

openHPI: **Wie funktioniert eine Suchmaschine?**
Ranking: Rankingfaktoren

Dr. Ralf Krestel
Hasso-Plattner-Institut
Universität Potsdam

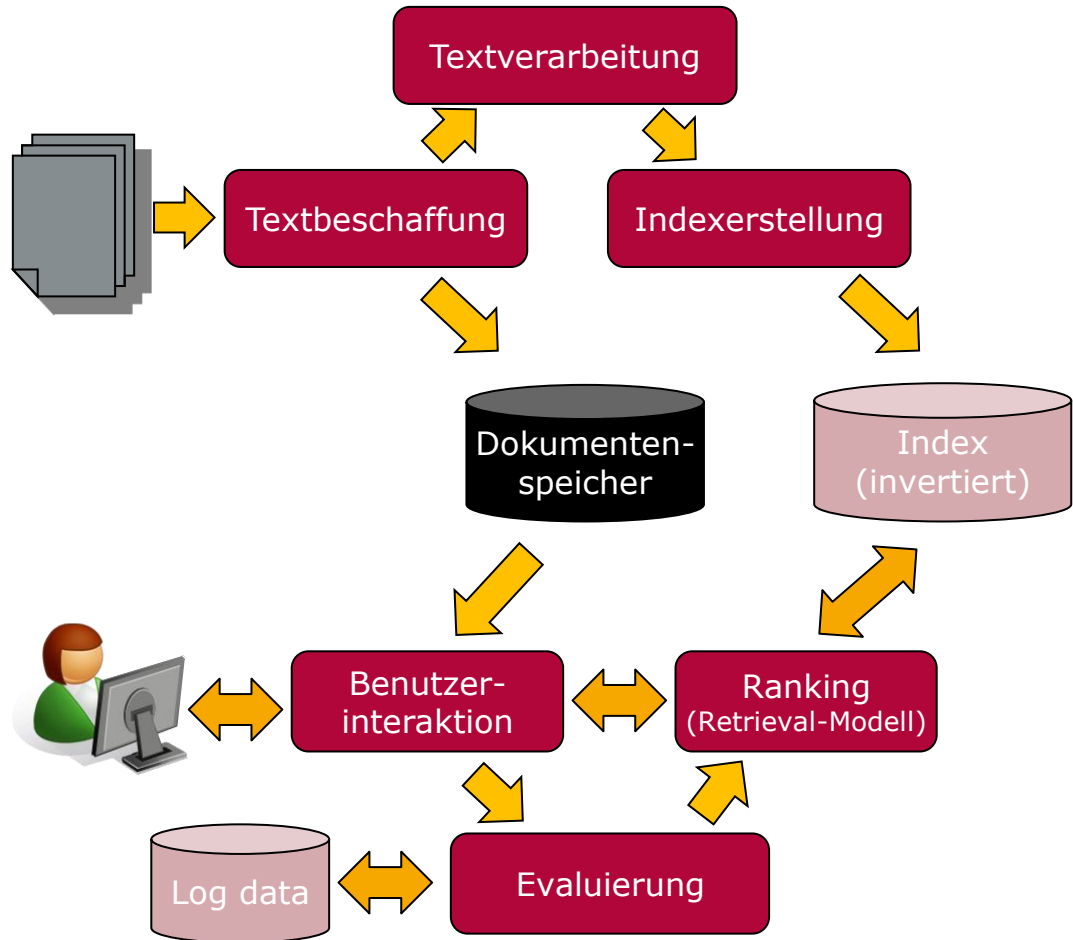
■ Grundlegende Bausteine

□ Indexierung

- Textbeschaffung
- Textverarbeitung
- Indexerstellung

□ Anfrage

- Benutzerinteraktion
- Ranking
- Evaluierung



- Ranking beruht auf Relevanz der Ergebnisse
- Relevanz subjektiv (Informationsbedürfnis)
- → Es gibt **nicht** das **eine, perfekte Ranking**

- Suchmaschine muss Kompromiss finden, um möglichst viele Nutzer zufriedenzustellen
 - Z.B. Abwägen zwischen Relevanz und Aktualität

- Viele Faktoren können/sollten berücksichtigt werden
 - Suchmaschinen allgemein
 - Textspezifische Faktoren
 - Aktualität
 - Personalisierung
 - Websuche
 - Popularität
 - Lokalität
 - Technische Faktoren

- Neben **Häufigkeit** eines Wortes auch **Struktur** der Dokumente interessant
 - Kommt Suchwort im
 - Titel vor?
 - Oder in einer Überschrift?
 - Am Anfang oder am Ende?
 - Kursiv oder fett gedruckt?
- **Synonyme** des Suchbegriffs
- **Reihenfolge** der Wörter kann auch wichtig sein
 - Paris Hilton vs. Hilton Paris





■ Google Startseite vom 11.09.2001

- „Unter Personalisierung versteht man die Anpassung eines Objekts an die Bedürfnisse eines Subjekts.“ (Riemer und Brüggemann 2009)
- Für Suchmaschinen
 - Anpassen der **Benutzeroberfläche**
 - Anpassen der **Suchergebnisse**
- Profil eines Nutzers
 - Explizite Informationen
 - Soziale Netzwerke
 - Anfragehistorie
 - Klickhistorie



