

University of Michigan * School of Information
SI 667
Foundations of Digital Curation
Winter 2025

Course Syllabus

Instructor: Jesse A. Johnston (jajohnst@umich.edu)
Office Hours: Thursdays, 10:00 a.m.-Noon ([book here](#)); NQ 3443
I am also available by appointment (in-person or online)
Location: North Quad 2185
Time: Thursdays, 1:00–3:50 p.m.
Website: See Canvas (<https://canvas.it.umich.edu/>)



This syllabus is licensed under a [Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License](#).

Course Description

From the catalog: The course will provide students with a foundational understanding of the core theoretical, conceptual, and practical problems in digital curation; students will put these insights into practice through hands-on projects. 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. Students interested in careers in this growing field will need a strong grounding in the fundamentals of digital preservation and migration; an understanding of the data practices and needs of researchers; trends and obstacles to data sharing and reuse; and experience making data fit for preservation and range of different users and uses. This course draws on current research in information science, archival and library service, and cognate fields.

Topics to be addressed include:

- key concepts in data and digital curation;
- the data practices and needs of researchers;
- trends in (and obstacles to) data sharing and reuse;
- data representation, organization, preservation, and access;
- digital curation project management;
- digital curation funding, policy development, and long-term thinking.

Learning Outcomes

After completing the course, you should be able to:

- Describe the intellectual foundations 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, librarianship).
- Describe 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, labs, institutions), and domains (humanities, sciences, etc).

Course Schedule & Topics

Note: These are the likely sequence of topics and readings, however, the sequence of some units may change. The schedule will be updated to reflect any changes in the weekly topics.

Wk #	Date	Topic/Due	Lab/Activity (if applicable)
1	1/9	Introductions, Foundations & Definitions	Comparing definitions of “digital curation”
2	1/16	OAIS	CLI (re)introduction
3	1/23	Digital Objects / File Formats	File Inspection, CLI Navigation
4	1/30	Digital Objects / Significant properties & Materialities	Embedded metadata, CLI for inspecting files
5	2/6	Metadata for Preservation	File Characterization w/ Siegfried & Brunnhilde; CoreTrustSeal
6	2/13	Curating Scientific & Quantitative data	Tabular data (w/ CLI & spreadsheets)
7	2/20	Curating Humanities data <i>Due: Assignment 1</i>	Text processing fundamentals (w/ CLI)
8	2/27	Accessioning and Arranging/Data Packaging	File packaging (w/ BagIt)
--	3/6	NO CLASS - SPRING BREAK	NO CLASS - SPRING BREAK

9	3/13	Digital Forensics & BitCurator - legacy media, reporting	Bit Curator
--	3/20	<i>Abbreviated class - possible BitCurator Q&A</i>	GUEST LECTURE - Trevor Owens
10	3/27	Data use & reuse	DMPTool
11	4/3	Data sharing & ethics <i>Due: Assignment 2</i>	Grant Writing
12	4/10	Research data management	Data wrangling with OpenRefine
13	4/17	Wrap Up; Project report-outs <i>Due: Assignment 3 (presentation)</i>	
--	4/24	<i>Due: Assignment 3 (written proposal), any other overdue assignments</i>	

Course Schedule & Readings

Note: Readings preceded by [SKIM] or [REFERENCE] or [RESOURCE] are offered as additional materials, which you may find useful to consult while completing assignments or desire access to in the future. Note that each indication suggests you do not need to read it in detail during the term, but rather:

- *[SKIM] - take a brief look; to skim a reading, it is useful to read the title, abstract, and any introductory material; then, look over the general structure (perhaps glance any sections that seem particularly interesting); finally, look at any conclusions or recommendations*
- *[REFERENCE] - this is something that might be useful if you need to find more information about a specific aspect of a topic or may find useful later on; for example, there is a resource on the UNIX file system, which you don't need to read but could be useful if you want to learn more about a particular aspect of the file system or a specific command. If you find yourself using UNIX in the future, you might want to return to this later on.*
- *[RESOURCE] - these are included in addition to the regular readings for each course module and may contain additional information or useful content about a topic, but it is not expected that you read each one. May be suitable for coming back to in future.*

