LIS 668
Digital Curation and Collections

Instructor: Dorothea Salo (please call me "Dorothea")   salo@wisc.edu, 4261 Helen C. White Hall
Office hours: by appointment   Course URL: https://canvas.wisc.edu/courses/169053
Special course attributes: Tier T   Pinboard URLs: https://pinboard.in/u:dsalo/t:668,
Course modality: Online   https://pinboard.in/u:dsalo/t:digitization

Introduction

Course description

Students completing this course will earn three credit hours. One credit is the learning that takes place in at least 45 hours of learning activities, which include time in lectures or class meetings, in person or online, labs, exams, presentations, tutorials, reading, writing, studying, preparation for any of these activities, and any other learning activities.

This course has no prerequisites or co-requisites.

This course explores core concepts and new developments in digital curation, preservation and digital collections. Topics include: digitization of various media; digital preservation; media archeology; basics of research data management; digital collection technologies and workflows; intellectual-property issues; metadata as applied in digital collections; digital collections planning and evaluation; trusted digital repositories; funding of digital collection projects and sustainability.

Students who successfully complete this course will learn to:

➢ Assess, plan for, manage, and execute a small-scale data-management or digital-preservation project.
➢ Assess digital or to-be-digitized data for preservability; make yes-or-no accessioning decisions.
➢ Appropriately manage intellectual-property issues related to data management and digital archiving.
➢ Understand (and where relevant, apply) technological, economic, and social models of digital preservation and sustainability.
➢ Understand forms, formats, and lifecycles of digital data across a wide breadth of contexts.
➢ Evaluate software and hardware tools relevant across the data lifecycle.
➢ Construct a current-awareness strategy; assimilate substantial amounts of relevant writing.
➢ Self-sufficiently acquire technical knowledge.

This course is designed to assess student progress in the following iSchool program-level outcomes: 1, 2, 4, 5, 7

Course Policies

I aim to make this course as accessible as possible to all students. Students seeking accommodations for lecture or assignments must obtain a McBurney Center Faculty Notification Letter. For more information, see `https://mcburney.wisc.edu/apply-for-accommodations/`**.**

Preferred name/pronouns: It is sometimes the case that a student's legal name or gender assigned at birth are reported to me on official documents in a form not in keeping with that student's preferred name or gender expression. Please let me know, as you are comfortable, about your preferences. My pronouns are she/her/hers. UW-Madison also permits students to indicate a preferred name: `https://registrar.wisc.edu/preferred_name.htm`

Contacting me
READ THE SYLLABUS before asking a question, please; the syllabus may answer it! For any difficulty with the course that is not private or confidential, please use the Canvas questions forum; I *will not answer such questions by email*. Please also do your best to assist your classmates.

Should you see dead links (it does happen, usually with no notice), weird due dates, or other syllabus problems, please bring them up in the appropriate Canvas forum.

Textbooks

There are no required textbooks for this course. Reading material is largely drawn from the web.

Assignments

Grading scale

Please review policies in the iSchool MA Handbook on final course grades below B.

A 94-100 Outstanding work. Student performance demonstrates full command of course materials. Work shows a degree of synthesis and creativity that surpasses course expectations.
AB 88-93 Very good work. Student performance demonstrates thorough knowledge of course materials. Work shows a degree of synthesis and creativity that is superior.
B 82-87 Good work. Student performance demonstrates the ability to meet designated course expectations. Overall work is at an acceptable level.
BC 77-81 Marginal work. Student performance demonstrates incomplete understanding of course materials. Or student fails to meet deadlines.
C 72-76 Unsatisfactory work. Student performance demonstrates inadequate understanding of course materials. Or student fails to meet deadlines.
D 67-71 Very unsatisfactory work. Student performance demonstrates inadequate understanding of course materials. Or student fails to meet deadlines.
F 66 and below Completely unsatisfactory work. Student performance demonstrates very inadequate understanding of course materials and serious lack of competence on site. Or student misses many deadlines.

Due dates

Due dates below are specified by module (mostly for my reference). Specific due dates/times are in the Canvas calendar.

Assignment	Final-grade %	Due (actual due date in Canvas)
Weekly planning-project assignments	70%	Each week
Final planning-project report	10%	Final day of course
Other assignments (see Canvas)	20%	As listed in Canvas

Late assignments will be penalized one final-grade percentage point per day or fraction thereof late. I will allow revision and resubmission at my sole discretion and on my schedule only; any student resistance will remove the opportunity.

