IMT-540: Information Systems, Architectures & Retrieval
Winter 2005
Master of Science in Information Management
Information School
University of Washington

Introduction to user-centered information system development processes. Overview of Web-based, stand-alone, and distributed search engines, database management systems, architectures, and retrieval models. Prerequisite: IMT 510.

Communication tools
Course website
http://www.ischool.washington.edu/efthimis/courses/imt540/index.htm
Find the weekly readings, assignments and announcements

Course Listserv
imt540w05@u.washington.edu
Registered students are subscribed automatically using their UW mail account. This is the official medium of communication between instructors and students. Course related announcements and answers to student questions are posted on this list.

Anonymous Feedback
Submit questions and comments to the instructors anonymously. Answers will be posted to the course listserv.

Credit Hours
5 (3 lecture hours; 2 lab hours; 10 outside hours)

Meeting times
Lecture & Lab
Friday, 430-820pm, Mary Gates Hall (MGH) 271
Labs will be held in either MGH 271 or MGH 430

Extra Lab
Friday, 8:30-9:30pm, optionally available for self-directed student team work. Use MGH Commons or MGH 430

Instructors
Efthimis N. Efthimiadis, Associate Professor
330S Mary Gates Hall
Office hours: By appointment
efthimis AT u.washington.edu
Tel: 206-616-6077

David G. Hendry, Assistant Professor
330J Mary Gates Hall
Overview
Information systems come in many flavors. Relational Database Management Systems and the applications that are built on top of them are crucial for operating most businesses and large websites. Object-Oriented Database Management Systems are often used to model, store, and access engineering data in such applications as Geographic Information Systems and Engineering Design. Information Retrieval Systems are best used for storing and retrieving weakly structured documents such as web pages, personal documents, and books. Question and Answering systems seek to return answers to queries rather than documents. Text and data mining applications seek to identify patterns in the data that, for example, may reveal fraud or emerging sales trends. Content Management Systems are used to store and access semi-structured data with technologies based on the Extensible Markup Language. These types of information systems, and many others, have a great influence on the success of individual businesses, the economy, and the ultimately the quality of our lives. Learning about all these different kinds of information systems is a life’s work.

This course focuses on two types of information system: Information Retrieval Systems and Database Management Systems. You will learn about the technical underpinnings of these systems and examine the kinds of problems that each type of system is best suited. You will also learn how to apply basic techniques in user-centered design to shape the design of an information system. Finally, in the most exciting, and challenging, part of this course you will explore how these three perspectives can be integrated.

Readings
Readings are posted on the course website. Readings are to be done the week ahead of class. Typically, you will be asked to read 3-4 papers each week. In general, two of the papers will be relatively basic and two will be more advanced. The more advanced papers are intended for students who already have significant experience with a topic. Additional readings and books are often cited for your own independent study.

Learning
Aims
The general aims of this course are to
1. Develop knowledge for the user-centered and system-centered perspectives of information systems.
2. Develop skills in application user-centered design methods.
3. Develop knowledge for the basic functional and operational features of Information Retrieval and Database Management Systems.

Objectives
On the successful completion of this course, you should be able to:
1. Discuss the strengths and weaknesses of the following design processes: Waterfall model, rapid prototyping, participatory design, goal-based design, and scenario-based design.
2. Create a use scenario, persona, and paper prototype of a system.
3. Create a task analysis and perform a simple usability evaluation.
4. Describe the functions and organization of database management systems.
5. From simple problem statements, derive SQL statements for querying, updating and creating databases.
6. Sketch a three-tier architecture, explaining the functional requirements of each tier and identify each tier in Microsoft Access and in a Web Application built with open source tools.
7. Create or understand an entity-relationship diagram for a small system.
8. Describe the function and organization of an information retrieval (IR) system, including documents, document collections, terms, queries, matching, ranking, and results.
9. Explain how an inverted file works
10. Describe the difference between Boolean and ranked retrieval.
11. Know the formula for the Zipf distribution and recognize its curve in a data set.
12. Critically discuss the differences between relational database and information retrieval systems.

**Assessment**

**Assignments**

<table>
<thead>
<tr>
<th>Task Description</th>
<th>Due (Week)</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Examining Google search</td>
<td>#2</td>
<td>10%</td>
</tr>
<tr>
<td>2. Scenario analysis and usability evaluation</td>
<td>#4</td>
<td>20%</td>
</tr>
<tr>
<td>3. IR processes and evaluation</td>
<td>#6</td>
<td>20%</td>
</tr>
<tr>
<td>4. SQL &amp; data modeling</td>
<td>#8</td>
<td>10%</td>
</tr>
<tr>
<td>5. IR versus DBMS</td>
<td>#10</td>
<td>10%</td>
</tr>
</tbody>
</table>

Assignment #2 is done in pairs.

