Course Description

This course examines the decision space of preservation. It explores the levels of control and decision making that need to be addressed in digital preservation and curation. We will explore the context of preservation planning and actions and enable students to develop active approaches to securing the access to digital collections over time by formulating specific objectives, evaluating possible alternatives and specifying concrete steps of action. In doing so, the course will discuss and critically evaluate a systematic method for decision making and planning and explore the limitations of currently available technologies as well as possible approaches to continuously assessing and improving organizational capabilities as new technologies, risks and opportunities emerge. Finally, it will discuss longevity as a design concern in Information Systems.

Goals and Objectives

Students having completed this course should be capable of:

- Formulating an awareness of environmental aspects of relevance for digital longevity and specify how a curation and preservation process needs to monitor these aspects (as demonstrated in assignment 4)
- distinguishing the levels of control and decision making that need to be addressed in digital preservation and analyze the decision space in a given organizational context (as demonstrated in assignments 2 and 4)
- conducting a practical analysis and assessment of a given digital collection and its key technical properties (as demonstrated in assignment 1)
- Conceptualizing, formalizing, and documenting the preservation context in an organizational environment to enable planning of actions (as demonstrated in assignment 2)
- Formulating specific preservation objectives and explore strategies to evaluate possible ways to achieve these objectives (as demonstrated in assignments 2 and 3)
- Empirically evaluate the quality and suitability of technical approaches to object-level preservation such as migration and emulation tools given specific circumstances (as demonstrated in assignment 3)
- Demonstrating a working experience with cutting-edge preservation tools and articulating their limitations (as demonstrated in all assignments, but particularly in assignment 4)
• Playing an active role in promoting, assessing and improving operational preservation in organizations (as demonstrated in assignments 1 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, and a crucial skill for the curator is effective decision making and control of those activities that aim to secure the understandability and authenticity of digital materials over time. Students completing this course will gain an understanding of the concepts, practices, and methods of decision making in digital preservation and explore some of the crucial interfaces between curation and IT, one of the “diverse horizons of information disciplines” (Program Outcome 1). Through a range of in class activities and course assignments they will engage in aspects of digital preservation 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… of information” and an appreciation of “the impact of such developments” on their professional roles in this process. Students will recognize that an active decision making and leadership role in 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 three hours each week. Each class will consist of a lecture, a team activity or guest conversation period, and a discussion period. The lecture portion of the class will introduce the key concepts and cover material that is not available in published literature. There will normally be a 15 minute break after 1 hour. This will be followed by a team-based activity. The last section of the class will involve discussion of the team-based activity and the readings using questions from blogs of class members as a starting point, and sometimes an additional (shorter) lecture.

On their own time, students must complete weekly course readings and complete written assignments, and maintain a blog.

Prerequisites

Required: INF2122H, Introduction to Digital Preservation and Curation
Recommended: INF1003H Information Systems, Services and Design

Course Materials

Textbooks/Readings
This course does not have a textbook. Most readings (both required and recommended) are available online, and links are included in the syllabus.

Website/Resources
“Course materials and resources aimed at helping students with assignments and key concepts will be made available online, through Blackboard (http://portal.utoronto.ca). 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 2-3 days of lecture.

TBD: A set of additional readings on related work in
- Governance, processes, responsibility
- Governance, Risk and Compliance
- Policy and formalization
- Significant properties
- Software quality and evaluation
- Characterization, content profiling and technical challenges
- Benchmark data sets, forensics, quality assurance and ground truth
- Future research challenges

Evaluation

The essential goal and premise of the assignments is to complete a preservation plan in teams of three, progressing through four stages as the course progresses. For each stage, a log book should be kept documenting what is being done, how long it takes, and what the major difficulties are. Additionally, each student individually maintains a weekly blog.

<table>
<thead>
<tr>
<th>Assignment</th>
<th>Weight</th>
<th>Due</th>
</tr>
</thead>
<tbody>
<tr>
<td>A content profile</td>
<td>20%</td>
<td>Week 5</td>
</tr>
<tr>
<td>Plan part 1</td>
<td>25%</td>
<td>Week 10</td>
</tr>
<tr>
<td>Plan stage 3: completed</td>
<td>35%</td>
<td>Week 12</td>
</tr>
<tr>
<td>Weekly blog</td>
<td>20%</td>
<td>weekly</td>
</tr>
</tbody>
</table>

Assignment 1: A Content Profile
Due Week 5: Weight 20%
Create a content profile for a personal or professional data collection.
- Characterize files using fits and/or tika
- Setup and use c3po
- Analyze and visualize the information using c3po
- Identify 3 sets of objects that are problematic, interesting, risky and/or valuable and for which you would like to create a preservation plan. Describe preservation intent for at least one of them.
- Export content profiles of the problematic sets
- The deliverable is a report on the process, including visuals on key properties as created by c3po, and a set of 3 XML files corresponding to the exported content profiles.
(In a feedback session, we will discuss the sets and choose a set of objects for which to create a plan.)