- Im Webkontext sehr wichtig, da keine Eingangsqualitätskontrolle
- Nur sinnvoll im Zusammenhang mit textbasierten Faktoren
- **Linkbasiert** (später im Kurs) und **Nutzungsbasiert**
- Auswerten des **Nutzerverhaltens**
 - Interaktion mit der Suchmaschine durch Log-Dateien
 - **Klicks** auf Suchergebnisse im Ranking
 - **Verweildauer** auf angeklickten Seiten
 - Analyse des Nutzerverhaltens außerhalb der Suchmaschinenseite
 - Toolbars
 - Eigene Browser
 - Webprotokolle
 - Analysedienste

Ranking: Rankingfaktoren

Lokalität

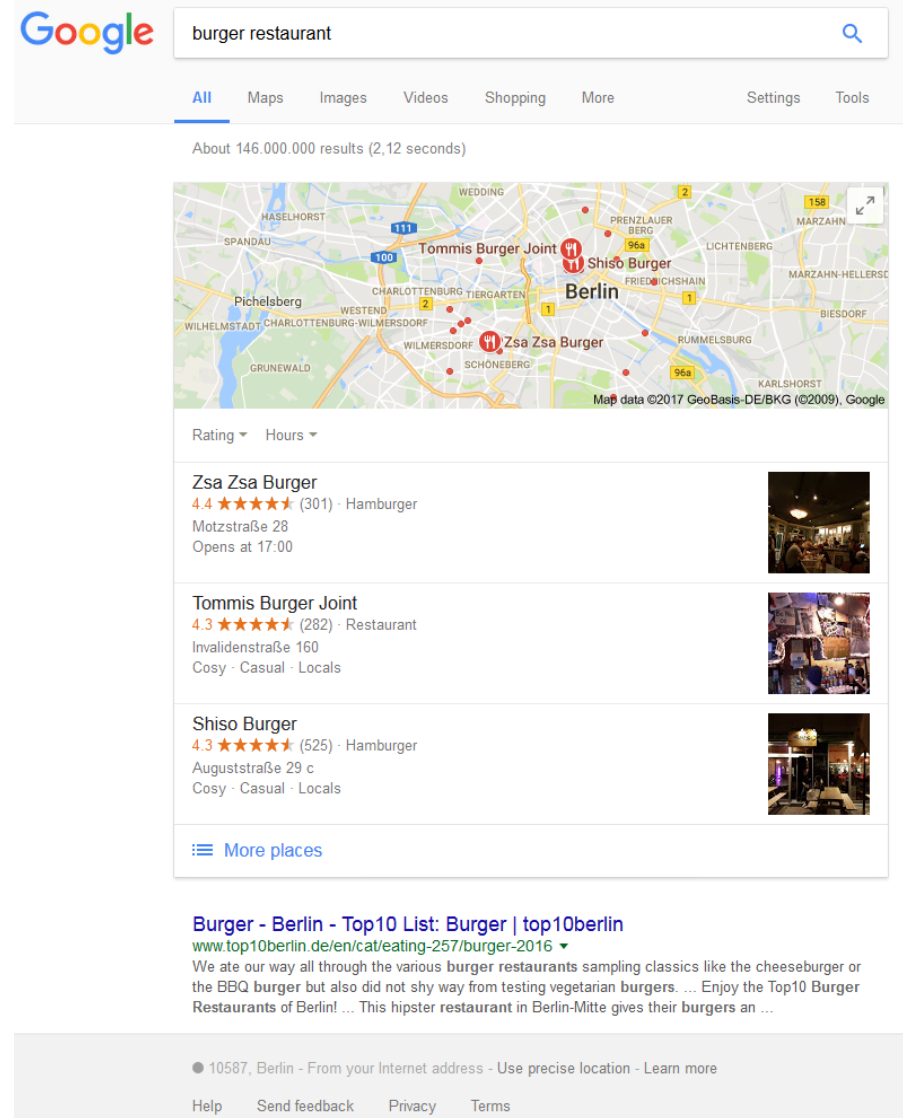
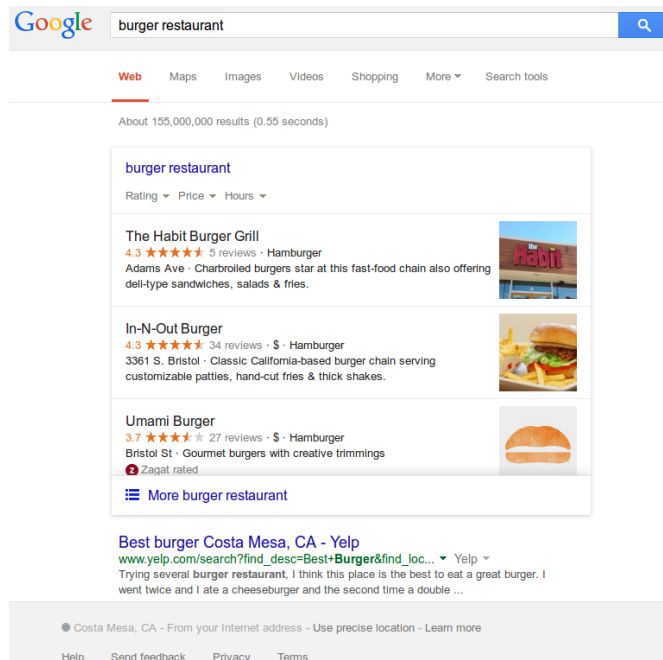
■ Standort des Nutzers

□ Land/Sprache

- Z.B. „Bundeskanzler“ bei google.de vs. google.at

□ Region/Stadt

- Z.B. „burger restaurant“



- Erreichbarkeit, Verfügbarkeit
 - Ladegeschwindigkeit
 - Webseitenprogrammierung
- Zugangsbeschränkungen
 - Z.B. Darstellung auf mobilen Endgeräten
 - Bezahlshranken

