LIS 563 – Digital Libraries  
Spring 2009 Dr. Jianqiang Wang

General Course Information

Instructional Staff

Dr. Jianqiang Wang  
536 Baldy Hall  
University at Buffalo  
Email: jw254@buffalo.edu  
Tel: 716-645-2412 ext. 1167

Schedule

Lecture: 3:00pm - 5:40pm, Thursday, 117 Baldy.  
Office hours: 2:00pm - 3:00pm, Thursday, and by appointment.

Official Catalog Description

Digital libraries are defined as electronic libraries in which geographically scattered users access diverse repositories of electronic objects, including scientific, business, and government datasets, networked text, images, maps, sounds, videos, merchandise catalogs, hypertext, hypermedia, and multimedia compositions. This course will cover professional issues related to digital libraries, technical aspects, tools for creation and use of digital products and library and public policy.

Course Goals

The goals of LIS 563 are to present a theoretical and practical introduction to:

- The purposes, functions, and users of digital libraries
- Tools for creation, maintenance, and implementation of digital libraries
- Examination and research of contemporary issues and directions for digital libraries
- Processes of creating, populating, implementing, and maintaining a digital library.
- Project management and team work process for digital libraries projects.
Recommended Books and Readings

There is no required textbook for the course. We will use reading materials from a variety of sources. Some of them are available online whereas others online in print. However, since many parts of the following book will be used, I suggest you possess a copy of it if possible:


In addition, we will use some sections of the following two books:


Additional readings from other books, journals, and conference proceedings may be required or recommended. See “Weekly Readings” document for detail.

Prerequisites

LIS506 or instructor’s permission.

Grading

Course grades will be assigned based on homework assignments, questions about the topics covered in each week, team project, and class participation. Scores for each of these course requirements will be assigned on a 100 point scale. The final grade for each student will be computed by combining the score of each component as follows:

<table>
<thead>
<tr>
<th>Component</th>
<th>Percentage</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Homework Assignments</td>
<td>30%</td>
<td>10% each</td>
</tr>
<tr>
<td>Questions, Answers, &amp; Discussions</td>
<td>30%</td>
<td>Each week from Week 2 - 14</td>
</tr>
<tr>
<td>Team Project</td>
<td>40%</td>
<td>Project outcome, report, and presentation</td>
</tr>
</tbody>
</table>

The conversion from a score grade to a letter grade will follow the rule defined by the Graduate School at UB:

<table>
<thead>
<tr>
<th>95-100: A</th>
<th>90-94: A-</th>
<th>87-89: B+</th>
<th>84-86: B</th>
<th>80-83: B-</th>
</tr>
</thead>
<tbody>
<tr>
<td>77-79: C+</td>
<td>70-76: C</td>
<td>60-69: D</td>
<td>0-59: F</td>
<td></td>
</tr>
</tbody>
</table>
Since there are multiple evaluation items and they will be graded with letter grades, they have to be converted into scores on a 100-point scale. The follow schema will be used for such conversion:

<table>
<thead>
<tr>
<th>Grade</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>A+</td>
<td>100</td>
</tr>
<tr>
<td>A</td>
<td>97.5</td>
</tr>
<tr>
<td>A-</td>
<td>92.5</td>
</tr>
<tr>
<td>B+</td>
<td>88.5</td>
</tr>
<tr>
<td>B</td>
<td>85.5</td>
</tr>
<tr>
<td>B-</td>
<td>82</td>
</tr>
<tr>
<td>C+</td>
<td>78.5</td>
</tr>
<tr>
<td>C</td>
<td>73</td>
</tr>
<tr>
<td>D</td>
<td>65</td>
</tr>
<tr>
<td>Others</td>
<td>0</td>
</tr>
</tbody>
</table>

**Homework Assignments**

There will be three homework assignments throughout the semester. The main purpose of the homework assignments is for students to further consolidate what they have learned from the class and the readings and to deepen their understanding of a couple of main digital libraries topics that each student chooses. Each homework assignment deserves 10% of the final grade. Homework assignments will be submitted electronically through the Digital Dropbox in the course’s blackboard (i.e., UBLearns). Detailed information about each homework assignment can be found in the “Homework Assignments” document.

