://www.w3.org/TR/2014/NOTE-rdf11-primer-20140624/>