

Introduction to the Creation and Management of Digital Libraries

Spring 2013

Time: Wednesdays 6:30 - 8:20

Place: Bobst PC1 / Avery East

Instructor: Brian Hoffman (bh@nyu.edu)

Syllabus Version: 0.1

Course Description

This class will introduce students to the practical social and technical challenges involved in building digital libraries. We will explore the foundational concepts necessary for planning and executing successful digital library projects, both by reading articles written by digital library practitioners and by doing hands-on exercises inside and outside the classroom. The class will cover both technical and non-technical digital library topics, including intellectual property, digitization, XML-encoded metadata, and digital repository design. No previous technical experience is required.

Goals and Learning Objectives

This class is intended to introduce you to the core concepts, technologies and best practices of the digital library field. It should serve to give you a foundation from which to build professional competency as a planner, manager, or technician participating in digital library projects.

Palmer School Learning Objectives: This course is designed to primarily address Learning Objective 2, "Utilize a broad range of systems and technologies to manage and deliver information." Within this Learning Objective, the class will focus in turn on each of the five parts of this Learning Objective:

- A. use professional standards to organize, manage, preserve, evaluate and deliver information resources in a variety of formats;
- B. explain and apply systems for organizing and structuring information and knowledge, such as cataloging, classification and other metadata formation standards;
- C. search, retrieve and synthesize information from a variety of systems and sources; and
- D. evaluate information systems and technologies based on functionality, usability, cost, and quality;
- E. build information systems and/or records used in such systems.

The course material will also be relevant to parts C and D of Learning Objective 3, "Deliver user-centered services and cost-effective programs tailored to the needs of increasingly diverse populations." However, the emphasis will remain on Learning Objective 2.

Assignments

There are 3 assignments over the course of the semester. However, you will need to work on these concurrently, since assignment 3 will require the application of skills and concepts we will start learning early in the semester. Grading rubrics for each assignment will be distributed on the first day of class.

Assignment 1: Response Paper

A 3 page (850-950 words) paper in which you make an argument for or against the importance of traditional librarianship to the creation, management, and use of digital libraries. There is no right answer to this question, but whichever position you choose to take must be backed up with sound reasoning and / or data. If you are writing a *for* paper, you will need to explain why purely digital "systems for organizing and structuring information" require one or more aspects of librarianship (e.g., classification, selection, integrated public services). If you are writing an *against* paper, you need to explain how the systems of traditional librarianship can be replaced by modern web technologies.

Assignment 2: PechaKucha

Each student will present a short PechaKucha style talk on a technology or methodology relevant to digital libraries. Your talk may cover a particular technical standard or semantics (e.g., METS, OAI-ORE, OAIS, RDF, LOC "Bagit", PREMIS); a system other than Omeka (e.g., DSpace, Fedora, Blacklight, Drupal); a methodology (e.g., Agile Project Management, Use Cases, Rapid Prototypes); or a service model (e.g., Software as a Service / SAAS, Trusted Digital Repositories, Scientific Data Curation). Your talk should explain what problem(s) the concept or technology was designed to solve, how well it does or doesn't actually solve such problems, how it works, and what competitors or alternatives also exist. Talks should be 5-6 minutes and use 10-20 slides.

Assignment 3: Digital Library / Archival Description

Over the course of the semester, each student will either A) build a digital collection using the Omeka system (<http://omeka.org/>) or B) create an archival arrangement using the ArchiveSpace system (<http://www.archivespace.org/>). We'll discuss the merits of each option in class. Students choosing either option will need to identify a collection to digitize and / or describe using one of these systems (digitization applies only to option A). You may choose a scholarly collection to which you can obtain access, or can select original materials (a friend's artwork, a family photo album, etc.). Students will work on their projects during the semester and will present their project on the last day of class.

Note: Students should not choose option B unless they are familiar with archival

arrangement (EAD, DACS) or able to research these standards during the semester.

Grading

Class Attendance and Participation: 25%

Assignments: 3 @ 25% each

Assignments will be graded according to rubrics documented separately. (Grading rubrics will be handed out during the first class.) Attendance for all classes is mandatory, and planned absences should be scheduled well in advance of the class to be missed.

Unplanned absences should not happen except in the case of emergency or illness.

Regular participation in class discussions is required for full credit for participation.

Readings

Selected texts available online or through NYU / LIU journal subscriptions. See class schedule.

Class Schedule and Readings

Note: The second half of the syllabus (following spring recess) may be revised considerably. Any revisions will be distributed prior to spring recess.

January 23 - Introduction and Background: Overview of Digital Library

Introductions. Discuss the goals of the class, class assignments and expectations.

There are no readings for the first day of class.

January 30 - What is a Digital Library?

Discussion of the history and development of digital libraries, and of the definition of a digital library.

Readings for 1/30:

Harter, Stephen P. "What is a Digital Library? Definitions, Content, and Issues"

<http://php.indiana.edu/~harter/korea-paper.htm>

Borgman, C.L. (1999). What are digital libraries? Competing visions. *Information Processing & Management*, 38(3), 227-243. In G. Marchionini & A. Fox (eds.), *Special Issue: Progress Toward Digital Libraries*

(Available through NYU / LIU electronic subscriptions)

Bush, V., 1945. As We May Think. *The Atlantic Monthly*, 176(1), 101-108.