**Assignment 2: A preservation plan part 1**
Due week 10, Weight: 25%
Building on assignment 2, start creating a preservation plan and complete the evaluation of at least two candidate actions: *No action* and one alternative action.

- Create a preservation plan in state 4:
  1. Basis defined
  2. Content profile defined
  3. Objective tree created
  4. Candidates specified (including no action)
- Evaluate candidates according to requirements and criteria
- Document results in the plan
- Suggest best action and complete the preservation plan
- The deliverable is the preservation plan in XML and PDF, the log book, and a short summary of progress.

**Assignment 3: Completed preservation plan**
Due Week 12: Weight 35%
- Review decision making process and revise if necessary
- Specify monitoring conditions and identify information sources for each
- Publish the preservation plan
- The deliverable is the preservation plan in XML and PDF, the log book, a summary of progress, a report summarizing the experience of planning, reflecting on the activities, roles, skills, expertise and challenges, and individual confidential self and peer evaluation reports.

**Assignment 4: Weekly blog**
Due Weekly (considered weekly but evaluated week 12): Weight 20%

You will maintain a weekly blog from week 2 to week 12. Your contributions will (and should) vary in terms of length and topic, but try to keep your posts focused on the topic of the upcoming week and brief (up to 200 words). 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. In your contribution you should pose a question which arises from the readings and is relevant to the focus of that week’s discussion. We will mine these questions each week to help shape the discussion period of each class. Your blog post must be completed by noon the day before class meets.

*You can skip three blog entries, i.e. 8 blog entries will be evaluated.*
Guidelines for Assignments

All written assignments for this course must be submitted through Blackboard and handed in before the start of lecture except for the weekly blog posts which are due at 12 noon the day before the week’s class. 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.

The required format for Assignments 2, 3 and 4 is as follows:
- 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.

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.

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 (Which can be found here: http://www.sgs.utoronto.ca/informationfor/students/track/extsn.htm).”

Grading

“Grading for this course will follow the iSchool’s official Guidelines to Grade Interpretation of letter grades, as well as the University’s policy on Graduate Grading and Evaluation Practices. These sources define grades in the A range as “excellent” and grades in the B range as “good.” Please refer to the guidelines for detailed descriptions of these categories. Assignments will be graded and returned within 2-3 weeks of submission.” The blog posts will be given a mark at the end of the term.

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.”

Students with a Disability or Health Consideration

“Students with diverse learning styles and needs are welcome in this course. If you have a disability or health consideration that may require accommodations, please feel free to approach the instructor and/or the Accessibility Services Office (http://www.accessibility.utoronto.ca/) as soon as possible. The Accessibility Services staff is available by appointment to assess specific needs, provide referrals and arrange appropriate accommodations. The sooner arrangements are made - the quicker we can assist you.”
Writing Support

“The SGS Office of English Language and Writing Support provides free writing support to graduate students. Services are designed for both native and non-native speakers of English, and include non-credit courses, single-session workshops, individual writing consultations, and online resources. Students are encouraged to use these services as needed.”

Academic Integrity

“The iSchool has a strict zero-tolerance policy on plagiarism, as defined in section B.I.1. (d) of the University’s Code of Behavior on Academic Matters. Before you embark on your first writing assignment, please make sure that you:

- Consult the University’s site on Academic Integrity: http://www.utoronto.ca/academicintegrity/
- Acquaint yourself with the Code and Appendix “A” Section 2; http://www.governingcouncil.utoronto.ca/policies/behaveac.htm
- Review the material you covered in Cite it Right;
- Consult the site How Not to Plagiarize: http://www.writing.utoronto.ca/advice/using-sources/how-not-to-plagiarize”

In Class Activity Teams

Each week we will have one team-based activity for about an hour to allow us to engage in practical ways with the concepts and knowledge that we are covering. We will breakout into teams of three to 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, brings it to class, puts the software if any on it, and returns it to the INFORUM after class. (Groups would be wise to appoint alternatives in case of illness.)