**History Places Project**

History Places is a group project worth 30% of your final grade. You will work in groups of 4.

There are three deliverables for this project: 1) A draft poster presentation; 2) A final poster presentation; and 2) A short report.

**Important dates**

<table>
<thead>
<tr>
<th>Deliverable</th>
<th>Due (Week)</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Draft poster presentation</td>
<td>#9</td>
<td>5%</td>
</tr>
<tr>
<td>Final Poster presentation</td>
<td>#10</td>
<td>10%</td>
</tr>
<tr>
<td>Short report</td>
<td>#10</td>
<td>15%</td>
</tr>
</tbody>
</table>
**Grading criteria**

Work in this course will be graded to criteria. In other words, you won't be graded on a curve. Each deliverable is designed to test your achievement against one or more of the learning objectives. Different assignments emphasize different learning objectives.

General grading information for the University of Washington is available at: [http://www.grad.washington.edu/stsv/gradpol3.htm#Grading](http://www.grad.washington.edu/stsv/gradpol3.htm#Grading)

**Standard cover sheet**

To protect your privacy when assignments are returned and to facilitate communication, submitted work must have a cover sheet. The cover sheet must include the following information and be formatted nicely:

- Course name
- Quarter, program, department, and university
- Assignment name
- Your name and e-mail address
- A date
- A web site address (if relevant).

Staple the assignment pages to the cover sheet.

**Right to revise**

The syllabus is a living document. The instructors will revise this syllabus by adding course material on the course website as the quarter progresses. The most up-to-date version of the syllabus is on this website.
Class schedule

Week 1: Introduction
Topics
Greetings; Introduction to Class; MSIM student mixer

Read

When reading these (old) articles we'd like you to view them in perspective. As you read, identify the "information system" issues raised and the solutions proposed. Reflect on the proposed solutions and compare them to present time issues and proposed solutions. Ponder a bit on the issues by removing the "technology" and focusing on the problems. How do the issues/problems of then and now compare?

Week 2: User-Centered Design, Part I
Topics
Introduction to User-Centered Design and Usability; Design Methodologies; Introduction to Scenario-Based Design; Requirements analysis with Scenarios and Personas; System conceptualization and design with Paper Prototyping

Read

Background

Lab
Develop a scenario and paper prototype
**Week 3: User-Centered Design, Part II**

*Topics*

Introduction to Task-Based Design; Requirements analysis with Task Analysis; Usability Evaluations

*Read*

   [https://www.ischool.washington.edu/efthimis/uwnetid/courses/readings/task-analysis.chapter.pdf](https://www.ischool.washington.edu/efthimis/uwnetid/courses/readings/task-analysis.chapter.pdf)
   Chapter 9, Usability testing on 10 cents a day
   Chapter 10, Usability testing: The movie
   [https://www.ischool.washington.edu/efthimis/uwnetid/courses/readings/krug09.pdf](https://www.ischool.washington.edu/efthimis/uwnetid/courses/readings/krug09.pdf)

*Background*


*Lab*

Conduct a Usability Evaluation

**Week 4: Information Retrieval, Part I**

*Topics*

Introduction to IR; Overview of the components of an IRS; Queries, Documents, Indexing; Theories & Models in IR (Retrieval Techniques) for text, hypermedia, web.

*Read*

   [http://www.infotoday.com/searcher/may01/liddy.htm](http://www.infotoday.com/searcher/may01/liddy.htm)
   Chapter 1, Overview
   Chapter 2, Extracting lexical features

*Background*

1. Search Engine Features For Webmasters
3. Search Engines by Search Features
5. A look at how Google's monopoly, algorithms, and privacy policies are undermining the Web. [http://www.google-watch.org/](http://www.google-watch.org/)
**Lab**  
Information Retrieval Toolbox: Indexing

**Week 5: Information Retrieval, Part II**

**Topics**
- Continuation of IR topics; Web IR, including link analysis and crawlers

**Read**

**Lab**  
Information Retrieval Toolbox: Matching

**Week 6: Database Management Systems, Part I**

**Topics**
- Database functions; Relational Model; Three-tier Architecture, Structured Query Language (SQL)

**Read**
   - Chapter 1, Introduction to Databases  
   - Chapter 2, Database Environment  
   - Chapter 3, The Relational Model  
   [https://www.ischool.washington.edu/efthimis/uwnetid/courses/readings/cb_01.pdf](https://www.ischool.washington.edu/efthimis/uwnetid/courses/readings/cb_01.pdf)  
   [https://www.ischool.washington.edu/efthimis/uwnetid/courses/readings/cb_02.pdf](https://www.ischool.washington.edu/efthimis/uwnetid/courses/readings/cb_02.pdf)  
   [https://www.ischool.washington.edu/efthimis/uwnetid/courses/readings/cb_03.pdf](https://www.ischool.washington.edu/efthimis/uwnetid/courses/readings/cb_03.pdf)