Digitization and digital preservation planning project

This semester-long small-group (four people max per group) project takes you through the entire project-planning process for digitization of heterogeneous materials, as well as digital preservation of the results of digitization and of born-digital collecting. You will do almost all the work in weekly stages (no end-of-semester all-nighters!), building on and refining your earlier work as you go. Your final deliverable will be a project-planning report and production manual, similar to:

> ➢ https://www.gwbhs.org/documents/2012/11/si-644-digitization-project-plan.pdf/ (hereafter "Bloomfield")
> ➢ https://web.archive.org/web/20120703181811/http://www.jisc.ac.uk/media/documents/programmes/digitisation/econtent/heritagedigi_pp.pdf (hereafter "JISC")
> ➢ http://indigitization-toolkit.sites.olt.ubc.ca/files/2018/03/2018-Indigitization-Toolkit-v04.pdf (hereafter "Indigitization")

Everyone is encouraged to ask questions, request and receive feedback on drafts, and share useful information related to this project via its dedicated Canvas forum. I see absolutely no point in information-hoarding here—the point is everyone learning!

Each weekly deliverable is worth five final-grade points (= 70 total) and is due on the first day of the subsequent week (so, Week 1's deliverable is due the first day of Week 2). Assuming deliverables are turned in on time and represent honest effort, you will only lose points for gross avoidable errors that would damage your project (e.g. choosing JPEG or mp3 as an archival master file format).

Week 1: Describe the **community or event** defining your collection, and **describe and enumerate the collection** itself. Requirements:

- ➢ The community or event must either exist in the present day or be recent enough to have an online presence (website and/or social media). It *must* involve a group or demographic on the less-empowered side of a societal power differential. (This is important because I will ask you to consider community engagement and ethics in future weeks, and power is a tremendously important aspect of both.) Choosing a community or event that a group member has direct experience with is encouraged but not required. You will need to research the community/event via its public footprint, so do not choose one that is largely underground, and *triply do not* choose one where you will have to damage anyone's privacy, safety, or sovereignty during the project. You may choose a community/event with existing archives, but avoid ones where the analog archives have already been mostly or entirely digitized.
    - ➢ Feel free to run ideas past the Canvas project forum before you start your writeup.
- ➢ The collection must be decidedly heterogeneous, but need not be immense. (Think "two archival boxes of A/V and digital media"—in my work with RADD, that's a surprisingly common amount in many contexts—plus papers and online presence.) It must include (invent as needed!): one or more pieces of realia, one or more textual documents, one or more analog photographic images, at least two different analog audio media, at least two different film or video media (digital video on a physical carrier, such as mini-DV, is fine here), at least two digital data carriers, at least one public social-media account, and at least one website.
    - ➢ If you are having difficulty coming up with this list, I suggest asking friends or relatives who used or collected mid-to-late-20th-century audiovisual media as well as late-20th/early-21st-century digital media what they have —or have thrown away.

Now, write up a background statement (à la JISC section 1), collection-importance statement (à la Bloomfield sections 1.1 and 1.2), and a collection inventory. For the inventory, a list of media with titles is fine, no need for extensive metadata. In fact, it's even fine if you don't even know what's on some of the media yet—that's very common! I reserve the right when grading this to add a few objects to your collection, but you won't lose points if I do.

Week 2: Create a statement of any ethical concerns relating to the project and materials therein, and (as best you can tell; you may need to refine this later) how you plan to resolve them, in light of the ethical issues discussed in class this week. Make a community-engagement plan, relative to the communities represented in the project materials, that will carry you through all phases of the project.

Week 3: How do you anticipate involving community members/event participants/subject-matter experts in your project? What documentation will they need in order to do the work? (Do not write the documentation! That is much too much work for one week! Do come up with an organized list, perhaps an outline, of what they will need to know. Skim Indigitization sections D through G for ideas.)

Week 4: Expand your Week 1 collection-importance statement into a business case for digitizing and preserving it. Next, find at least three grant programs that you could feasibly apply to for digitization funding. (N.b. not digital preservation funding!)

Week 5: Write a condition statement for the collection, emphasizing at-risk materials in it for which digitization is the correct preservation option. Next, determine as best you can the copyright status of materials in the collection, what that means for your project, and as appropriate, what takedown policy should be applied. As appropriate, suggest Traditional Knowledge labeling.

Week 6: Determine and list appropriate file formats and minimum-capture standards/settings for *archival master files* and *access copies* for all objects in your collection that you will be digitizing, à la Bloomfield section 4.1 or Indigitization sections D through G. (Ignore the digital-data media for now—we'll get to that later—but do include any digital video on physical carriers you have.) Given this and your inventory from Week 1, estimate your storage needs (assuming a single copy of all files —you will obviously revise this later!) for the digitized files.

Week 7: Locate appropriate *descriptive* metadata standards for all digitized objects in your collection, initially seeking the most specific, fittest-for-purpose standards available. Include appropriate controlled vocabularies (CVs) for any part of the record where they apply. Evaluate these standards for feasibility, given your staffing levels and anticipated community participation. Finally, decide on one descriptive standard you will use for the entire collection, determine the required minimum record (CV use included), and create that minimum record for at least three collection objects per group member. Revise your Week 3 work to include anything about metadata you did not initially anticipate (see Indigitization section H).

