

Ranking mit lucene

January 29, 2010

- Indizieren Sie die Dokumente der Übung Scoring und suchen Sie für die Fragen der letzten Übung Scoring mittels Vektorspace-Modell
- Benutzen Sie dabei einen Analyser, der nur Whitespaces trennt.
- Benutzen Sie `Explanation`
- Verwenden Sie den `QueryParser` über Kommandozeilen-Eingaben
- Vergleichen Sie die Ergebnisse mit den manuellen Ergebnissen und erklären Sie die Abweichungen

- Indizieren Sie Sie Dokumente und machen Sie dabei folgende Änderungen:
 - Boosten Sie Dokument 3 mit dem Faktor 2
 - Welche Änderungen ergeben sich jetzt bei der Suche nach "Krebs"
- Ändern Sie die Frage "Krebs OR Adenokarzinom" so, dass Sie als ersten Treffer ein Dokument mit dem Begriff "Krebs" erhalten