

## Course Number & Name: LIS 531H - Information Retrieval

- **Description:** This course addresses all aspects of the automatic collection, retrieval, matching, ranking, and presentation of machine-readable information resources (“information retrieval” or IR). Students study the techniques most commonly associated with full text retrieval and search engine design, by examining retrieval models (Boolean, set, algebraic, probabilistic), text analysis (for concept clustering, term weighting), and text and multimedia languages and properties (such as SGML and file formats). In addition, students review indexing and searching techniques, user interface designs, relevance feedback, evaluation, and developing techniques, such as information visualization.
- **Prerequisites:** LIS 415 and 488, TOR
- **Audience (level, environment/setting):** All doctoral students and Master’s students interested in web-based information sources, digital library construction, digital archives, e-business, and data structures.
- **Student Learning Outcomes** [by number]: 1, 2, 6, 10
- **Topics/List of Lectures:**
  - IR vs. Data Retrieval; demonstration of parsing and retrieval software applications
  - Retrieval models (basic, Boolean, vector, probabilistic, fuzzy set, latent semantic indexing, neural networks)
  - Retrieval evaluation
  - Query languages
  - Query operations & relevance feedback
  - Text and Multimedia languages and properties
    - information theory and modeling natural languages
    - Markup languages
    - Multimedia formats
  - Text operations
    - lexical analysis
    - stemming, index terms, and thesauri
  - Indexing and searching
  - User interfaces and visualizations
  - Searching the Web
  - Libraries & Bibliographical Systems
  - Digital libraries
- **Suggested Textbooks/Readings**
  - Baeza-Yates, R., & Ribeiro-Net, B. (1999). *Modern Information Retrieval*. Reading, MA: Addison Wesley/ACM Press.

- van Rijsbergen, C. J. (19??). *Information Retrieval*. London: Butterworths.
  - Korfhage, R.R. (1997.) *Information Storage and Retrieval*. New York: John Wiley & Sons.
  - Witten, I. H., Moffat, A., & Bell, T. C. (1999). *Managing Gigabytes*. San Francisco: Morgan Kaufmann.
- 
- **Course History:**
    - March 2005: Presented at curriculum meeting?
    - April 2005: Voted in at faculty meeting?
- 
- **Document prepared by:** Gerry Benoit, 3/05