**Lab**  
Structured Query Language

**Week 7: Database Management Systems, Part II**

**Topics**
- Introduction to Data Modeling and Entity-Relationship Modeling and Enhanced-Entity Relationship Modeling

**Read**
  - Chapter 11, Entity Relationship Modeling  
  - Chapter 12, Enhanced Entity Relationship Modeling [Optional]


Lab
Modeling Exercise

Week 8: Integration/Application ??
Topics
Guest Lecture
Read
• One paper related to guest lecture

Lab
History Places – A double lab to work on project

Week 9: Integration/Application
Topics
Guest Lecture
Read
• One Paper related to guest lecture

Lab
History Places – A double lab to work on project

Week 10: Conclusion
Topics
Review & Further Directions

Poster Show of History Places
Student services
J. D. Miller,
Director of Graduate Student Services
470E Mary Gates Hall
mrmiller@u.washington.edu
Tel: (206) 685.6698

Please note: If you have any concerns about a course please contact the instructors directly, efthimis AT u.washington.edu , dhendry AT u.washington.edu

If you are still not satisfied with the response that you receive, you may contact Joseph Janes, the Associate Dean for Academics in 370 Mary Gates Hall, by phone at : (206) 616-0987, or by e-mail at jwj@u.washington.edu.

You may also contact the Graduate School at G-1 Communications Building, by phone at (206) 543-5900, or by e-mail at efeetham@u.washington.edu

Academic accommodations
To request academic accommodations due to a disability, please contact Disabled Student Services: 448 Schmitz, 206-543-8924 (V/TTY). If you have a letter from DSS indicating that you have a disability which requires academic accommodations, please present the letter to me so we can discuss the accommodations you might need in the class.

Academic accommodations due to disability will not be made unless the student has a letter from DSS specifying the type and nature of accommodations needed.

For additional information, see Statements to Ensure Equal Opportunity and Reasonable Accommodation, downloaded March 5, 2003, http://www.washington.edu/admin/eoo/ eoost.html

Academic honesty
The essence of academic life revolves around respect not only for the ideas of others, but also their rights to those ideas and their promulgation. It is therefore essential that all of us engaged in the life of the mind take the utmost care that the ideas and expressions of ideas of other people always be appropriately handled, and, where necessary, cited. For writing assignments, when ideas or materials of others are used, they must be cited. The format is not that important—as long as the source material can be located and the citation verified, it’s OK. What is important is that the material be cited. In any situation, if you have a question, please feel free to ask. Such attention to ideas and acknowledgment of their sources is central not only to academic life, but life in general.

Please acquaint yourself with the University of Washington's resources on academic honesty:
• http://depts.washington.edu/grading/issue1/honesty.htm
Students are encouraged to take drafts of their writing assignments to the Writing Center for assistance with using citations ethically and effectively. Information on scheduling an appointment can be found at:


**Copyright**

All of the expressions of ideas in this class that are fixed in any tangible medium such as digital and physical documents are protected by copyright law as embodied in title 17 of the United States Code. These expressions include the work product of both: (1) your student colleagues (e.g., any assignments published here in the course environment or statements committed to text in a discussion forum); and, (2) your instructors (e.g., the syllabus, assignments, reading lists, and lectures). Within the constraints of "fair use", you may copy these copyrighted expressions for your personal intellectual use in support of your education here in the iSchool. Such fair use by you does not include further distribution by any means of copying, performance or presentation beyond the circle of your close acquaintances, student colleagues in this class and your family. If you have any questions regarding whether a use to which you wish to put one of these expressions violates the creator's copyright interests, please feel free to ask the instructor for guidance.

**Privacy**

To support an academic environment of rigorous discussion and open expression of personal thoughts and feelings, we, as members of the academic community, must be committed to the inviolate right of privacy of our student and instructor colleagues. As a result, we must forego sharing personally identifiable information about any member of our community including information about the ideas they express, their families, life styles and their political and social affiliations. If you have any questions regarding whether a disclosure you wish to make regarding anyone in this course or in the iSchool community violates that person's privacy interests, please feel free to ask the instructor for guidance.

Knowing violations of these principles of academic conduct, privacy or copyright may result in University disciplinary action under the Student Code of Conduct.

**Student Code of Conduct**

Good student conduct is important for maintaining a healthy course environment. Please familiarize yourself with the University of Washington's Student Code of Conduct at:

- [http://www.washington.edu/students/handbook/conduct.html](http://www.washington.edu/students/handbook/conduct.html)