**Meeting Time:** August 25- October 13, 2014  
**Location:** Web  
**Credits:** 1.5  
**Perquisite:** None  
**Instructor:** Ayoung Yoon  
**E-Mail:** ayyoon [at] email [dot] unc [dot] edu  
**Course Web Site:** [https://sakai.unc.edu/](https://sakai.unc.edu/)

**Course Description and Objectives**

This course aims to provide knowledge of the role and potential of the Web as a source for archival collection development, as it has already been part of archival collections for many government, educational, and cultural institutions. Archivists need not only to react to the needs of these many institutions but also to be proactive to these changes and their needs for collection development.

Through reading materials on a variety of topics, issues, and challenges surrounding Web archiving and preservation, analyzing Web archives, and performing hands-on work using a Web archiving tool, students will gain insight on relevant issues arising from the nature and characteristics of the Web and how to make correct use of the Web in their archival work. This course will expose students to existing and emerging tools for capturing Web content, with an emphasis on laboratory practice using Web crawlers (Archive-it). Students will also learn about current preservation formats of Web-based content.

Upon completing this course, students will be able to:

1. Discuss the role and the potential of the Web as information and characteristics of the Web for archiving and preservation.
2. Be familiar with the tools and appropriate techniques for preservation of information delivered through the “surface” Web (static Web pages, blogs, etc.) and information that is part of the “deep” Web (e.g., databases, authenticated resources, etc.).
3. Recognize the challenges of Web archiving.
4. Become proficient at setting up a Web crawl using Archive-It.
5. Create a Web archive.
6. Increase their awareness of legal and policy constraints on Web archiving.
7. Be familiar with standards and best practices for sustainability of archived Web content.

**Email**

Please check the course listserv on a regular basis if not daily. This will be our primary means of communication.

The address is: inls690_208W14@sakai.unc.edu

If you have a question of general interest to the class such as “What do you mean by X in assignment Y?” please send this to the class list and I will answer it so that everyone can benefit.

**How we will conduct “Class”**

While online courses afford great convenience, they also demand extra effort from both instructors and students. Because there is no face time for lectures, discussions, group work, or other activities, all this must take place through the course site on Sakai. This involves extensive writing and creation of slides, videos, and other media we will use to communicate our ideas and questions. The syllabus, assignments, and many of the readings (unless available online) will be posted on Sakai. Each week I will provide slides or notes on important points, but much of the work of the course will take place in the forums via discussion of the readings, videos, and slides, and your own work (small assignments).

**Online Etiquette (Netiquette) Guideline**

- [http://www.indiana.edu/~icy/netiquette.html](http://www.indiana.edu/~icy/netiquette.html)
- [http://www.brighthub.com/education/online-learning/articles/26946.aspx](http://www.brighthub.com/education/online-learning/articles/26946.aspx)

**Readings and Sources**

Readings are on Sakai or links on the World Wide Web. It is expected that students will have read the materials before class, as we will be referring to them in lectures and in the exercises.

This class has one required Textbook. This textbook is available electronically via Sakai.


Other useful sources for the class:
• PoWR, the Preservation of Web Resources Handbook. The JISC-PoWR Team. (2008).
• Web Archiving. Alex Bal. UKOLN, University of Bath.
• International Internet Preservation Consortium (IIPC). http://netpreserve.org/about/index.php
• International Web Archiving Workshop (IWAW). http://bibnum.bnf.fr/ecd/!

Assignments and Grading

Grades will be based on class attendance and participation and a series of assignments.

<table>
<thead>
<tr>
<th>Assignment</th>
<th>Due Date</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class Participation and Work Log (Sakai Forum)</td>
<td>Ongoing</td>
<td>15%</td>
</tr>
<tr>
<td>Review of Web archives</td>
<td>Week 1. (Aug 28)</td>
<td>30%</td>
</tr>
<tr>
<td>Web Crawl Project</td>
<td>Week 2.</td>
<td>55%</td>
</tr>
<tr>
<td>a. Archive-It training</td>
<td>Week 2.</td>
<td>5%</td>
</tr>
<tr>
<td>b. Project description and selection / scope update</td>
<td>Week 3. (Sep 11)</td>
<td>10%</td>
</tr>
<tr>
<td>c. Metadata update</td>
<td>Week 4. (Sep 18)</td>
<td>10%</td>
</tr>
<tr>
<td>d. Quality control</td>
<td>Week 5 (Sep 25)</td>
<td>5%</td>
</tr>
<tr>
<td>e. Preservation plan</td>
<td>Week 6. (Oct 2)</td>
<td>10%</td>
</tr>
<tr>
<td>f. Project and tool evaluation</td>
<td>Week 8. (Oct 15)</td>
<td>10%</td>
</tr>
<tr>
<td>g. Issue Tracking (Sakai Forum)</td>
<td>Ongoing</td>
<td>5%</td>
</tr>
</tbody>
</table>
**Class Participation and Work Log (15% of total grade)**

Students are expected to complete all required readings. This is important, as a portion of the assignments (or exercises) will be based on readings.

