InformationRetrieval

Books

- Introduction to Information Retrieval, Manning, Raghavan & Schütze, 2007

- Managing Gigabytes [KevinBurton] I can vouch for M.G. as I have a copy and it's a GREAT book. Should be called Managing Terabytes. Its not light reading by any means and you'll probably have to sit down with each chapter for a bit. ([DavidSpencer] – I agree, great thorough book, must have)

- Modern Information Retrieval, Ricardo Baeza-Yates and Berthier Ribeiro-Neto, 1999
- Foundations of Statistical Natural Language Processing, Chris Manning and Hinrich Schütze, 1999
- Readings in Information Retrieval
- Mining the Web - Discovering Knowledge from Hypertext Data by Soumen Chakrabarti, Morgan-Kaufmann. A good book that covers all the aspects of web and text mining.

Web links

- Introduction to Information Retrieval, Manning, Raghavan & Schütze, 2007
  This book is available online in PDF form.
- Big list of links to IR resources: http://www-csli.stanford.edu/~schuetze/information-retrieval.html

- Inquery Query help [PaulElschot]
  This engine has a more elaborated query language than Lucene. However, Lucene supports most of the mechanisms used by the Inquery operators. Recommended: the section on Query Operator Types that makes the distinction between Belief List Operators and Proximity List Operators.

- Course on Information Retrieval [JoaquinDelgado]
  A very solid (and free) online course on "intelligent information retrieval" with focus on practical issues prepared by Prof. Mooney (Univ. of Texas), a well known expert in Machine Learning and IR.

- http://finance.groups.yahoo.com/group/mg/
- http://www.sims.berkeley.edu/~hearst/irbook/
- course notes from the Stanford course on IR: http://www.stanford.edu/class/cs276/handouts/lecture1.pdf to http://www.stanford.edu/class/cs276/handouts/lecture16.pdf
- http://www.dcs.gla.ac.uk/~tombrosa/AIS/SMART-tutorial/weights.html
- http://trec.nist.gov/pubs/trec3/t3_proceedings.html
  TREC (Text REtrieval Conference) 3 Proceedings, including Salton's paper on SMART.
- http://www.soi.city.ac.uk/~ser/idf.html
  The Spärck Jones / Robertson IDF page. Karen Spärck Jones is the author of the paper which introduced IDF.

User Discussions on Java-User

- http://www.gossamer-threads.com/lists/lucene/java-user/43042