UNIVERSITY OF MICHIGAN

University of Michigan
School of Information

SI 615
Seminar on Digital Libraries

SYLLABUS

Term:	Winter 2008
Meeting Time:	Mondays, 9:00 am to 12:00 pm
Location:	311 West Hall
Website:	https://ctools.umich.edu/portal
Credits:	3
Instructor:	Paul Conway

Description

This is a special topics seminar focusing on the current state of "digital libraries" broadly defined. The seminar is multi-disciplinary in focus and in method, covering the history of the idea, its manifestation as projects and programs in academic, non-profit, and research settings, and the suite of policy issues that influence their development and growth. The concept of the digital library will serve as an intellectual construct within which to explore the related concepts of scholarly communication, digital preservation, cyberinfrastructure, representation, and information technology standards. Given the seminar format, students will be expected to master a diverse literature, to participate actively in the discussion of issues, and to take steps, collectively and individually, to advance our understanding of future directions of digital libraries.

Objectives
- Understand the history and development of digital libraries as an international phenomenon
- Explore the literature, key players, and significant digital library programs
- Specify the critical skills required to build and maintain digital libraries
- Define a set of research questions associated with knowledge representation in digital libraries
- Establish a broad context issues and challenges facing digital libraries

Readings

Required readings average 150 to 250 pages per week, with optional reading determined by each student's interests and knowledge, as well as the relevance of a given topic to course projects and final reports. All required readings are either on the World Wide Web (WWW), accessible through the CTools site for the course, or available through University Reserves' Electronic Reserve Service. http://www.lib.umich.edu/reserves

Resources

Weekly lecture slides, additional resources for class assignments and weekly discussion topics will be posted on CTools by start of class. Students wishing to follow the lecture with the slides can download them from CTools. The CTools Portal URL is: http://ctools.umich.edu

Grading

- Class participation   25%
- Literature scan   15%
- Image representation project   25%
- Research report   35%

Academic Integrity

Academic honesty and responsibility is fundamental to our scholarly and professional community. Students are responsible for maintaining high standards of conduct while engaged in course work, research, dissertation or thesis preparation, and other activities related to academics and their profession. It is expected that students will abide by the provisions of the Rackham Graduate School Policy Statement on Academic and Professional Integrity:
http://www.rackham.umich.edu/StudentInfo/Publications/GSH/html/APPC.html#1

Students with Disabilities

Any student who feels that he/she may need an accommodation for any sort of disability, please see us during office hours or email us to make an alternative appointment.

Classroom Etiquette

Students are encouraged to bring notebook computers to class and to use them actively as learning tools. Students should:
… use laptops for taking notes, conducting research required for activities, and other specific classroom tasks as assigned by the instructor. During class, students should not check e-mail, chat, IM, play games, or perform other off-task activities.
… engage in class activity as actively as they would in any other class. The computer should not become a barrier to one-on-one interaction, but instead should help facilitate the exchange of ideas and engagement in classroom contact.
… demonstrate sensitivity to others. Students should not display screen images, including wallpapers and screen savers, that might be distracting or offensive to other members of the class.

Office Hours

Monday 12:30 pm - 2:30 pm or by appointment

Students should try to take advantage of at least one office hour session. The instructor is available and willing to advise on project topics, specialized readings, and contacts in the digital library field.

Week 1: Introduction and Overview of Themes

Required Readings:

Borgman. [2000] *Gutenberg to Global Information Infrastructure*, Chapter 2, pp. 33-52.

Levy. [2003] "Documents and Libraries, in Bishop, et al. *Digital Library Use*, Chapter 2, pp. 25-42.

Resources:

Fox, & Urs. [2002] "Digital Libraries." *Annual Review of Information Science and Technology* 36. Information Today, Inc.

Bearman. [2007] "Digital Libraries." *Annual Review of Information Science and Technology* 41. Information Today, Inc.

