



LIS 2672: Technologies and Services for Digital Data

Instructor:

Daqing He, PhD

School of Information Sciences, University of Pittsburgh

Phone: 412-624-2477

E-mail: dah44@pitt.edu

Office: Room 618, Information Science Building

Office Hours: by appointment

Graduate Student Assistants:

TA: Yu Chi – E-mail: yuc73@pitt.edu

Office hours: Room 617 Information Science Building

CourseWeb URL: <http://courseweb.pitt.edu>

I. Course Description:

The goal of this course is to prepare students to work creatively and productively in digital environment. With the focus on the technologies and services around digital data in web information context, this course includes the development of theoretical knowledge and a practical understanding of digital data. Taking a social-technical perspective, this course will help students to develop a broad understanding of digital data rather than solely view digital data as technical achievements. The course consists of three broad modules: overview of digital data, technologies supporting the life cycle of digital data, and services built around digital data in current networked participatory daily and professional environments.

II. Course Goals

Upon finishing this course, the students should be able to

- a. engage academic discussions on digital data both as a concept as well as artifacts
- b. evaluate and adopt major technologies of digital data in different stages of the life cycle
- c. examine the benefits and limitations of various services around digital data, and participant in providing the services
- d. review a practical problem associated with digital data, and develop a valid solution to solve it.

III. Recommended books and Readings

There is no required textbook for this class, instead there will be about 3-4 required readings each week (about 60 – 100 pages). Most of the readings are conference and journal articles, however, some parts of the following books will be used as well:

1. Borgman, Christine L. “Big Data, Little Data, No Data: Scholarship in the Networked World.” MIT press, 2015. [Borgman15]

2. Hey, Tony, Stewart Tansley, and Kristin Michele Tolle, eds. "The Fourth Paradigm: Data-Intensive Scientific Discovery." Vol. 1. Redmond, WA: Microsoft Research, 2009. <http://research.microsoft.com/en-us/collaboration/fourthparadigm/> [Hey09]
3. Borgman, Christine L. "Scholarship in the Digital Age." MIT press, 2007. [Borgman07]
4. Christine L. Borgman, "From Gutenberg to the Global Information Infrastructure: Access to Information in the Networked World". MIT Press, 2001. [Borgman01]
5. William Arms. "Digital Libraries". MIT Press 2000. Available online <http://www.cs.cornell.edu/wya/DigLib/MS1999/index.html> [Arms00]

There will be about 3-4 required readings each week. You will be asked to submit a reading note each week before the class to respond to the issues raised in 3 of the required readings. This can be informal in style – even bulleted lists can be used when appropriate, however, the response should clearly indicate the context, including the part of the text that triggered your questions. Do not summarize the readings. Instead, discuss your thoughts, ideas, and questions related to them. Put the responses for each week’s readings into the blog you created, and you should do so by 11:59pm of the Sunday before the class (i.e., Week 2 Sunday evening post Week 3 class readings). As described below, 10 responses are required as part of your final grade, each of which counts for .5 participation point.

Readings will generally be available via CourseWeb (if available in electronic format) and on reserve in the IS Library. I will communicate each week which readings are required both in class and on CourseWeb. Additional readings may be added as needed.

IV. Course Schedule Summary

Unit/Date	Topic	In-Class Lab Session	Assignments and Term Projects
Unit 1 / Sept 9	Overview 1: Introduction to Digital Data	Lab 1: Types of online digital data	<i>Assignment 1 is Out</i>
Unit 2 / Sept 16	Overview 2: Infrastructure for Digital Data	Lab 2: Online Seeking and Online Collaboration with Google document	<i>Term Project Introduction</i>
Unit 3 / Sept 23	Overview 3: Open Source Database and Digital Repository Systems	Lab 3: CMS tools, Omeka	<i>Assignment 1 is Due Assignment 2 is Out</i>
Unit 4 / Sept 30	Technology 1: Representing Digital Data	Lab 4: DSpace	
Unit 5 / Oct 7	Technology 2: Describing Digital Data	Lab 5: Describing data with XML	<i>Assignment 2 is Due</i>
Unit 6 / Oct 14	Technology 3: Storing Digital Data	Lab 6: Web scraping tools	<i>Project Initial Proposal</i>

