SummerOfCode2011ProjectRanking

Implementing State of the Art Ranking for Lucene

Google Summer of Code project

Abstract

Lucene employs the Vector Space Model (VSM) to rank documents, which compares unfavorably to state of the art algorithms, such as BM25. Moreover, the architecture is tailored specifically to VSM, which makes the addition of new ranking functions a non-trivial task. This project aims to bring state of the art ranking methods to Lucene and to implement a query architecture with pluggable ranking functions.

For more information, visit issue LUCENE-2959 in Jira.

Subpages

- A short overview of Terrier's scoring architecture
- JIRA issues opened for this project
- Notes and Q&A