Week 8: Locate and budget for equipment that will digitize your collection to the archival-master standards you identified in Week 6. Show your work with screenshots. (I will give you some leeway on obsolete A/V decks—their availability can be sporadic.) Do your best to locate and cost out outsourcing vendors who can digitize to your standards. (Some vendors are very cagey about pricing. You don't have to call or email to get costs if the vendor doesn't list them; just find out what you can!) N.b. continue to ignore your collection's digital media! We get to it in Week 11.

Week 9: Sketch out a workflow for this digitization project (with variations by object type, obviously). Don't forget project-initiation tasks such as creating documentation for SMEs or community members! You may assume that at least some work can be done concurrently (we will discuss in class). Roughly estimate the time needed to complete the project, as well as labor costs. Word to the wise: there is no Magic Metadata Fairy. Indigitization sections A and B may be helpful.

Week 10: Draft a digital-preservation plan and associated technical requirements for this collection that reaches NDSA Level 3 on storage (but ignore the old-media question until next week), Level 2 on fixity, Level 2 on information security, Level 1 on metadata, and Level 2 on file formats. You will refine this in subsequent weeks; do your best for now! Revise your storage estimate from Week 6 in light of your current plan.

Week 11: In your digital-preservation plan, resolve the outstanding old-media question from last week. Add the resulting additional storage capacity to your storage estimate (originally Week 6, revised last week). Complete your Week 8 budget by adding equipment for handling the digital-data media in your collection—remember to consider write-blocking as needed.

Week 12: Augment your Week 9 workflow with a preservation and accessioning plan for the social-media account(s) and website(s) you listed in Week 1. Include post-project maintenance/recapture planning. This should not take you too long, so start reorganizing, refining, and polishing your report.

Week 13: Evaluate Omeka, ContentDM, CollectionSpace, and one other package of your choice (see useful spreadsheet by Ashley Blewer et al. at `https://docs.google.com/spreadsheets/d/1cXOug3qM0pNNeD_wssiVEv9c0W1Y5I1VDTnSPTk7fb4/edit?ts=5ca9a929#gid=0`) as *access software for end-users* of your collection. (In other words, preservation features are at most a nice bonus, not requirements, here.) Especially investigate how well these systems mesh with the variety of file formats in your collection, and the metadata decisions you made in Week 7.

Week 14: Add a digital-preservation accessioning workflow to your digitization workflow from Week 9. Keep polishing!

Turn in the final version of your report by the last day of the course. I will grade it on mechanical correctness (yes, spelling and grammar matter), clarity, reasonableness, and persuasiveness to the Powers That Hold The Pursestrings.

Other assignments
Most Canvas modules will contain surveys, short reflections, hands-on work, or discussions that ask you to apply what you learn from readings and lecture. These are graded on a participation basis—do the things on time demonstrating reasonable effort, get the points.

Reading schedule

Unit 1: Context and human factors

Module 1: Digitization and digital preservation in context
Seelye. "Patricia Battin, lightning rod in a library war, dies at 89." `https://www.nytimes.com/2019/06/05/obituaries/patricia-battin-dead.html` (Pay attention to and don't always believe the rhetoric; it's Typical Journalist awful. That said, Battin was brilliant and amazing and all of us who do this work owe her a debt of gratitude.)
Meyer. "The museum of the future is here." `https://www.theatlantic.com/technology/archive/2015/01/how-to-build-the-museum-of-the-future/384646/`
Goldstein. "How state CIOs should preserve digital records." `https://statetechmagazine.com/article/2018/11/how-state-cios-should-preserve-digital-records`
Padilla and Steeves. "Data librarianship: a path and an ethic." `https://acrl.ala.org/dh/2018/04/04/repro/`
Appel et al. "Creating a foundation for digital asset management and preservation." `https://www.swarthmore.edu/sites/default/files/assets/documents/libraries/trico_damp_report.pdf` (Through page 7.)
Redwine. "Personal digital archiving." `https://www.dpconline.org/docs/technology-watch-reports/1460-twr15-01/file` (Executive summary, sections 4, 6, and 9.)

Lavoie. "The OAIS reference model: introductory guide." `http://dx.doi.org/10.7207/twr14-02` (pp. 1-6 and 21-23 only, for now.)

NEDCC. "NEDCC's Mission." `https://www.nedcc.org/about/mission` (I strongly suggest poking around the top-level navigation as well, to get a sense of the breadth of work NEDCC does.)