Week 2: History and International Perspectives

Required Readings:

Dempsey. [2006] "The (Digital) Library Environment: Ten Years After."
http://www.ariadne.ac.uk/issue46/dempsey/

Stephenson. [1996] "Mother Earth Mother Board."
http://www.wired.com/wired/archive/4.12/ffglass.html?topic=&topic_set=

Woldering. [2004] "The European Library: Integrated Access to the National Libraries of Europe."
http://www.ariadne.ac.uk/issue38/woldering/

Resources:

Fox. [1993] Digital Library Sourcebook. http://fox.cs.vt.edu/DLSB.html

Digital Library Federation. http://www.diglib.org/

ERPANET: European Resource Preservation and Access Network.
http://www.erpanet.org/index.php

JISC: Joint Information Systems Committee. http://www.jisc.ac.uk/

PADI: Preserving Access to Digital Information. http://www.nla.gov.au/padi/

Week 3: Martin Luther Kind Day – No Class

Week 4: Content Landscapes: How Much Information

Required Readings:

Lyman and Varian. [2003] *How much information? 2003.*
http://www2.sims.berkeley.edu/research/projects/how-much-info-2003/

OCLC. [2004] "2004 Information Format Trends: Content, Not Containers,"
http://www.oclc.org/reports/2004format.htm

O'Toole. [1994] "On the Idea of Uniqueness."

Powell. [2003] "Mapping the JISC IE [Information Environment] Service Landscape."
http://www.ariadne.ac.uk/issue36/powell/intro-alternative.html

Resources:

Digital Library Federation. Registry of Digital Masters. http://www.diglib.org/collections/reg/reg.htm

Digital Library Federation. DLF Aquifer. http://www.diglib.org/aquifer/

JISC Collections: delivering resources online. http://www.jisc-collections.ac.uk/

Week 5: Research Agendas

Required Readings:

Anderson and Choudhury. [2003] "A Usability Research Agenda for Digital Libraries."
http://dkc.jhu.edu/usability_1.html

Hedstrom. [2003] *It's About Time: Research Challenges in Digital Archiving and Long-Term Preservation.* [Attached]

Hedstrom. [2002] "The Digital Preservation Research Agenda."
http://www.clir.org/pubs/abstract/pub107abst.html

Levy. [1998] "Heroic Measures." http://www.acm.org/pubs/citations/proceedings/dl/276675/p152-levy/

Lynch and Garcia-Molina. [1995] "Interoperability, Scaling, and the Digital Libraries Research Agenda." http://www-diglib.stanford.edu/diglib/pub/reports/iita-dlw/main.html

Waters and Garrett. [1996] *Preserving Digital Information.* http://www.rlg.org/ArchTF/

Resources:

UM Collaboratory Working Groups. http://www.si.umich.edu/UMDL/EU_Grant/home.htm

EU-NSF Working Group on Metadata. Metadata for Digital Libraries: a Research Agenda.
http://www.ercim.org/publication/ws-proceedings/EU-NSF/metadata.html

National Science Foundation. Digital Libraries Initiative, Phases One and II. http://www.dli2.nsf.gov/
http://www.dli2.nsf.gov/dlione/

Week 6: Interface and Metadata Infrastructure

Required Readings:

Borgman. [2000] "Making Digital Libraries Easier to Use," pp. 143-163.

Campbell. [2000] "The Case for Creating a Scholars Portal to the Web."
http://www.arl.org/newsltr/211/portal.html

Caplan and Guenther. [2005] "Practical Preservation: the PREMIS Experience."
http://www.loc.gov/standards/premis/caplan_guenther-librarytrends.pdf

Dahlquist, et al. [2005] "Integrating Digital Libraries and Electronic Publishing in the DART Project."
http://doi.acm.org/10.1145/1065385.1065411

Lavoie. [2004] *Open Archival Information System: Introductory Guide.*
http://www.dpconline.org/graphics/reports/index.html

Lynch. [2000] "Authenticity and Integrity in the Digital Environment."
http://www.clir.org/pubs/reports/pub92/lynch.html

McDonough [2006] "METS"

Resources:

National Digital Information Infrastructure Preservation Program.
http://www.digitalpreservation.gov/index.html

Flecker. [2000] "Harvard's Library Digital Initiative: Building a First Generation Digital Library Infrastructure." *D-Lib Magazine* 6 (November 2000).
http://www.dlib.org/dlib/november00/flecker/11flecker.html

Higgins and Semple. [2006] *OAIS Five-year Review*. http://www.dcc.ac.uk/events/oais-fyr-2006/.

