

SI 667 001 WN 2019

Foundations of Digital Curation [syllabus last updated 9 Jan]

Instructor: Andrea Thomer [athomer@umich.edu]

Time: Monday & Wednesdays, 1-2:30

Location: NQ 2185

Office hours: Wednesday 2:30-4:30 (but please [sign up for a 20 minute slot](#)). Other times possible by appointment.

Weekly schedule of topics and readings

Course Description

Digital curation is the active involvement of information professionals in the management of digital data through its lifecycle of interest and usefulness to science, scholarship and education. This course will provide you with a foundational understanding of the core concepts and practical problems in digital curation; you will put these insights into practice through hands-on projects with real-world datasets throughout the semester.

Topics to be addressed include:

- the fundamentals of digital preservation and migration;
- the data practices and needs of researchers;
- key concepts in data and digital curation
- trends in (and obstacles to) data sharing and reuse
- data representation, organization, preservation, and access
- fit for preservation and a range of different users and uses. This course draws on current research in information science, archival and library science, and cognate fields.

Learning objectives:

By the end of this class, you will be able to:

- Describe the intellectual foundation of digital curation, and relate the work and core concepts of digital curation to those in related domains of study (e.g. information organization, archival science, systems analysis)
- Describe the key goals, priorities and activities of work in digital curation.
- Define key concepts, topics and entities in digital curation (e.g. data, datasets, digital objects, lifecycles, collections, provenance, data reuse, data reproducibility, fitness for use).

- Compare and critique different data lifecycle models.
- Assess metadata and data standards to gauge their appropriateness for use with different formats and domains of digital objects
- Perform simple curatorial tasks on digital objects to support access and use over time.
- Evaluate and draft digital curation strategies for user groups at a range of time scales (e.g. near term or long term), organizational structures (e.g. individual scholars, research labs, institutions), and domains (humanities, sciences, etc).

Assignments

At the beginning of the semester, you will be introduced to a range of real-world datasets and digital collections; you will work (largely in groups) with one of these collections over the course of the semester. Some of the potential data collections include

- Enron emails: <https://www.cs.cmu.edu/~enron/> (Links to an external site.)Links to an external site.
- Treebank: <https://catalog ldc.upenn.edu/ldc99t42> (Links to an external site.)Links to an external site. <- main page is currently and inexplicably down! However, has been archived in the [Internet Archive \(Links to an external site.\)Links to an external site.](#), and a small fair-use subset is available [here \(Links to an external site.\)Links to an external site.](#). You'd focus on working with and curating the subset.
- UFO sightings database: <http://www.nuforc.org/webreports.html> (Links to an external site.)Links to an external site.
- The Pettigrew papers, including their [afterlife as a teaching dataset \(Links to an external site.\)Links to an external site.](#) in past workshops I've run: <https://openarchaeologydata.metajnl.com/articles/10.5334/4f913ca0cbb89/> (Links to an external site.)Links to an external site.
- From Morphosource, 3D scan of a flying fox fossil that may or may not be from the U-M Vert Paleo collection: https://www.morphosource.org/Detail/SpecimenDetail/Show/specimen_id/9177 (Links to an external site.)Links to an external site.
- Tate 2014 Collection Data - note that you would be primarily curating metadata about these art collections, not the images: <http://www.tate.org.uk/about-us/digital/collection-data> (Links to an external site.)Links to an external site.
- A dataset of your choice from Documenting the Now - <http://www.docnow.io/catalog/> (Links to an external site.)Links to an external site.
- American Sign Language Lexicon Video Dataset (note: this is perhaps a more challenging option): <http://csr.bu.edu/asl/asllvd/annotate/index.html> (Links to an external site.)Links to an external site.
- A subset of TEI files from the [Walt Whitman archive \(Links to an external site.\)Links to an external site.](#): <https://github.com/whitmanarchive> (Links to an external site.)Links to an external site.
- A paleoecology dataset from Neotoma, such as dataset 2662: <http://data-dev.neotomadb.org/2662> (Links to an external site.)Links to an external site.

- Other suggestions welcome!