<http://www.theatlantic.com/magazine/archive/1945/07/as-we-may-think/303881/3/>

Licklider, J.C.R. *Libraries of the Future*. 1965 (Excerpt)

February 6 - Omeka Introduction; The basic LAMP stack

Explore the Omeka application in the lab. Walk through the basic steps of creating a collection and loading materials and metadata. We'll review the basic landscape of "LAMP"-style applications like Omeka (and Wordpress, Drupal, etc.).

Readings for 2/6:

Read all the documents in the "Getting Started" section of the Omeka documentation. Come to class prepared with questions and / or comments on the documentation.
<http://omeka.org/codex/Documentation>

Recommended:

Visit Wikipedia to familiarize yourself with the following: The **L**inux operating system
The **A**pache Web Server; The **M**ySQL Relational Database; The **P**HP programming language. These are basic, ubiquitous web technologies that you will see again and again.

February 13 - Digitization & Digital Conversion

Learn the basics of digital conversion, focussing on imaging but also considering audio and video conversion.

*Note: Date may change / swap depending on scheduling with Digital Library Technology Services.

Readings for 2/13:

Besser, Howard. *Introduction to Imaging*. Getty Electronic Publications Online.
http://www.getty.edu/research/publications/electronic_publications/introimages/index.html

Read the Introduction and all of Part 1.

February 20 - Markup Languages, XML

Readings for 2/20:

Yott, Patrick. "Introduction to XML." *Cataloging and Classification Quarterly*. Vol. 40, no. 3/4, p. 213-235.

<http://www.columbia.edu/cu/libraries/inside/units/bibcontrol/osmc/yott.pdf>

Pitti, Daniel V. "Introduction to XML"
<http://www.iath.virginia.edu/~dvp4c/xmlintro.html>

Discussion Questions:

Why does Daniel Pitti claim that XML is "not a markup language"?

February 27 - Modeling and Describing Objects with Metadata

*****Assignment 1 is due*****

Diving further into using metadata standards to describe and model information resources.

Readings for 2/27:

McDonough, Jerome. "METS: standardized encoding for digital library objects"
<https://www.ideals.illinois.edu/bitstream/handle/2142/177/ME?sequence=2>

Discussion Questions:

Why does McDonough describe the METS standard as "open"?

What does McDonough mean when he mentions the "lack of structural constraints METS imposes on a digital object"?

March 6 - The Open Archival Information System model (OAIS)

Repositories, Digital Libraries, and Content Management Systems. Discuss the OAIS reference model.

Readings for 3/6:

"Pink Book": Chapter 2: "OAIS Concepts" (pp 2-1 through 2-12)
http://ddp.nist.gov/refs/650x0p11_OAIS_pink_book.pdf

March 13 - SPRING RECESS / NO CLASS

March 20 - Assignment 2: Student Presentations

See Assignment 2 explanation.

****Send me your slides prior to class meeting time (Powerpoint, Open Office, or Keynote format) ****

March 27 - Omeka / ArchivesSpace Workshop

We will use this class meeting to address specific questions you may have about your project for assignment 3.

April 3 - Sharing, Harvesting and Transforming Metadata

Readings for April 3:

Zeng, Marcia Lei and Lois Mai Chan. "Metadata Interoperability and Standardization – A Study of Methodology Part I"

<http://www.dlib.org/dlib/june06/chan/06chan.html>

April 10 - Discovery Systems (Federated Search, Faceted Search / Browse, OAI-PMH)

Readings for April 10:

Hearst, Marti. "The Design of Search User Interfaces."

http://searchuserinterfaces.com/book/sui_ch1_design.html

April 17 - Identification, Citation and Permanence

Readings for April 17:

Tonkin, Emma. "Persistent Identifiers: Considering the Options."

<http://www.ariadne.ac.uk/issue56/tonkin/>

Berners-Lee, Tim. "Cool URIs don't change."

<http://www.w3.org/Provider/Style/URI>

Kunze, John A. "Towards Electronic Preservation Using ARK Identifiers."

<https://confluence.ucop.edu/display/Curation/ARK> (See link in "Additional Information" box.)

Lynch, Clifford. "Identifiers and Their Role in Networked Information Applications."

<http://www.arl.org/bm~doc/identifier.pdf>

April 24 - Institutional Repositories & Scholarly Communication

Readings for April 24:

Brown, Laura; Griffiths, Rebecca; Rascoff, Matthew. "University Publishing In A Digital Age."

[http://www.ithaka.org/ithaka-s-r/strategyold/
Ithaka%20University%20Publishing%20Report.pdf](http://www.ithaka.org/ithaka-s-r/strategyold/Ithaka%20University%20Publishing%20Report.pdf)

Lynch, Clifford. "Institutional Repositories: Essential Infrastructure for Scholarship in the Digital Age"

http://scholarship.utm.edu/21/1/Lynch,_IRs.pdf

<http://www.arl.org/resources/pubs/br/br226/br226ir.shtml>

May 1 - MAKE UP DAY / NO CLASS

May 8 - Final Presentations

Present the work you have done for Assignment 3. The format of this presentation is entirely up to you.