INF 2122H
Digital Preservation and Curation

Winter Semester, 2017
Instructor: Professor Christoph Becker
Time: Monday, 1pm to 4pm
Venue: BL 507

Course Description

We are all aware that digital technologies are changing every aspect of our lives. They have transformed how we create, play, work, and share. Ensuring the long term access to digital materials generated through our use of these technologies has become a cornerstone of the information professions. There is an expectation that information professionals are cognisant of the challenges to the preservation and curation of digital materials and the approaches to addressing them. This course will enable students to acquire just that knowledge.

This course examines the creation, curation, conservation, and preservation of digital materials in both the public and private sectors and enables students to develop an appreciation of the principles of management of digital information in the context of digital longevity. Students gain an understanding of the organizational, technical, social, and economic challenges encountered when enabling the long-term availability of digital materials. It provides an introduction to key models, workflows (from pre-ingest to dissemination), policies, characteristics of digital repositories, standards, metadata, annotation, audit and certification, technical approaches from hardware preservation to emulation, and future research challenges that need to be addressed if the preservation landscape is to be transitioned out of an arts and craft mode.

Course Objectives and Learning Outcomes

Students having completed this course should be capable of:
- Summarizing the challenges and approaches to digital preservation within the context of institutions working with documents and data (demonstrated in class activities and assignments 1, 2, and 4);
- describing core research issues and directions in the area of digital curation and preservation (partially demonstrated in assignment 1 and more fully in assignment 4);
- organising practical activities (e.g. format identification, preservation options identification and assessment) related to digital preservation (demonstrated in class activities and assignment 2, 3 and 4);
- implementing workflow modelling, metadata definition, and ingest process management (demonstrated in class activities and assignment 1, 2 and 4);
- using knowledge of digital object assessment and digital repositories to manage curation and preservation processes (demonstrated in class activities and assignment 3 and 4);
- applying a working knowledge of the techniques and practices that underlie digital curation (demonstrated in assignments 2, 3 and 4);
- characterize preservation scenarios and identify key issues in given environments (demonstrated in assignments 1 and 3 and in class activities); and
- explaining the issues of authenticity, integrity, and reliability in relation to digital preservation and curation (demonstrated in assignment 3 and 4).

Relationship of Course Objectives to MI Program Outcomes

Digital Preservation and curation is an essential challenge and activity of memory institutions, public sector organizations and commercial enterprises. Students completing this course will understand and be conversant with the fundamental concepts, practices, and methods of digital curation and preservation (Program Outcome 1). Through a range of in class activities and course assignments they will engage in aspects of preservation and curation that will enable them to understand how preservation research can be approached (Program Outcome 3). The course will contribute to their progress towards achieving Program Outcome 5 by giving them an "understanding of the application of new technological developments to the preservation" and an appreciation of "the impact of such developments" on contemporary society, the record we pass to posterity, and their professional roles in this process. Students will recognize that digital preservation and curation requires continued ability renewal and demands regular engagement in "life-long intellectual growth beyond graduation" (Program Objective 6).

Class Format

The class will meet for up to three hours each week. Each class will normally consist of a lecture, a team activity or guest conversation period, and a discussion period. The lecture portion of the class will introduce key concepts and sometimes cover material that is not available in published literature. There will normally be a 15 minute break after the lecture. This will be followed by a team-based activity, presentation, or guest lecture. The last part will involve discussions of the team-based activity and the readings that include questions from blogs of class members, and sometimes an additional (shorter) lecture. On their own time, students must complete weekly course readings and complete written assignments.

Prerequisites

INF1003 or INF1342.

Course Materials

Textbooks/Readings
This course does not have a textbook, but students wishing to have access to a textbook should consider: Adrian Brown (2013). *Practical Digital Preservation: a how-to guide for organizations of any size*. Facet Publishing, London. [Inforum: 025.840285642 B877P - Course reserves – link]. Most readings (both required and recommended) are available online, and links are included in the syllabus.
Another option is Harvey, D. R., & Oliver, G. (2016). *Digital curation* (2nd ed.). Facet Publishing, London (on course reserve at the Inforum - link).
Recommended for a broad and insightful perspective on research data is Christine Borgman (2015). Big Data, Little Data, No Data: Scholarship in the Networked World. MIT Press, on course reserve at the Inforum.

Website/Resources
Course materials and resources will be made available online, through Blackboard (http://portal.utoronto.ca). Most readings are available through the UTL system. Those that are not will be available on Blackboard. Students are responsible for keeping up to date with these

online resources and are expected to log into Blackboard during the first week of class to enroll for email notices. Please be sure to check Blackboard periodically for new materials, announcements, updates and other important information. When PowerPoint slides are used, they will be made available online within 1-2 days of the lecture.

Evaluation

	Assignments	*Weight*	*Due*	*Requirements*
1	Blog posts	20%	4 times	Due in weeks 1, 5, 9, 11 and 12 at 9am on the day of class
2	Tool presentation (teams of 3)	20%	In class, weeks 8-10	Slots will be available for signing up to tool demonstrations of 15 minutes plus discussion
3	A preservation case: intent statement	25%	Feb 28	Paper up to 1500 words, excluding any tables, images, or bibliography
4	Approaches to Preservation	35%	April 5	Paper up to 3500 words, excluding any tables, images, or bibliography.

Assignment 1: Blog posts
You will post a series of related blog entries building a discussion of a particular preservation scenario that is of interest to you. The blog will be made available on Blackboard.
You can post these entries any time before the due date. You can skip one post, i.e. (the best) four posts will be graded at the end term.
Aim to address the brief questions outlined below. Your contributions will (and should) vary in terms of length, but try to keep your posts focused on the topic and brief (up to 400 words).

- **Week 1**: What is your background? Do you bring specific **interests** in digital preservation and digital curation to the course? Which kinds of digital resources are you most interested in engaging with through the lens of digital curation?
- **Week 5**: Which specific digital curation **scenario** is of particular interest to you? Which particular **threats** can you identify for the preservation scenario that interests you? Which **properties** of the digital resources are significant, to whom, and why?
- **Week 9**: Which administrative, descriptive, technical and preservation-oriented **metadata** are relevant in your scenario and why? Which metadata schemes should be considered?
- **Week 11**: Choose one full paper from a recent conference in digital curation and preservation such as IPRES or IDCC, or the journal IJDC. Review it in your blog post and connect it to what we have discussed in the course.
- **Week 12**: Which **systems and tools** did you identify that are applicable and useful in your preservation scenario? How do these address particular concerns across the digital curation lifecycle?

