

**Übungen zu Information Retrieval, Wintersemester 2011/2012**

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

**Übungsblatt 4**

Abgabe bis **14.11.2011, 12.00 Uhr**

---

**Aufgabe 9: Klassifikationen**

Es seien die folgenden Begriffe gegeben:

- Kontextoperatoren
  - Trunkierung
  - Maskierung
  - Grundformreduktion
  - Stammformreduktion
  - Stoppworteliminierung
  - Tokenisierung
  - Klassifikationen
  - Thesaurus
  - Facettenklassifikation
  - ACM-CCS
  - Ontologien
  - Yago
  - Freebase
  - WordNet
- (a) Entwickle zunächst eine Facettenklassifikation für diese Begriffe. Verwende 3–4 Facetten.
- (b) Entwickle als Alternative zur Klassifikation (a) eine monohierarchische Klassifikation der Begriffe mit mindestens 8 Klassen und Unterklassen. Vergleiche die Schwierigkeiten, die bei den unterschiedlichen Ansätze auftreten.
- (c) Ergänze die Liste der Begriffe um andere Begriffe aus dem Bereich der Wissensrepräsentation. Musst Du Deine Klassifikationen anpassen?

20 Punkte

**Aufgabe 10: Thesaurus**

Entwickle einen Thesaurus-Ausschnitt für das Gebiet „Studentenleben“. Er soll mindestens 15 Deskriptoren enthalten und mindestens 3 Ebenen tief sein (bezüglich der hierarchischen Relationen). Benutze zur Strukturierung die Relationen *Oberbegriff*, *Unterbegriff*, *Verwandter Begriff*, *Benutze Synonym* und *Benutzt für Synonym*.

Stelle den Thesaurus in geeigneter Weise alphabetisch dar.

10 Punkte