

# LIS 637-001 Information Technology<sup>1</sup>

## Course Syllabus

*Last updated on January 10, 2006*

Spring 2006

School of Library and Information Science  
University of Kentucky

### Instructor

Kwan Yi

Office: 518H King Library South

E-Mail: [kwany.yi@uky.edu](mailto:kwany.yi@uky.edu)

Office Hour: Wednesday 10:00-11:00 & Thursday 14:00-15:00

Or by appointment

## ***COURSE DESCRIPTION***

Study fundamental concepts and practical methods on information technologies related to the field of library and information science.

The course consists of three components: (1) Text Analysis; (2) Information Representation; (3) Digital Library. In text processing, theoretical principles and practical methods on text manipulation and processing will be covered. In information representation, the concept, function, and methods of information representation using metadata, focusing on XML, will be studied. In digital library, general concepts, issues, and trends on Digital Libraries (DLs) will be covered with emphasis on the DL applications and projects.

## ***COURSE OBJECTIVES***

The two-fold objectives of this course are to expose the students to both conceptual and theoretical knowledge and practical experiences on Information Technologies.

Upon completion of this course, students are expected:

- To be able to familiar with fundamental concepts, techniques, and issues in Text Analysis, Information Representation, and Digital Libraries
- To gain practical experiences on Perl and XML
- To understand Perl programming, XML applications, and the process of building digital libraries

---

<sup>1</sup> The content is subject to change without notice.

## ***COURSE CONTENT***

### ***Overview of Topics & Assignments***

<b><i>Date</i></b>	<b><i>Topics</i></b>	<b><i>Quiz &amp; Due</i></b>
Jan. 11	Module 1. Course Introduction - Perl 1: Data Type	
Jan. 18	Module 2. Characteristics of Text - Perl 2: Data Operator; Function	
Jan. 25	Module 3. Text Indexing - Perl 3: Data structure	PA#1
Feb. 01	Module 4. Text Representation - Perl 4: Control structure	
Feb. 08	Module 5. Text Applications - Perl 5: Regular Expression	PA#2
Feb. 15	Module 6. Perl 6: Subroutine & File I/O	Quiz 1.
Feb. 22	Module 7. Introduction to Metadata - XML 1: XML Fundamentals	PA#3
Mar. 01	Module 8. Metadata: DC - XML 2: Document Type Definition	
Mar. 08	Module 9. Metadata: SGML family - XML 3: XML Schemas	XA#1
Mar. 15	<i>Academic Holiday (Spring Vacation)</i>	
Mar. 22	Module 10. Metadata: TEI/RDF - XML 4: CSS	XA#2
Mar. 29	Module 11. Introduction to Digital Library (DL) - Greenstone 1	Quiz 2.
Apr. 05	Module 12. Development of DL - Greenstone 2	XA#3
Apr. 12	Module 13. DL Projects - Greenstone 3	
Apr. 19	Module 14. Special topic or Guest Speaker	
Apr. 26	Project Presentations	Group Project

### ***Weekly Schedule of Topics, Reading, & Assignments***

- Jan. 11            **Course Introduction**  
                      **Perl 1: Introduction to Perl; Data Type**  
*Required Reading:*  
                      [PerlTutorial] Chapter 1; Chapter 2; Chapter 3.1.1-3.1.3
- Jan. 18            **Characteristics of Text: Compression; Collocations; Zipf's Law**  
                      **Perl 2: Data Operator; Function**  
*Required Reading:*  
                      [http://en.wikipedia.org/wiki/Zipf%27s\\_law](http://en.wikipedia.org/wiki/Zipf%27s_law)  
                      [http://en.wikipedia.org/wiki/Heaps%27\\_law](http://en.wikipedia.org/wiki/Heaps%27_law)  
                      [Korfhage] Chapter 2.5-2.7  
                      [Meadow03] Chapter 3.1-3.3  
                      [PerlTutorial] Chapter 4.1-4.3  
*Reference:*  
                      [Baeza-Yates99] Chapter 6.1-6.3; 7.4-7.5
- Jan. 25            **Text Indexing & Operation: Term Weighting; Lexical Analysis**  
                      **Perl 3: Data structure**  
*Required Reading:*  
                      [Harter]  
                      [Korfhage] Chapter 5.1-5.9  
                      [Meadow03] Chapter 9.2, 9.4  
                      [PerlTutorial] Chapter 3.2-3.3  
*Reference:*

