

LIS 3600: Seminars of Information Technology and Systems

Class time:Fridays 12:00pm – 2:50pm or negotiated between students and instructorLocation:411 IS Building

Instructor:

Daqing He, PhD School of Information Sciences, University of Pittsburgh Phone: 412-624-2477 E-mail: dah44@pitt.edu Office: Room 618, Information Science Building

I. Course Description:

This course offers an examination of the latest development in the literature of information technology and systems, with the emphasis on Web based information access systems and their corresponding technologies. Dividing the literature into the areas of theoretical foundations, support technologies and real applications, latest published articles are drawn from the three areas to form the pool of readings. Students will be assigned papers from the pool and will lead the discussions about the assigned papers. The course aims to help PhD students to develop a broad understanding of modern Web based information access systems, and to develop skills on literature updates, critical thinking, and scholarly communication.

This course is open only to SIS PhD students.

II. Course Goals

Upon finishing this course, the students should be able to

- a. identify major literature sources in the field of Web based information access systems.
- b. understand the state of the art of Web based information access technologies, systems and their applications.
- c. conduct individual or group based research discussions

III. Reading Pool

The following articles are examples of papers put into the reading pool by the instructor. The actual pool will be expanded according to the background and research interests of the enrolled students, and the further development of the literature.

Theoretical Foundations

- 1. Tefko Saracevic. (1999) Information Science. Journal of the American Society for Information Science, 50(12), 1051-1063.
- 2. Fuhr, N. (1992). Probabilistic Models in Information Retrieval. The Computer Journal, 35(3),243--255. http://citeseer.ist.psu.edu/fuhr92probabilistic.html
- 3. Oard, D., 2003. 'When You Come to a Fork in the Road, Take It!', Keynote speech at SIGIR 2003 Workshop on the future of Cross-Language Information Retrieval research in Tampere, Finland
- 4. Stefano Mizzaro. (1999) How Many Relevances in Information Retrieval? Interacting With Computers, 10(3), 305-322
- 5. T. Saracevic, Relevance: A review of and a framework for thinking on the notation in information science. Journal of the American society for information science. 26, 321-343. Or in Reading in Information Retrieval, Karen Sparck Jonese and Peter Willett (ed).
- 6. G. Salton and C. Buckley. Term-Weighting Approaches in Automatic Text Retrieval. Information Processing and Management. 24.513-523. 1988. Or in Reading in Information Retrieval, Karen Sparck Jonese and Peter Willett (ed).
- S.E. Robertson. The Probability Ranking Principle in IR. Journal of Documentation. 33.294-304. 1977. Or in Reading in Information Retrieval, Karen Sparck Jonese and Peter Willett (ed).
- 8. N.J. Belkin, R.N. Oddy and H.M. Brooks. ASK for Information Retrieval. Part 1. Background and Theory. Journal of Documentation. 38.61-71. 1982. Or in Reading in Information Retrieval, Karen Sparck Jonese and Peter Willett (ed).

Support Technologies

- 9. Gerard Salton and Chris Buckley. Improving Retrieval Performance by Relevance Feedback. Journal of the American Society for Information Science, 41(4):288-297, 1990.
- 10. Jinxi Xu and W. Bruce Croft: Query Expansion Using Local and Global Document Analysis. SIGIR 1996: 4-11
- 11. Nie, J.Y., Simard, M., Isabelle, P., & Durand, R. 1999, 'Cross-Language Information Retrieval based on Parallel Texts and Automatic Mining of Parallel Texts in the Web', In the proceeding of 22nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval SIGIR-99, pp. 74-81 Berkeley.
- 12. Pirkola, A., 1998, The effects of query structure and dictionary setups in dictionary-based crosslanguage information retrieval. In Proceedings of the 21st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 55-63. ACM Press.
- 13. F. Diaz, D. Metzler, 2006. Improving the estimation of Relevance Models using Large External Corpora. In Proceedings of SIGIR 2006 154-162.
- 14. R. White et al. 2005. A Study of Factors Affecting the Utility of Implicit Relevance Feedback. In proceedings of SIGIR 2005 35-42.
- 15. X. Shen, et al. 2005 Context-Sensitive Information Retrieval Using Implicit Feedback. In proceedings of SIGIR 2005 43-50
- 16. X. Shen and C. Zhai 2005 Active Feedback in Ad Hoc Information Retrieval. In proceedings of SIGIR 2005 59-67
- 17. Improving pseudo-relevance feedback in web information retrieval using web page segmentation group of 17 S Yu, D Cai, JR Wen, WY Ma the 12th international conference on World Wide Web, 2003 - portal.acm.org

Real Applications

- Daqing He, Douglas W. Oard, Jianqiang Wang, Jun Luo, Dina Demner-Fushman, Kareem Darwish, Philip Resnik, Sanjeev Khudanpur, Michael Nossal, Michael Subotin, Anton Leuski. (2003) <u>Making MIRACLEs: Interactive Translingual Search for Cebuano and Hindi.</u> ACM Transactions on Asian Language Information Processing (TALIP), 2(3):219-244.
- 19. Lynette Hirschman and Robert Gaizauskas. (2001) Natural Language Question Answering: The View from Here. Natural Language Engineering, 7(4), 275-300.
- 20. Ellen M. Voorhees. (2003) Evaluating the Evaluation: A Case Study Using the TREC 2002 Question Answering Track. (2003) Proceedings of the 2003 Human Language Technology Conference and the North American Chapter of the Association for Computational Linguistics Annual Meeting (HLT/NAACL 2003).
- 21. M. Chau, J. Qin, Y.Zhou, et al. (2005) SpidersRUs, Automated Development of Vertical Search Engines in Different Domains and Languages. Proceedings of JCDL 2005. 110-112

IV. Course Format

This course consists of a serial of seminars. In the first two weeks, the instructors will lead the discussions. Then students are required to pick up one paper from each of the three areas, and present and lead discussions for the assigned week. Depending on students' participations, the quality of the presentations and discussions, they will receive letter grades when they successfully finish this course.