Unit 7 / Oct 21	Technology 4: Analyzing Digital Data	Lab 7: Data Visualization	<i>Assignment 3 is Out</i>
Unit 8 / Oct 28	Technology 5: Discovering and Accessing Digital Data	Lab 8: Database Access with SQL	<i>Assignment 4 is Out</i>
Unit 9 / Nov 4	Technology 6: Preserving Digital Data	Lab 9: Web Archiving, Metadata for preservation	<i>Assignment 3 is Due</i>
Unit 10 / Nov 11	Exam		<i>Project Initial Presentation</i>
Unit 11 / Nov 18	Service 1: Digital Data and People	Lab 11: User-Centered Design	<i>Assignment 4 is Due</i>
Unit 12 / Nov 25	Service 2: Digital Repositories, Data Centers, and Data Security	Lab 12: Password and Security	
Unit 13 / Dec 9	Overview 4: Future		<i>Student Term Project Final Presentation</i>

V. Detailed Course Schedule

Unit 1 Overview 1: Introduction to Digital Data

Objectives: After this class, you should be able to

- restate main characteristics of digital data
- identify different communities working on digital data and the related services
- develop your own learning goals on this digital data course.

Required Readings

1. Atkinson, M. (2013). The Digital - Data Revolution. The DATA Bonanza: Improving Knowledge Discovery in Science, Engineering, and Business, 15-36.
2. McGlone, J. (2013). Looking Under the Hood: A View of the Digital Projects Librarian in the Academic Library. The New Academic Librarian: Essays on Changing Roles and Responsibilities, 67.
3. Gitelman, L. (Ed.). (2013). Raw data is an oxymoron. MIT Press. Ch1. Data before fact.
4. Borgman, C. (2015). Big Data, Little Data, No Data. MIT Press. Chapter 2. What are Data.

Lab Materials:

1. LIS2672_Lab1.pptx
2. LIS2672 Lab1_worksheet.docx

Unit 2 Overview 2: Infrastructure for Digital Data

Objectives: After this class, you should be able to

- explain the concepts of various infrastructures, including information infrastructure, data infrastructure and scholarly infrastructure
- examine ideas and major components of various infrastructure design

Required Readings:

1. Bowker, Geoffrey C., Karen Baker, Florence Millerand, and David Ribes. "Toward information infrastructure studies: Ways of knowing in a networked environment." In *International handbook of internet research*, pp. 97-117. Springer Netherlands, 2010.
2. Atkinson, M. (2013). *The Data - Intensive Survival Guide*. *The DATA Bonanza: Improving Knowledge Discovery in Science, Engineering, and Business*, 37-60.
3. Berman, Francine. "Got data?: a guide to data preservation in the information age." *Communications of the ACM* 51, no. 12 (2008): 50-56.
4. Sandra Payette, Christophe Blanchi, Carl Lagoze, Edward A. Overly. "Interoperability for Digital Objects and Repositories, The Cornell/CNRI Experiments", *D-Lib Magazine*, May 1999, Volume 5 Issue 5. <http://www.dlib.org/dlib/may99/payette/05payette.html>. (pay attention to the discussion of interoperability and extensibility)

Background readings:

5. ARMS, Chapter 2, <http://www.cs.cornell.edu/wya/DigLib/MS1999/Chapter2.html>
6. Hussein Suleman and Edward A. Fox. "A Framework for Building Open Digital Libraries", *D-Lib Magazine*, December 2001. Volume 7 Number 12. <http://www.dlib.org/dlib/december01/suleman/12suleman.html>.
7. Pandey, Satyendra C., and Andrew Dutta. "Role of knowledge infrastructure capabilities in knowledge management." *Journal of knowledge management* 17, no. 3 (2013): 435-453.