[Baeza-Yates99] Chapter 7.1-7.2

- Feb. 01      **Text Representation: Attribute/Value; Data Structure; File Structure**  
**Perl 4: Control structure**  
*Required Reading:*  
    [Korfhage] Appendix B  
    [Meadow03] Chapter 6.1-6.4  
    [Pao89]  
    [PerlTutorial] Chapter 5.6-5.7
- Feb. 08      **Text Applications**  
**Perl 5: Regular Expression**  
*Required Reading:*  
    [Meadow03] Chapter 9  
    [Lesk97] Chapter 2.7 & 5.2  
    [PerlTutorial] Chapter 9.1-9.3
- Feb. 15      **Quiz 1.**  
**Perl 6: Subroutine & File I/O**  
*Required Reading:*  
    [PerlTutorial] Chapter 5.3; Chapter 8.1-8.3
- Feb. 22      **Introduction to Metadata: Fundamental Concepts; History; Examples**  
**XML 1: XML Fundamentals**  
*Required Reading:*  
    [Gilliland-Swetland98]  
    [Witten03] Chapter 5.4-5.6; Pp. 431-450  
    [Ray03] Chapter 1-2  
*Reference:*  
    [Galnares]
- Mar. 01      **Metadata: Dublin Core**  
**XML 2: Document Type Definition (DTD)**  
*Required Reading:*  
    [DCMES]  
    [Ray03] Chapter 3-4  
    [Weibel99]  
*Reference:*  
    [DTD1]  
    [DTD2]
- Mar. 08      **Metadata: SGML Family**  
**XML 3: XML Schemas**  
*Required Reading:*  
    [Flynn97]  
    [Marcoux97]  
    [Pfaffenberger98] Pp. 53-75 & 173-189  
    [Ray03] Chapter 4  
    [Witten03] Chapter 5.1-5.3

- Mar. 15      *Academic Holiday (Spring Vacation)*
- Mar. 22      **Metadata: TEI/RDF**  
**XML 4: Cascading Style Sheets (CSS)**  
*Required Reading:*  
[Barnard97]  
[Ray03] Chapter 5  
[RDFprimer]
- Mar. 29      **Quiz 2**  
**Introduction to Digital Library (DL)**  
**Greenstone 1: Introduction; Installation**  
*Required Reading:*  
[Lesk97] Pp. 1-25.  
[Borgman99]  
[Coffman99]  
[Witten03] Chapter 1.1, 1.2, 1.4, Chapter 2, 6.1-6.2
- Apr. 05      **Development of DL: Architectural and Design Challenges; Issues of Library Resources and Services in DL**  
**Greenstone 2: Building collection; Importing**  
*Required Reading:*  
[Garrett93]  
[Witten03] Chapter 6.3-6.4  
*Reference:*  
[Bates98]
- Apr. 12      **DL Projects: DL Initiative Projects; Digitalization in LC**  
**Greenstone 3: Archived documents; Configuration**  
*Required Reading:*  
[Witten03] Chapter 6.5-6.8  
*Reference:*  
[DLI2]  
[IFLA]  
[LoC]
- Apr. 19      **Special topic or Guest Speaker**
- Apr. 26      **Project Presentations**

#### **PERL Resources**

[PerlDoc] used as a reference  
[PerlInstall] for installing a Perl program  
[PerlCD] 6 Perl books available through the Internet  
[PerlQuickRef] Perl Regular Expression Quick Reference  
[PerlTutorial] used as a supplementary to lecture

#### **XML Resources**

[Ray03] used as a supplementary and reference to lecture  
[XMLTutorial\_1] used as a supplementary to lecture  
[XMLTutorial\_all] used as a supplementary to lecture

### **Greenstone Resources**

[Greenstone] for the installation of the GREENSTONE software and the supporting documents  
[Witten03] introduction to how to build using the Greenstone

## Referenced Texts

- *2-Hour Reserved in the Young Library:*  
☞ [Ray03] Ray, Erik T. (2003). Learning XML. Sebastopol, CA: O'Reilly.