Mesa County Libraries. "Digital conversion station." `https://mesacountylibraries.org/services/conversionstation/`

Module 2: Ethics

Topics: What do GLAMs typically digitize, and why? What about commercial organizations and governments? What born-digital data is typically preserved, who by, for how long, and why? What does that mean about the corpus of available digital information? When should analog information not be digitized, or if digitized, not be made publicly accessible? How should GLAMs navigate difference?

Ziegler. "Digitization selection criteria as anti-racist action." `https://journal.code4lib.org/articles/14667`

Caswell. "Diversifying the digital historical record." `http://www.dlib.org/dlib/may17/caswell/05caswell.html`

Vines et al. "The availability of research data declines rapidly with article age." `https://doi.org/10.1016/j.cub.2013.11.014` (Highlights and summary required; rest optional. Title says it all, really.)

Gallagher. "Judge shocked to learn NYPD's evidence database has no backup." `https://arstechnica.com/information-technology/2017/10/nypd-database-that-tracks-seized-evidence-and-cash-has-no-backup/`

Mattson. "Queer histories, videotape, and the ethics of reuse." `https://www.centerforthehumanities.org/blog/queer-histories-videotape-and-the-ethics-of-reuse`

Robertson. "Not all information wants to be free: ethical considerations for digitization." `http://tararobertson.ca/wp-content/uploads/2016/08/code4libNYS.pdf`

Jules, Summers, and Mitchell Jr. "Ethical considerations for archiving social media content generated by contemporary social movements." `https://www.docnow.io/docs/docnow-whitepaper-2018.pdf`

Rowell and Cooksey. "The archive of hate: ethics of care in the preservation of ugly histories." `https://www.ladyscience.com/essays/archive-of-hate-ethics-of-care-in-the-preservation-of-ugly-histories` (Content alert: discussion of the Ku Klux Klan.)

Punzalan and Caswell. "Critical directions for archival approaches to social justice." `https://escholarship.org/uc/item/1xq5k4kg`

Kim. "The Passamaquoddy reclaim their culture through digital repatriation." `https://www.newyorker.com/culture/culture-desk/the-passamaquoddy-reclaim-their-culture-through-digital-repatriation`

Module 3: Community engagement; training and documentation

Topics: OAIS's concept of "designated community." Actual communities relevant to digitization and digital-preservation processes (OAIS gets this dangerously wrong!). Working alongside communities; community outreach and engagement; community archives; post-custodial archives. Crowdsourcing. Approaching communities across power differentials and oppression. Documentation and its importance; doing documentation usefully. Data-management plans.

Lavoie. "Overview of the OAIS Reference Model." `http://dx.doi.org/10.7207/twr14-02` (Section 5.2.3, "Consumers and the Designated Community.")

Kilbride. "Minding the gaps." `https://www.dpconline.org/blog/minding-the-gaps` (Scroll down to the "Preservation, Access, Impact" section.)

Schreiner. "Teaching personal digital archiving through community digitization." `https://saaers.wordpress.com/2018/02/12/teaching-personal-digital-archiving-through-community-digitization/`

Mischo et al. "An analysis of data management plans in University of Illinois [NSF] grant proposals." `https://escholarship.umassmed.edu/cgi/viewcontent.cgi?article=1060&context=jeslib`

Taylor. "Scholar sourcing, crowdsourcing, and community sourcing." `http://laurientaylor.org/2014/02/08/scholar-sourcing-crowdsourcing-and-community-sourcing/`

Earhart. "Can we trust the university? Digital humanities collaborations with historically exploited cultural communities." `http://hdl.handle.net/1969.1/173030`

Terras. "Digitisation's most wanted." `https://melissaterras.blogspot.com/2014/05/digitisations-most-wanted.html`

Werner. "How to destroy special collections with social media." `http://sarahwerner.net/blog/2015/07/how-to-destroy-special-collections-with-social-media/`

Bowen. "RTFM? How to write a manual worth reading." `https://opensource.com/business/15/5/write-better-docs`

Losh. "Teach, don't tell." `http://stevelosh.com/blog/2013/09/teach-dont-tell/`

For reference: "The hitchhiker's guide to documentation." `https://docs-guide.readthedocs.io/en/latest/`

Module 4: Sustainability: staff and funding

> *Topics: What does "sustainability" actually mean in practice? How do organizations get to "sustainable" from wherever they are? What are the costs and staffing needs around digitization and digital preservation? Staffing and budgeting anti-patterns and how they happen. Why grants can pay for digitization, but not digital preservation. Making a business case for digitization and/or digital preservation. Assessment as part of a business case.*

Marquez and Downey. "Service design." `https://quod.lib.umich.edu/w/weave/12535642.0001.201?view=text;rgn=main`

Marcum. "Due diligence and stewardship in a time of change and uncertainty." `https://sr.ithaka.org/publications/due-diligence-and-stewardship-in-a-time-of-change-and-uncertainty/`

DeMeo. "The politics of digitization." `https://modelviewculture.com/pieces/the-politics-of-digitization`

Rosenthal. "Digital Preservation Network is no more." `https://blog.dshr.org/2019/01/digital-preservation-network-is-no-more.html`

"Sustainable economics for a digital planet." `http://blueribbontaskforce.sdsc.edu/biblio/BRTF_Final_Report.pdf` (Pay close attention to who digital-preservation stakeholders are and what their incentives and disincentives look like.)