Lab Material: LIS 2672-Lab#2.ppt

Google

Docs: https://docs.google.com/document/d/1TEwAGg1TX_jULv6_tzwES69hnt17ddX6mxa4kSUzKII/edit

Unit 3 Overview 3: Open Source Database and Digital Repository systems

Objectives: After this class, you should be able to

- understand the basics of database systems such as Access and MySQL
- understand the basics of digital repository systems such as DSpace and Fedora DL systems
- know the basic procedures for installing databases, as well as DSpace and Fedora
- evaluate practical considerations for adopting one open source system

Required Readings:

1. MacKenzie Smith et al. "DSpace: An Open Source Dynamic Digital Repository" D-Lib Magazine, January 2003 <http://www.dlib.org/dlib/january03/smith/01smith.html>
2. Kucsma, J., Reiss, K., & Sidman, A. "Using Omeka to build digital collections: The METRO case study". D-Lib magazine, March/April 2010. <http://webdoc.sub.gwdg.de/edoc/aw/d-lib/dlib/march10/kucsma/03kucsma.html>
3. Lagoze, C., Payette, S., Shin, E., & Wilper, C. (2006). Fedora: an architecture for complex objects and their relationships. International Journal on Digital Libraries, 6(2), 124-138.
4. Goutam Biswas and Dibyendu Paul. An evaluative study on the open source digital library softwares for institutional repository: Special reference to Dspace and greenstone digital library. International Journal of Library and Information Science Vol. 2(1) pp. 001-010, February, 2010. <http://academicjournals.org/IJLIS/PDF/pdf2010/Feb/Biswas%20and%20Paul.pdf>

Background readings:

1. Steve Suehring, MySQL Bible. <http://hcgl.eng.ohio-state.edu/~mupparthy.1/Book.pdf>.
2. Madalli, D. P., Barve, S., & Amin, S. (2012). Digital preservation in open-source digital library software. The Journal of Academic Librarianship, 38(3), 161-164.

Hands-on points, DSpace, and Omeka:

Slides: 2672-u3-CMStools.pdf

Data: handondb1.csv

Lab: LIS2672 Lab3_worksheet.

docx; Scenic.jpg; Book cover1.jpg; Book cover2.jpg

Unit 4 Technology 1: Representing Digital Data

Objectives: After this class, you should be able to

- grasp essential components for representing digital data
- critic common formats for representing digital data
- evaluate common approaches for converting non-digital data to digital data
- understand the digital data representation in the context of whole digital data life cycle

Required Readings:

1. ARMS. Chapters 9. <http://www.cs.cornell.edu/wya/DigLib/MS1999/Chapter9.html>.
2. Clifford Lynch, "Identifiers and Their Role In Networked Information Applications". <http://www.arl.org/bm~doc/identifier.pdf>
3. Norman Paskin. "Digital Object Identifier (DOI) System". Encyclopedia of Library and Information Sciences. <http://www.doi.org/overview/080625DOI-ELIS-Paskin.pdf>

Background Readings:

4. Sam Sun, Larry Lannom, and Brian Boesch. "Handle System Overview", <http://www.handle.net/rfc/rfc3650.html>.

Lab: LIS2672_Lab#4.pdf

; LIS2672 Lab4_worksheet.docx; Book Cover.jpg; Test PDF.pdf

Unit 5 Technology 2: Describing Digital Data

Objectives: After this class, you should be able to

- outline the importance of metadata in representing and organizing digital data into collections
- understand xml-based schema in describing various digital data

Required Readings:

1. R. Gartner. Metadata for digital libraries: state of the art and future directions 2008. www.jisc.ac.uk/media/documents/techwatch/tsw_0801pdf.pdf
2. Anne J. Gilliland. Introduction to Metadata, pathways to Digital Information: 1: Setting the Stage 2008 http://www.getty.edu/research/publications/electronic_publications/intrometadata/pdf.html
3. Doug Tidwell, Introduction to XML http://www.ibm.com/developerworks/xml/tutorials/xmlintro/?S_TACT=104AHW06
4. XML Schema Tutorial <http://www.w3schools.com/Schema/default.asp>

Background Readings:

5. Stuart L. Weibel, "Border Crossings: Reflections on a Decade of Metadata Consensus Building", D-Lib Magazine, Volume 11 Number 7/8, July/August 2005 <http://www.dlib.org/dlib/july05/weibel/07weibel.html>
6. Duval, Erik, Wayne Hodgins, Stuart Sutton, and Stuart L. Weibel. "Metadata principles and practicalities." *D-lib Magazine* 8, no. 4 (2002): 16. <http://www.dlib.org/dlib/april02/weibel/04weibel.html?>
7. Disciplinary Metadata. <http://www.dcc.ac.uk/drupal/resources/metadata-standards>
8. Martin Bryan. Introducing the Extensible Markup Language (XML) <http://burks.bton.ac.uk/burks/internet/web/xmlintro.htm>
9. Uche Ogbuji. A survey of XML standards: Part 1. January 2004. <http://www-128.ibm.com/developerworks/xml/library/x-stand1.html>
10. Wickett, Karen M., Antoine Isaac, M. Doerr, Katrina Fenlon, Carlo Meghini, and Carole L. Palmer. "Representing Cultural Collections in Digital Aggregation and Exchange Environments." *D-Lib Magazine* 20, no. 5 (2014): <http://www.dlib.org/dlib/may14/wickett/05wickett.html>