Schedule of Lecture Topics and Readings

<table>
<thead>
<tr>
<th>Week 1</th>
<th>Introduction to Preservation Planning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Read:</td>
<td>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</td>
</tr>
</tbody>
</table>

#### Lecture topics:
- Digital Preservation and Preservation Planning
- The nature of digital, fragility and levels of threats, threat models and mitigation
- What is a preservation plan? What is in a preservation plan? (activity)
- The difference between policies and plans
- Governance, Management, Operations
- Capabilities, processes, activities, responsibilities, and skills: What do we need to know?
- Overview of course modules and assignments
- How the topics from INF2122 feed into this course and how this course will build on them

### Week 2
#### Approaches, models, methods, and tools
Object Management Group (2010). The Business Motivation Model v1.1

#### Lecture topics
- Levels of control
- Preservation capabilities, models and assessment methods
- Preservation planning methods, tool support and automation

#### Recommended Optional Readings:

### Week 3
#### Content profiling (what we have)
Becker, Faria & Duretec: Scalable Preservation Intelligence for Information Longevity. OCLC Systems and services 2013 (accepted for publication)

#### Tool websites:
- [https://code.google.com/p/fits/](https://code.google.com/p/fits/)
- [http://ifs.tuwien.ac.at/imp/c3po](http://ifs.tuwien.ac.at/imp/c3po)
Syllabus for INF 2308H: Special Topics: Digital Preservation Planning

**Week 4**

**Preservation Context and Management (where we are)**


Lecture topics:
- Ends and Means: The decision space of preservation planning
- Policies and the context of preservation planning
- Understanding preservation intent and the preservation case
- Formalizing preservation goals and objectives

**Week 5**

**Preservation goals and objectives (what we want)**

Reading:

Lecture topics:
- High-level goals
- Authenticity, Reliability and Significant properties
- From high-level goals to measurable criteria
- Types of decision criteria, quality models, and measurement methods
- Formal representation of objectives as an interface to automated systems

**Week 6**

**What to evaluate and how (can we get there?)**

Reading:

- [http://openplanets.github.io/scout/](http://openplanets.github.io/scout/)

This is the most technical session and will be organized like a tutorial
- What is (in) a content profile?
- Formats, object characteristics, technical processes and tools
- How to aggregate and analyze
- Tool support: Using fits and c3po
- Scout and policy compliance
<table>
<thead>
<tr>
<th>Week 7</th>
<th><strong>Decisions and deployment</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Lecture topics:</strong></td>
<td></td>
</tr>
<tr>
<td>• How to decide?</td>
<td></td>
</tr>
<tr>
<td>• Multi-criteria decision making in practice</td>
<td></td>
</tr>
<tr>
<td>• Objective evidence and subjective assessment</td>
<td></td>
</tr>
<tr>
<td>• Do nothing is a real option</td>
<td></td>
</tr>
<tr>
<td>• Example cases</td>
<td></td>
</tr>
<tr>
<td>• Workflows for data processing</td>
<td></td>
</tr>
<tr>
<td>• Taverna and other means of constructing workflows</td>
<td></td>
</tr>
<tr>
<td>• Component profiles and workflow specification in Taverna</td>
<td></td>
</tr>
<tr>
<td>• Quality of Service and how to specify it</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Week 8</th>
<th><strong>What to monitor and how</strong></th>
</tr>
</thead>
</table>

Faria et al (2012) *A novel design for a preservation watch system*, ICADL, Taipei |
| **Lecture topics:** |  
| • Monitoring compliance, risks, and opportunities |  
| • What to monitor and how |  
| • Limits of existing methods and tools |  
| • A preservation lifecycle |  

**Recommmended optional reading:**  
### Week 9
**Case studies and lessons learned**

**Reading:**
- Christoph Becker and Andreas Rauber (2011) Four cases, three solutions: Preservation Plans for Images. Tech report, Vienna University of Technology

**Lecture topics:**
- The simple case? Preserving images and sound
- The complex case? Preserving video games
- Some lessons learned
- How long does it take to create a preservation plan? Experiences from case studies
- Comparison of cases and approaches

### Week 10
**Review of preservation plans. Roles and responsibilities**

- Presentation and discussion of preservation plans
- Activity: Peer-review draft preservation plans of the teams
- Present content sets, goals and objectives, candidates and evaluation methods
- Review planning steps and team work activity and discuss difficulties
- Review: What activities are carried out in planning? What are the skills for each of them? Who should participate?

### Week 11
**Capability assessment and improvement**

**Reading:**

**Lecture topics:**
- How does an organization go about assessing what they are doing and identifying needs and opportunities for improvement?
- Standard DP frameworks provide (limited) support
- Standard assessment frameworks are not always applicable
- Recent work in DP to address this issue

### Week 12
**Practical challenges, emerging approaches and future research directions**

**Reading:**

**Lecture topics:**
- Roundup and lessons learned
- Information longevity and sustainable computation
- Longevity as a design concern in Information Systems
Contacting the Instructor & Supporting Instructor

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

Acknowledgement and citation. The structure of this syllabus is derived from Prof. Seamus Ross’ Digital Preservation and Curation course 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 Website/Resources, Guidelines for Assignments, Grading, Ground Rules, Students with a Disability or Health Consideration, Writing Support, and Academic Integrity are in quotes because they are taken directly from her Syllabus—there were no better words to say what she had said so eloquently.”