Hughes. "What a digital organization looks like." `https://medium.com/doteveryone/what-a-digital-organisation-looks-like-82426a210ab8` (How would you recognize a "digital superhero" job from a job ad, so you avoid applying?)

Teperek. "What I wish I'd known at the start—setting up an RDM service." `https://unlockingresearch-blog.lib.cam.ac.uk/?p=1535`

Skim: Maemura et al. "A survey of organizational assessment frameworks in digital preservation." `http://eprints.gla.ac.uk/187623/1/187623.pdf#page=126` (Point being, there are A LOT of ways to assess a service. Cynically: pick the one that gets you the respect and resources you need.)

Skim: "Selection and digital collection policy" `https://docs.google.com/document/d/1yMNteAj_T8JUlohBOf81kcx5z8OII4WOvM5WNNl-DNU` and my response, starting bottom p. 1 of `https://uwmadison.app.box.com/s/fdqkephvz3nj2wb9mb7wbxihmlucjm7m`

For reference: Digital Preservation Coalition. "Step by step guide to building a business case." `http://wiki.dpconline.org/index.php?title=Step_by_step_guide_to_building_a_business_case` and "Template for building a business case." `http://wiki.dpconline.org/index.php?title=Template_for_building_a_business_case`

Unit 2: Digitization

Module 5: Appraisal and selection

> *Topics: Bases (yes, plural) for appraising and selecting materials for digitization and/or digital preservation. Copyright, fair use, and digitization. Appraisal and selection in different organizational contexts; records scheduling. "Significant properties" in OAIS. Primary and secondary value; evidential and informational value. Appraisal and selection policies.*

Bettivia. "Mapping significance properties in OAIS." `https://www.asist.org/files/meetings/am15/proceedings/submissions/posters/288poster.pdf`

Niu. "Appraisal and selection for digital curation." `https://doi.org/10.2218/ijdc.v9i2.272`

UK Data Service. "Collections development selection and appraisal criteria." `https://ukdataservice.ac.uk/media/455175/cd234-collections-appraisal.pdf` (Read this for what a policy could contain and look like—not so much for the details. The Appraisal Grid may be handy for this week's project work!)

Dryden. "The role of copyright in selection for digitization." `https://doi.org/10.17723/aarc.77.1.3161547p1678423w`

Summers. "Periscope: an appraisal story." `https://news.docnow.io/periscope-an-appraisal-story-532695257c1c#.xr87kjvoh`

Kung and Campbell. "What not to keep: not all data has future research value." `https://journals.library.ualberta.ca/jchla/index.php/jchla/article/view/26048`

Project Stand. "S.A.V.E methodology." `https://standarchives.com/s-a-v-e-methodology/`

Module 6: File formats and digitization standards

> *Topics: Master vs. access/service file formats. File size vs. quality tradeoff. Lossy vs. lossless compression. Acceptable master file formats and quality settings for images, audio, film/video. Text digitization; OCR vs. transcription; markup; TEI. Accessibility for digitized materials. Digitization standards and guidelines. Digitization for preservation. MPLP digitization.*

Perrin. "Digital [image] file basics." On Canvas. (MAKE SURE YOU UNDERSTAND the lossless/lossy format distinction. It is important in all types of digitization. Also, don't miss the TEI anecdote at the end.)

"TEI Libraries SIG Manifesto." `https://wiki.tei-c.org/index.php/TEI_Libraries_SIG_Manifesto`

Tanner. "Text capture and optical character recognition 101." `https://simon-tanner.blogspot.com/2015/06/text-capture-and-optical-character.html`

Miller. "All text considered: a perspective on mass digitizing and archival processing." `https://doi.org/10.17723/aarc.76.2.6q005254035w2076`

Arnott. "Sharing with all: accessibility and historical resources." `http://modernhypatia.info/digital-commonwealth-2018/`

Skim, compare, evaluate for completeness/correctness/usability, ask questions on Canvas about:

"Guidelines." (This is the "FADGI" everybody talks about.) `http://www.digitizationguidelines.gov/guidelines/`

"ALCTS minimum digitization capture recommendations." `http://www.ala.org/alcts/resources/preserv/minimum-digitization-capture-recommendations`

"Wisconsin Heritage Online digital imaging guidelines." `https://recollectionwisconsin.org/wp-content/uploads/2017/08/WHO-Scanning-Quick-Guide.pdf`

