

## Wer sucht, der findet...

### Eine Anleitung zur Suche im Web

Too much time in the web?



© 1995 WWW.DICKYBOONIES.COM

www.google.at



## Informationsexplosion oder falsche Suchstrategie?

Resultate 1 - 10 von ungefähr 1,500,000 Suchdauer: 0.19 Sekunden

beides!

FH Fachhochschul-Studiengang  
INFORMATIONSBERUFE  
Informations & Knowledge Management

## Suchstrategien...

... mit Konzept ans Suchen

FH Fachhochschul-Studiengang  
INFORMATIONSBERUFE  
Informations & Knowledge Management

## Suchansätze

- **intuitive Suche:**  
direkte Suche nach Web-Sites  
(durch Probe einer URL)
- **Schneeballsystem:**  
kumulative Suche
- **„Bergsteigen“:**  
systematische Suche
- **strukturierte Suche:**  
Begriffe ordnen
- **taktische Suche:**  
zentrale Seite als Ausgangspunkt

© Berthold Meier

FH Fachhochschul-Studiengang  
INFORMATIONSBERUFE  
Informations & Knowledge Management

## »best practice«

1. Thema definieren
  - Stichworte, Sprache, Zeitraum
2. Abfragephrasen (query sets) überlegen
3. Quellen auswählen
  - Blick in online-Hilfe sowie Sitemap
4. Suche durchführen
5. Ergebnis liefert Rückschuß
  - Stichworte richtig gewählt
  - Suchwerkzeug geeignet

FH Fachhochschul-Studiengang  
INFORMATIONSBERUFE  
Informations & Knowledge Management

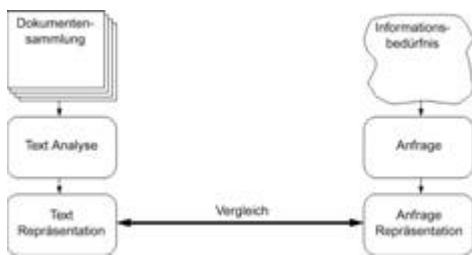
## Suchmaschinen...

... oder: wie man die Nadel(n) im Heuhaufen findet

## Allgemeines

- Inhalte automatisiert gesammelt
- basiert auf einer Datenbank, die zwei „Listen“ enthält:
  - Stichwort-Index
  - Seiten-Index
- ein Suchroboter „crawler“
  - füttert die Datenbank
  - und wertet die Benutzeranfragen aus

## Funktionsweise einer Suchmaschine



© Dieter Merkl

## Aufnahme in den Index

- Suchroboter
- Link-Verfolgung
- Seite indizieren
- Wortextraktion
  - Analyse von Meta-Tags
  - Entfernen von Stopwörtern
  - Ermittlung der Position im Text
- Liste von Stichwörtern aktualisieren

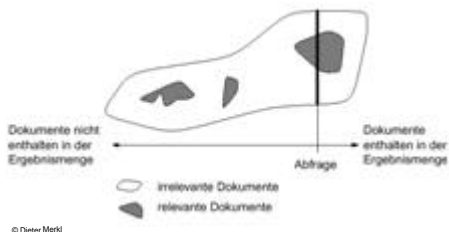
## Erstellung einer Trefferliste

- Auswerten der Suchanfrage (Query)
- Vergleich des Index mit Suchanfrage
- Relevanzbewertung
  - Wörter im Titel, Meta-Zeilen, ...
  - Anzahl der Links auf einer Seite
  - Aktualität
- Relevance Feedback

## Information Retrieval

- Werkzeuge und Methoden zur Recherche u. Informationssuche selbst
- Maßzahlen:
  - **Recall** = Summe der gefundenen relevanten Dokumente / Summe der relevanten Dokumente in der Datenbank
  - **Precision** = Summe der gefundenen relevanten Dokumente / Summe der gefundenen Dokumente

## Bewertung von Suchergebnissen



© Dieter Merkl

FH Fachhochschul-Studiengang  
INFORMATIONSBERUFE  
Informations & Knowledge Management

## Beispiele

- international:
  - Google (<http://www.google.com>)
  - Altavista (<http://www.altavista.com>)
  - Hotbot (<http://www.hotbot.com>)
  - Lycos (<http://www.lycos.com>)
  - AlltheWeb (<http://www.alltheweb.com>)
- deutschsprachig:
  - Astronaut (<http://www.astronaut.at>)
  - Klammer@ffe (<http://www.klammeraffe.at>)
  - Web.de (<http://suche.web.de/search/>)
  - Fireball (<http://www.fireball.de>)

FH Fachhochschul-Studiengang  
INFORMATIONSBERUFE  
Informations & Knowledge Management

## erweiterte Suche

- Boole'sche Operatoren
- Trunkierung/Maskierung
- Suchschrittverknüpfung
- mehrsprachige Suche
- Einschränkung auf Domains
- Suche nach Medien-Elementen
  - Bild, Video, Audio

FH Fachhochschul-Studiengang  
INFORMATIONSBERUFE  
Informations & Knowledge Management

