

# LIS618

## Online Information Retrieval Techniques

last revised: 2004-02-16

The latest version of this document is available on the web at <http://openlib.org/home/krichel/lis618p04s/>.

### **Course Description**

This course will introduce the students to the theory of information retrieval and its application in large-scale commercial database system and on the WWW.

### **Course objectives**

On completing this course, students

- will have been introduced to information retrieval models;
- will have been introduced to several commercial database systems and be aware of their strengths and weaknesses;
- will have been introduced to expert search strategies with web search engines and databases.

### **Prerequisites**

Students should have a basic command of the Microsoft Windows operating system because the machines in the lab run on this operating system.

### **Instructor**

Thomas Krichel  
Palmer School of Library and Information Science  
C.W. Post Campus of Long Island University  
720 Northern Boulevard  
Brookville, NY 11548-1300  
krichel@openlib.org  
work phone: +1-(516)299-2843  
private phone: +1-(718)507-1117

### **Class structure**

Classes will be held at in the computer lab of the Palmer School on the CW Post campus of of LIU, between 12:30 and 16:30. There will be a mixture of lectures and hands-on work in the lab. Provisional class details are:

- 0 2004-01-25 Introduction to the course and to Information Retrieval
- 1 2004-02-01 preparing to search and Dialog 1
- 2 2004-02-08 IR performance and Dialog 2
- 3 2004-02-15 vector model and Nexis I
- 4 2004-02-29 other issues in information retrieval and Nexis II
- 5 2004-03-07 Google and Amazon
- 6 2004-03-14 IR in P2P, relational data, OpenURL and full-text databases  
the course web site.

Slides for all classes are downloadable from

**Class mailing list**

A mailing list for this class has been set up. Students who wish to stay informed are encouraged to sign up.

**Readings**

The powerpoint slides of the instructor are the reading. The slides may point to other sources of reference as required. Database practice makes for a master searcher, only reading about them is not getting students very far. The session on Google will make use of a new book on Google hacks published by O'Reilly and Associates.

Some slides on teaching Dialog have been given to the instructor, mail him if you want a copy.

Some reference questions to work on are available.

**Assessment**

Each student will have to prepare a search exercise and report as detailed in the first lecture. This report must not exceed 5 pages. Appendices will be permitted, but may not get read. It will count for 50% of total grade. It is due on March 21. The remaining 50% will come from quizzes held at the start of each lecture except the first. Quizzes will last around 10 minutes and focussed on a precise short answer. The worst quiz performance will be discarded when the average is being computed.