

**Information Retrieval - Sommer 2006**  
Dipl.-Inform. Sascha Kriewel, LF 137  
kriewel@is.informatik.uni-duisburg.de

## Übungsblatt 4                      Repräsentation von Textinhalten

---

### **Aufgabe 9:    Klassifikationen und Webkataloge**

Betrachte den Webkatalog unter <http://www.dmoz.org/>. Das Open Directory Project benutzt ein Klassifikationssystem, um die katalogisierten Webseiten zu ordnen und zugreifbar zu machen.

Untersuche die Klassifikation in Hinblick auf die in der Vorlesung genannten Merkmale und Kriterien. Um was für eine Art von Klassifikation handelt es sich? Versuche auch, Schwächen der konkreten Klassifikation herauszuarbeiten.

Angenommen, Du suchst nach Dokumenten zur Artussaga. Versuche zunächst diese über die Klassifikation zu finden. Benutze anschliessend die Suche. Welche Dokumente hast Du bei der Suche zusätzlich gefunden, bzw. welche Dokumente aus der Klassifikation fehlen im Ergebnis der Suche? Welches Herangehen war insgesamt erfolgreicher?

### **Aufgabe 10:    Thesauri**

**Abgabe bis 11. Mai 2006, 12 Uhr:**

[iruebg-abgaben@is.informatik.uni-duisburg.de](mailto:iruebg-abgaben@is.informatik.uni-duisburg.de)

Entwickle einen Thesaurus-Ausschnitt für das Fachgebiet „Informatik“. Er soll mindestens 20 Deskriptoren (z.B. „Theoretische Informatik“, „Betriebssysteme“, „Algorithmen“, etc.) 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.

### **Aufgabe 11:    RDF**

Lies die Einführung zu RDF unter <http://www.w3.org/TR/rdf-primer/> und versuche, die folgenden Begriffe zu erklären:

- Was ist ein *blank node*?
- Was versteht man unter *reification* (Reifizierung)?

Benutze die graphische Notation aus der Vorlesung, um die folgenden Aussagen in RDF-Schema zu modellieren.

„Kraftfahrzeuge sind auf Personen zugelassen. PKW, Bus und Mini-Van sind Arten von Kraftfahrzeugen. Ein Kraftfahrzeug hat eine Marke, eine Farbe und

ein Kennzeichen. Eine Person hat einen Namen und eine Ausweisnummer.”

Modelliere dazu auch folgende Aussagen:

„Der rote GMC Savana mit dem Kennzeichen 'E-DU 1111' ist ein Mini-Van.  
Der Wagen ist Hans Müller zugelassen.”