

---

## CSE 454 Advanced Internet and Web Services

---

### **Credits**

5.0 (5 hrs lecture/meetings)

### **Lead Instructor**

Daniel Weld

### **Textbook**

None

### **Course Description**

Design of Internet search engines, including spider architecture, inverted indices, frequency rankings, latent semantic indexing, hyperlink analysis, and refinement interfaces. Construction of scalable and secure web services. Datamining webserver logs to provide personalized and user-targeted services. Large project.

### **Prerequisites**

CSE 326 or CSE 332.

### **CE Major Status**

Selected Elective

### **Course Objectives**

Understand the intellectual foundations of Web Search and Internet technologies including information retrieval, data mining, and cryptography. Be able to understand, build, and debug a Web Search Engine.

### **ABET Outcomes**

- (a) an ability to apply knowledge of mathematics, science, and engineering
- (b) an ability to design and conduct experiments, as well as to analyze and interpret data
- (c) an ability to design a system, component, or process to meet desired needs within realistic constraints such as economic, environmental, social, political, ethical, health and safety, manufacturability, and sustainability
- (d) an ability to function on multi-disciplinary teams
- (e) an ability to identify, formulate, and solve engineering problems
- (f) an understanding of professional and ethical responsibility
- (g) an ability to communicate effectively
- (h) the broad education necessary to understand the impact of engineering solutions in a global, economic, environmental, and societal context
- (i) a recognition of the need for, and an ability to engage in life-long learning

(j) knowledge of contemporary issues

(k) an ability to use the techniques, skills, and modern engineering tools necessary for engineering practice

## **Course Topics**

- Internet history & future
- Architecture of an information extraction system
- Supervised learning, logistic regression & application to relation extraction
- Parsing, POS tags, named entity recognition & other features for extraction
- Crawling the Web
- Information retrieval models, index construction & compression
- Search engine query processing & Alta Vista case study
- Hyperlink analysis & fast external-memory Pagerank computation
- Case study: NYU's knowledge-base population system
- Mining unstructured healthcare data
- Computational advertising
- Crowdsourcing
- Cryptography & practical internet security