You are free to add links to materials, tools, articles and so forth, but you must ensure that they relate clearly to the issues at hand and that you add comments and explanation of their relevance. Your blog post must be completed by 9am on the day the class meets, and you are encouraged to include particular questions you would like to see raised or addressed. Make sure to read others' blog entries and questions and use them to help shape the discussion period of class.

Assignment 2: Tool demonstration (in teams of 3)

In teams of 3, you will present, explain and demonstrate one particular software tool or system that supports an identified task or range of tasks in the digital curation lifecycle. The presentations take place in class in the weeks 8-10. The process of signing up to presentation slots and tools will be managed on Blackboard. A list of tools will be suggested, but other tools can be proposed and presented if they are approved in advance by the instructor. Contextualizing the tool or system clearly in the digital curation lifecycle and range of concepts in the course is an important aspect of the demonstration. The presentation must clearly identify the purpose that the tool or system addresses; shortly outline its main characteristics; include a short, concise live demonstration of key functions of the tool or system in action; and conclude with a short summary of strengths and limitations, and (if applicable) a comparison to other available tools that support similar tasks. Note that presentation time is limited to 15 minutes, so make sure to time your presentation very carefully to complete within the allotted time slots.

In week 11, all tool demos should be present for an informal in-class 'tool fair' to allow everyone to interact with tools they are interested in.

Assignments 3 and 4: Evaluating Approaches to Digital Preservation

Your essays will in two parts examine the problem of preserving a complex digital object.

In Assignment 3, you will develop the case for preserving the object and characterize the object and the scenario of preserving it succinctly to clarify *what* should be preserved; *by whom* and *for whom*; for *how long*; and *why*. You will develop a scenario around a particular object provided, a digital chest x-ray and the mechanism for managing and displaying it. You will be provided with access to a digital chest x-ray, the current software needed to view it on a computer running a recent version of the Microsoft windows operating system, and the permission to use the x-ray and a DICOM viewer but not to distribute either to others. Alternatively, you can propose a specific digital object in a particular scenario other than this, but in this case you must get approval upfront to ensure your scenario provides a topic of sufficient complexity but still is a feasible subject for your essay.

You are encouraged to make assumptions about any parameters not defined here and document them explicitly. For example, you could regard the x-ray as a clinical health record; a teaching object used in a graduate course; a personal record; a sample of a research data set; a museum object; or something else altogether? Your essay should reflect on the ethical, legal and security issues of this particular type of resource. You should ground your discussion in relevant literature, but you also must engage specifically with the digital object provided and its particular characteristics, and perform some technical analysis on the object itself to identify its file format and requirements for performance, as well as gain an understanding of its usage possibilities.

In Assignment 4, you take this scenario as a starting point to evaluate possible preservation options and recommend *how* the object should be preserved. You should consider specifically the various preservation actions that are applicable to the object (ranging from print-to-paper/film, migration, emulation, to hardware/software preservation), discuss measurement of authenticity over time, and use all relevant factors to recommend an action to take in the concrete preservation case you developed. (You can revise the scenario developed in Assignment 3 if needed to address new insights that emerge).

Guidelines for Assignments

All written assignments for this course must be **submitted in a PDF format through Blackboard** at the corresponding due date. (Where no time is given, the assignments are due at 23:59). All assignments should be written as clearly and cleanly as possible (i.e. watch the typos, grammar, hanging sentences, etc.), in a formal but accessible academic language. The "look and feel" should be professional, and figures and tables should, where possible, be included in the text, not at the end of the document.
Blog entries must be posted in the normal manner as text, not as attachments to the post.
The required format for other written assignments is as follows:
-   A **PDF** format (not MS Word).
-   Typed, 1.5 space, 11 or 12 point font, one-inch margins, page numbers in the upper or lower right hand corner.
-   Align paragraphs in a standard way and avoid superfluous indentation.
-   No cover page required, but be sure to include your name & student number on page 1.
-   Total word count should be indicated at the end of the essay.
-   Use of footnotes/endnotes is permitted.
-   Any citation style is fine as long as it is consistently applied.

NOTE: Assignments that do not meet a minimum standard (in terms of legibility, formatting, and proofreading) will be returned for re-submission, with late penalties in full effect.

Normally, students will be required to submit their course essays electronically to Turnitin.com for review of textual similarity and detection of possible plagiarism. In doing so, students will allow essays to be included as source documents in the Turnitin.com database, where they will be used solely for the purpose of detecting plagiarism. Turnitin.com services are described on the Turnitin.com website.

Images
"Students can include copyrighted images in their assignments as long as they follow the *Canadian Copyright Act*'s current exceptions for fair dealing, in that the images must only be used for the purposes of criticism or review, and each image must be accompanied by:
        (a) the source; *and*
        (b) the name of the author(s) (if given in the source)."

Acceptable Secondary Sources
"As graduate students, you will be expected to use a majority of academic (i.e. peer reviewed) sources when writing your term paper. Students are very much allowed, but not at all limited, to use course readings and other sources referenced in lectures in their own papers. Additional sources and relevant journals that are recommended by the instructor are also acceptable. However, students are strongly encouraged to track down those resources that are best suited to their specific area of interest or inquiry, rather than rely too heavily on those provided in class.

For cutting edge information, news, announcements, etc., popular press articles are of course acceptable. But these should be used to supplement or update rather than replace peer reviewed sources, and should *never* be used to explain a theoretical concept. They should also come from credible, verifiable sources, who have the credentials (whatever these may be) to back up their

claims. Often these articles point to underlying scholarly articles in peer reviewed journals or conferences, students are encouraged to pursue."

Late Papers
"Unless a formal extension has been negotiated with the instructor in advance of the due date, late assignments (defined here as an assignment submitted *after* the deadline) will be penalized by one full letter grade per week (e.g. from A to A-), for a maximum of two weeks. After that point, late assignments will no longer be accepted. Furthermore, late papers will not receive detailed feedback or comments."

