

## Aufgabenblatt 10: Imaging, Korpusbasierte Methoden

### Aufgabe 1 Imaging-Verfahren nach Crestani und Rijsbergen

- a) Wie wird im System von Crestani und Rijsbergen versucht die Probleme, die sich bei logikbasierten IR Systemen ergeben, zu lösen?
- b) Wie werden Terme in diesem System charakterisiert?
- c) Welche Daten werden in dem System benötigt?
- d) Was liefert das System als Ergebnis?

### Aufgabe 2 Der Assoziative Ansatz

- a) Welche Annahme liegt dem assoziativen Ansatz zugrunde?
- b) Was ist eine Term-Term-Matrix?
- c) Was ist ein assoziativer Thesaurus, wie hängt er mit einer Term-Term-Matrix zusammen?
- d) Wie kann eine Term-Term-Matrix aus einer Dokument-Term-Matrix berechnet werden? Wie können die Einträge in einer Term-Term-Matrix interpretiert werden?

### Aufgabe 3 Kookurrenzmethoden im IR

- a) Wie kann ein assoziativer Thesaurus im Information Retrieval genutzt werden?
- b) Welche Probleme wurden dabei beobachtet?
- c) Wie können diese Probleme angegangen werden?

### Aufgabe 4 Retrieval in mehreren Sprachen

- a) Was versteht man unter "Cross Language Retrieval"?
- b) Wie können assoziative Methoden im Cross Language Retrieval eingesetzt werden? Welche Daten werden dafür benötigt?
- c) Wie wurden diese Methoden von Sheridan und Ballerini angewendet und überprüft?

### Aufgabe 5 Automatische Indexierung

- a) Wie können assoziative Verfahren zur automatischen Indexierung genutzt werden?
- b) Welche Vorteile haben sie gegenüber Gewichtungsmethoden?
- c) Wie können sie zur Indexierung mit einem kontrollierten Vokabular verwendet werden?