CARLI. "Digital projects 101: a resource guide." `https://www.carli.illinois.edu/products-services/contentdm/digital_projects_101` (scroll down to "Creating Digital Objects" section for file format and quality guidelines)

UMass Amherst Libraries. "Guidelines for digitization." `https://www.library.umass.edu/assets/Digital-Strategies-Group/Guidelines-Policies/UMass-Amherst-Libraries-Best-Practice-Guidelines-for-Digitization-20110523-templated.pdf`

Module 7: Metadata

Topics: Descriptive, structural, administrative, technical metadata. Content vs. structure metadata standards. So-called "preservation metadata;" PREMIS, OAIS. The 1:1 problem. File-embedded metadata; EXIF. Dublin Core. OAI-PMH and metadata harvesting. MODS, METS, PCDM. Medium- or content-specific metadata: VRA Core, PBCore, FGDC CSDGM (look, I don't make this stuff up), Darwin Core, MIX, etc. Rights metadata; Creative Commons; Traditional Knowledge labels; RightsStatements.org. Metadata crosswalking. ALWAYS DO DATES IN ISO 8601.

NOAA Coastal Services Center. "Sailing through coastal metadata with Metadata Bob." `https://permanent.access.gpo.gov/lps89915/Bob_guide.pdf` (Just… roll with it, okay?)

Mattson. "Creating metadata by hand." `https://pushcartcatalog.wordpress.com/2017/05/25/creating-metadata-by-hand-musings-on-the-limits-of-automation-in-archives/`

Urban. "The 1:1 principle in the age of linked data." `https://dcpapers.dublincore.org/pubs/article/view/3707`

Lavoie. "The OAIS reference model: introductory guide." `http://dx.doi.org/10.7207/twr14-02` ("Representation information" and "Preservation description information" pp. 16-19)

Eckard. "A primer on PREMIS and PREMIS Rights." `https://archival-integration.blogspot.com/2016/02/a-primer-on-premis-and-premis-rights.html`

Briney. "It started with a tweet…" `http://dataabinitio.com/?p=893` (Future research data managers: click through and read her article also. Finding Easter egg optional.)

Skim, reading the ones most relevant to you (you may want to watch lecture first):

"DCMI metadata terms." `https://www.dublincore.org/specifications/dublin-core/dcmi-terms/` (Sections 1 through 3.)

MODS user guidelines "Lite." `https://www.loc.gov/standards/mods/userguide/lite.html`

Compare "METS Primer" `https://www.loc.gov/standards/mets/METSPrimer.pdf` with "Portland Common Data Model" `https://github.com/duraspace/pcdm/wiki`

DCAT. `https://www.w3.org/TR/vocab-dcat/` (Sections 1 and 4.)

"Darwin Core quick reference guide." `http://rs.tdwg.org/dwc/terms/`

FGDC "[CSDGM] Metadata quick guide." `https://www.fgdc.gov/metadata/documents/MetadataQuickGuide.pdf`

"An introduction to VRA Core." `https://www.loc.gov/standards/vracore/VRA_Core4_Intro.pdf`

PBCore. "Tutorials." `https://pbcore.org/tutorials` (I suggest "PBCore Structure" and "PBCore Controlled Vocabularies."

"UC guidelines for born-digital archival description." `https://github.com/uc-borndigital-ckg/uc-guidelines/blob/master/CompletePDF_UCGuidelinesForBorn-DigitalArchivalDescription_v1.pdf`

Module 8: Digitization equipment. Film, audio, and video digitization. 3D-object digitization.

Topics: Sourcing equipment. When consumer-grade equipment is good enough. When to outsource. Maintaining equipment. Scanners. Telecines, film scanners, DSLR rigs. Audio interfaces. Video-capture equipment (cards, boxes). Mini-DV/HDV and Firewire. Dealing with common media-repair and degrading-media situations. 3D imaging.

McKenzie. "A race against time to preserve university media collections." `https://www.insidehighered.com/news/2018/06/29/universities-urged-save-media-collections-it%e2%80%99s-too-late` (Multiply this problem by local history everywhere, every archives everywhere, private collections...)

Manus. "What resolution should I use?" `https://blogs.loc.gov/thesignal/2012/12/what-resolution-should-i-use-part-1/` and `https://blogs.loc.gov/thesignal/2013/01/what-resolution-should-i-use-part-2/` and `https://blogs.loc.gov/thesignal/2013/03/what-resolution-should-i-use-part-3/`

Whalen. "The expanding digitization universe." `https://www.library.ucla.edu/blog/preservation/2019/02/06/the-expanding-digitization-universe`

Salo and Hocking. "Digital Reformatting and Data Rescue with RADD and the PROUD and PRAVDA Kits." `https://kula.uvic.ca/articles/10.5334/kula.19/`

Kenan Science Library. "3D imaging." `https://library.unc.edu/science/makerspace/3d-imaging/`

Suggested for folks with an interest: Piepenburg. *Digitizing audiovisual and nonprint materials.* On print reserve.