Extensions
"Extensions on assignments within the term must be negotiated in advance, and require supporting documentation (e.g. doctor's note). Students must email requests for extensions to the instructor at least 24 hours prior to the due date. Exceptions will only be made in extenuating circumstances. Extensions beyond the end of the term in which a course is taken are subject to the guidelines established by the School of Graduate Studies."

Grading

Please consult the iSchool's Grade Interpretation Guidelines and the University Assessment and Grading Practices Policy. These documents will form the basis for grading in the course.

Ground Rules

"Each student in this course is responsible for keeping up with the course materials, which includes (all) the required course readings, as well as topics, debates, and concepts discussed in class. Students are expected to attend lectures and to take their own lecture notes. You are expected to participate in class discussions, and are encouraged to use your laptops/mobile devices during class to look up relevant information that will contribute to the discussion in a *meaningful way*. At all times, however, remember to be respectful of the instructor and of your classmates – turn your phone function off, turn off the sound on your computer, and be sure not to browse any websites that may be offensive or illegal, or that might be deemed irrelevant to the task of taking this course. Students should arrive on time and are expected to stay for the duration. If you miss a class, you are responsible for obtaining any information or materials given in class, either from your classmates or online. Unauthorized recording of the lectures is not permitted."

Accommodations

Students with diverse learning styles and needs are welcome in this course. If you have a disability or a health consideration that may require accommodations, please feel free to approach me and/or the Accessibility Services Office as soon as possible. The Accessibility Services staff is available by appointment to assess needs, provide referrals and arrange appropriate accommodations. The sooner you let us know your needs, the quicker we can assist you in achieving your learning goals in this course.

Writing Support

As stated in the iSchool's Grade Interpretation Guidelines, "work that is not well written and grammatically correct will not generally be considered eligible for a grade in the A range, regardless of its quality in other respects". With this in mind, please make use of the writing support provided to graduate students by the SGS Office of English Language and Writing Support. The services are designed to target the needs of both native and non-native speakers and all programs are free. Please consult the current workshop schedule for more information.

Academic Integrity

Please consult the University's site on Academic Integrity. The iSchool has a zero-tolerance policy on plagiarism as defined in section B.I.1.(d) of the University's Code of Behaviour on Academic Matters. You should acquaint yourself with the Code. Please review the material in Cite it Right and if you require further clarification, consult the site How Not to Plagiarize. Cite it Right covers relevant parts of the U of T *Code of Behaviour on Academic Matters* (1995). It is expected that all iSchool students take the Cite it Right workshop and the online quiz. Completion of the online Cite it Right quiz should be made prior to the second week of classes. To review and complete the workshop, visit the orientation portion of the iSkills site: uoft.me/iskills

In Class Activity Teams

In most weeks, we will have one team-based activity for up to an hour to allow us to engage in practical ways with the concepts and knowledge that we are covering. We will be about 35 students in the class. We will breakout into teams of four each in the first class and that will be your team for the term. Each week your team will appoint a team leader, a rapporteur, and technology guru for the following week. The role of the team leader is to manage the team activity (e.g., to ensure that you work as a team, to ensure that you manage your time effectively so you can complete the task in the time available). The role of the rapporteur will be to summarize the key points and to act as the key person for sharing the reflections of each team with the whole class during the discussion part of the class. The role of the technology guru is to ensure that in the weeks that you will require access to technology that he/she signs out a laptop from the Inforum if needed, brings it to class, install the software if any is needed, and returns it to the Inforum after class. (Groups would be wise to appoint alternatives for the technology guru in case of illness.)

Websites that Provide Useful Resources and Tools

The Digital Curation Centre (http://www.dcc.ac.uk/)
The Digital Preservation Coalition (http://www.dpconline.org/)
DigiPres Commons http://www.digipres.org/, a community-maintained website
The DCC glossary: http://www.dcc.ac.uk/resource/glossary/
Open Preservation Foundation. http://openpreservation.org/
Research Data Alliance http://www.rd-alliance.org
Research Object: http://www.researchobject.org

Websites of recent research projects:

4C: http://4cproject.eu/
ARCOMEM http://www.arcomem.eu/
BenchmarkDP: http://benchmark-dp.org
Blockchain@UBC: http://blockchainubc.ca/ (ongoing)
BlogForever http://blogforever.eu/
CASPAR http://www.casparpreserves.eu/
DigitalPreservationEurope (DPE) http://www.digitalpreservationeurope.eu/
E-ARK: http://www.eark-project.com/ (ends 2017)
ERPANET http://www.erpanet.org/
KEEP http://www.keep-project.eu/ezpub2/index.php?/eng
LiWA http://www.liwa-project.eu/
PARSE.Insight http://www.parse-insight.eu/
PERICLES: http://pericles-project.eu/ (ends 2017)
Planets http://www.planets-project.eu/
PrestoPRIME http://www.prestoprime.org/
PRESTOSPACE http://prestospace.org/
SCAPE http://www.scape-project.eu/
SHAMAN http://shaman-ip.eu/shaman/
TIMBUS http://www.teco.edu/research/projects/timbus/index.html
Wf4Ever http://www.wf4ever-project.org/

Contacting the Instructor & Supporting Instructor

Prof Christoph Becker is available by email: christoph.becker@utoronto.ca. Usual response time: within 2 working days.

Schedule of Lecture Topics and Readings

Week 1: January 9	**Introduction to Digital Preservation and Curation**
	Required Reading: Yakel, Elizabeth (2007). Digital Curation. OCLC Systems & Services: International Digital Library Perspectives 23, no. 4 (November 6, 2007): 335–40. doi:10.1108/10650750710831466. Higgins, S. (2011). Digital Curation: The Emergence of a New Discipline. International Journal of Digital Curation, 6(2), 78–88. http://doi.org/10.2218/ijdc.v6i2.191 UNESCO Charter on the preservation of Digital Heritage. Retrieved from http://portal.unesco.org/en/ev.php-URL_ID=17721&URL_DO=DO_TOPIC&URL_SECTION=201.html **Recommended readings:** National Science Board (2005). Long-Lived Digital Data Collections: Enabling Research and Education in the 21st Century. National Science Foundation. *(Read sections 1-3)* Ross, S. (2006). Approaching digital preservation holistically. In A. Tough and M. Moss (Eds.), *Information Management and Preservation.* (pp. 115-153).