PREMIS: Preservation Metadata Maintenance Activity. http://www.loc.gov/standards/premis/

METS: Metadata Encoding and Transmission Standard. http://www.loc.gov/standards/mets/

Week 7: Representation of Image/ Text in Visual Digital Libraries

Required Readings:

Barney et al. [2005] "Ordering Chaos" doi:10.1093/llc/fqi002 [access through Mirlyn]

Bolter & Grusin. [1996] "Remediation."
http://muse.jhu.edu/journals/configurations/v004/4.3bolter.html

Sassoon. [1998] "Photographic Meaning in the Age of Digital Reproduction."
http://www.sl.nsw.gov.au/lasie/prepdf.htm

Smith. [2001] *Strategies for Building Digitized Collections*.
http://www.clir.org/pubs/abstract/pub101abst.html

Summit on Digital Tools for the Humanities. [2003] http://www.iath.virginia.edu/dtsummit/

Resources:

National Archives and Records Administration. [2004] *NARA Guidelines for Digitizing Archival Materials for Electronic Access*. http://www.archives.gov/research/arc/digitizing-archival-materials.html

Handout on digital imaging guidelines [on CTools]

SPRING BREAK

Week 8: Cyberinfrastructure

Required Readings:

American Council of Learned Societies. [2006] *Our Cultural Commonwealth*. http://www.acls.org/cyberinfrastructure/cyber.htm

Arms et al. [2006] "A Research Library Based on the Historical Collections of the Internet Archive." http://www.dlib.org/dlib/february06/arms/02arms.html

Association of Research Libraries. [2006] "Long-term Stewardship of Digital Data Sets in Science and Engineering." http://www.arl.org/info/frn/other/ottoc.html

Berman and Brady [2005] "Final Report: NSF SBE-CISE Workshop on Cyberinfrastructure and the Social Sciences," http://vis.sdsc.edu/sbe/

Messerschmit. [2003] "Opportunities for Research Libraries in the NSF Cyberinfrastructure Program," http://www.arl.org/newsltr/229/cyber.html

National Science Foundation. [2003] *Revolutionizing Science and Engineering Through Cyberinfrastructure.* http://www.nsf.gov/od/oci/reports/toc.jsp

NSF Cyberinfrastructure Council. [2006] *NSF's Cyberinfrastructure Vision for 21st Century Discovery.* http://www.nsf.gov/od/oci/ci-v7.pdf

Resources:

National Science Foundation. Cyberinfrastructure Office. *Reports and Workshops Relating to Cyberinfrastructure and Its Impacts.* http://www.nsf.gov/od/oci/reports.jsp

Data Document Initiative. http://www.ddialliance.org/index.html

Economic and Social Data Services. [2005] *Guide to Good Practice: Data Management.* http://www.esds.ac.uk/support/guides/A4.pdf

Lord and Macdonald. [2003] *e-Science Curation Report*. http://www.jisc.ac.uk/uploaded_documents/e-ScienceReportFinal.pdf

Week 9: Scholarly Communication and Digital Library Use

Required Readings:

Borgman. [2003] "Designing Digital Libraries for Usability," pp. 117-168.

Campbell. "The Case for Creating a Scholars Portal to the Web."
http://www.arl.org/newsltr/211/portal.html

Farb. [2006] "Libraries, licensing and the challenge of stewardship."
http://www.firstmonday.org/issues/issue11_7/farb/index.html

Levy. [1998] "Heroic Measures." http://www.acm.org/pubs/citations/proceedings/dl/276675/p152-levy/

Rusbridge. [2005] "Information Life Cycle and Curation." Powerpoint. [Attached]

Waters. [2006] "Managing Digital Assets in Higher Education."
http://www.arl.org/newsltr/244/assets.html

Resources:

First Monday. [2006] Openness: Code, Science, and Content. Website.
http://www.firstmonday.org/issues/issue11_7/index.html

Association of Research Libraries. New Models for Scholarly Communication. Website.
http://www.arl.org/osc/models/index.html

Create Change. Website. http://www.createchange.org/

Week 10: Institutional Repositories and Preservation

Required Readings:

Crow. [2002] "The Case for Institutional Repositories: A SPARC Position Paper."
http://www.arl.org/sparc/IR/ir.html.