**Team Project**

The term project is designed for students to integrate and extend knowledge acquired throughout the course and to apply that knowledge to solve a problem of substantial scope. Students are required to work in groups of 6 or 7 people. Experience suggests that successful teams require expertise in design, implementation, and project management.

Your task is to propose, plan and develop a prototype digital collection, using Open Source software (e.g., Greenstone, DSpace). Your digital library should address the need of a group of real users, and it should include at least three collections and at least one other media format in addition to text (minimally 20 documents in each collection). Detailed descriptions of the requirements and evaluation of the project can be found in “Team Project” document.

**Weekly Questions, Answers, & Discussions**

Starting from the second week, each student is required to submit at least one questions related to the subjects covered in that week’s required or optional readings. Questions should be posted in the thread of “Weekly Questions & Discussions” on the UBLearns’ Discussion Board by 10:00pm, Tuesday so that people will have enough time to read and respond to them before the class. Also, each student should provide answers to or discussions about at least one question submitted by others. Your answers or discussions should be posted following the questions(s) on the Discussion Board by 12:00pm, Thursday of each week. You should answer or discuss a question or questions that have not been responded by others, unless you can provide a different response. Please focus your questions on the content of the readings. Questions regarding grammatical errors or typos in the readings are far less interesting. In each class, we will spend about half an
hour discussing the questions and answers. Therefore, this evaluation item also includes class participation.

**Administrative rules**

**Homework assignment format** Unless otherwise specified, your homework assignment must be submitted electronically through UB Learns. Homework assignments in handwritten form will not be accepted. Pay close attention to specific requirements for each homework assignment.

**Late submission of assignments** Due to extreme causes (such as illness and traffic accidents) can be accommodated if supporting documents (doctor appointment slips, police records of car accidents, etc.) are provided to the instructor. In this case, no grade reduction will be applied. Late submission due to other less extreme causes (such as breakdown of your home computer, busy work schedule, etc.) will be accepted only if the submission is no later than one week. In this case, a 10% reduction of grade per day will be applied. Late assignments due to extreme causes must be submitted before the last day of class for consideration of grading.

**Incomplete grades** may be granted in cases of illness or other difficult circumstances. Incompletes must be requested in writing by the student by filling out and submitting to the professor a "Request for Grade of Incomplete" obtained from the department office. An “incomplete grade” request is subject to approval by either the instructor or the Admissions & Academic Standards Committee.

**Academic integrity** It is expected that you will behave in an honorable and respectful way as you learn and share ideas. Therefore, *recycled papers, work submitted to other courses, and major assistance in preparation of assignments without identifying and acknowledging such assistance* are not acceptable. All work for this class must be original for this class. Be forewarned: faculty members do talk to each other and as information specialists we can spot plagiarism and track it down. Please be familiar with the University and the School policies regarding plagiarism. Read the [Academic Integrity Policy and Procedure](#) for more information.

**Special accommodation** Any student with a disability that will require accommodation under the terms of federal regulations must present an accommodation request to the Office of Disability Services (25 Capen Hall). Unless a letter from that office is received by the instructor, no special assistant or treatment will be provided by the instructor.
# Weekly Topics, Assignments & Guest Speakers

<table>
<thead>
<tr>
<th>Week</th>
<th>Date</th>
<th>Topics</th>
<th>Assignment due before class</th>
<th>Guest speaker</th>
</tr>
</thead>
</table>
| 1    | 1/15 | Course Overview  
Introduction to DL | | |
| 2    | 1/22 | Methodologies & Tools | Stacy Person | |
| 3    | 1/29 | Models & Architectures | Project Proposal | |
| 4    | 2/5  | Representation of Digital Objects | | |
| 5    | 2/12 | Metadata | HW1 | |
| 6    | 2/19 | XML and Markups | | |
| 7    | 2/26 | Digital Preservation | TBD | |
| 8    | 3/5  | Access in DL (1) | | |
| 9    | 3/12 | Access in DL (2) | HW2 | |
| 10   | 3/19 | No class (spring recess) | | |
| 11   | 3/26 | Interaction & Evaluation | Interim Report | Valerie Nesset |
| 12   | 4/2  | Security & Economics | | |
| 13   | 4/9  | Social Issues | | |
| 14   | 4/16 | Legal Issues & Future of DL | HW3 | |
| 15   | 4/23 | Project Presentation | Final Report | |
Team Project

Introduction

The team project is designed for students to integrate and extend knowledge acquired throughout the course and to apply that knowledge to solve a problem of substantial scope. Students are required to work in groups of 6 or 7 people. Experience suggests that successful teams require expertise in design, implementation, and project management.