Oxford: Chandos Press. [Available on the course blackboard site.]

Supplementary Sources:
Duranti, L. (1995). Reliability and authenticity: The concepts and their implications. *Archivaria 39*, 1–10. [e-article] http://journals.sfu.ca.myaccess.library.utoronto.ca/archivar/index.php/archivaria/article/view/12063

Carlson, S., & Anderson, B. (2007). What Are Data? The Many Kinds of Data and Their Implications for Data Re-Use. Journal of Computer-Mediated Communication, 12(2), 635–651. http://doi.org/10.1111/j.1083-6101.2007.00342.x

Clifford A. Lynch, "Jim Gray's Fourth Paradigm and the Construction of the Scientific Record," The Fourth Paradigm, edited by Tony Hey. Microsoft Research, 2009. (Available online)

Gantz, J. F. (2008). The diverse and exploding digital universe: An updated forecast of worldwide information growth through 2011. (International Data Corporation (IDC), white paper, pp. 1-15). Retrieved from http://www.emc.com/collateral/analyst-reports/diverse-exploding-digital-universe.pdf

Team Digital Preservation. (2009, 1 May). Digital Preservation and Nuclear Disaster: An Animation. [Video from Youtube]. Retrieved from http://www.youtube.com/watch?v=pbBa6Oam7-w

National Library of Australia (2003). UNESCO Guidelines for the preservation of digital heritage. Retrieved from http://unesdoc.unesco.org/images/0013/001300/130071e.pdf

Tibbo, H. (2003). On the nature and importance of archiving in the digital age. Advances in Computers *57*, 1–67. [e-article] http://dx.doi.org.myaccess.library.utoronto.ca/10.1016/S0065-2458(03)57001-2

Garrett, J. & Waters, D. (1996). Preserving digital information: Report of the task force on archiving of digital information. (The Commission on Preservation and Access and RLG). Retrieved from http://www.rlg.org/legacy/ftpd/pub/archtf/final-report.pdf.

Hedstrom, M. (1997). Digital preservation: A time bomb for digital libraries. *Computers and the Humanities 31*(3), 189-202. Retrieved from http://www.uky.edu/~kiernan/DL/hedstrom.html (also in UNC Library E-Journals)

Ross, S. (2000). Changing trains at Wigan: Digital preservation and the future of scholarship. (National Preservation Office occasional paper). Retrieved from http://www.bl.uk/blpac/pdf/wigan.pdf

Ross, S. (2012). Digital Preservation, Archival Science and Methodological Foundations for Digital Libraries. New Review of Information Networking, 17(1), 43–68. https://doi.org/10.1080/13614576.2012.679446

Ross, S., & Hedstrom, M. (2005). Preservation research and sustainable digital libraries. International Journal on Digital Libraries, 5(4), 317–324.

Week 2: January 16	**The curation of web archives** *Web Archiving*

Guest lecture, activity and discussion

Required Readings:
Niu, J. (2012). An Overview of Web Archiving. D-Lib Magazine, 18(3/4).
https://doi.org/10.1045/march2012-niu1
Milligan, I. (2016). Lost in the Infinite Archive: The Promise and Pitfalls of Web
Archives. International Journal of Humanities and Arts Computing, 10(1),
78–94. https://doi.org/10.3366/ijhac.2016.0161

Recommended Readings:
Read one of the case study reports available at
http://buddah.projects.history.ac.uk/

Supplementary Materials:
These very recent overview reports on the subject are a good starting point:
Maemura, E. (2016, November 1). Web Archiving: Past, Present and Future.
Retrieved from http://dci.ischool.utoronto.ca/?p=1306
Truman, G. (2016). Web Archiving Environmental Scan. Retrieved from
https://dash.harvard.edu/handle/1/25658314
Thompson, S. D. (2016). Preserving Social Media (DPC Technology Watch
Report No. 16–1). Digital Preservation Coalition. Retrieved from
http://dx.doi.org/10.7207/twr16-01

A public media perspective:
Lepore, J. (2015). What the Web Said Yesterday. *The New Yorker.* Retrieved
December 22, 2016, from
http://www.newyorker.com/magazine/2015/01/26/cobweb
(Keep an eye on the discussions about the End of Term archive in public media.)

For overviews and discussions of particular broad issues:
Masanès, J. (2006). Web archiving: Issues and methods. In Web Archiving. (pp.1-
53). Springer Berlin Heidelberg. Retrieved from
http://link.springer.com/book/10.1007/978-3-540-46332-0/page/1
Banos V., Kim Y., Ross S., Manolopoulos Y.: CLEAR: a credible method to
evaluate website archivability, 10th International Conference on
Preservation of Digital Objects (iPRES'2013), Lisbon, 2013
Brügger, N. (2013). Web historiography and Internet Studies: Challenges and
perspectives. New Media & Society, 15(5), 752–764.
https://doi.org/10.1177/1461444812462852
Niu, J. (2012). Functionalities of Web Archives. D-Lib Magazine, 18(3/4).
https://doi.org/10.1045/march2012-niu2
Brown, A. (2006). Archiving Websites: A Practical Guide for Information
Management Professionals. London: Facet Publishing.