In addition, it is important to work on your small assignments each week in order to complete your term project (Web crawl project) on time, and not to do all the work at the last minute.

*Work log*: a Work Log board in Forums (Sakai). You should describe what you have done for the week and how much time you put into that work. If you run into any issue, please fill the issue tracking form as well. (Also read others’ issue tracking in case your colleagues have the same issue and have resolved it.)

*Questions* and *General Discussion* boards in Forums (Sakai): Students are encouraged to use these spaces for sharing their thoughts on certain topics, providing reflections on any of readings, and raising any questions about anything relevant to the course or the tool. I’ll also raise some questions for you to think and respond in *General Discussion* board (may not every week).

**Assignments**

* All assignments must be turned in through the Sakai class website (except for your work log and issue tracking). Late submissions will not be accepted unless students have consulted with the instructor prior to the late submission.

1. **Review of Web archives (30%): Due (Aug 28)**
   - A brief paper (2-3 pages, font-size 12, double-spaced)
   - Instructions will be distributed a week before the deadline.

2. **Web Crawl Project (55% of total grade)**

Students will work on term projects throughout the semester.

   - **Archive-It Training** (Watch the video in week 2)

The project is to develop a Web archive using Archive-It. All students are required to watch the Archive-It training video run by an Archive-It specialist. This is important, as you need to learn about the tool in order to develop your own Web archive throughout the semester.

   - **Small assignments** (weekly exercises)
     - Project description and selection / scope update: Due Week 3 (Sep 11)
     - Metadata update: Due Week 4 (Sep 18)
Students will complete small tasks each week based on topics of the week for their Web archives. The class assignments will lead students to follow a step-by-step process to develop a Web archive, and thus it is important to be on track with the small assignments and not get delayed.

Instructions will be distributed a week before the deadline of each assignment. You should write a one-page report (font size 12) each week about what you have worked on using Archive-It, following the instructions for each week.

- **Issue Tracking:** On going

In addition to the weekly report, students should document issues that they experience during the Web archiving development. It is likely that your colleagues will also have the same issues that you experience, so put your issue log on the Sakai forum, named “issue tracking” so that everyone in the class can see what issues others have and help (by replying to each other’s posts) to resolve some of the issues. An issue log form is available on the Sakai forum (see “sample issue tracking”). If you run into any issue, check others’ issue tracking. Your issue is not likely something that only you are encountering. Collaborative efforts are always welcome to resolve the issue.

* **Special Needs:** If you need an accommodation for a disability or have any other special need, please make an appointment to discuss this with me. I will be most able to address special circumstances if I know about them early in the semester. My contact information are listed at the beginning of this syllabus.

**Important note on plagiarism**

Unless otherwise specified in an assignment, all submitted work must be your own, original work. Any experts from the work of others must be clearly identified as a quotation, and a proper citation provided. Be aware of the University of North Carolina policy on plagiarism. All cases of plagiarism (unattributed quotation or paraphrasing) of anyone else's work, (e.g. from published materials) will be officially reported and dealt with according to UNC policies (Instrument of Student Judicial Governance, Section II.B.1. and III.D.2, http://instrument.unc.edu).

**Evaluation**

Based on UNC Registrar Policy for graduate-level courses (http://registrar.unc.edu/AcademicServices/Grades/ExplanationofGradingSystem/index.htm#grad), both assignment and semester grades will be H, P, L or F. Few students will obtain an "H," which signifies an exceptionally high level of performance (higher than an "A" in an A-F systems). The following is a more detailed breakdown used for class
assignments:

H  Superior work: complete command of subject, unusual depth, great creativity or originality

P+  Above average performance: solid work somewhat beyond what was required and good command of the material

P  Satisfactory performance that meets course requirements (expected to be the median grade of all students in the course)

P-  Acceptable work in need of improvement

L  Unacceptable graduate performance: substandard in significant ways

F  Performance that is seriously deficient and unworthy of graduate credit

F  An unacceptable performance. The F grade indicates that the student's performance in the required exercises has revealed almost no understanding of the course content.

According to UNC Registrar Policy, undergraduate grades are based on the following definitions:

A  Mastery of course content at the highest level of attainment that can reasonably be expected of students at a given stage of development. The A grade states clearly that the students have shown such outstanding promise in the aspect of the discipline under study that he/she may be strongly encouraged to continue.

B  Strong performance demonstrating a high level of attainment for a student at a given stage of development. The B grade states that the student has shown solid promise in the aspect of the discipline under study.

