

# Konzepte des Information Retrieval

Dauer: 20 Minuten

## Vorbereitungsphase

Wie lange sollte man sich auf die Klausur vorbereiten?

2-3 Tage

Wie hast du dich vorbereitet (allein, in der Gruppe)?

überwiegend allein

Welche Vorlesungen gingen der Prüfung voraus / waren hilfreich?

Konzepte des Information Retrieval

Welche Literatur/Skripte waren hilfreich?

Skript und Übungsbögen

Gibt es allgemeine Tipps, die bei der Vorbereitung helfen könnten?

- Skript sehr genau durcharbeiten.
- Sich evtl. erst einmal eine schriftliche Zusammenfassung machen
- Anhand der Zusammenfassung Detailfragen, mit Hilfe des Skriptes, klären und verstehen
- Überlegen, welche Fragen gestellt werden könnten
- Ganz wichtig: Überblick behalten und Zusammenhänge verstehen.

Wo lagen Deiner Meinung nach besondere Schwierigkeiten der Klausur?

Im allgemeinen gibt es keine große Schwierigkeit, wenn man sich gut vorbereitet.  
evtl. Detailfragen zu gewissen Bereichen, bei mir Fuzzy-Modell :)

## Verlauf der Prüfung

Wie verlief die Prüfung?

Fragen wurde aufeinander aufgebaut. 1. Frage nach IR-Systeme, 2. Frage nach Parsen/Pre- Process-Methoden  
3. Frage nach Index usw.

Prüfer geht auf Antworten ein und versucht so einen Faden durch das Thema zu ziehen.

Formeln wurden nicht direkt nachgefragt jedoch, nach Anbieten meinerseits, angenommen und bewertet.

Wie reagierte der Prüfer, wenn Fragen nicht sofort beantwortet wurden?

- Gibt Tipps und Hinweise.
- Formuliert seine Frage nochmal um und lenkt in eine bestimmte Richtung

Dein Kommentar zur Prüfung:

Eine sehr angenehme Prüfung. Sehr nette Prüfer/Prüfling-Atmosphäre. Nimmt einen sofort die Nervosität.

Dein Kommentar zur Benotung:

fair :) Obwohl ich bei manchen Sachen Schwierigkeiten hatte es auszudrücken und nicht sofort drauf kam, hat es für ein "sehr gut" gereicht.

Welche Fragen wurden konkret gestellt?

- Was ist ein IR-System? Wie aufgebaut?
- Was genau macht man beim "Pre-Process"?

- Was sind Stopwords und wie findet man sie? (Zipf-Verteilung)
- Wie kodiert man Dokumente? (Index)
- Welche Methoden gibt es dort?
- tf\*idf: Was genau geschieht da?
- Wie berechnet man die Ähnlichkeit von Dokumenten?
- Gibt es Probleme bei gewissen Ähnlichkeitsberechnungen? (eukld. Abstand)
- Propabilistisches Modell: Wie berechnet man dort den Rang der Dok.?
- Unterschied zu tf\*idf?
- Fuzzy-Modell: Was geschieht dort? Korrelationsmatrix usw.
- WWW-Suchmaschinen: Webcrawler, PageRank,
- Wie funktioniert PageRank genau: Formel aufschreiben und erklären
- Kurze Frage zu Multimedia-Retrieval: Was geschieht dort? Unterschied zu Dokumenten-Retrieval?