A number of publications aim to measure the disappearance of websites:
Klein, M., Sompel, H. V. de, Sanderson, R., Shankar, H., Balakireva, L., Zhou, K.,
& Tobin, R. (2014). Scholarly Context Not Found: One in Five Articles

	Suffers from Reference Rot. PLOS ONE, 9(12), e115253. https://doi.org/10.1371/journal.pone.0115253 Jackson, A. (2014). What is still on the web after 10 years of archiving? Retrieved from http://britishlibrary.typepad.co.uk/webarchive/2014/10/what-is-still-on-the-web-after-10-years-of-archiving-.html Wren, J.D. (2004). 404 not found: The stability and persistence of URLs published in MEDLINE. Bioinformatics 20(5), 668-672. Retrieved from http://bioinformatics.oxfordjournals.org/cgi/reprint/20/5/668.pdf Ainsworth, S. G., AlSum, A., SalahEldeen, H., Weigle, M. C., & Nelson, M. L. (2012). How Much of the Web Is Archived? arXiv:1212.6177 [Cs]. Retrieved from http://arxiv.org/abs/1212.6177 *Technology-focused articles on web archiving systems and infrastructure:* Van de Sompel, H., Nelson, M. L., Sanderson, R., Balakireva, L. L., Ainsworth, S., & Shankar, H. (2009). Memento: Time Travel for the Web. arXiv:0911.1112 [Cs]. Retrieved from http://arxiv.org/abs/0911.1112 Kasioumis N., Banos V., Kalb H. (2013). Towards building a blog preservation platform, World Wide Web Journal Special Issue on Social Media Preservation and Applications, Springer, DOI 10.1007/s11280-013-0234-4 Risse, T., Demidova, E., Dietze, S., Peters, W., Papailiou, N., Doka, K., … Spiliotopoulos, D. (2014). The ARCOMEM Architecture for Social- and Semantic-Driven Web Archiving. Future Internet, 6(4), 688–716. https://doi.org/http://dx.doi.org.myaccess.library.utoronto.ca/10.3390/fi6040688
Week 3: January 23	**Lifecycles, Models, Scenarios, and Abilities** ***Guest lecture and discussion*** **Required Readings:** Higgins, S. (2008). The DCC curation lifecycle model. *International Journal of Digital Curation 3*(1), 134-140. [e-journal] http://simplelink.library.utoronto.ca/url.cfm/330573 Lee, C. (2009, June 17). Matrix of digital curation knowledge and competencies. [website]. Retrieved from http://ils.unc.edu/digccurr/digccurr-matrix.html *DigCurV Practitioner, Manager, and Executive Lenses* (Available on the course Blackboard site). **Supplementary Resources:** *This detailed study includes predictions on job prospects in digital curation:* National Research Council. Preparing the Workforce for Digital Curation. Washington, DC: The National Academies Press, 2015. (Available online) *An interesting discussion of research data curation curricula:* Palmer, C. L., Weber, N., Muñoz, T., & Renear, A. (2013). Foundations of Data Curation: The Pedagogy and Practice of "Purposeful Work" with Research Data - Archive Journal Issue 3. Archive Journal, (3). Retrieved from http://www.archivejournal.net/issue/3/archives-remixed/foundations-of-data-curation-the-pedagogy-and-practice-of-purposeful-work-with-

	research-data/ *Canadian developments relevant for research data curation:* Social Sciences and Humanities Research Council (2013). Capitalizing on Big Data: Toward a Policy Framework for Advancing Digital Scholarship in Canada. Consultation report. (Available online) *Tri-Agency Statement of Principles on Digital Data Management*: http://www.science.gc.ca/default.asp?lang=En&n=83F7624E-1 *More from the DigCCurr project:* Christoph A. Lee and Helen Tibbo (2011). Where's the Archivist in Digital Curation? Exploring the Possibilities through a Matrix of Knowledge and Skills. *Archivaria* 72 (Fall 2011): 123-168 (UTL) Lee, C., Tibbo, H. R., & Schaefer, J.C. (2012). Defining what digital curators do and what they need to know: The DigCCurr Project. In Proceedings of the *ACM IEEE Digital Preservation & Access Fall, 2012 Joint Conference on Digital Libraries*. (18-23 June 2007). Vancouver, British Columbia, Canada. Retrieved from http://www.ils.unc.edu/digccurr/jcdl2007_paper.pdf

Week 4:
January 30
	Information Encoding, Representation, and File Formats
	Team Activity: File Type Identification Challenge
	(Each group will be given a set of files to identify. The names will be meaningless and the file extensions will be deleted. The objective will be to work out what types of files they are, to save them with meaningful names and appropriate extensions, and to explain what tools and strategies your team used to complete the task in coherent, consistent, robust, and documented way.)
	
	Required Readings:
	Adrian Brown (2008, August). "Selecting File Formats for Long-Term Preservation". Retrieved from
	http://www.nationalarchives.gov.uk/documents/selecting-file-formats.pdf
	Pearson, D. &Webb, C. (2007). Defining File Format Obsolescence: A Risky Journey. IJDC 2008, Vol. 3, No. 1, pp. 89-106. Retrieved from
	http://www.ijdc.net/index.php/ijdc/article/view/76
	
	Sites to Review:
	• PRONOM. http://www.nationalarchives.gov.uk/PRONOM/Default.aspx
	• Sustainability of digital formats.
	http://www.digitalpreservation.gov/formats/index.shtml
	• Have a look at the visual explanations of file formats at
	http://imgur.com/a/MtQZv#0
	   o Portable Network Graphics (PNG) v1.03
	   o TIF: Tagged Image File Format v1.00
	   o GIF: the Graphics Interchange Format v1.00
	   o JFIF: the JPEG File Interchange Format