## Referenced Articles and Resources

- *Freely available online:*
  - ☞ [Barnard97] Barnard, David T., & Ide, Nancy M. (1997). The Text Encoding Initiative: Flexible and extensible document encoding. Journal of the American Society for Information Science, 48, 622-28.
  - ☞ [Bates98] Bates, Marcia J. (1998). Indexing and access for digital libraries and the Internet: Human, database, and domain factors. Journal of the American Society for Information Science, 49, 1185-1205.
  - ☞ [Coffman99] Coffman, S. (1999). Building Earth's Largest Library: Diving into the future. Searcher, 7 (3). Available:  
<http://www.infoday.com/searcher/mar99/coffman.htm>.
  - ☞ [Baca98] Baca, Murtha, ed. *Introduction to Metadata: Pathways to Digital Information*. (Los Angeles, CA: Getty Research Institute, 1998). Available:  
[http://www.getty.edu/research/conducting\\_research/standards/intrometadata/index.html](http://www.getty.edu/research/conducting_research/standards/intrometadata/index.html).
  - ☞ [DCMES] Dublin Core Metadata Element Set (DCMES). Available:  
<http://dublincore.org/documents/dces/>.
  - ☞ [DLI2] DLI2 -- Digital Libraries Initiative Phase 2. Available:  
<http://www.nsf.gov/pubs/1998/nsf9863/nsf9863.htm>.
  - ☞ [DTD1] Available: <http://www.xmlfiles.com/dtd>.
  - ☞ [DTD2] Available: <http://www.w3schools.com/dtd/default.asp>.
  - ☞ [EAD] Available: <http://www.loc.gov/ead>.
  - ☞ [Galnares] Available: <http://www.cs.njit.edu/~galnares/Metadata.html>.
  - ☞ [Garrett93] Garrett, John. "Digital Libraries: The Grand Challenges." EDUCOM Review, July/August 1993, 28 (4). Available:  
<http://www.ifla.org/documents/libraries/net/garrett.txt>.
  - ☞ [Gilliland-Swetland98] Gilliland-Swetland, Anne. J. "Setting the Stage." In *[Baca98]*.
  - ☞ [Greenstone] *Greenstone Digital Library Software*. Available:  
<http://www.greenstone.org>.
  - ☞ [IFLA] IFLA -- Digital Libraries: Cataloging and Indexing of Electronic Resources. Available: <http://www.ifla.org/II/catalog.htm>
  - ☞ [LoC] Library of Congress -- American Memory: Historical Collections for the National Digital Library. Available: <http://memory.loc.gov/>
  - ☞ [NISO04] *Understanding Metadata*. National Information Standards Organization. (2004). Available:  
<http://www.niso.org/standards/resources/UnderstandingMetadata.pdf>.
  - ☞ [PerlDoc] Available: <http://www.perldoc.com>
  - ☞ [PerlInstall] Available: <http://www.perl.com/download.csp>
  - ☞ [PerlCD] Available: <http://www.unix.org.ua/oreilly/perl/index.htm>
  - ☞ [PerlQuickRef] Available: <http://www.erudil.com/pdf/preqr.pdf>
  - ☞ [PerlTutorial] Available: <http://www.cbkihong.com/download/perl tut.pdf>
  - ☞ [RDFprimer] Available: <http://www.w3.org/TR/rdf-primer>
  - ☞ [Weibel99] Weibel, Stuart. The State of the Dublin Core Metadata Initiative. D-Lib Magazine 5, 4 (April 1999). Available:  
<http://www.dlib.org/dlib/april99/04weibel.html>

