

PageRank

Page Rank is *part* of how Google scores search results. Lucene cannot easily use Page Rank however as you have to post-process the sites you've indexed, make calculations on how the sites are linked together, and then update the index with this info.

Papers:

[The Anatomy of a Large-Scale Hypertextual Web Search Engine](#)

[The PageRank Citation Ranking: Bringing Order to the Web](#)

Citeseer is a great repository of technical papers and will show papers that describe or refer to Page Rank: <http://citeseer.ist.psu.edu/>

There is a java implementation of page rank: <http://jung.sourceforge.net/api/1.4.1/edu/uci/ics/jung/algorithms/importance/PageRank.html>