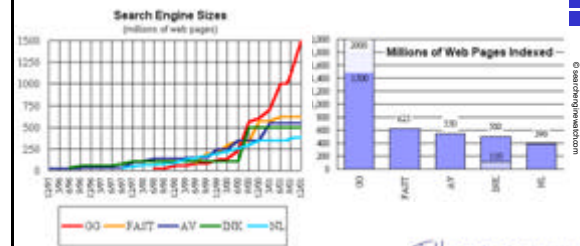
## erweiterte Suche (Google)

FH Fachhochschul-Studiengang  
INFORMATIONSBERUFE  
Informations & Knowledge Management

## erweiterte Suche (Altavista)

FH Fachhochschul-Studiengang  
INFORMATIONSBERUFE  
Informations & Knowledge Management

## Datenbank-Größen der Suchmaschinen



FH Fachhochschul-Studiengang  
INFORMATIONSBERUFE  
Informations & Knowledge Management

## Webkataloge...

... die Inhaltsverzeichnisse des  
World Wide Web

## Allgemeines

- von Redakteuren manuell erstellt
- Einteilung in hierarchisches Kategorien-System (Themen)
- Kurzrezensionen
- Vorteil:
  - „intellektuelle“ Informationsauswahl
- Problem:
  - Aktualisierung und Link-Überprüfung

## Beispiele

- <http://www.yahoo.com>
- <http://dmoz.org>
- <http://www.vlib.org>
- <http://www.bubl.ac.uk/link/>

## Portale...

... laden zum Informations-  
Bummel ein

## Allgemeines

- Navigationsstruktur und graphische Aufmachung nach Anbieter sehr unterschiedlich
- Suchfunktion
- Sitemap
- Personalisierung: Anpassung des Informationsangebotes an persönliche Präferenzen

## Beispiele

- <http://www.orf.at>
- <http://www.ptt.at>
- <http://www.aon.at>
- <http://www.schooltalk.at>
- <http://www.fhf.at>
- <http://www.cnn.com>
- <http://www.sourceforge.net>
- <http://www.ebay.com>

## Beispiel-Recherche

## Beispiel-Recherche

Thema: „11. September“

## Mögliche Suchkonzepte

World Trade Center, NYC, 11. Sept. 2001

- Terror, Angriff, Osama Bin Laden, Al-Qaeda
- Medien, Bericht, Reportage, Manipulation, Sensationslust
- Menschen, Verletzte, Tote, Leid, Trauer, Obdachlosigkeit, psychische Betreuung
- Wirtschaft, Folgen, Aktien, Absturz

## sonstige Quellen...

... für »Sonderwünsche«

## Zeitungsarchive

- Artikelsuche
  - nach Rubrik
  - in einem bestimmten Zeitraum
  - nach Autor
- tlw. kostenlose Anmeldung erforderlich
- Beispiele:
  - <http://www.derstandard.at>
  - <http://www.profil.at>
  - <http://www.derspiegel.de>
  - <http://www.time.com/>

## Newsgroups & Mailinglisten

- elektronische Nachrichtenforen
  - Newsgroups (weltweit):
    - gruppiert nach Themenbereichen:
      - rec. (Recreation), comp. (Computer), sci. (Naturwissenschaften), soc. (Soziales, Kultur)
    - und Ländern:
      - de., at., us. usw.
  - Mailinglisten (spezielle Interessensgebiete)

## Beispiele

### Newsgroups

- <http://groups.google.com/>
- <http://netnews.web.de/>
- <http://www.findolin.de>

### Mailinglisten

- <http://www.meta-list.net>
- <http://www.liszt.com/>
- <http://www.lisde.de>

## Adreßdatenbanken, Stadtpläne

### Adreß-/Telefonnummernsuche

- Super Pages (<http://www.superpages.at>)
- Telekom Austria -Tel.buch (<http://www.etb.at>)
- Four11 (<http://www.four11.com>)
- Internet Address Finder (<http://www.iaf.net>)

### Landkarten, Straßensuche

- map-o-rama (<http://www.maporama.com>)
- Plan.at (<http://www.plan.at>)

## Domainnamen-DB

- Suche nach diversen Top-Level- und länderspezifischen Domains
- Verzeichnisse:
  - <http://www.nic.at>
  - <http://www.ripe.net>
  - <http://www.uwhois.com>

## Spezialwerkzeuge

## Meta-Suchmaschinen

- versch. Suchmaschinen / Portale etc. werden über ein Suchinterface abgesehen
  - Vorteil: übergreifende Recherche
  - Nachteil: spezielle Suchsyntax der einzelnen Datenquellen berücksichtigt?
- Bsp.: Copernic2000 (<http://www.copernic.com>)

## Suchagenten

- durchsuchen eigenständig das Netz
- Datenbasis:
  - Stichworte
  - Interessenprofil des Nutzers
  - Anzahl der Links auf eine Seite
  - Anzahl der Zugriffe
- Beispiel: Alexa (<http://www.alexa.com>)