

# Internet-Suchmaschinen: Beispielfragen

Norbert Fuhr

17. Juli 2013

- Einführende Fragen: Erzählen sie 3 Minuten etwas zu folgendem Thema. Gehen sie dabei auf die zu lösenden Probleme und die behandelten Lösungsansätze ein.
  - Wie funktioniert eine ISM (Internet-Suchmaschine)?
  - Wissensrepräsentation für Textinhalte
  - Evaluierung im IR
  - IR-Modelle
  - Information Seeking und Searching
  - Gestaltung von Benutzerschnittstellen für IR-Systeme
- Wissensrepräsentation:
  - Was ist der Unterschied zwischen Syntax, Semantik und Pragmatik? Auf welcher Ebene spielt sich IR ab? Auf welcher Ebene arbeiten die betrachteten Repräsentationen?
  - Welche Probleme der Freitextsuche werden durch welche Methoden beim informati-schen Ansatz (/beim computerlinguistischen Ansatz) versucht zu lösen?
  - Welche Probleme der Freitextsuche werden durch die vorgestellten Verfahren nicht gelöst?
  - Freitextsuche vs. Dokumentationssprachen
  - Welche der vorgestellten Verfahren erhöhen die Precision, welche den Recall?
  - Nennen Sie Beispiele für die drei Arten von Dokumentationssprachen
  - Ontologien: Was sind die wesentlichen Konstrukte, wie funktioniert Vererbung?
- Evaluierung:
  - Welche Maße gibt es?
  - Wie kann man den Recall abschätzen?
  - Wie entsteht eine Recall-Precision-Kurve?
  - Was ist der Unterschied zwischen System- und Benutzer-orientierten Maßen?
  - Welche Rolle spielt bei letzteren das Abbruchverhalten?
  - Was ist Diversity ranking, und wie kann man es evaluieren?
- Nicht-Probabilistische Modelle
  - Welches sind die Vor- und Nachteile des Booleschen Retrieval?

- Was ist die Grundidee des Vektorraummodells?
- Was versteht man unter Relevance Feedback, welche Modelle unterstützen es?
- Was versucht man bei Relevance Feedback im Vektorraum zu optimieren?
- Wieso verwendet Rocchio nicht den 'optimalen Fragevektor'?
- Probabilistische Modelle:
  - Wie funktioniert Relevance Feedback?
  - Was heißt Optimalität beim PRP? Wie kann man das Ranking nach fallender Wahrscheinlichkeit rechtfertigen?
  - Was ist die Grundidee des interaktiven PRP?
  - Was ist die Grundidee des BIR-Modells, und wie kann man die benötigten Parameter schätzen?
  - Inwiefern erweitert BM25 das BIR-Modell, welche Parameter werden hier zusätzlich berücksichtigt?
  - Was ist die Grundidee von statistischen Sprachmodellen?
  - Was ist die Grundidee von learning to rank?
  - Wie kann man diversity Ranking erzeugen?
- Interaktives Retrieval
  - Welche Ebenen kann man bei der Architektur von IR-Systemen unterscheiden?
  - Was ist der Unterschied zwischen Seeking und Searching?
  - Erläutern Sie das Ellis/Meho/Tibbo-Modell, und wie man es in IR-Systemen unterstützen kann?
  - Was ist das klassische Modell für den Suchprozess, und was sind seine Nachteile?
  - Was besagt die ASK-Hypothese?
  - Was versteht man unter Polyrepräsentation, wie kann ein System sie unterstützen?
  - Wozu dienen Taxonomien für Web-Suchaktivitäten?
  - Welche Arten von Suchaktivitäten kann man unterscheiden? Nennen Sie jeweils Beispiele!
  - Welche Stufen der Systembeteiligung gibt es? Nennen Sie Beispiele bei ISM!
- Benutzerschnittstellen
  - Welche Arten von Eingabeschnittstellen gibt es?
  - Nennen Sie Grundregeln für Anfrageschnittstellen
  - Nennen Sie Beispiele für Visualisierungen für die verschiedenen Elemente der Suche
  - Was versteht man unter Mixed-Initiative-Interfaces? Nennen Sie Beispiele