Week 1 (meeting Jan 9, discussion up to Jan 14): Foundations

- Yakel, Elizabeth, "Digital Curation," *OCLC Systems & Services: International Digital Library Perspectives* 23, no. 4 (2007): 335–40.
<https://doi.org/10.1108/10650750710831466>.
- Hedstrom, Margaret et al., "New Imperatives in Digital Curation and Its Workforce" (Ch 1), from *Preparing the Workforce for Digital Curation* (National Academies Press, 2015).
- Karen Weise, "A.I., the Electricians, and the Boom Towns of Central Washington," *New York Times*, 25 December 2024, at
<https://www.nytimes.com/2024/12/25/technology/ai-data-centers-electricians.html>
(note free subscription available with your U-M email, see
<https://www.csg.umich.edu/subscriptions>; and via MLibrary,
<https://search.lib.umich.edu/databases/record/46372>).
- Pranshu Verma, "In the shadows of Arizona's data center boom, thousands live without power," *Washington Post*, 23 December 2024, at
<https://www.washingtonpost.com/technology/2024/12/23/arizona-data-centers-navajo-power-aps-srp/>.
- [SKIM] Costis Dallas, "Digital curation beyond the "wild frontier": a pragmatic approach," *Archival Science* 16 (2015): 421–457. DOI: [10.1007/s10502-015-9252-6](https://doi.org/10.1007/s10502-015-9252-6).
- [RESOURCE] Melvin Vopson, "The world's data explained: how much we're producing and where it's all stored," *The Conversation* (4 May 2021), at
<https://theconversation.com/the-worlds-data-explained-how-much-were-producing-and-where-its-all-stored-159964>.

Week 2 (meeting Jan 16, discussion up to Jan 21): OAIS

- Brian Lavoie, "OAIS Introductory Guide (2nd Edition)," DPC Technology Watch Reports (2014), <http://dx.doi.org/10.7207/TWR14-02>.
- Doug Bierend, "The Hackers Who Recovered NASA's Lost Lunar Photos," *Wired* (23 April 2014),
<http://www.wired.com/2014/04/lost-lunar-photos-recovered-by-great-feats-of-hackerdom-developed-at-a-mcdonalds/>.
- Owens, Trevor, "Beyond Digital Hype and Digital Anxiety" (Ch 1), from *The Theory and Craft of Digital Preservation* (pre-print draft, 2017), <https://osf.io/5cpjt/>.
- Hedstrom, Margaret et al., "The Current State of Digital Curation" (Ch 2), from *Preparing the Workforce for Digital Curation* (National Academies Press, 2015),
http://www.nap.edu/download.php?record_id=18590.
- [REFERENCE] Corrado & Moulaison, "The OAIS Reference Model" (Ch 3, pp. 43–54), in *Digital Preservation for Libraries, Archives, and Museums*.
- [REFERENCE] Reference Model for an Open Archival Information System (OAIS). 2012. Recommended Practice. CCSDS 650.0-M-2 (Magenta Book) Issue 2 June 2012. Consultative Committee for Space Data Systems: Washington, DC. Available at:
<http://public.ccsds.org/publications/archive/650x0m2.pdf>.

Week 3 (meeting Jan 23): Digital Objects 1: File Formats & Boundaries

- Owens, "Artifact, Information, or Folklore: Preservation's Divergent Lineages" (Ch 2), "Understanding Digital Objects" (Ch 3), from *The Theory and Craft of Digital Preservation* (pre-print draft, 2017), <https://osf.io/5cpjt/>.
- Bertram Lyons, "What is the Chemistry of Digital Preservation?" *AVP Blog*, 18 February 2016, at <https://www.weareavp.com/what-is-the-chemistry-of-digital-preservation/>.
- Library of Congress (n.d.) *Sustainability of Digital Formats*, at <https://www.loc.gov/preservation/digital/formats/intro/intro.shtml>. READ "Overview," and SKIM "Formats, Evaluation Factors, and Relationships", the Content Categories and at least one Format Description.
- [REFERENCE] Wikipedia, "Byte," at <https://en.wikipedia.org/wiki/Byte>.
- [REFERENCE] David S.H. Rosenthal "Format Obsolescence: Assessing the Threat and the Defenses." *Library Hi Tech* 28, no. 2 (2010): 195-210. doi: [10.1108/07378831011047613](https://doi.org/10.1108/07378831011047613).
- [REFERENCE] Paul Murrell, "5.2 Plain Text formats" (read sections 5.2 up to 5.2.7), from *Introduction to Data Technologies*, at <https://www.stat.auckland.ac.nz/~paul/ItDT/HTML/node38.html>.
- [REFERENCE] "Regular Expressions," *Library Carpentry Introduction to Data* at <https://librarycarpentry.github.io/lc-data-intro/01-regular-expressions.html>.
- [REFERENCE] Anthony Cocciolo (2014) "Unix Commands and Batch Processing for the Reluctant Librarian or Archivist," *code4lib Journal* 2014, no 23, at <http://journal.code4lib.org/articles/9158>.

Tools: Hex Viewer (VS Code Extension)

- "Hex Editor" VS Code extension:
<https://marketplace.visualstudio.com/items?itemName=ms-vscode.hexeditor>
(ms-vscode.hexeditor)

Week 4 (meeting Jan 30): Digital Objects 2: Significant Properties & File Format Risks

- Owens, Trevor, "The Craft of Digital Preservation" (Ch 5), "Managing Copies & Formats" (Ch 7), from *The Theory and Craft of Digital Preservation* (pre-print draft, 2017), <https://osf.io/5cpjt/>.
- Matthew Kirschenbaum, "The Textual Forensics of Mystery_House.dsk," in *Mechanisms* (MIT Press, 2008), 111-158.
- Doug Reside, "'No Day But Today': A look at Jonathan Larson's Word Files," *New York Public Library blog* (22 April 2011), at <https://www.nypl.org/blog/2011/04/22/no-day-today-look-jonathan-larsons-word-files>.
- [SKIM] Jordan Eppink, "A Brief History of the GIF (So Far)," *Journal of Visual Culture* 13 (2014): 298-306. DOI [10.1177/1470412914553365](https://doi.org/10.1177/1470412914553365).
- [REFERENCE] Knight, Gareth, and Maureen Pennock, "Data without Meaning: Establishing the Significant Properties of Digital Research," *International Journal of Digital Curation* 4 (2009), DOI: <https://doi.org/10.2218/ijdc.v4i1.86>.

