

Übungen zu Information Retrieval, Wintersemester 2011/2012

Matthias Jordan (LF 136)
Sprechstunde nach Vereinbarung
mjordan@is.inf.uni-due.de

Übungsblatt 7

Abgabe bis **5.12.2011, 12.00 Uhr**

Aufgabe 14: Sprachmodelle

Vergleiche die Indexierung mittels BIR-Modell (und BM25) mit der Indexierung mittels der beiden Sprachmodelle aus der Vorlesung: Was sind Gemeinsamkeiten? Wo liegen die Unterschiede?

5 Punkte

Aufgabe 15: Experimentelle Bestimmung von Websuchmaschinen-Features

Unterschiedliche Websuchmaschinen implementieren unterschiedliche Features in ihrer textbasierten Suche. Entwirf ein Experiment, um für eine Websuchmaschine Deiner Wahl herauszufinden, ob sie folgende Features unterstützt:

- Grundform-/Stammformreduktion
- Kompositabehandlung
- Synonyme
- einfache Nominalphrasen (“computer research facility“ – Wortreihenfolge eindeutig, Abstand 0)
- komplexere Nominalphrasen (“facility for the research on computers“ - Wortreihenfolge mehrdeutig, größerer Abstand)
- Rechtschreibkorrektur
- Cross language retrieval (Dokumente werden auch gefunden, wenn die Suchterme im Dokument in einer anderen Sprache vorkommen als in der Anfrage)

Führe das Experiment für eine Websuchmaschine Deiner Wahl durch, beschreibe Dein Verfahren und berichte die Ergebnisse.

20 Punkte