Lab materials: LIS2672_Lab#5.pdf

; LIS2672 Lab5_worksheet.docx ; book_info.xlsx; book_infor.xml; Course_reservelist.xml

Unit 6 Technology 3: Storing Digital Data

Objectives: After this class, you should be able to

- Understand the basic steps of converting digital data to databases or searchable collections
- recall and explain the basic components of storing digital data in databases or document repositories

Required Readings:

1. Database. <http://en.wikipedia.org/wiki/Database>

2. Entity relationship model in database: http://en.wikipedia.org/wiki/Entity-relationship_model
3. Manning, Christopher D., Prabhakar Raghavan, and Hinrich Schütze. Introduction to information retrieval. Vol. 1. Cambridge: Cambridge university press, 2008. Section 2.1 and 2.2. <http://www-nlp.stanford.edu/IR-book/>

Lab materials: LIS2672_Lab#6.pdf
; LIS2672 Lab6_worksheet.docx

Unit 7 Technology 4: Analyzing Digital Data

Objectives: After this class, you should be able to

- restate the basic strategies for data analytics
- utilize basic analytics and visualization tools

Required Readings:

1. Heer, J., Bostock, M., & Ogievetsky, V. (2010). A tour through the visualization zoo. Commun. Acm, 53(6), 59-67. Available at: <http://delivery.acm.org/10.1145/1750000/1743567/p59-heer.pdf>
2. Yoo, D., Bell, P., Webster, B., Keegan, M., Porteus, M., & Schillinger, J. (2014). Our favorite Pew Research Center data visualizations from 2014. Pew Research Center. Available at: <http://www.pewresearch.org/fact-tank/2014/12/29/our-favorite-pew-research-center-data-visualizations-from-2014/>

Lab worksheets: LIS2672_dataviz_LAB1.docx
; LIS2672_dataviz_LAB2.pdf

Unit 8 Technology 5: Discovering and Accessing Digital Data

Objectives: After this class, you should be able to

- conduct querying to digital data in databases
- conduct querying to digital data in document collections
- understand various methods for discovering digital data on the web
- evaluate various retrieval methods for accessing digital data

Required Readings:

1. Introduction to the SQL Database Query By Ron Plew and Ryan Stephens. <http://www.informit.com/articles/article.aspx?p=29661>.
2. David Hawking , Web Search Engines: Part 1 and Part 2 IEEE Computer, June 2006. http://www.computer.org/portal/site/computer/menuitem.5d61c1d591162e4b0ef1bd108bcd45f3/index.jsp?&pName=computer_level1_article&TheCat=1055&path=computer/homepage/0606&file=thingswork.xml&xsl=article.xsl&jsessionId=G3nJ1tNBGT7hHN4hLhySlzJ4JnrH42kkmWrNhdrllpkGnMp2nC8N!-1182277384 and http://www.computer.org/portal/site/computer/menuitem.5d61c1d591162e4b0ef1bd108bcd45f3/index.jsp?&pName=computer_level1_article&TheCat=1055&path=computer/homepage/0606&file=thingswork.xml&xsl=article.xsl&jsessionId=G3nJ1tNBGT7hHN4hLhySlzJ4JnrH42kkmWrNhdrllpkGnMp2nC8N!-1182277384

[sp?&pName=computer_level1_article&TheCat=1055&path=computer/homepage/0806&file=howthings.xml&xsl=article.xsl&](#)

3. Norbert Lossau, "Search Engine Technology and Digital Libraries: Libraries Need to Discover the Academic Internet" D-Lib Magazine, June 2004, Volume 10 Number 6. <http://www.dlib.org/dlib/june04/lossau/06lossau.html>
4. OAI for Beginners - the Open Archives Forum online tutorial. <http://www.oaforum.org/tutorial/>