- [REFERENCE] Niamh Murphy, "A Beginner's Guide to Brunnhilde: Reasons for Using the Brunnhilde Software," *Digital Preservation Coalition Blog* (16 February 2022), <https://www.dpconline.org/blog/blog-niamh-murphy-brunnhilde-reasons>.

Tools: RegEx

- Reference: [Library Carpentry useful regex characters](#) & what they do
- Reference: [RegEx one-pager](#) from an MIT linguistics lab
- For testing regexes: <https://regex101.com/>

Week 5 (meeting Feb 6): Metadata for Preservation & Management

- Owens, Trevor, "Challenges & Opportunities of Digital Preservation" (Ch 4), from *The Theory and Craft of Digital Preservation* (pre-print draft, 2017), <https://osf.io/5cpjt/>.
- Jesse A. Johnston, "Metadata," in *Digital Preservation: A Critical Vocabulary*, release 1 (2024), at <https://digital-preservation-a-critical-vocabulary.pubpub.org/pub/2ib9hvvg/release/1>.
- Priscilla Caplan, *Understanding PREMIS*, Library of Congress (2009), <http://www.loc.gov/standards/premis/understanding-premis.pdf>.
- [SKIM] Library of Congress, "Jerry McDonough: Digital Preservation Pioneer," at <https://www.digitalpreservation.gov/series/pioneers/mcdonough.html>.
- [LAB REVIEW] Tim Walsh, "Data-Driven Reporting and Processing of Digital Archives with Brunnhilde," *Practical Technology for Archives* no. 8 (July 2017), now only available via the Internet Archive's Wayback Machine at https://web.archive.org/web/20170725074906/https://practicaltechnologyforarchives.org/issue8_walsh/.
- [REFERENCE] *PREMIS Data Dictionary version 3.0* (2015) at <http://www.loc.gov/standards/premis/v3/premis-3-0-final.pdf> and <http://www.loc.gov/standards/premis/>.
- [REFERENCE] Library of Congress, "METS: An Overview & Tutorial" at <http://www.loc.gov/standards/mets/METSOverview.v2.html>.
- [REFERENCE] For METS details, see *<METS> Primer and Reference Manual* at <https://www.loc.gov/standards/mets/METSPrimer.pdf> (v. 1.6 rev. 2010) and <http://www.loc.gov/standards/mets/>.
- [REFERENCE] Coyle, Karen, "Understanding Metadata and Its Purpose," *Journal of Academic Librarianship* 31 (2005): 160–163. DOI: <https://doi.org/10.1016/j.acalib.2004.12.010>.

Tools: working with XML

- "XML Tools" VS Code Extension: <https://marketplace.visualstudio.com/items?itemName=DotJoshJohnson.xml> (DotJoshJohnson.xml)

- “XML” by Red Hat VS Code Extension:
<https://marketplace.visualstudio.com/items?itemName=redhat.vscode-xml>
(redhat.vscode-xml)
- Terminal (bash): optional tool is [xmlstarlet](#) (not planning to cover in class, but if you ever want or need to work with XML on the command line, this is a current option)

Week 6 (meeting Feb 13): Domain Practices 1: Quantitative and Empirical

- Renear, A. H., Sacchi, and Wickett (2011) “Definitions of dataset in the scientific and technical literature,” *Proceedings of the American Society for Information Science and Technology*. DOI: [10.1002/meet.14504701240](https://doi.org/10.1002/meet.14504701240).
- Christine Borgman (2015) “Provocations” and “What are Data?” Chs 1 & 2, from *Big Data, Little Data, No Data* (MIT Press). DOIs: 10.7551/mitpress/9963.003.0005 & 10.7551/mitpress/9963.003.0006
- Jane Greenberg, Hollie C. White, and Ryan Scherle, “A Metadata Best Practice for a Scientific Data Repository,” *Journal of Library Metadata* 9 (2009): 194-212. DOI: [10.1080/19386380903405090](https://doi.org/10.1080/19386380903405090).
- Cornell Research Data Management Service Group, “Guide to writing ‘readme’ style metadata,” at <https://data.research.cornell.edu/content/readme> (as of 2024).
- [SKIM] Sarah Higgins, “The DCC Curation Lifecycle Model,” *International Journal of digital Curation* 1 (2008), 135-140. DOI: <https://doi.org/10.2218/ijdc.v3i1.48>.
- [REFERENCE] Karl W. Broman & Kara H. Woo “Data Organization in Spreadsheets,” *The American Statistician*, 72 (2018), at [10.1080/00031305.2017.1375989](https://doi.org/10.1080/00031305.2017.1375989)
- [REFERENCE] Hadley Wickham, “Tidy Data,” *Journal of Statistical Software* 59 (2014). At <https://www.jstatsoft.org/article/view/v059i10>. See also Hadley’s manual on “Data Tidying” in R: <https://r4ds.hadley.nz/data-tidy.html>.
- [REFERENCE] Francine Berman, “Got Data? A Guide to Data Preservation in the Information Age,” *Communications of the ACM* 51 (2008): 50-56. DOI: [10.1145/1409360.1409376](https://doi.org/10.1145/1409360.1409376).