	Recommended Optional Readings: Rosenthal, D. (2010). Format obsolescence: Assessing the threat and the defenses. *Library Hi-Tech, 28*(2), 195-210. [e-article] http://dx.doi.org.myaccess.library.utoronto.ca/10.1108/07378831011047613 Todd, M. (2009, October). Technology Watch Report: File Formats for Preservation. DPC Technology Watch Series Report 09-02. Retrieved from http://www.dpconline.org/component/docman/doc_download/375-file-formats-for-preservation **Supplementary Resources:** Petrov, P. & Becker, C. (2012*). Large-scale content profiling for preservation analysis.* In: 9th International Conference on Preservation of Digital Objects (IPRES 2012), Toronto. Retrieved from http://publik.tuwien.ac.at/files/PubDat_213858.pdf Esteva, M., Xu, W., Jain, S. D., Lee, J. L., & Martin, W. K. (2011). Assessing the Preservation Condition of Large and Heterogeneous Electronic Records Collections with Visualization. International Journal of Digital Curation, 6(1), 45–57. https://doi.org/10.2218/ijdc.v6i1.171 Xu, W., Esteva, M., Jain, S. D., & Jain, V. (2013). Interactive visualization for curatorial analysis of large digital collection. Information Visualization, 1473871612473590. https://doi.org/10.1177/1473871612473590 Kim, Y. & Ross, S. (2011, December 5-7). Digital forensics formats: Seeking a digital preservation storage format for web archiving. International Digital Curation Conference (IDCC 2011). Retrieved from http://eprints.gla.ac.uk/62565/1/62565.pdf
Week 5: February 6	**Significant Properties and Characteristics of Digital Objects** *Team Activity: Describe significant properties of a digital object (e.g. a computer game, or a digital document). (Each team will be assigned an example digital object to work with.)* **Required Readings:** Hockx-Yu, H. & Knight, G. (2008). What to preserve?: Significant properties of digital objects. *International Journal of Digital Curation 3*(1), 141-153. [e-article] doi:10.2218/ijdc.v3i1.49 Knight, G. & Pennock, M. Data without meaning: Establishing the significant properties of digital research. *International Journal of Digital Curation* 4(1) (2009): 159-174. [e-article] doi:10.2218/ijdc.v4i1.86 Giaretta, D. (2011). Advanced Digital Preservation. Springer Heidelberg, New York. Available online through http://go.utlib.ca/cat/8214530. *Read chapter 13, section 6 on significant properties.* **Supplementary Sources:** Grace, S. & Knight, G. & Montague, L. (2009). INSPECT Final Report. Retrieved from http://www.significantproperties.org.uk/inspect-finalreport.pdf Dappert, A. & Farquhar, A. (2009). Significance is in the eye of the stakeholder. In Agosti, M., Borbinha, J., Kapidakis, S., Papatheodorou, C., & Tsakonas,

	G. (Eds.), *Proceedings of the 13th European conference on Research and advanced technology for digital libraries (ECDL'09)*. (pp. 297-308). Wetherby: The British Library. Retrieved from http://www.planets-project.eu/docs/papers/Dappert_Significant_Characteristics_ECDL2009.pdf Mike Coyne, David Duce, Bob Hopgood, George Mallen & Mike Stapleton (2007). The Significant Properties of Vector Images. JISC Digital Preservation Programme: Study on the Significant Properties of Vector Images. (Available from the JISC website) Yeo, Geoffrey. "'Nothing is the Same as Something Else': Significant Properties and Notions of Identity and Originality." *Archival Science* 10 (2010): 85-116.
Week 6: February 13	**Tools and approaches to Digital Preservation** *Team Activity: Team based comparison of emulation and migration* **Required Reading:** Adrian Brown (2013) Practical Digital Preservation: a how-to guide for organizations of any size. Facet Publishing, London. Chapter 8: "Preserving digital objects" (scan available via Blackboard). The State Records Authority of New South Wales. *Effectively manage the migration of your digital records.* Last accessed December 2016. **Website to review:** The POWRR Tool Grid v2. Available at http://www.digipres.org/tools/ **Supplementary Sources:** Guttenbrunner, M. & Rauber, A. (2012). Evaluating Emulation and Migration: Birds of a Feather? Proceedings of the 14th Conference on Asian Digital Libraries (ICADL'12), pages 158-167, Taipei, Taiwan, November 12-15 2012. Retrieved from http://link.springer.com/chapter/10.1007%2F978-3-642-34752-8_22#page-1 Rothenberg, J. (1998). Avoiding technological quicksand: Finding a viable technical foundation for digital preservation. (CLIR 77). Retrieved from http://www.clir.org/pubs/reports/rothenberg/contents.html Ross, S. & Gow, A. (1999). Digital archaeology? Rescuing neglected or damaged digital collections. (pp. 1-94). London & Bristol: British Library and JISC. Retrieved from http://www.ukoln.ac.uk/services/elib/papers/supporting/pdf/p2.pdf Kirschenbaum, M.G., Ovenden, R., & Redwine, G. (2010). Digital forensics and born-digital content in cultural heritage collections. (pp. 1-93). Washington, D.C.: Council on Library and Information Resources. Retrieved from: http://www.clir.org/pubs/abstract/reports/pub149, and [Inforum: 070.5797 K61D – Course Reserves - Check availability] Lee, C.A., Kirschenbaum, M., Chassanoff, A., Olsen, P., & Woods, K. (2012). BitCurator: Tools and Techniques for Digital Forensics in Collecting Institutions, *D-Lib Magazine 18*(5/6). doi:10.1045/may2012-lee. Retrieved from http://www.dlib.org/dlib/may12/lee/05lee.html

February 20	Reading week, no class
Week 7: February 27	**Strategy and planning** *Team Focus: Preservation planning* **Required Readings:** Becker, C., Hannes, K., Guttenbrunner, M., Strodl, S., Rauber, A., & Hofman, H. (2009). Systematic planning for digital preservation: Evaluating potential strategies and building preservation plans. *International Journal of Digital Libraries 10*(4), 133-157. http://resolver.scholarsportal.info.myaccess.library.utoronto.ca/resolve/14325 012/v10i0004/133_spfdpepsabpp.xml or http://dx.doi.org/10.1007/s00799-009-0057-1 Parliamentary Archives (2009). A Digital Preservation Policy for Parliament. Retrieved from http://www.parliament.uk/documents/upload/digitalpreservationpolicy1.0.pdf National Library of Australia. (2008). Digital preservation policy. (3rd ed). Retrieved from http://www.nla.gov.au/policy/digpres.html **Recommended Readings:** *This article is quite helpful for the assignment in its attempt to describe intentions:* Colin Webb, David Pearson, Paul Koerbin (2013) 'Oh, you wanted us to preserve that?!' Statements of Preservation Intent for the National Library of Australia's Digital Collections. *D-Lib Magazine*, January/February 2013, Volume 19, Number 1/2. http://www.dlib.org/dlib/january13/webb/01webb.html **Supplementary sources:** *On making preservation planning decisions…* Kulovits, H., Rauber, A., Kugler, A., Brantl, M., Beinert, T. & Schoger, A. (2009) From TIFF to JPEG 2000? Preservation Planning at the Bavarian State Library Using a Collection of Digitized 16th Century Printings. *Dlib Magazine* Volume 15 Number 11/12. Retrieved from http://www.dlib.org/dlib/november09/kulovits/11kulovits.html *….on large collections*: Christoph Becker , Luis Faria , Kresimir Duretec (2014). Scalable decision support for digital preservation. *OCLC Systems & Services: International digital library perspectives* 2014 30:4 , 249-284 *Additional preservation planning case studies and analysis:* Mark Guttenbrunner, Christoph Becker and Andreas Rauber, (2010) *Keeping the Game Alive: Evaluating Strategies for the Preservation of Console Video Games. International Journal of Digital Curation*, Vol 5, No 1. (Available online) Becker, C., & Rauber, A. (2011). Decision criteria in digital preservation: What to measure and how. *Journal of the American Society for Information Science and Technology*, 62(6), 1009–1028.