Green and Gutmann. [2006] "Building Partnerships Among Social Science Researchers, Institution-based Repositories and Domain Specific Data Archives." [preprint]

Lagoze, et al. [2006] "Fedora"

Lavoie and Dempsey. [2004] "Thirteen Ways of Looking at…Digital Preservation."
http://www.dlib.org/dlib/july04/lavoie/07lavoie.html

Lynch and Lippincott. [2005] "Institutional Repository Deployment in the United States as of Early 2005." http://www.dlib.org/dlib/september05/lynch/09lynch.html

Keller, Reich, and Herkovic. [2003] "What is a library anymore, anyway?"
http://www.firstmonday.org/issues/issue8_5/keller/index.html

Rusbridge. [2006] "Excuse Me… Some Digital Preservation Fallacies?"
http://www.ariadne.ac.uk/issue46/rusbridge/

Rusbridge et al. [2005] "The Digital Curation Centre: A Vision for Digital Curation." [Attached]

Resources:

Lee, et al. [2002] "The State of the Art and Practice in Digital Preservation."
http://nvl.nist.gov/pub/nistpubs/jres/107/1/cnt107-1.htm

Lynch. [2003] "Institutional Repositories: Essential Infrastructure for Scholarship in the Digital Age."
http://www.arl.org/newsltr/226/ir.html .

NEDCC. [2007] "Digital Preservation Readiness Webliography."
http://www.nedcc.org/resources/leaflets/6Reformatting/08DigitalPreservationReadiness.php

Research Libraries Group. Digital Repository Certification. Web Page.
http://www.rlg.org/en/page.php?Page_ID=580

Research Libraries Group and National Archives and Records Administration. "Audit Checklist for Certifying Digital Repositories." http://www.rlg.org/en/page.php?Page_ID=20769

Week 11: Case Study of UM Deep Blue and Blue Stream

Required Readings:

Readings list distributed during week 8

Week 12: Sustaining Digital Libraries

Required Readings:

Digital Curation Center. http://www.dcc.ac.uk/index

Lynch. [2005] "Where Do We Go from Here? The Next Decade in Digital Libraries."
http://www.dlib.org/dlib/july05/lynch/07lynch.html

Lagoze. [2005] "What Is a Digital Library Anymore, Anyway? Beyond Search and Access in the NSDL." http://www.dlib.org/dlib/november05/lagoze/11lagoze.html

Research Libraries Group. [2002] *Trusted Digital Repositories*.
http://www.rlg.org/legacy/longterm/repositories.pdf

Wheatley. [2004] *Institutional Repositories in the Context of Digital Preservation.*
http://www.dpconline.org/graphics/reports/index.html

Resources:

TRAC: Trustworthy Repositories Audit & Certification: Criteria and Checklist.
http://www.crl.edu/content.asp?l1=13&l2=58&l3=162&l4=91

DRAMBORA: Digital Repository Audit Method Based on Risk Assessment.
http://www.repositoryaudit.eu/

NESTOR: Network of Expertise in Digital Long-term Preservation (Germany)
http://www.langzeitarchivierung.de/index.php?newlang=eng

ASSIGNMENTS

Class Participation (25%)

The overall success of the seminar depends on the active participation of all members of the class. Class participation is a sizable portion of the grade. Students should attend all classes and be prepared to enter into class discussions and to raise questions reflecting their reading and interests. Students are also expected to complete all required readings in advance of class. This is especially important since a portion of the class sessions will be discussions about the readings.

On the first class, each student will sign up with two other students to lead the discussion on the required readings for a given week. This is an opportunity to read one set of assigned readings more deeply than average, to work collaboratively on identifying themes or critical ideas, and to demonstrate understanding of one component of the course.