Tools: working with CSV

- “Rainbow CSV” VS Code Extension:
<https://marketplace.visualstudio.com/items?itemName=mechatroner.rainbow-csv>
(mechatroner.rainbow-csv)
- How to save data in a CSV format using Excel:
<https://librarycarpentry.org/lc-spreadsheets/05-exporting-data/index.html>
- “Opening a CSV File in Excel,” Federal Communications Commission, at <https://www.fcc.gov/general/opening-csv-file-excel>
- How to open CSV files in Excel (Microsoft):
<https://support.microsoft.com/en-us/office/import-or-export-text-txt-or-csv-files-5250ac4c-663c-47ce-937b-339e391393ba>

Week 7 (meeting Feb 20): Qualitative Data & Curation

- Christof Schöch, "Big? Smart? Clean? Messy? Data in the Humanities," *Journal of Digital Humanities* 2 (2013), at <http://journalofdigitalhumanities.org/2-3/big-smart-clean-messy-data-in-the-humanities/>.
- Katie Rawson & Trevor Muñoz, "Against Cleaning," available at <http://curatingmenus.org/articles/against-cleaning/>.
- Donald Waters, "The Emerging Digital Infrastructure for Research in the Humanities," *International Journal on Digital Libraries* (2022). DOI: [10.1007/s00799-022-00332-3](https://doi.org/10.1007/s00799-022-00332-3)
- [REFERENCE] Christine Borgman, "Data Scholarship in the Humanities," in *Big Data, Little Data, No Data: Scholarship in the Networked World* (MIT Press, 2015), 161-202.

Week 8 (meeting February 27): Accessioning and Arranging/Data Packaging

- Trevor Owens, "Arranging and Describing Digital Objects" (Ch 8), from *The Theory and Craft of Digital Preservation* (pre-print draft, 2017), <https://osf.io/5cpjt/>.
- Rhee (2022) "A New Lifecycle Model Enabling Optimal Digital Curation," *Journal of Librarianship and Information Science*, 1-26, doi: [10.1177/09610006221125956](https://doi.org/10.1177/09610006221125956).
- Phillips, M., et al. "The NDSA Levels of Digital Preservation: An Explanation and Uses" (2013), whitepaper posted at https://www.digitalpreservation.gov/documents/NDSA_Levels_Archiving_2013.pdf.
- [REFERENCE] John Kunze, et al. (2018), "The BagIt File Packaging Format (V1.0)," RFC 8493, Internet Engineering Task Force, at <https://www.rfc-editor.org/rfc/rfc8493>.

Tools: working with BagIt

- [REFERENCE] BagIt python module <https://github.com/LibraryOfCongress/bagit-python>

March 6 - Spring Break - NO MEETING

Week 9 (meeting March 13): Digital Forensics & BitCurator

- Corinne Rogers, "From time theft to time stamps: mapping the development of digital forensics from law enforcement to archival authority," *International Journal of Digital Humanities* (2019): 13–28. DOI: [10.1007/s42803-019-00002-y](https://doi.org/10.1007/s42803-019-00002-y)
- Sam Meister and Alexandra Chassanoff, "Integrating Digital Forensics Techniques into Curatorial Tasks: A Case Study," *International Journal of Digital Curation* 9 (2014): 6-16. DOI: [10.2218/ijdc.v9i2.325](https://doi.org/10.2218/ijdc.v9i2.325).
- Adrian Brown, "Digital Preservation Guidance Note 3: Care, handling and storage of removable media," (2008), The National Archives (DPGN-03). Available via Canvas.

- [RESOURCE] Julianna Barrera-Gomez and Ricky Erway, "Walk This Way: Detailed Steps for Transferring Born-Digital Content from Media You Can Read In-house," 2013, OCLC Research. Available via Canvas & at <https://www.oclc.org/content/dam/research/publications/library/2013/2013-02.pdf>.
- [RESOURCE] Christopher A. Lee, et al, "From Bitstreams to Heritage: Putting Digital Forensics into Practice in Collecting Institutions," whitepaper for *BitCurator*, September 2013. Available via Canvas & BitCurator.net
- [REFERENCE] Smithsonian Time-Based Media & Digital Art, "Tableau T8u Write-blocker," at <https://www.si.edu/tbma/tableau-t8u-write-blocker>.
- [REFERENCE] "How to Understand the Filesystem in Linux," Digital Ocean (2013), <https://www.digitalocean.com/community/tutorials/how-to-understand-the-filesystem-layout-in-a-linux-vps>

March 20 - Guest Lecture by Trevor Owens (location will be at the Rackham Building - next door to North Quad)

Note: Possible abbreviated Q&A session or work session for first hour of class.