https://doi.org/10.1002/asi.21527
Christoph Becker and Andreas Rauber (2011) *Four cases, three solutions: Preservation Plans for Images*. Tech report, Vienna University of Technology
Christoph Becker & Andreas Rauber (2011) Preservation Decisions: Terms and Conditions apply. Challenges, Misperceptions and Lessons Learned in Preservation Planning. In: *ACM/IEEE Joint Conference on Digital Libraries (JCDL 2011)*. June 13-17, Ottawa, Canada (DOI)

On preservation policy and strategy:
Digital Preservation Working Group at the Parliamentary Archives (2008). A Digital Preservation Strategy for Parliament. Retrieved from
http://www.parliament.uk/documents/upload/digital-preservation-strategy-final-public-version.pdf
Beagrie, N., Semple, N., Williams, P. & Wright, R. (2008). Digital preservation policies study. Part 1: Final report October 2008. Retrieved from http://www.jisc.ac.uk/media/documents/programmes/preservation/jiscpolicy_p1finalreport.pdf
Smith, M. & Moore, R. (2008). Digital archive policies and trusted digital repositories. *The International Journal of Digital Curation 2*(1), 92-101. Retrieved from doi:10.2218/ijdc.v2i1.16
DigitalPreservationEurope (DPE). D3.2 Repository Planning Checklist and Guidance. (Thematic Priority: IST-2005-2.5.10). Retrieved from http://www.digitalpreservationeurope.eu/publications/reports/Repository_Planning_Checklist_and_Guidance.pdf
Bishoff, L. (2007). Planning for digital preservation: A self-assessment tool. NEDCC. Retrieved from http://www.nedcc.org/assets/media/documents/DigitalPreservationSelfAssessmentfinal.pdf

On planning processes and governance:
Kulovits, H., Becker, C., & Rauber, A. (2012). Roles and Responsibilities in Digital Preservation Decision Making. In The Memory of the World in the Digital Age: Digitization and Preservation (pp. 35–42). Retrieved from http://www.scape-project.eu/wp-content/uploads/2013/11/UNESCO_MOW2012_Kulovits_RolesResponsibilities.pdf
ISACA (2007). COBIT4.1 Executive Summary, Framework. Retrieved from http://www.isaca.org/Knowledge-Center/cobit/Documents/COBIT4.pdf
Christoph Becker, Goncalo Antunes, Jose Barateiro, Ricardo Vieira and Jose Borbinha (2011). Control Objectives for DP: Digital Preservation as an Integrated Part of IT Governance. ASIST-AM 2011, New Orleans.

Week 8: March 6	**Archival Information Systems requirements, models, systems, and services.** ***Tool demonstrations*** **Required Readings:** Lavoie, B. (2014). The Open Archival Information System (OAIS) Reference

	Model: Introductory Guide (2nd Edition) . Digital Preservation Coalition Technology Report. http://dx.doi.org/10.7207/TWR14-02 Trusted Digital Repositories: Attributes and Responsibilities. An RLG-OCLC Report. Mountain View, CA: RLG, 2002. *(Review sections 1-3)* http://www.oclc.org/resources/research/activities/trustedrep/repositories.pdf **Product websites to review:** • RODA Community – Repository of Authentic Digital Objects. http://www.roda-community.org/ • Archemativa - https://www.archematica.org/wiki/Main_Page • Preservica - http://preservica.com/ • Islandora - http://islandora.ca/ **Supplementary Sources:** Rosenthal, D. S. H., Robertson, T., Lipkis, T., Reich, V., & Morabito, S. (2005). Requirements for digital preservation systems, a bottom-up approach. *D-Lib Magazine 11*(11). doi:10.1045/dlib.magazine. [e-article] Available at http://www.dlib.org/dlib/november05/rosenthal/11rosenthal.html Consultative Committee on Space Data Systems. (2012). Reference model for an open archival information system (OAIS). (Magenta Book. CCSDS 650.0-M-2). [This Recommendation has been adopted as ISO 14721:2003]. Retrieved from http://public.ccsds.org/publications/archive/650x0m2.pdf
Week 9: March 13	**Metadata** *Tool demonstrations* **Required Readings:** Lavoie, B. & Gartner, R. (2013). Preservation Metadata (2nd edition). DPC Technology Watch Report 13-03. http://dx.doi.org/10.7207/twr13-03 **Recommended readings:** PREMIS Editorial Committee. (2015). PREMIS Data Dictionary for Preservation Metadata, Version 3.0. Retrieved from http://www.loc.gov/standards/premis/v3/premis-3-0-final.pdf (Read Introduction; skim other sections to get an overview.) See http://www.loc.gov/standards/premis/v3/ for other PREMIS documents. **Supplementary Sources:** *The following two articles are interesting because they illustrate how working systems use and combine metadata standards.* Dappert, A. & Farquhar, A. (2011). Implementing metadata that guide digital preservation services. *International Journal of Digital Curation 6*(1), 238-254. Retrieved from http://www.ijdc.net/index.php/ijdc/article/viewFile/176/245 Dappert, A. & Enders, M. (2008).Using METS, PREMIS and MODS for Archiving eJournals. *D-Lib Magazine 14*(9-10). Retrieved from http://www.dlib.org/dlib/september08/dappert/09dappert.html