Your task is to propose, plan and develop a prototype digital collection, using Open Source software (e.g., Greenstone, DSpace). Your digital library should address the need of a group of real users, and it should include at least three collections and at least one other media format in addition to text (minimally 20 documents in each collection).

Digital libraries are complex in nature and vary in content, supporting technological infrastructure, organization mechanism, presentation style, and the user group to serve. However, generally each digital library project contains the following key processes of development activities, for each of which a team may want to assign one or two persons leading the effort:

- System requirements analysis / user liaison
- Copyright and content development
- Digitization
- Organization and metadata
- Interface design
- Technological infrastructure development

Requirements for the Proposal

A project proposal from each team will be turned in on Jan 29 in Week 3. The proposal should include the names of the team member and a title of the project in the beginning section. In addition, your proposal should describe briefly the following elements:

- What content will be contained in the digital library that your team will build?
- Who will be the potential users of it?
- Why is it important to develop such a digital library?
- What systems/software will likely be used? (could be several alternatives from that you will further make some choices)?
- How do you plan to complete the project? That is, a sketchy project plan including team member responsibilities and a time-table.

The main purpose of the proposal is for the instructor to assess whether the suggested project by each team is feasible. It also prompts students to think about the project as
early as the beginning of the semester, because a satisfying project does take the whole semester to complete.

**Requirements for the Interim Report**

An interim project report is due on March 26 in Week 11. The interim report is to assure that each team is making the appropriate progress. Therefore, the interim report should explain how much has been done to, for example, each of the six major development activities. Of course, different team may plan to complete the project in a different way. For that reason, you should report your project progress with an eye on your plan that is included in your proposal. In addition to what you have done with your project, you should also report any major problems you’ve encountered and how you solved or plan to solve them.

**Requirements for the Final Report**

Along with a working digital library system on April 23 in Week 15, each team will submit a final project report. Your project report should include the following elements:

- **Problem statement**: what problems does your digital library intend to solve and why is it important?
- **Overview and scope**: what information content is included and provided by your digital library? Is there any similar existing digital library? If so, what’s the major contribution of your project? Or in other words, what differs your digital libraries from others?
- **User requirements**: what are the users? How do you find out their requirements for the system that you developed? If you have used interviews or surveys, this part should include a summary of your interview/survey. Your survey/interview protocol should better be included in an appendix.
- **Content development and copyright consideration**: what collections of information objects are included in and provided by your digital library? How did you come up with them? Included here will be document converting activities such as digitization, audio/video editing and formatting; what technologies/software did you use and what’s your rationale of choosing them?
- **Conceptual design**: independent on the actual technology or software that you used, your conceptual design of the digital library describes what data elements and information objects and included and how they interact with or are related to each other. It also shows how the system works, usually with diagrammatic illustrations.
- **Schemas for representing and organizing digital objects**: this is indeed part of your conceptual design. Since they are among the most important things of digital libraries, they are emphasized specifically here. This will include a detailed description of the metadata schemas that you used for representing the information objects and your rationale of choosing/developing and using such schemas.
• Interface and interaction: you should describe clearly the features of the interfaces and dialogues through which users communicate with your digital library. Tools like dialogue diagrams may be helpful.
• Technological consideration and development/adoption: digital libraries are featured with computer and networking technology. Therefore, you should explain clearly what technology you planned and actually developed and/or adopted for your digital library. This may include, but not limit to, scanning and optical character recognition, image processing, audio/video editing, XML, web publishing, indexing, database, web server, as well as programming.
• Evaluation: how did you do to make sure your digital library serves what it is supposed to? What methods and measures did you use to assess it? What did you find and conclude? What works and what can be further improved?
• Maintenance and future of the digital library: you should describe briefly whether your digital library can be a true valuable resource for its intended users? Why and why not?