Week 10 (meeting March 27): Data Use and ReUse; Grantwriting

- Pasquetto, I. et al. (2017). "On the Reuse of Scientific Data." *Data Science Journal* 16:8, pp. 1-9, DOI: <https://doi.org/10.5334/dsj-2017-008>.
- Dietrich, D., Adamus, T., Miner, A., & Steinhart, G. (2012). "De-Mystifying the Data Management Requirements of Research Funders." *Issues in Science and Technology Librarianship* no. 70. <https://doi.org/10.29173/istl1556>.
- Bess G. DeFarber, "Strategies for Completing Application Components" (Ch 6), from *Collaborative Grant-Seeking: A Practical Guide for Librarians* (Rowman & Littlefield, 2016), pp. 99-128.
- [REFERENCE] Sarah Higgins, "The DCC Curation Lifecycle Model," *International Journal of Digital Curation* 1 (2008): 135-140. DOI: [10.2218/ijdc.v3i1.48](https://doi.org/10.2218/ijdc.v3i1.48).

Tools: [grant planning spreadsheet](#), [DMP \(Data Management Plan\) Tool](#)

- Johnston, Funding opportunity planning template (reusable), <https://bit.ly/rln-2022-grant-workshop-materials>.
- DMP Tool: <https://dmptool.org/>

Week 11 (meeting April 3): Data Ethics & Access

- Carroll, S., et al. (2020) "The CARE Principles for Indigenous Data Governance," *Data Science Journal* 19, pp. 1-12. DOI: <https://doi.org/10.5334/dsj-2020-043>. See also <https://www.gida-global.org/care>.

- Wilkinson, M., et al. (2016) "The FAIR Guiding Principles for Scientific Data Management and Stewardship," *Scientific Data* 3, 160018. DOI: <https://doi.org/10.1038/sdata.2016.18>.

Week 12 (meeting April 10): Research Data Management; Data Wrangling

- Andrikopoulou et al., "Research Data Management (RDM) and the nEvolving Identity of Academic Libraries and Librarians: A Literature Review," *New Review of Academic Librarianship* 28 (2022): 349–365, DOI: [10.1080/13614533.2021.1964549](https://doi.org/10.1080/13614533.2021.1964549).
- Owens, "Enabling Multimodal Access and Use" (Ch 9) and "Tools for Looking Forward" (Ch 10), from *The Theory and Craft of Digital Preservation* (pre-print draft, 2017), <https://osf.io/5cpjt/>.
- Hedstrom, Margaret et al., "Current and Future Demand for a Digital Curation Workforce" (Ch 3), from *Preparing the Workforce for Digital Curation* (National Academies Press, 2015), http://www.nap.edu/download.php?record_id=18590.

Tools: working with OpenRefine

- [RESOURCE/REFERENCE] Library Carpentry, "Open Refine," at <https://librarycarpentry.org/lc-open-refine/>. DOI: <https://doi.org/10.5281/zenodo.3266144>.

Job Resources: [ALA Job List](#); [SAA career center](#); [ArchivesGig](#); [Google Jobs](#); LinkedIn; ZipRecruiter;

Week 13 (meeting April 17): Project report-outs / grant pitches

- Presentation slides and class presentation of your grant proposal due on **April 17**.
- Final written proposal (in the style of IMLS two-pager) due on **April 24**.

Assignments & Activities

More details about each assignment and activity will be posted on the course Canvas site. At a high level, your learning will be based on:

- **Engagement/Discussions (10%).** Generally, each week will include a structured discussion, activity, or other component, which is designed to help you work through ideas and concepts relevant to the week. These are generally designed for in-person engagement, and synthesis is demonstrated through integrating your personal thoughts on the readings, contributions or questions in class, and sharing in "share out" group discussions. Beyond the weekly meetings, the main venue for engagement is the online discussion forums. Each person should sign up for one week of initiating the discussion (choose a reading, summarize it and note questions for further thoughts), which will be 2 points; during other weeks, you should contribute to the

discussion in a substantive way, and each discussion contribution is 1 point. This area constitutes up to 10 points of your overall grade.