	van Ballegooie, M. & Duff, W. (2006). Re: Archival Metadata. [website manual: DCC Digital Curation Manual, (Ross, S. & Day, M (Eds.)]. Retrieved from http://www.dcc.ac.uk/resource/curation-manual/chapters/archival-metadata
Week 10: March 20	**Trust and risk models** *Tool demonstrations* **Required Readings:** Ross, S. & McHugh, A. (2006). The role of evidence in establishing trust in repositories. *D-Lib Magazine 12*(7/8). Retrieved from http://www.dlib.org/dlib/july06/ross/07ross.html McHugh, A., Ross, S., Innocenti, P., Ruusalepp, R., & Hofman, H. (2008). Bringing Self-assessment Home: Repository Profiling and Key Lines of Enquiry within DRAMBORA. International Journal of Digital Curation, 3(2), 130–142. https://doi.org/10.2218/ijdc.v3i2.64 Vermaaten, S., Lavoie, B., & Caplan, P. (2012). Identifying threats to successful digital preservation: The SPOT model for risk assessment. *D-Lib Magazine 18*(9/10). Retrieved from http://www.dlib.org/dlib/september12/vermaaten/09vermaaten.html **Websites to review:** • Data Seal of Approval: http://datasealofapproval.org/en/ **Supplementary Sources:** *For the development of audit and certification standards:* OCLC/CRL. (2007, February). Trustworthy repositories audit & certification (TRAC): Criteria and checklist. Retrieved from http://www.crl.edu/PDF/trac.pdf Consultative Committee for Space Data Systems. (2011). Audit and Certification of Trustworthy Digital Repositories: Recommended Practice. (Magenta Book. CCSDS 652.0-M-1). Retrieved from http://public.ccsds.org/publications/archive/652x0m1.pdf) Nestor Working Group. (2009, November). Catalogue of Criteria for Trusted Digital Repositories. (Version 2). Retrieved from http://nbn-resolving.de/urn:nbn:de:0008-2010030806 *For DRAMBORA:* Digital Curation Centre and DigitalPreservationEurope, (February 2007), DCC and DPE Digital Repository Audit Method Based on Risk Assessment, *v1.0*, retrieved Nov 2013 from http://www.repositoryaudit.eu/download (Available on Blackboard. Focus on Part I and Part III, but review Part II to understand the basic concepts and stages of the assessment process.) *For a different risk management perspective:* José Barateiro, Gonçalo Antunes, Filipe Freitas, José Borbinha (2010) Designing Digital Preservation Solutions: A Risk Management-Based Approach. *International Journal of Digital Curation* Vol. 5, No. 1, pp. 4-17.

	doi:10.2218/ijdc.v5i1.140 http://www.ijdc.net/index.php/ijdc/article/view/143 *For a comprehensive review of assessment frameworks, including ISO 16363, see* Maemura, E., Moles, N., & Becker, C. (forthcoming). Organizational Assessment Frameworks for Digital Preservation: A literature review and mapping. *Journal of the Association for Information Science and Technology.* Preprint retrieved from https://tspace.library.utoronto.ca/handle/1807/73869
Week 11: March 27	**Key issues and case studies; sustainability** We will review key issues in digital curation, discuss them in the context of concrete cases, and talk about sustainability – the 'capacity to endure'. **Tool fair:** The tools demonstrated previously will be shared in the style of a fair to allow everybody to get their hands on some of the tools they found most interesting. **Required Reading:** Choose one full paper from a recent conference in digital curation and preservation such as IPRES or IDCC (2014 or later), or in the IJDC. Review its contribution to digital curation theory and practice in your blog post in relation it to what we have discussed in the course. (Both research and practice papers qualify as "full papers", just not posters, demos, short papers, panels etc.) *Skim the following:* The Blue Ribbon Task Force on Sustainable Digital Preservation and Access. (2010). Sustainable Economics for a Digital Planet: Ensuring Long-Term Access to Digital Information. Retrieved from http://brtf.sdsc.edu/biblio/BRTF_Final_Report.pdf **Supplementary Sources:** *The 4C project created a number of insightful documents about costs, benefits and sustainability that build on the BRTF task force report, in particular:* 4C project (2015). Investing in Curation: A Shared Path to Sustainability. Retrieved from http://www.4cproject.eu/roadmap 4C project (2015). The Digital Curation Sustainability Model. Retrieved from http://www.4cproject.eu/dcsm *For cost estimation models, start here:* 4C project: Summary of Cost Models. http://4cproject.eu/summary-of-cost-models 4C project: Evaluation of Cost Models and Needs & Gaps Analysis http://4cproject.eu/d3-1
Week 12: April 3	**Current and emerging trends** We will close by reviewing emerging trends and the direction of the field in practice and research.

Required Readings:
NDSA: National Digital Stewardship Agenda 2015. Available online at
http://www.digitalpreservation.gov/ndsa/documents/2015NationalAgenda.pdf. See http://www.digitalpreservation.gov/ndsa/nationalagenda/

Recommended Readings:
Dallas, C. (2016). Digital curation beyond the "wild frontier": a pragmatic approach. *Archival Science*, 16(4), 421–457.
https://doi.org/10.1007/s10502-015-9252-6

Supplementary Sources:
Milic-Frayling, N. (2013). Sustainable Computation—Foundation for Long Term Access to Digital. *Open Research Challenges in Digital Preservation* workshop at the 10th International Conference on Preservation of Digital Objects (IPRES 2013), Lisbon, Portugal. Retrieved from
http://digitalpreservationchallenges.files.wordpress.com/2013/09/2013milicfrayling.pdf

Statement of Acknowledgement of Traditional Land.
We wish to acknowledge this land on which the University of Toronto operates. For thousands of years it has been the traditional land of the Huron-Wendat, the Seneca, and most recently, the Mississaugas of the Credit River. Today, this meeting place is still the home to many Indigenous people from across Turtle Island and we are grateful to have the opportunity to work on this land.

Acknowledgement and citation. This syllabus is in part built on Prof. Seamus Ross' earlier syllabus. In it, he extended his "[t]hanks to Prof. Sara Grimes for providing her Research Course Syllabus as a model. The text in sections … *Guidelines for Assignments* and *Ground Rules* are in quotes because they are taken directly from her Syllabus—there were no better words to say what she had said so eloquently."