Module 9: Digitization workflows

Topics: Workflows for outsourcing; RFPs, QA/QC. In-house workflows. Time estimation. The "real time" peril of A/V digitization.

Hagedorn et al. "Old wine in new bottles." `https://www.lib.umich.edu/blogs/library-tech-talk/old-wine-new-bottles-our-efforts-migrating-legacy-materials-hathitrust`

Handel. "Data migration, digital asset management, and microservices at CUNY TV." `https://blogs.loc.gov/thesignal/2016/03/data-migration-digital-asset-management-and-microservices-at-cuny-tv/`

"UCLA Library special collections digitization toolkit." `https://www.library.ucla.edu/special-collections/programs-projects/digital-projects-special-collections` (Click through everything. Yes, everything! Especially everything in Part II!)

"Digitizing video for long-term preservation: an RFP guide and template." `https://web.archive.org/web/20160620183516/http://library.nyu.edu/preservation/VARRFP.pdf`

Click through several of the digitization workflows on Library Workflow Exchange: `http://www.libraryworkflowexchange.org/category/digitization/`

Unit 3: Digital preservation

Module 10: Models, standards, risk management

Topics: Lifecycle models in digital preservation. Risk management and threat models in digital preservation. The OAIS model; why "OAIS compliance" is utterly meaningless but vendors hype it anyway. NDSA Levels of Preservation. ISO 16363; Trusted Digital Repository certification; why almost all of you can largely ignore them. Ransomware (and by extension, weak information security) as a growing and worrisome risk.

"They paid nearly a half million in ransom. Where's the data?" `https://www.news18.com/news/world/lake-city-they-paid-nearly-a-half-million-in-ransom-wheres-the-data-2220743.html`

Rosenthal et al. "Requirements for digital preservation: a bottom-up approach." http://www.dlib.org/dlib/november05/rosenthal/11rosenthal.html

Lavoie. "The OAIS reference model: introductory guide." `http://dx.doi.org/10.7207/twr14-02` (The whole thing, finally!)

NDSA. "Levels of digital preservation." `https://ndsa.org/activities/levels-of-digital-preservation/`

Orbis Cascade Alliance. "Digital preservation step by step." `https://ndsa.org/activities/levels-of-digital-preservation/`

CRL. "Certification report on the Hathi Trust Digital Repository." `http://www.crl.edu/sites/default/files/d6/attachments/pages/CRL%20HathiTrust%202011.pdf`

CRL. "Updated CLOCKSS audit 2018." `https://www.crl.edu/reports/clockss-audit-report`

Rosenthal. "TRAC audit: process." `https://blog.dshr.org/2014/08/trac-audit-process.html`

Rosenthal: TRAC audit: lessons." `https://blog.dshr.org/2014/08/trac-audit-lessons.html`

Module 11: Digital forensics; media archaeology

Topics: Equipment for digital forensics and media archaeology; FREDs and why you don't need them; Kryoflux; FC5025; WiebeTech/CRU. Disk images; why and how to create them; why you might not create them; retrieving information from them. Remanence. Write-blocking and how to do it for common media. Testing for sensitive information; redaction; BulkExtractor. "Bit preservation" and why it is often not enough; software emulation; format migration.

"[Digital] Preservation storage criteria, version 2." `https://osf.io/2epjx/`

Prater. "How to talk to IT about digital preservation." `http://digital.library.wisc.edu/1793/78844`

Bailey. "Protect your data: file fixity and data integrity." `https://blogs.loc.gov/thesignal/2014/04/protect-your-data-file-fixity-and-data-integrity/`

Wolverton. "Digital forensics: from the crime lab to the library." `https://www.nature.com/news/digital-forensics-from-the-crime-lab-to-the-library-1.19998`

Hester. "The quest for a universal translator for old, obsolete computer files." `https://www.atlasobscura.com/articles/how-to-open-old-computer-files`

Prael. "Centralized accessioning support for born digital archives." `http://journal.code4lib.org/articles/13494`

Salo material on disk imaging; on Canvas.

Garfinkel. "Bulk extractor." `https://www.forensicswiki.org/wiki/Bulk_extractor`

Humor, but not wrong: Jeong. "The data formats of Star Wars suck (spoilers)." `https://www.vice.com/en_us/article/9a3mmp/the-disk-formats-of-star-wars-rogue-one-spoilers`

Module 12: Web, email, and social-media archiving

Topics: The WARC file format; related tools. The Internet Archive and Archive-It; perma.cc; their missions and limitations. Documenting the Now; twarc. Privacy law and social-media archiving; GDPR; "right to be forgotten." Email archiving and its discontents.