- **Assignments (60%).** A series of assignments throughout the course constitutes the bulk of your course grade. Each assignment can be up to 20 points. Assignment points will be assessed according to the rubrics provided for each assignment. Assignment details will be provided later in the course, but these will be according to particular topics:
 - **Assignment 1: Dataset analysis (individual) - Due Feb 20.** This assignment will be a written essay that analyses curation needs of a dataset. You will have the option to choose from a series of lightly curated, publicly available datasets, and your essay will discuss the nature of the dataset, its provenance, what metadata is present or needed for preservation, and based on the identification of a user community, what curation activities you would propose to undertake were this dataset accessioned by your organization for long-term preservation.
 - **Assignment 2: BitCurator exercise (pair or small group) - Due April 3.** This assignment will combine hands-on technical activity and written reflection. Your basic task will be to use legacy hardware to create an accessionable data package of digital information from legacy, removable media (floppy disks). We will demonstrate all of the steps required - some of the software is available to you on your own laptop, but we are setting up a physical workstation to promote working with legacy digital media.
 - **Assignment 3: Data curation grant proposal (small group) - Due April 17/24.** This group assignment will be a chance for you to combine your technical skills and understandings with development of a professional competence: grant writing. Your group will write a proposal for a small grant to accomplish the curation of a dataset, or a larger grant that proposes an event or longer-term curation project dealing with one of the course datasets.
- **Lab activities (30%).** This portion of the course activities will be structured around in-class demonstrations and discussion, which will be followed by reflection questions or problems to solve on your own. All files and tools will be available during class or via UMSI-provided equipment. As a general rule, these will alternate weeks and be assigned or demoed on Thursday, and work will be due on Tuesday (so you have the weekend to work on things). Lab assignments will generally be worth 5 points and you can earn 30 points in this section of the grade.

Grading

Following Covid-19 and learning about the various [equity issues that exist around grading](#), I have endeavored to use a less complex grading approach (it does not involve slicing and dicing +/-, percentages, or elaborate point structures). This approach aims to emphasize engagement with the course ideas, and emphasize the effort put toward learning, understanding, and communicating the topics covered in the course. For each component of the grade, I will define and communicate with you what I consider to be “complete” (including what is required, how to submit, and when it must be done); complete assignments are evaluated as

explained below. This grading approach is intended to be more pragmatic and based on the philosophy that your grade reflects a professional credential that indicates you have a competence in the course's subject matter. In my estimation, the following evaluations reflect this:

Letter	Total Course Points	Description
A	>=90	You have done an excellent job, the assignment meets all the stated criteria and/or requirements, it was complete and on time, and you have demonstrated competency or engagement with the material (hopefully you've also learned something of it and gained confidence along the way).
B	80-90	You have done the work, but something was lacking (for example, the work was late, it was incomplete based on the assignment expectations, etc).
C	<80	Some of the work was done, but some of the work was missing or did not meet expectations.

Please do ask questions about this approach, particularly if the expectations and requirements for a particular assignment or portion of your grade are not clear. I will explain or provide more information as needed.

Additional Course Policies & Information

Materials

To the extent possible, the course adopts readings and resources that are available as open-access or free tools available on the web. In some cases, resources may require login, but all materials should be freely available to students using their University credentials. The instructor will provide reasonable information to allow you to locate items (e.g., DOIs, URNs or URLs); in some cases, hard-to-find items may be posted for individual educational usage on the course Canvas site.

We will undertake in-class activities, so you should plan to use a laptop regularly (and actively) in class. Tablets and phones may be useful for taking notes, but as a caution, these devices may not support all of the tasks and tools that we will be using. All software used will be accessible as free and open source software, via virtualized environments provided by UMSI, or via workstations set up and operated by UMSI. Instructions on accessing these will be distributed after the course begins, as will information about any specific software and tools that you may need.

Course information, readings (or information on where to find them), assignments, and other course resources are available via Canvas (link above).

Flexibility of Assignment Submission Deadlines

Please take time at the beginning of the term to review the assignment due dates and other cyclical deadlines, such as discussion forum posts and Labs. I expect you to honor these deadlines to the extent possible. That said, I am also aware that schedules planned are often not possible to maintain to the letter. Thus, any assignment can be submitted up to two days late with no penalty or questions asked. Keep in mind that you will receive more benefit from some assignments by completing them at the same time as your colleagues (discussion posts, for example, are more useful if they can be part of the discussion versus an ex post facto reply). If lateness becomes a recurrent pattern, I may ask you about it but am willing to work with you on developing a schedule that works for you.

The main exception to the late assignment policy is the end of the course. Due to the university's calendar and expectations that I submit grades on time for students to graduate, etc, I do not offer flexibility for the end of term deadline (noted in the Course Schedule).

Accessibility

I am dedicated to providing an accessible classroom environment and accessible learning tools. If you have any questions or requests that I am not aware of, please do speak with me. Here are some of the things that I am thinking about in this area:

- Reducing the amount of navigation options in the Canvas site to only those that are necessary. If you have questions or are in need of functions that are not opened, please let me know.
- Providing accessible slides and demonstrations during class: I have generally reduced the number and amount of images in my slides, I favor text, use high contrast, and adopt consistent slide layouts. I will also make slides directly available via Google slides so that you can see them on your own device, download them, or annotate them for yourself.
- Flexibility around deadlines. The course schedule as posted is what I plan to follow. In most cases, I ask that you adapt to this schedule, but I understand that in some cases you may need extra time, have a need for additional submission options or timelines, or have other requests. If so, please let me know your needs or requests in advance.