- 📄 [XMLTutorial\_1] Available:  
[http://www.spiderpro.com/ebooks/kickstart\\_tutorial\\_xml.pdf](http://www.spiderpro.com/ebooks/kickstart_tutorial_xml.pdf)
- 📄 [XMLTutorial\_all] Available: <http://www.w3schools.com/xml/default.asp>

- *Available at the course folder in the IT lab:*

- 📄 [Baeza-Yates99] Baeza-Yates, Ricardo, & Berthier Ribeiro-Neto. (1999). *Modern Information Retrieval*. New York: ACM Press.
- 📄 [Borgman99] Borgman, Christine L. (1999). What are digital libraries? *Information Processing and Management*, 35, 227-243.
- 📄 [Flynn97] Flynn, Peter. (1997). W[h]ither the Web? The extension or replacement of HTML. *Journal of the American Society for Information Science*, 48, 614-21.
- 📄 [Harter] Harter, Stephen P. (1988). *Statistical Approaches to Automatic Indexing*. In 'Document Retrieval Systems' London, UK: Taylor Graham Publishing. Pp. 81-98.
- 📄 [Marcoux97] Marcoux, Yves, & Sevigny, Martin. (1997). Why SGML? Why now? *Journal of the American Society for Information Science*, 48, 584-92.
- 📄 [Korfhage] Korfhage, Robert R. (1997). *Information Storage and Retrieval*. New York: John Wiley & Sons.
- 📄 [Lesk97] Lesk, M. (1997). *Practical Digital Libraries: Books, Bytes & Bucks*. San Francisco: Morgan Kaufmann.
- 📄 [Meadow03] Meadow, C. T., R. B. Boyce, & D. H. Kraft. (2003). *Text Information Retrieval Systems*. San Diego: Academic Press.
- 📄 [Pao89] Pao, M.L. (1989). Database Systems. In: *Concepts of Information Retrieval*. Englewood, CO: Libraries Unlimited. Pp 146-159.
- 📄 [Pfaffenberger98] Pfaffenberger, Bryan. (1998). *Web Publishing with XML in Six Easy Steps*. San Francisco: Morgan Kaufmann. Pp. 53-75 & 173-189
- 📄 [Witten03] Witten, I. H. & David Bainbridge. (2003). *How To Build a Digital Library*. San Francisco: Morgan Kaufmann.
- 📄 [Witten99] Witten, I. H., A. Moffat, & T. C. Bell. (1999). *Managing Gigabytes: compressing and indexing documents and images*. San Francisco: Morgan Kaufmann. Chapter 3.2

## ***EVALUATION***

Evaluation will be graded based on the total score you get on the following components of the course and the grading scale below:

Quizzes:	40%
Assignments:	30%
Group Project:	30%
Participation and miscellaneous exercises:	5%

Grading scale:

100-90	A
89-80	B
79-70	C
69 or less	Fail

No late submission will be accepted except unavoidable situations prior. The only acceptable circumstances for an extension or a delay without penalties include personal illness, illness in the immediate family (physician's certificate required), and other situations in emergency approved to instructor.

## ***GENERAL INFORMATION***

University of Kentucky values academic integrity. Therefore, all students must understand the meaning and consequences of cheating, plagiarism and other academic offences of the UK student code (<http://www.uky.edu/StudentAffairs/Code>), and are required to follow them.

If you have a disability please contact the instructor to arrange a time to discuss your situation.

## ***RULE OF ABSENCE***

The only excused absence with reasonable and provable situation will be acceptable. Otherwise, failure to appear on exams will result in F for the corresponding exam.

## ***RULE OF THE NON-TRANSFERABLE***

A same work prepared and submitted by a student for this course is not allowed to be transferable to other courses. Or vice versa. That is, a work is prohibited from being graded for multiple courses.