Background Readings:

5. Information Retrieval in Digital Libraries, by Bruce R. Schatz, Science Vol 275, 1997 <http://www.canis.uiuc.edu/archive/papers/science-irdl-journal.pdf> (too old)
6. Searching the Web: General and Scientific Information Access, by Steve Lawrence and C. Lee Giles, IEEE Communications 3791) 1999. <http://citeseer.ist.psu.edu/lawrence99searching.html>

Lab: LIS2672_Lab8_worksheet(1).docx

Unit 8 Technology 6: Preserving Digital Data

Objectives: After this class, you should be able to

- understand the roles of metadata in digital data preservation
- review technology developed for preserving digital data

Required Readings:

1. Berman, Francine. "Got data?: a guide to data preservation in the information age." Communications of the ACM 51, no. 12 (2008): 50-56.
2. Margaret Hedstrom "Research Challenges in Digital Archiving and Long-term Preservation" http://www.sis.pitt.edu/~dlwkshop/paper_hedstrom.pdf
3. [Brian F. Lavoie, The Open Archival Information System Reference Model: Introductory Guide.](#) http://www.dpconline.org/docs/lavoie_OAIS.pdf
4. Justin Littman. Actualized Preservation Threats: Practical Lessons from Chronicling America. D-Lib Magazine July/August 2007. <http://www.dlib.org/dlib/july07/littman/07littman.html>

Background Readings:

5. Trusted Digital Repositories: Attributes and Responsibilities, An RLG-OCLC Report. May 2002. <http://www.rlg.org/longterm/repositories.pdf>
6. Preserving Access to Digital Information (PADI), National Library of Australia. <http://www.nla.gov.au/padi/>
7. Digital Preservation Tutorial: <http://www.library.cornell.edu/iris/tutorial/dpm/foundation/oais/>
8. Jeff Rothenberg, "Ensuring the Longevity of Digital Information" <http://www.clir.org/pubs/archives/ensuring.pdf>
9. UNESCO, "Guidelines for the Preservation of Digital Heritage" http://portal.unesco.org/ci/en/ev.php-URL_ID=8967&URL_DO=DO_TOPIC&URL_SECTION=201.html

Lab materials: LAB#9.pdf; LIS2672_Lab9_worksheet_V1.docx; book.jpg; GNC_Biotin.jpg ; body lotion.jpg

Unit 10 Service 1: Digital Data and People

Objectives: After this class, you should be able to

- examine the various approaches for enabling people to manipulate digital data

Required Readings:

1. Arms chapter 8. <http://www.cs.cornell.edu/wya/DigLib/new/Chapter8.html>. This is useful if you want to learn really basic of interaction.
2. Rob Kling and Margaret Elliott "Digital Library Design for Usability" <http://www.cSDL.tamu.edu/DL94/paper/kling.html>
3. Tefko Saracevic, "Evaluation of digital libraries: An overview" http://www.scils.rutgers.edu/~tefko/DL_evaluation_Delos.pdf.
4. Marti Hearst, The Design of Search User Interfaces, Chapter 1 of Search User Interfaces, Cambridge University Press 2009, http://searchuserinterfaces.com/book/sui_ch1_design.html
5. Foster, Elvis C., and Shripad V. Godbole. "Database User Interface Design." In Database Systems, pp. 119-127. Apress, 2014.

Lab slides: Lab#11.pptx

Lab URLs: 1)ICPSR: <https://www.icpsr.umich.edu/icpsrweb/landing.jsp>

2)openICPSR: <https://www.openicpsr.org/>;

3)D-Scholarship: <http://d-scholarship.pitt.edu/>;

4)figshare: <http://figshare.com/>

Unit 12 Service 2: Data Repositories, Centers and Data Security

Objectives: After this class, you should be able to

- Explain the user-centered design for digital repositories, data centers
- Explain the elements of access management, associated issues and related technologies
- Employ appropriate technologies for ensuring the security of existing digital libraries

Required Readings:

1. Lynch, Clifford A. "Institutional Repositories: Essential Infrastructure for Scholarship in the Digital Age" ARL, no. 226 (February 2003): 1-7. <http://www.arl.org/bm~doc/br226ir.pdf>
2. Barroso, Luiz André, Jimmy Clidas, and Urs Hölzle. "The datacenter as a computer: An introduction to the design of warehouse-scale machines." Synthesis lectures on computer architecture 8, no. 3 (2013): 1-154. Chapter 1: Introduction. <http://www.cs.berkeley.edu/~rxin/db-papers/WarehouseScaleComputing.pdf>
3. Tankard, Colin. "Big data security." Network security 2012, no. 7 (2012): 5-8. <http://www.sciencedirect.com/science/article/pii/S1353485812700636>
4. Gross, Ralph, and Alessandro Acquisti. "Information revelation and privacy in online social networks." In Proceedings of the 2005 ACM workshop on Privacy in the electronic society, pp. 71-80. ACM, 2005. <http://dl.acm.org/citation.cfm?id=1102214>

Lab material: LIS2672 Lab 12.pptx

Unit 13 Overview 4: Future

Objectives: After this class, you should be able to

- explain the major achievements in the past 10 years
- create you own view about the valid directions of the further development of technology and services to digital data in next 10 years and longer.

Required Readings:

1. Clifford Lynch, "Where Do We Go From Here? The Next Decade for Digital Libraries", D-Lib Magazine, Volume 11 Number 7/8 July/August 2005, <http://www.dlib.org/dlib/july05/lynch/07lynch.html>
2. Knowledge lost in Information. Report of the NSF Workshop on Research directions for digital libraries <http://www.sis.pitt.edu/~dlwkshop/report.pdf>
3. Higgins, Sarah. "Digital curation: the emergence of a new discipline." *International Journal of Digital Curation* 6, no. 2 (2011): 78-88. <http://web.simmons.edu/~watkins/digitalcuration.pdf>
4. Stephen M. Griffin, "Funding for Digital Libraries Research: Past and Present" D-Lib Magazine, Volume 11 Number 7/8, July/August 2005 <http://www.dlib.org/dlib/july05/griffin/07griffin.html>

Background Readings:

5. President's Information Technology Advisory Committee: Panel on Digital Libraries, *Digital Libraries: Universal Access to Human Knowledge*, PITAC February 2001 <http://www.itrd.gov/pubs/pitac/pitac-dl-9feb01.pdf>

VI. Assessment

Participation 12%

Class attendance is required for success in this course, as material will be covered in class that is not included in the readings. Participation is based on active participation in on-class/online discussions (4 participation points), and off-class contribution to each week's "my reading questions" before the class and "my muddiest points" after the class. The detail of assess contribution to "my reading questions" is stated in section IV. Your muddiest points can be submitted with any questions regarding the issues covered during the class. The deadline for posting the muddiest is 11:59pm of the Saturday evening after the class (Week 2 Saturday evening post Week 2 class muddiest points). Again, 10 responses are required as part of your final grade, each of which counts .5 participation point.

Class participation 2%.

If you must miss a class, please notify the teaching assistant, and make arrangements to obtain course notes and handouts. Makeup exams will not be offered except under extreme circumstances.

Assignment 32%

There are total four assignments, each of which will count 8% in the final course score. You are required to make a clear presentation about your ideas, and the essay should be about one or two pages.

The deadline of submitting each assignment is before the noon of the due date. Each 24 hours delay will have 40% deduction of the maximal score. No submission later than 2 days will be accepted except in the case of emergencies and personal disasters.

Exam 28%

The exam will be offer either as a sitting in for 90 minutes or work at home, and covers all the topics taught in the weeks before it. Common exam questions include multiple choices, short definitions, and discussion questions.

Previous year's exam questions and answers can be located at the past exam section in the Courseweb.

Term Project 28%

Please see section VII for detail description of term project.

Course Grading Scale:

The final grade depends on the percentage of points you have earned, and the definition of letter grades is:

- $90 \leq A^- < 93$, $93 < A \leq 97$, $97 < A^+ \leq 100$
- $80 \leq B^- < 83$, $83 < B \leq 87$, $87 < B^+ < 90$
- $70 \leq C^- < 73$, $73 < C \leq 77$, $77 < C^+ < 80$
- $60 \leq D < 70$,
- $F < 60$

VII. Term Projects

Introduction:

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 2 people. Experience suggests that successful teams require expertise in design, implementation, and project management.

Your task is to propose, plan and develop a prototype data collection system, using Open Source software (e.g., MySQL, DSpace, Omeka, etc.). Your data collection system should address the need of a group of real users, and it should include at least three types of data, which include numeric, textual and other media.