Engagement

Participation in class activities and providing timely responses on assignments is expected. If you are unable to participate in the course or meet any of the required assignments and due dates, please let the instructor know as soon as possible.

That said, do not come to class if you are sick! Lecture portions of class should be recorded automatically and will be available via Canvas. As of the beginning of class, I plan to lecture

without a mask when I am in front of the classroom, but I will keep in touch with the current campus guidelines and health situation, while to the extent possible making decisions informed on the data and guidance that we receive.

Part of the engagement component is being an active member of our learning community. Throughout the semester, various opportunities to demonstrate this engagement and community care will be offered. One of these is to be an authentic participant (I understand that each of us has various comfort levels and social interaction styles, but since dialogue and interaction is part of most professional positions, I do invite you to use the course time as a place to practice or model active engagement), think about the readings and respond to them, ask questions when things are not clear. We will also have a collaborative course notes document; consider being a community scribe for at least one week - it is helpful to the class, and it is another way to show your engagement.

Discussions. In order to learn, we must be open to the views of people different from ourselves. In this time we share together over the semester, please honor the uniqueness of yourself and that of your fellow classmates, welcome and appreciate the opportunity we have to learn from one another, and let us get to know each other as colleagues. Fruitful discussion requires civil and engaged dialogue, which includes respecting each others' opinions and refraining from personal attacks or demeaning comments of any kind.

In this course, I am dedicated to providing and fostering a welcoming, supportive, and safe learning environment for everyone, regardless of background or identity. While participating in this class, you should also abide by these principles (developed from the [Carpentries Code of Conduct](#), see also the CRLT "[Guidelines for Classroom Interactions](#)") and accept that feedback and resolution may be addressed in consultation with the instructor. Any form of behavior that excludes or intimidates someone else in the course is a violation of these principles. In order to foster a positive and professional learning environment, I encourage the following:

- Use welcoming and inclusive language
- Be respectful of different viewpoints and experiences
- Gracefully accept constructive criticism
- Focus on what is best for us all as a community of learners
- Show courtesy and respect towards other community members
- If you believe someone is violating these principles, speak with the instructor.

These are principles and values, not rules. However, it is important that we all take part in practicing them. If you have concerns about them or see areas where they are being ignored, please chat with the instructor.

Screen Etiquette in Class

You are encouraged to bring laptops or portable digital devices to class, on the assumption that you will use it as an active learning tool. Therefore, you should:

- Use laptops for taking notes, conducting research required for in-class activities, or for other specific classroom tasks assigned. During class, limit checking email, chatting with others on non-course-related topics, IMing, gaming, or otherwise engaging in off-task activities. In part this is to limit distraction to your classmates, but keep in mind that this sort of activities will divide your attention and can distract you from the learning process.
- Engage in-class activity actively. Your computer (or other devices) should not become a barrier to one-on-one interaction, instead use it to facilitate engagement and the exchange of ideas during class.
- Demonstrate sensitivity to others. Do not display screen images, including wallpapers, lock screens, or screensavers, which might be distracting or offensive to others in class.

Attendance & Engagement

Given the recent shifts to online instruction, hybrid courses, and the possibility that pandemic- or health-related issues may prevent us from meeting in person every session, this is a developing area. Attendance, in the sense of engaging with all of the units and course topics, is expected. If you are unable to attend class or cannot complete asynchronous work for some reason, please get in touch with the instructor. In the event of attendance problems, the instructor will address them on a case-by-case basis.

Because of the wide variation in individual needs and accommodations, there is no hard and fast rubric for the engagement portion of the grade. The grade is for the most part determined by your participation in activities and online discussions, but if there are chronic problems or need for alternative arrangements, these may be reflected differently in each individual's grades and evaluations.

Communication

It is possible that you may not receive an immediate response, but the instructor will respond to all emails within 24 hours.

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.

Generative AI

Emerging generative AI tools can be useful (in general, and maybe in this class). Using generative AI in course activities and assignments is permitted, unless indicated otherwise (see below for specific prohibitions). If an assignment requires you to cite sources, you should include a note about your prompts and outputs if you use generative AI.

You should exercise caution. Generative AI tools assemble text and other content in ways that may or may not reflect reality, they may combine unexpected text snippets or facts, create plausible (but unreal) citations, and other artifacts, in ways that are difficult to anticipate or notice. You are responsible for content you appropriate from generative AI outputs. In the context of this class, I suspect that you may find it simpler to complete the assignments using your own knowledge and expertise rather than correcting output from AI tools. But who knows! You might also consider what it means if your ethics is indistinguishable from the median output of an AI tool.

Regarding particular usages:

1. **For Programming tasks:** In principle you may submit AI-generated code, or code that is based on or derived from AI-generated code, as long as this use is properly documented in the comments: you need to include the prompt and the significant parts of the response. AI tools may help you avoid syntax errors, but there is no guarantee that the generated code is correct. It is your responsibility to identify errors in program logic through examination or testing. Moreover, generated code, even if syntactically correct, may have significant scope for improvement, in particular regarding implementation in an actual workflow.

