

## **CSI 550 Fall 2009**

### ***Information Retrieval***

#### ***Reading topics***

NOTE: these are suggested topics; other topics are possible. Please contact me to select or propose a topic and to obtain reading materials.

1. Probabilistic retrieval techniques  
references: Robertson; Turtle & Croft in KSI&PW book
2. Natural language processing in IR  
Lewis & Sparck-Jones (from instructor) and Sparck-Jones in TS book
3. Automated summarization  
selection from Mani & Maybury book (from instructor)
4. Latent Semantic Indexing retrieval model  
Dumais et al.
5. Text Retrieval Conference  
selection from proceedings (from instructor)
6. Topic Detection and Tracking approaches  
selection from DARPA Broadcast News Workshop (from instructor)
7. Flavors of tf\*idf weighting  
article by Buckley
8. Pivoted weighting  
articles by Robertson and Singhal et al.
9. Document Clustering and Categorization  
selections from SIGIR proceedings
10. Information Filtering applications  
Selections from SIGIR proceedings, Arampathsis thesis.
11. SMART system  
Salton's article in KSI&PW
12. INQUERY system  
selection from KSI&PW and TREC proceedings, SIGIR proceedings
13. OKAPI system  
Robertson
14. Use of thesauri in IR, including Wordnet  
Voorhees (SIGIR), Smeaton (TS book)
15. Automatic derivation of thesauri  
selection from ACL, SIGIR
16. Information extraction systems

- Jacobs & Rau (SCISOR system)
- BBN's Identifinder (Weischedel)
- Sheffield's GATE (Cunningham)
- Selections from TS book (Riloff, Wilks)
- 17. FERRET semantic IR system
  - Mauldin
- 18. Cross-lingual IR, foreign language search
  - selection from SIGIR, Grefenstette's book
- 19. Query-expansion & relevance feedback
  - papers from instructor
- 20. Question-answering - classical
  - Lehnert, Woods' LUNAR, STUDENT, BASEBALL
- 21. Question-answering - modern
  - Kupiec's MURAX (TS book chapter 13); TREC QA papers;
  - Strzalkowski & Harabagiu book
- 22. Web search engines
  - Google, Bing, etc.
- 23. Image and/or video, multimedia retrieval
  - find own material
- 24. User Studies in IR
  - Selections from SIGIR, JASIS
- 25. Interactive IR
  - Belkin, SIGIR, TREC
- 26. Information fusion
  - Kantor & Ng
- 27. MEDLINE – searching medical literature
  - Own material
- 28. Digital Libraries
  - Proceedings of annual conference
- 29. Genetic sequence search – applications to bioinformatics
  - SIGIR'02 tutorial
- 30. Collaborative filtering
  - TBD