

13. Übungsblatt

Ausgabe: 26. Januar 2015 · Besprechung: 2. Februar und 4. Februar 2015

1 Suche

1. Warum sind relationale Datenbanken für die Suche in unstrukturierten Datenbeständen nicht geeignet?
2. Was ist das Ziel von Ranking? Wie können Rankings berechnet werden?
3. Erklären Sie hierzu die Idee des Vektorraummodells und des tf/idf Rankings!
4. Erklären Sie die Begriffe Recall and Precision! Wie hängen Sie zusammen?

2 PageRank

1. Erklären Sie die Funktionsweise des PageRank Algorithmus!
2. Welche Verarbeitungsschritte muss eine Suchmaschine auf gesammelte Dokumente anwenden, damit Suchanfragen darauf gestellt werden können?

3 Invertierte Dateien

1. Wie sind invertierte Dateien aufgebaut?
2. Wie können wir invertierte Dateien zur Beantwortung von Suchanfragen nutzen?
3. Warum werden invertierte Dateien oft sehr groß? Wie kann die Größe reduziert werden?