2. **To create a design, generate reading responses, analyze data, or do other “thinking” for you, DO NOT use AI tools** without written permission from the primary instructor. Indications of such use may be reported for investigation as academic misconduct, and subject to consequences like failing the assignment or failing the course depending on the scope and severity of the actions taken.

Your Health and Wellbeing

The life of learning and scholarship, and life on campus in general, can be stressful. If you're feeling stressed or stuck, remember that you are not alone, and there are many resources on campus that are here to support you. Students may experience stressors that can impact both their academic experience and their personal well-being. These may include academic pressure and challenges associated with relationships, mental health, alcohol or other drugs, identities, finances, etc.

If you are experiencing concerns, seeking help is a courageous thing to do for yourself and those who care about you. If the source of your stressors is academic, please contact me so that we can find solutions together. For personal concerns, U-M offers the following resources:

- [Counseling and Psychological Services \(CAPS\)](#) - confidential; 734-764-8312; for after-hours urgent support, call and press 0; counseling, workshops, groups and more. Ashley Evealitt, a CAPS counselor is embedded in UMSI, information about how to schedule an appointment with her can be found [here](#).
- [Dean of Students Office](#) - 734-764-7420; provides support services to students and manages critical incidents impacting students and the campus community
- [Ginsberg Center for Community Service Learning](#) - 734-763-3548; opportunities to engage as learners and leaders to create a better community and world
- [Maize and Blue Cupboard \(MBC\)](#) - 734-936-2794; Food pantry with groceries, kitchen and cooking supplies, personal and household items, and support
- [Multi-ethnic Student Affairs \(MESA\)](#) - 734-763-9044; diversity and social justice through the lens of race and ethnicity
- [Office of Student Conflict Resolution](#) - 734-936-6308; offers multiple pathways for resolving conflict
- [Office of the Ombuds](#) - 734-763-3545; students can raise questions and concerns about the functioning of the university.
- [Services for Students with Disabilities \(SSD\)](#) - 734-763-3000; accommodations and access to students with disabilities
- [UWill](#) - video, phone, chat, and message sessions with licensed mental health professionals at no cost, access provided by CAPS
- [Stressbusters](#) - mobile app, access provided by CAPS, allows to view, share, store and use campus event announcements, news alerts, student and department videos, one-touch dial emergency buttons, health tips, surveys and other features (for Android and iOS mobile devices). Offers a “sonic spa” feature with calming music and sounds.

- [Sexual Assault Prevention and Awareness Center \(SAPAC\)](#) - confidential; 734-764-7771 or 24-hour crisis line 734-936-3333; addresses sexual assault, intimate partner violence, sexual harassment, and stalking
- [Spectrum Center](#) - 734-763-4186; support services for LGBTQ+ students
- [Trotter Multicultural Center](#) - 734-763-3670; intercultural engagement and inclusive leadership education initiatives
- [University Health Service \(UHS\)](#) - 734-764-8320; clinical services include nurse advice by phone, day or night
- [Well-being for U-M Students website](#) - searchable list of many more campus resources
- [Wolverine Wellness](#) - 734-763-1320; provides Wellness Coaching, Collegiate Recovery Program, and much more

Disability Statement

The University of Michigan recognizes disability as an integral part of diversity and is committed to creating an inclusive and equitable educational environment for students with disabilities. Students who are experiencing a disability-related barrier should contact [Services for Students with Disabilities](#) (SSD) at <https://ssd.umich.edu/>; 734-763-3000; or ssdoffice@umich.edu. For students who are connected with SSD, accommodation requests can be made in Accommodate. If you have any questions or concerns please contact your SSD Coordinator or visit SSD's Current Student webpage. SSD considers aspects of the course design, course learning objects and the individual academic and course barriers experienced by the student. Further conversation with SSD, instructors, and the student may be warranted to ensure an accessible course experience. The instructional team will treat any information that you provide in as confidential a manner as possible.

Class Recordings

All sessions will be audio and video recorded to enable access for those who would like to consult recordings for reference or who cannot attend class in person on a given day. These recordings will not be made available publicly, but recordings of all sessions will be available on Canvas only to students registered for this class. As part of your participation in this course, your image or voice may be recorded. If you do not wish to be recorded, please contact the professor during the first week of class to discuss alternative arrangements. The camera only picks up the front of the room (instructor and slides), but this may require you to sit in a particular place in the room, outside the camera's view. Our classroom, NQ 2185 has a ceiling mic that picks up all voices in the room, although the most audible recording is from the instructor's microphone. Students may not copy and share the lecture videos with those not in the class, or upload them to any other online environment (this is a violation of the Federal Education Rights and Privacy Act (FERPA)).

Personal recordings are prohibited except with permission

Students are prohibited from recording/distributing any class activity without written permission from the instructor, except as necessary as part of approved accommodations for

students with disabilities. Any approved recordings may only be used for the student's own private use.