

Hans Uszkoreit Vorlesung  
Informationswissenschaft  
und Informationssysteme



# Übung zur Vorlesung im Sommersemester 2009

**Indexierung, Suche, Relevanz**

**Fachrichtung 4.7 Allgemeine Linguistik  
an der Universität des Saarlandes**

**Betreuung: Brigitte Jörg**

### **Aufgabe 68: Informationserschließung**

Die Erschließung von Dokumenten durch die Indexierung kann mittels Notationen, Stich- und Schlagwörter, sog. Deskriptoren bzw. Nichtdeskriptoren erfolgen. Beschreiben Sie kurz die Unterschiede in der jeweiligen resultierenden Terminologie bzw. dem Index.

### **Aufgabe 69: Indexierung / Koordinierung**

Welche Möglichkeiten zur Koordinierung der Indexterme kennen Sie?

### **Aufgabe 70: Indexierungsmethoden**

Indexierungsmethoden unterscheiden sich durch die im Index zugelassenen Terme. Nennen Sie verschiedene Indexierungsmethoden und beschreiben Sie das zugehörige Vokabular.

### **Aufgabe 71: Linguistische Indexierungsverfahren**

Im Rahmen der Computerlinguistik sind insbesondere die automatischen sowie linguistisch-basierten Indexierungsverfahren von Interesse. Beschreiben Sie in wenigen Worten die Unterschiede der einzelnen Verfahren hinsichtlich der Analyse (1) Morphologisch-lexikalisch (2) Morphosyntaktisch (3) Semantisch.

### **Aufgabe 72: IDX System**

Am DFKI ist im Moment das IDX System im Einsatz, was zur Generierung eines Index verschiedene linguistische Komponenten einsetzt. Das System durchläuft verschiedene Phasen:

- Phase 0: Grundformermittlung
- Phase M: Mehrwortbegriffe
- Phase B: Getilgte Teilwörter
- Phase 1: Strukturanalyse
- Phase 2: Dekomposition
- Phase 3: Dervation
- Phase G: Thesaurus
- Phase T: Übersetzung

Geben Sie für jede Phase zwei Generierungsbeispiele.

### **Aufgabe 73:      Boolesche Suche**

Suche basiert immer auf einen Index und ist daher abhängig von dessen Struktur und Inhalt. Zudem erfolgt Suche immer in Abhängigkeit von der gestellten Anfrage. Die Ergebnisse einer Suche sind daher immer von der Anfrage und dem Index abhängig. Das bekannteste Suchverfahren ist das Boolesche. Welche Suchanfragen lassen sich damit formulieren? Geben Sie jeweils Beispiele für möglichen Anfrage und kommentieren Sie kurz die Unterschiede in den Ergebnissen.

### **Aufgabe 74:      Fuzzy Suche**

Was unterscheidet eine Fuzzy-Anfrage von einer Booleschen Anfrage ?

### **Aufgabe 75:      Relevanz / Google PageRank**

Worauf basiert grundsätzlich das PageRankverfahren von Google?

### **Aufgabe 76:      Relevanz / Ranking**

Erklären Sie in eigenen Worten die Bedeutung von Relevanz im Kontext von Suchergebnissen.