UK Data Service. "The GDPR: research and archiving FAQs." `https://www.ukdataservice.ac.uk/media/621794/gdpr-faqs.pdf`

Madrigal. "Future historians probably won't understand our internet." `https://www.theatlantic.com/technology/archive/2017/12/it-might-be-impossible-for-future-historians-to-understand-our-internet/547463/`

Bustillos. "The internet isn't forever." `https://longreads.com/2018/02/20/the-internet-isnt-forever/`

Bixenspan. "When the Internet Archive forgets." `https://gizmodo.com/when-the-internet-archive-forgets-1830462131`

CLIR. "The future of email archives." `https://clir.wordpress.clir.org/wp-content/uploads/sites/6/2018/08/CLIR-pub175.pdf` (Appendixes optional but quite useful.)

Graham. "Twitter archiving, or TWARC." `https://github.com/hist3907b-winter2015/module3-wranglingdata/blob/master/twarc.md`

Merity. "Navigating the WARC file format." `https://commoncrawl.org/2014/04/navigating-the-warc-file-format/` (You can safely ignore WET files; they are not used in digital preservation.)

"About Archive-It." `https://archive-it.org/blog/learn-more/`

Webrecorder. "FAQ." `https://webrecorder.io/_faq`

For reference: "Awesome web archiving." `https://github.com/iipc/awesome-web-archiving` and "Social media data collection tools" `http://socialmediadata.wikidot.com/`

Module 13: Infrastructure: software, storage, services

Topics: Preservation storage; fixity; geographic replication; why IT will tell you "we back it up!" and why you should tell them "not good enough, buckos." Cloud storage. Appraisal and ingest software; BitCurator, Archivematica. Why digital-collections software is rarely if ever a preservation system. Digital-collections software; Omeka; ContentDM. Software specialized for certain kinds of content; OHMS. Institutional-repository software; DSpace. Preservation software and services; (C)LOCKSS; Preservica; MetaArchive Cooperative.

Prater. "How to talk to IT about digital preservation." (Yes, reread it!) `http://digital.library.wisc.edu/1793/78844`

Gates. "A place where you process: an introduction to Archivematica workflows." `https://ethan-gates.com/files/Archivematica_Workshop_2017.pdf`

Meister. "Using BitCurator" `https://confluence.educopia.org/display/BC/Using+BitCurator`

Bruhns. "[BitCurator] Tasks and tools overview." `https://confluence.educopia.org/display/BC/Tasks+and+Tools+Overview`

"Top reasons to use DSpace." `https://duraspace.org/dspace/about/features/`

MetaArchive Cooperative. "Distributed digital preservation." `https://metaarchive.org/how-ddp-works/`

Preservica. "Active digital preservation for libraries & museums." `https://preservica.com/heritage-digital-preservation-services/museums-libraries`

"Oral history metadata synthesizer [OHMS]." `http://nunncenter.org/ohms-info/`

Module 14: Digital-preservation workflows

Topics: BitCurator. Pros and cons of file normalization. BagIt; Bagger. E-discovery.

"Guest post: Walker Sampson on disk imaging workflow." `https://bitcurator.net/2015/02/04/guest-post-walker-sampson-on-disk-imaging-workflow/` See also "UC Boulder workflow map." `https://bitcuratorconsortium.org/system/files/workflows/University%20of%20Colorado%20Boulder%20Workflow%20Map%20-%20Born%20digital%20materials%20on%20physical%20media.pdf`

IETF. "The BagIt File Packaging Format." `https://tools.ietf.org/html/rfc8493` (Sections 1, 2, and 4.)

Memory and Whitney. "Restructuring and uploading zip files to the Internet Archive with bash." `https://saaers.wordpress.com/2018/07/24/restructuring-and-uploading-zip-files-to-the-internet-archive-with-bash/`

[UK] National Archives. "The application of technology-assisted review to born-digital records transfer." `http://www.nationalarchives.gov.uk/documents/technology-assisted-review-to-born-digital-records-transfer.pdf`

iSchool learning outcomes

iSchool learning outcomes	Course measurable outcomes
1. Students demonstrate understanding of societal, legal, policy, or ethical information issues.	Planning project weeks 2, 3, and 5.
2. Students apply principles of information organization.	Planning project week 7.
4. Students demonstrate understanding of professional competencies important for management of information organizations.	Planning project weeks 2, 3, 4, 9-11, and 14.
5. Students demonstrate competency with information technologies important to the information professions.	Planning project weeks 5, 6, 8, 10-14.
7. Students demonstrate understanding of issues surrounding marginalized communities and information.	Planning project weeks 2 and 3.