In groups or on their own, students will curate, assess and draft curation plans for their dataset/collection through the following assignments.

1. *Dataset case report and threat assessment(15%)*: Digital objects are rarely discrete objects; rather they often consist of a network of data files, contextualizing documentation and other ephemera (sometimes, non-digital ephemera). Create a “forensic graph” showing the agents, objects and relationships between the entities in your dataset/collection. Inventory the objects and agents in your dataset, and summarize your dataset in a case report. Conduct a threat assessment of two key file formats in your dataset. Turn in the graphs, case report, and threat assessment.
2. *Significant properties and community description (10%)*: Who might reuse your dataset? What aspects of the dataset need to be prioritized for curation for this community
3. *Fit for reuse (15%)*: curate some aspect of your dataset for reuse by your designated community, and create some metadata for your files.
4. *Data management plan (10%)*: Draw on your prior work and create a plan for your dataset's long term curation. Include cost analysis and staffing needs.
5. *Speculative curation fiction (10%)*: Draft a position paper, describing a vision for a digital archive 10-100 years into the future that would serve some aspect of your dataset's domain, topic, format, or other significant property.

You will additionally be graded on the following course participation criteria:

Labs/reading reflections (30%): Readings are assigned every week. Please read the assigned readings before class on Monday. Every Wednesday, we'll have a lab or discussion section. Labs will have a “reading reflection” component in which you're asked to reflect on the readings from that week/lecture. Standard reflections ask:

1. What did you find most interesting, compelling, intriguing, or thought provoking about this week's readings?
2. What was confusing, or what questions did they raise?

On occasion there may be a third question that relates the readings to the lab.

Completed labs and reflections will be due every Friday at midnight. **The lowest two lab grades will be dropped from your final grade.**

Attendance and participation (10%): Students are expected to attend class regularly, and participate in discussions and activities. We'll do a lot of in class activities and group work, so bring your laptops and be ready to work and collaborate. Egregious absences (> 4 or so) will result in a loss of all 10% of your participation grade.

I will evaluate student attendance, active participation and overall progress throughout the course of the semester. The following rubric will be used to assign a score mid-semester and again at the end of the semester. These will be averaged together to create the final score.

- 10 = Student has been an active participant in class discussions, bringing to the class insights from their interpretation of readings and lived experiences and is demonstrating an interesting grasp of the key concepts covered in class.
- 8 = Student has been an active participant in some of the class discussions and is demonstrating some gains in grasping key concepts covered in class.
- 6 = Student is occasionally active in class and is demonstrating some learning but it is clear they are not performing to their full capabilities
- 4 = Student has missed several classes and/or is not always active when attending class
- 0 = Student has missed more than four class sessions

General course policies

Email policy and office hours

I will try to answer emails within 72 hours of receiving them; however, if you have extensive questions about an assignment, please come to office hours or make an appointment to meet. If in-person office hours don't work, I am also available via Skype. I cannot guarantee a prompt response to last minute, just-before-the-deadline emails about assignments, so try to work ahead.

While I am happy to answer questions about assignments and look over small sections of what you're working on, I unfortunately cannot look over complete drafts of assignments. If you need referrals to copy editing services, please contact the Office of Student Affairs or your local academic library.

Grading criteria and late assignment policies

When completing your assignments, be sure to follow the instructions carefully, and pay special attention to the grading criteria. I will use the following scheme for assigning final grades: A = 93-100, A- = 90-92; B+ = 87-89, B = 83-86, B- = 80-82; C+ = 77-79, C = 73-76, C- = 70-72, D+ = 67-69, D = 63-67, D- = 60-62; F = <60.

Late assignments are not acceptable, although extensions are possible if discussed in advance. If you anticipate difficulties completing an assignment on time, talk with me as soon as possible so that alternate solutions can be discussed. **If we talk in person, and I do grant an extension, you must promptly email me a reminder of your new due date as soon as possible, otherwise I will likely forget that I granted it, and your extension will be reneged!**

In the absence of any prior arrangement with me, the policy is that for each day that an assignment is late, ten percent of the possible points will be deducted from the score. An assignment is one day late when the time it is due is passed, and it continues to be one day late until 24 hours later.