Requirements for the Presentation/Demonstration

On April 23 in Week 15, each team will give a 20-minute presentation and demonstration of its project. Due to the time constraint, you should focus on the live demo of your digital library at least as much as your development philosophy. The instructor or technical staff can help you set up computer and network for your demo, but you’ll need to initiate it early enough.

Technological/Web Infrastructure for the Project

The scope and size of the team project will probably prohibit a team from hosting its digital library in any individual students’ UB web space. The GSE technical department will provide web space and hosting for student projects of the course. In fact, Greenstone, a popular digital library software package, has been installed and set up in the GSE web server. If a team chooses to use it, the GSE staff can help create user accounts. Of course, if a team chooses some other digital library software or plans to develop its own, the GSE staff can help with the setup too.

Milestones for the project

Introduction of team project: January 15&22
Project team & proposal: January 29
Interim project report: March 26
Final project report: April 23
Project presentation: April 23

Format Requirements

Please follow one of the established styles for reference and citation when writing your project proposal, interim report, and final report (visit “Research, Writing, and Style
Guides” (http://www.aresearchguide.com/styleguides.html) for various existing styles). However, you are highly recommended to adopt the American Psychological Association APA style (the fifth edition of the *Publication Manual of the American Psychological Association* published by the American Psychological Association (2001)). “A Guide for Writing Research Papers” (http://webster.commnet.edu/apa/) is a wonderful online place to obtain the guidance for this style. For your project presentation/demo, use of PowerPoint is encouraged but not required.
Reading List

Week 1 - Introduction

Objectives: After this class, you should be able to
- restate different views and definitions of digital libraries
- identify causes for different communities working on digital libraries
- develop or adopt your own definition of digital libraries, and verify some self-defined digital libraries based on your definition.

Required Readings

Background Readings

Week 2 - Methodologies and Tools

Objectives: after this class, you should be able to:
- understand the importance and key phases in developing digital library systems
- understand the deliverables of each phase
- gain some basic understanding of Greenstone software package – how to install it and what major functions it provides
Required readings

   http://www.emeraldinsight.com/Insight/ViewContentServlet?Filename=Published/EmeraldFullTextArticle/Articles/1640230105.html

   http://www.emeraldinsight.com/10.1108/10650750410564655


Background readings


---

Week 3 - Models and Architectures

Objectives: After this class, you should be able to

- explain major components of digital libraries
- examine ideas behind various DL architecture
- explain major concepts of computer and networking technologies relevant to digital libraries

Required Readings:

   http://www.dlib.org/dlib/december01/suleman/12suleman.html

   http://www.dlib.org/dlib/february97/cnri/02arms1.html

   http://www.dlib.org/dlib/may99/payette/05payette.html. (pay attention to the discussion of interoperability and extensibility)

4. ARMS, Chapter 2, or “An Overview of World Wide Web,”  
   http://www.cs.cornell.edu/wya/DigLib/MS1999/Chapter2.html

Background Readings:

---

**Week 4 - Representation of Digital Objects**

Objectives: After this class, you should be able to
- outline common approaches for selecting materials for digitalization
- criticize various approaches for digitalizing, and representing digital objects
- evaluate practical considerations adopted for digitalization and representation in existing digital libraries

Required Readings:
1. LESK sections 2.1, 2.2, 2.7, and Chapter 3.

Background Readings:

---

**Week 5 - Metadata**

Objectives: After this class, you should be able to
- outline the importance of metadata in representing and organizing digital objects into collections
- create basic elements of Dublin Core for digital objects

Required Readings:
1. WITTEN. Section 2.2 and Section 5.4-5.7 (available in UBLearns course space.)

---

**Week 6 - XML and Markups**

Objectives: After this class, you should be able to
- explain the functionalities of the major components of XML suite
- create simple XML document for describing objects

Required Readings:
1. Martin Bryan. Introducing the Extensible Markup Language (XML)  
[http://burks.bton.ac.uk/burks/internet/web/xmlintro.htm](http://burks.bton.ac.uk/burks/internet/web/xmlintro.htm)
3. Extending you Markup: a XML tutorial by Andre Bergholz  
4. XML Schema Tutorial  
[http://www.w3schools.com/Schema/default.asp](http://www.w3schools.com/Schema/default.asp)

---

**Week 7 - Digital Preservation**

Objectives: After this class, you should be able to
- restate the new challenges in digital preservation
- explain the basic strategies for digital preservation.
- explain why OAIS reference model is needed, and how OAIS model can be used for preservation
- restate the purposes and functions of preservation metadata.