Literature Scan (15%)

The purpose of the individual project is to explore and find some interesting published literature on digital libraries and report on experiencing three important research resources.
- Use the search and browse tools of three resource collections on a particular topic of your choosing.
- Identify ten items of relevance to the topic (at least two from each source) from ACM Digital library. *ACM Digital Library* http://portal.acm.org/ [access via UM Library Search Tools]; *D-Lib Magazine* http://www.dlib.org/; and *Ariadne Magazine* http://www.ariadne.ac.uk/.
- Prepare full and consistent bibliographic citations for all ten items (using MLA, Chicago, or other appropriate format).
- Write between one- (1) and five- (5) sentence description/summary of each source.
- Write up to 1,000 words that states the topic of your search, describes the scope of the resources available on digital libraries, assesses their strength and weaknesses for digital library research, and reports on your experience using the resources,.

Image Representation Database (25%)

The purpose of this project is to explore the various ways that photographic images are represented through digital library interfaces and to become adept at recognizing such variation. A secondary purpose of the project is to begin compiling a reference and research collection of digitized images that illustrates representation variation systematically. The mechanism we will use to complete the project is the collaborative construction of a database of image elements following a fairly complex scheme. The scheme captures descriptive information about each image, codes each image for its representation style, annotates the image in ways that connect the visual product with its associated metadata, and codes the image for the "significant properties" that it displays (or contains).

Each student will be responsible for identifying and coding data in the database for a minimum of 100 images. Half of the images will be chosen from collections in one of ten possible large scale digital libraries [a list will be distributed in class]. Half of the images may be chosen from any online collection of digitized photographs anywhere in the world, as long as certain criteria are met relating to the scale of the undertaking and the organizational source of the images. Prior to the start of the project, the instructor will provide definitions of key terms, illustrative examples of various representation styles, and detailed coding instructions for building the database. The results of the exercise will be presented in class.

Grading will be based on quantitative and qualitative assessments. Quantitative scores will be assigned based on the number of images completed and the completeness of the data compiled. A qualitative score will be assigned based on the level of judgment exercised in coding representation style and other subjective elements of the combined database.

Final Report (35%)

The final report for the seminar is an opportunity to take a close look at a particular aspect of a digital library program or service and to describe that aspect in the larger context of published and informally posted information on the digital library. The paper will be between 3,000 and 5,000 words in length, will be properly documented with references and a bibliography, and will conform to a common structure. The point of departure for the paper is a close examination of one feature of a digital library from a group of large and diverse digital libraries. A list will be distributed in class after the second week. Each student will choose one digital library to examine. All of the students who choose a particular digital library will constitute one or more teams varying in size from 3 to 6 individuals. You are encouraged to work together as a team on your analysis. The team is required to prepare a single document listing published and unpublished documentation on the digital library.

Students will negotiate within their respective teams over the choice of a particular topic to investigate in greater detail. Topics relevant to this investigation include the following areas:
- Content domains, accession or selection criteria
- Digitization guidelines or collection development focus
- Metadata standards, processes, and products
- End user interface design, usability assessment, revision history
- End-user productivity tool development, including search tools, display tools, edit tools, etc.
- Underlying repository infrastructure, including software architecture, database design, etc.
- Collaboration with other institutions at the level of process, practice, standards, data sharing, etc.
- Repository/archive development, including perspectives on digital preservation versus digital access
- Staffing, cost analysis, benefit analysis, and other economic/budgetary issues.

The following issue is out of scope for this review: Copyright and intellectual property

Depending upon the size and individual interests of the group, the final report may parse the project in any number of ways. The final report must include at least the following five components:
- Executive summary with critical findings (1 to 2 paragraphs)
- Description of the origins, history, and purposes of the digital library (may be done collaboratively by the team for all reports from that team)
- Report on the issue investigation, with the following components
    - Definition of the issue
    - Assessment of state of the art for the issue
    - How the issue is represented in the digital library
    - Reconciling state of the art with reality
    - Trends and directions for the issue either in the digital library or more broadly
    - Challenges you faced in your investigation (e.g., lack of documentation, outdated interface, etc.)