Class attendance and participation are essential components of this course. As such, I expect everyone to come to class having read the required material, prepared to

critically engage the concepts and participate in activities. Students can miss one class without penalty; however, please contact me ahead of time if you need to miss more than that. If you miss more than three classes, I reserve the right to dock an additional 5% of your grade per absence.

It is also your responsibility to ask questions regarding concepts that you don't understand. Remember, asking questions is an excellent form of participation. Please do not be afraid to ask questions; no question is "dumb."

Code of conduct

I consider this classroom to be a place where you will be treated with respect, and I welcome individuals of all ages, backgrounds, beliefs, ethnicities, genders, gender identities, gender expressions, national origins, religious affiliations, sexual orientations, ability – and other visible and nonvisible differences. All members of this class are expected to contribute to a respectful, welcoming and inclusive environment for every other member of the class.

Preferred Name/Pronoun Policy

I will gladly honor your request to address you by an alternate name or gender pronoun. Please advise me of this preference early in the semester so that I may make appropriate changes to my records.

Disability accommodation

If you think you need an accommodation for a disability, please let me know at your earliest convenience. Some aspects of this course, the assignments, the in-class activities, and the way the course is usually taught may be modified to facilitate your participation and progress. As soon as you make me aware of your needs, we can work with the Services for Students with Disabilities (SSD) office to help us determine appropriate academic accommodations. SSD (734-763-3000; <http://ssd.umich.edu>) typically recommends accommodations through a Verified Individualized Services and Accommodations (VISA) form. Any information you provide is private and confidential and will be treated as such.

Student Mental Health and Wellbeing

The University of Michigan is committed to advancing the mental health and wellbeing of its students, while acknowledging that a variety of issues, such as strained relationships, increased anxiety, alcohol/drug problems, and depression, directly impacts students' academic performance.

If you or someone you know is feeling overwhelmed, depressed, and/or in need of support, services are available. For help, contact Counseling and Psychological Services (CAPS) at (734) 764-8312 and <https://caps.umich.edu/> during and after hours, on weekends and holidays or through its counselors physically located in schools on both North and Central Campus. You may also consult University Health Service (UHS) at (732) 764-8320 and <https://www.uhs.umich.edu/mentalhealthsvcs>, or for alcohol or drug concerns, see www.uhs.umich.edu/aodresources.

For a more comprehensive listing of the broad range of mental health services available on campus, please visit: <http://umich.edu/~mhealth/> ([Links to an external site.](#))[Links to an external site.](#) ([Links to an external site.](#))[Links to an external site.](#)

Statement on Academic Integrity

Collaboration

UMSI strongly encourages collaboration while working on some assignments, such as homework problems and interpreting reading assignments as a general practice. Active learning is effective. Collaboration with other students in the course will be especially valuable in summarizing the reading materials and picking out the key concepts. You must, however, write your homework submission on your own, in your own words, before turning it in. If you worked with someone on the homework before writing it, you must list any and all collaborators on your written submission. Each course and each instructor may place restrictions on collaboration for any or all assignments. Read the instructions carefully and request clarification about collaboration when in doubt. Collaboration is almost always forbidden for take-home and in class exams.

Plagiarism

All written submissions must be your own, original work. Original work for narrative questions is not mere paraphrasing of someone else's completed answer: you must not share written answers with each other at all. At most, you should be working from notes you took while participating in a study session. Largely duplicate copies of the same assignment will receive an equal division of the total point score from the one piece of work.

You may incorporate selected excerpts, statements or phrases from publications by other authors, but they must be clearly marked as quotations and must be attributed. If you build on the ideas of prior authors, you must cite their work. You may obtain copy editing assistance, and you may discuss your ideas with others, but all substantive writing and ideas must be your own, or be explicitly attributed to another. See the (Doctoral, MSI, BSI) student handbooks available on the UMSI intranet for the definition of plagiarism, resources to help you avoid it, and the consequences for intentional or unintentional plagiarism.