Required Readings:
1. Margaret Hedstrom “Research Challenges in Digital Archiving and Long-term Preservation”  
[http://www.dpconline.org/docs/lavoie_OAIS.pdf](http://www.dpconline.org/docs/lavoie_OAIS.pdf)
[http://www.dlib.org/dlib/july07/littman/07littman.html](http://www.dlib.org/dlib/july07/littman/07littman.html)
Background Readings:

Week 8 - Access in Digital Libraries (1)

Objectives: After this class, you should be able to
  • evaluate various retrieval methods developed in digital libraries, especially those ones that do not exist in traditional libraries

Required Readings:
1. LESK Chapter 4.

Background Readings:
Week 9 - Access in Digital Libraries (2)

Objectives: After this class, you should be able to
- examine the various approaches for meta-search and federate search in DL

Required Readings:
1. The Open Archive Initiative Protocol for Metadata Harvesting (read “Introduction” and “Definitions and Concepts”).
   http://www.openarchives.org/OAI/openarchivesprotocol.html
   http://www.libraryjournal.com/article/CA406012.html&
   http://www.infotoday.com/it/oct03/hane1.shtml
   http://www.dlib.org/dlib/april97/04lynch.html

Week 11 - Interaction and Evaluation

Objectives: After this class, you should be able to
- explain the designing process for the interaction between human and computers, especially digital libraries
- evaluate the usability design of existing digital libraries interfaces by considering the learned principles

Required Readings:
2. Rob Kling and Margaret Elliott "Digital Library Design for Usability"
   http://www.csdl.tamu.edu/DL94/paper/kling.html
4. Ben Shneiderman, Catherine Plaisant, "Designing the user interfaces" 4th ed. chapter 1. A good introduction about usability and its application in human computer interaction (available in UBLearns course space)

Background Readings:
6. Ben Sheiderman, Catherine Plaisant, "Designing the user interfaces" 4ed. chapter 14. An introduction about search interface and information visualization. (available in UBLearns course space)
7. Ben Sheiderman, Catherine Plaisant, "Designing the user interfaces" 4ed. chapter 2. An excellent introduction about some basic theories, guidelines and principles of designing a good interface (available in UBLearns course space)

Week 12 - Security and Economics

Objectives: After this class, you should be able to
- explain the elements of access management, associated issues and related technologies
- employ appropriate technologies for ensuring the security of existing digital libraries
- explain the models that people have developed for making digital libraries sustainable
- restate the long term implications related to the issue of access vs. ownership

Required Readings:
3. LESK, Chapter 10 “economics”

Background Readings:
5. digital library security http://bubl.ac.uk/link/d/digitallibrarysecurity.htm

Week 13 - Social Issues

Objectives: After this class, you should be able to
- Identify the social considerations to be included when designing and developing DL
- Apply related social considerations in designing and developing DL
- Explain social impacts of DL when considering certain topics in DL
• Explain social impacts of DL when considering different groups of people involved in DL
• Express certain view of the future of libraries

Required Readings:
   http://www.ifla.org/documents/libraries/net/ucladl01.pdf
   http://www.technologyreview.com/web/14408/?a=f
   http://www.dlib.org/dlib/july05/arms/07arms.html

________________________________________________________________________

Week 14 - Legal Issues and Future of Digital Libraries

Objectives: After this class, you should be able to
• Explain the definition and purpose of copyright and copyright protection
• Identify whether or not an activity in digital libraries violate copyright laws
• Know the controversial nature of privacy
• Explain the complex nature of privacy, and the need for a balance between privacy protection and the desire of building better systems
• explain the major achievements in the past 10 years
• create you own view about the valid directions of the further development of digital libraries in next 10 years and longer.

Required Readings:
1. LESK, Chapter 11
   http://www.dlib.org/dlib/july05/lynch/07lynch.html

Background Readings:
5. ARMS, Chapter 6  http://www.cs.cornell.edu/wya/DigLib/new/Chapter6.html
   http://www.dlib.org/dlib/july05/griffin/07griffin.html