In pursuing your research, you may use any published or posted information available on the digital library and you may interview or correspond with appropriate staff who may have information to share. I encourage this strategy.

Week 1 (January 7): Introduction and Overview of Themes
In this session we will explore perspectives on the idea of a library, digital or otherwise, and collaborate to establish a shared definition of a digital library. The session will review the structure of the courses and outline its principal underlying themes. Finally, this session cover the major sources of information on digital libraries.

Week 2 (January 14): History and International Perspectives
Digital libraries have a distinctive history, a consideration of which sheds light on the predispositions of those who design and build them. This session will contrast design/build perspectives on an international basis with what we do and do not know about the people who use them.

Week 3 (January 21): Martin Luther King Day – No Class

Week 4 (January 28): Content Landscapes: How Much Information
In spite of the fact that the manufacture of paper continues unabated, we are living in a digital world. This session will explore alternative scenarios for describing the landscape of digital content, using as a point of departure a seminal study that attempts to measure the volume of information.

Week 5 (February 4): Research Agendas
Sometimes it's useful to make lists. Research agendas for digital libraries, digital preservation capabilities, and the management of electronic records provide insight into the shifting perspectives on needs and requirements. This session will explore at least a decade's worth of research agendas against our emerging understanding of what digital libraries are.

Week 6 (February 11): Interface and Metadata Infrastructure
What you see isn't necessarily what you get. This session will explore the distinction between the user interface and the underlying structure of information used to describe, structure, and administer a digital archive or repository. We will work with several XML models of digital library metadata infrastructure and explore the elements of interface design that enable or constrain use.

Week 7 (February 18): Representation of Image/Text and Visual Digital Libraries
This session will look closely at several models of how humanities text and image resources are represented in digital form. A project on diagnosing and detecting image representation will be introduced.

SPRING BREAK (February 25)

Week 8 (March 3): Cyberinfrastructure
This session will introduce the concept of "cyberinfrastructure" and explore the technical and policy issues associated with science, social science, and humanities data and users.

Week 9 (March 10): Scholarly Communication and Digital Library Use
Models of scholarly communication (which vary somewhat by discipline) provide a framework for exploring access and use issues in digital libraries. This and next week's sessions together will force us to confront the ongoing tensions between the idealization of digital libraries and the barriers presented by the technical, organizational, and policy constraints.

Week 10 (March 17): Institutional Repositories and Preservation
This session will look closely at emerging models of institutional repositories, their relationship to emerging preservation requirements, including data modeling and trust systems.

Week 11 (March 24): Case Study of UM Deep Blue and Blue Stream
This session takes a close look at two technologies at the University of Michigan that support the university's digital repository efforts, including the ingest and processing of digital multi-media resources. The case study explores the extent to which multiple technologies are and are not yet integrated from the perspective of infrastructure and interface.

Week 12 (March 31): Sustaining Digital Libraries
Sustainability could turn out to be one of the principal challenges of digital libraries (an open question). This final session will explore a variety of incentive mechanisms associated with the design of digital libraries and their use.

Week 13 (April 7): Wrap-up and Research Paper Reports
This and the final session will provide opportunities for synthesis across seminar topics, taking as a point of departure the issues explored in student research papers. Each project will be allotted be allocated roughly five minutes for a succinct presentation on the core concept and key findings of the project. The quality of the presentation, measured in terms of clarity of communication, will be count for roughly 25% of the project grade.

Week 14 (April 14): Wrap-up and Research Paper Reports

DEADLINES and DUE DATES

January 7		First Class
January 21		No Class Today
January 28	5:00 pm	Literature Scan Due
February 4	5:00 pm	Individual DL Review Topics Due
February 11		Form Digital Library Review Teams
February 22	5:00 pm	Digital Library Review Bibliography Due
February 28		Spring Break
March 14	5:00 pm	Working outline of final report Due, if advice desired from instructor
March 21	5:00 pm	Image Representation Database Entries Complete
April 7		Final Report Presentations (clustered by investigation topic)
April 14		Final Report Presentations (clustered by investigation topic)
April 18	5:00 pm	Final Reports Due