Requirements to the Report

The outcome of the term project includes two components:

1. a final powerpoint presentation, which includes:
 - a. Problem statement
 - b. Digital library overview and scope
 - c. User requirements
 - d. Conceptual design
 - e. Schemas for representing and organizing digital objects.
 - f. Data collections
 - g. Sample information access scenarios
2. a running system.

Milestones for the project:

Introduction of term project:	September 16
Team formation deadline:	September 30
Project proposal presentation:	November 11
Final project presentation:	December 9

When writing either your assignment essays, please follow one of the established styles for reference and citation (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.comnet.edu/apa/>) is a wonderful online place to obtain the guidance for this style.

VIII. CourseWeb Information:

CourseWeb is a Web-based system using BlackBoard software that facilitates course-related communication as well as distribution of course materials and grades. You can access CourseWeb at <http://courseweb.pitt.edu> . You must log in with your University Computer Account – this is the one that goes with your ‘pitt.edu’ e-mail address. If you do not have a Pitt account, please contact Computing Services (CSSD) at 412-624-HELP [4357] to find out how to get one. Course-related e-mail will be sent to your Pitt e-mail account. If you do not read e-mail on your Pitt account, you are responsible for forwarding any e-mail received on your Pitt account to the e-mail address that you use. See <http://accounts.pitt.edu/> for information on managing your Pitt account and forwarding e-mail. If you have trouble logging in to CourseWeb, you may need to log in to the accounts website above to activate your Pitt e-mail account. Call 412-624-HELP with any problems relating to your account.

IX. Course Policies

Ground rules for class discussion

On-class interaction and discussion will be an important means of learning in this course, therefore, it is important that we work together to create a constructive environment by observing these rules:

- You should participate in the discussion of ideas.
- You should respect diverse points of view.
- You should aware the diverse backgrounds of peers.
- You may not belittle or personally criticize another individual for holding a point of view different than your own
- Your use of language should be respectful of other individuals or groups

Plagiarism

It is expected that the work you submit in this course will be your own. While collaboration is allowed for the course project, it should be approved in advance and the nature of each contribution should be specified in the project proposal and the final submission.

The following statement is taken from *The Teaching Assistant Experience: A Handbook for Teaching Assistants and Teaching Fellows at the University of Pittsburgh* (A.P. Haley and J.M. Nicoll, eds.)]

Plagiarism means submitting work as your own that is someone else’s. For example, copying material from a book or other source without acknowledging that the works or ideas are someone else’s and not

your own is plagiarism. If you copy an author's words exactly, treat the passage as a direct quotation and supply the appropriate citation. If you use someone else's ideas, even if you paraphrase the wording, appropriate credit should be given. You have committed plagiarism if you purchase a term paper or submit a paper as your own that you did not write¹.

Plagiarism is a violation of the University of Pittsburgh's standards on academic honesty, and violations of this policy are taken seriously. **From the *Guidelines on Academic Integrity: Student and Faculty Obligations and Hearing Procedures* (effective September, 1995):**

A student has an obligation to exhibit honesty, and to respect the ethical standards of the historical profession in carrying out his or her academic assignments. Without limiting the application of this principle, a student may be found to have violated this obligation if he or she:

- Presents as one's own, for academic evaluation, the ideas, representations, or words of another person or persons without customary and proper acknowledgment of sources.
- Submits the work of another person in a manner which represents the work to be one's own. [Quotation ellipsed.]²

Special Needs

Students with disabilities who require special accommodations or other classroom modifications should notify the instructor and the University's Office of Disability Resources & Services (DRS) no later than the 2nd week of the term. Students may be asked to provide documentation of their disability to determine the appropriateness of the request. DRS is located in 216 William Pitt Union and can be contacted at 648-7890 (Voice), 624-3346(Fax), and 383-7355(TTY). Students who must miss an exam or class due to religious observances must notify the instructor ahead of time and make alternative arrangements.

¹ B. G. Davis, *Tools for Teaching* (San Francisco: Jossey-Bass, 1993), 300.

² University of Pittsburgh, *Guidelines on Academic Integrity: Student and Faculty Obligations and Hearing Procedures* (Pittsburgh: University of Pittsburgh, 1995), 7-8.