C  A totally acceptable performance demonstrating an adequate level of attainment for a student at a given stage of development. The C grade states that, while not yet showing unusual promise, the student may continue to study in the discipline with reasonable hope of intellectual development.

D  A marginal performance in the required exercises demonstrating a minimal passing level of attainment. A student has given no evidence of prospective growth in the discipline; an accumulation of D grades should be taken to mean that the student would be well advised not to continue in the academic field.

F  For whatever reason, an unacceptable performance. The F grade indicates that the student's performance in the required exercises has revealed almost no understanding of the course content. A grade of F should warrant an advisor's questioning whether the
student may suitably register for further study in the discipline before remedial work is undertaken.

**AB**
Absent from final examination, but could have passed if exam taken. This is a temporary grade that converts to an F* after the last day of class for the next regular semester unless the student makes up the exam.

**FA**
Failed and absent from exam. The FA grade is given when the undergraduate student did not attend the exam, and could not pass the course regardless of performance on the exam. This would be appropriate for a student that never attended the course or has excessive absences in the course, as well as missing the exam.

**IN**
Work incomplete. This is a temporary grade that converts to F* at the end of eight weeks into the next semester unless the student makes up the incomplete work.

**W**
Withdrew passing. Entered when a student drops after the six-week drop period.

### Course Schedules

**Week 1. (Week of Aug 25):** Introduction & Basic concepts in Web Archiving; Surface and Deep Web

**Review of Web archives assignment DUE on Aug 28**

- Basic concepts in Web Archiving
  - What is web archiving?
  - What do we need to know about Web for archiving it?
  - Surface Web and deep Web

### Required readings

- Masanès, Chapters 1 and 9

### Additional readings

- Smith, Elizabeth H. Lost in Cyberspace: Have Archives a Future? Paper delivered at the *Australian Society of Archivists Conference,* Melbourne, 19 August 2000. (not in UNC library)


---

**Week 2. (Week of Sep 1): Archiving different types of Web materials**

**Watch video “Archive-It training”**

- Types of materials on the Web
- Who does archives Web? Why they do?
- Web and organizational memory
- Archiving personal web & social media

**Required readings**


**Additional readings**

Week 3. (Week of Sep 8): Exploration of existing web archives / Selection & Scoping

Project description and selection / scope update Due on Sep 11

- Review of existing web archives
  - What are the different approaches current web archives take?
- Selection policy and criteria
- Different selection approaches: domain, topic or event, media type and genre based

Required readings

- Please review at least 3 from the following examples and think about their different approaches
  - UK Web Archive: [http://www.webarchive.org.uk/ukwa/](http://www.webarchive.org.uk/ukwa/)
  - Topical archives developed using two different approaches
  - Media, form, and genre based
    - U.S. Fish and Wildlife services, Digital Media Archives: [http://images.fws.gov/](http://images.fws.gov/)

- Masanès, Chapter 3

Additional readings


Week 4. (Week of Sep 15): Acquisition and Collection Methods / Metadata and Description

Metadata update Due on Sep 18

• Dynamics of websites and different technology
• What information should be provided? (Metadata)
• How are users of web archives’ needs different from users of active websites?

Required readings


Additional readings

• Masanès, Chapter 4

Week 5. (Week of Sep 22): Quality Control and Post Capture Processing

**Quality Control Due on Sep 25**

• Why is post-collection processing necessary?
• Different methods of quality control (types of tests)
• Issues and challenges

Required readings


Additional readings

• Martin, C., Lasfargues. F., and Medjkoune, L. What if Web Archiving were as reliable as pushing a simple button? IMF, France. Available at http://www.museumsandtheweb.com/mw2011/papers/what_if_web_archiving_were_as_reliable_as_push

Week 6. (Week of Sep 29): Preserving Web Sites

**Preservation plan due on Oct 2**

• Challenges of preserving websites (technical, financial, and organizational)
• Strategies (passive vs. active; emulation vs. migration)
• Significant properties of web
Required readings

- Masanès, Chapter 8
- PoWR, the Preservation of Web Resources Handbook. The JISC-PoWR Team. (2008).

Additional readings


Week 7. (Week of Oct 6): Access and Use / Legal and Ethical Issues

- Web archives access tools
- Who uses web archives? How are web archives used?
- Intellectual properties / Privacy
- Content reliability

Required readings

- Masanès, Chapter 2 and 6
- Charlesworth, A. Legal issues relating to the archiving of Internet resources in the UK, EU, USA and Australia. JISC, The Wellcome Trust, University of Bristol. (2003).
Additional readings


Week 8. (Week of Oct 13): Evaluation of the project

Evaluation of the project and tool Due on 15