

# ***Unbeschränkter Zugang zu Wissen?***

*Leistungsfähigkeit und Grenzen von Suchdiensten  
im Web. Zwischen informationeller Absicherung und  
manipulierter Information*

*Joachim Griesbaum*



Informationswissenschaft Universität Konstanz

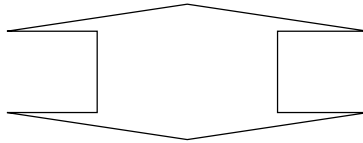
## ***Inhalt***

- ▶ **Google & Co.: Pfadfinder oder Flaschenhalse im Datenschungel des Internet?**
- ▶ Suchdienstetypen - Erschließung der Wissensbestände, Sortierverfahren, Rankingkriterien
- ▶ Ausprägung des Suchdienstemarktes
- ▶ Ergebnis & Schlussfolgerungen



## Google & Co.: Pfadfinder oder Flaschenhalse des Internet?

Mit Hilfe des Internet ist für einen Großteil der Bevölkerung Zugriff auf Information in bisher nicht gekanntem Ausmaß möglich. Theoretisch steht eine Vielzahl von Suchdiensten bereit, um benötigte Informationen zu erschließen.



Faktisch werden meist nur wenige marktführende Suchmaschinen und Webverzeichnisse verwendet. Die informationelle Absicherung der Nutzer wird somit in vielen Fällen von den populären Suchdiensten präterminiert.



## Google & Co.: Pfadfinder oder Flaschenhalse des Internet?

Google ist gegenwärtig der dominierende Suchdienst

- ▶ Die „American Dialect Society“ nominierte das „Verb“ „google“ als nützlichstes Wort des Jahres 2002.  
[<http://www.americandialect.org/woty.html>]
- ▶ Goedart Palm, **Die Welt ist fast alles, was Google ist**. Zur Macht der Suchmaschinen 28.03.2002.  
[<http://www.heise.de/tp/deutsch/special/auf/12187/1.html> 28.03.2002]

Frage nach der **Macht** und **Auswirkungen**, auf funktionaler Ebene nach den **Fähigkeiten** und **Grenzen** der dominierenden Suchdienste.



## Google & Co.: Pfadfinder oder Flaschenhalse des Internet?

- ▶ Im Folgenden wird versucht das grundsätzliche Leistungsvermögen der dominierenden Suchdienste anhand einer Analyse der Verfahren bzw. Kriterien der Indexierung und Ergebnissortierung zu skizzieren.
  - ▶ Auf welche Inhalte kann theoretisch zugegriffen werden?
  - ▶ Welche Eigenschaften sind für die Sichtbarkeit in den Suchergebnisse maßgeblich?



## Inhalt

- ▶ Google & Co.: Pfadfinder oder Flaschenhalse im Datendschungel des Internet?
- ▶ **Suchdienstetypen - Erschließung der Wissensbestände, Sortierverfahren, Rankingkriterien**
- ▶ Ausprägung des Suchdienstemarktes
- ▶ Ergebnis & Schlussfolgerungen

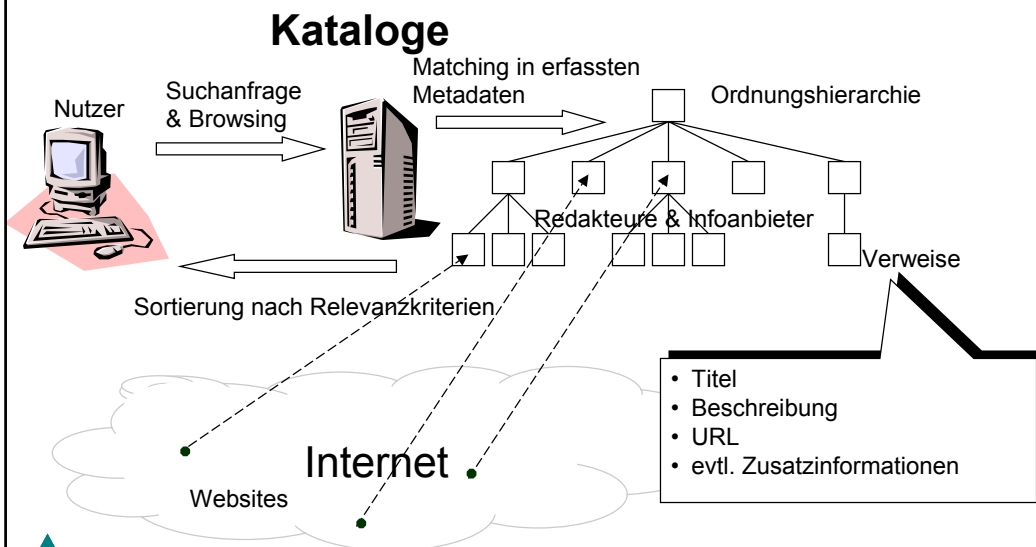


## Suchdienstetypen

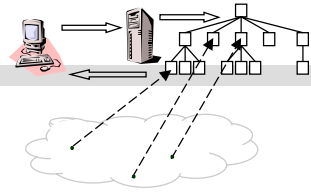
- ▶ Kataloge
- ▶ Suchmaschinen
- ▶ Pay-per-Click-Suchdienste



## Suchdienstetypen - Kataloge



## Suchdienstetypen - Kataloge

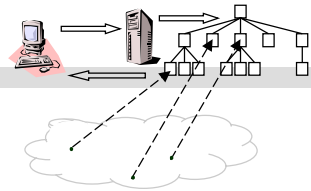


### Erschließung der Wissensbestände

- ▶ redaktionelle Begutachtung ausgewählter Sites
- ▶ Titel, URL, Beschreibungstext, kein Volltext
- ▶ Qualität & Zahlungsbereitschaft
- ▶ Erschließen nur einen sehr kleinen Teil des Internets



## Suchdienstetypen - Kataloge



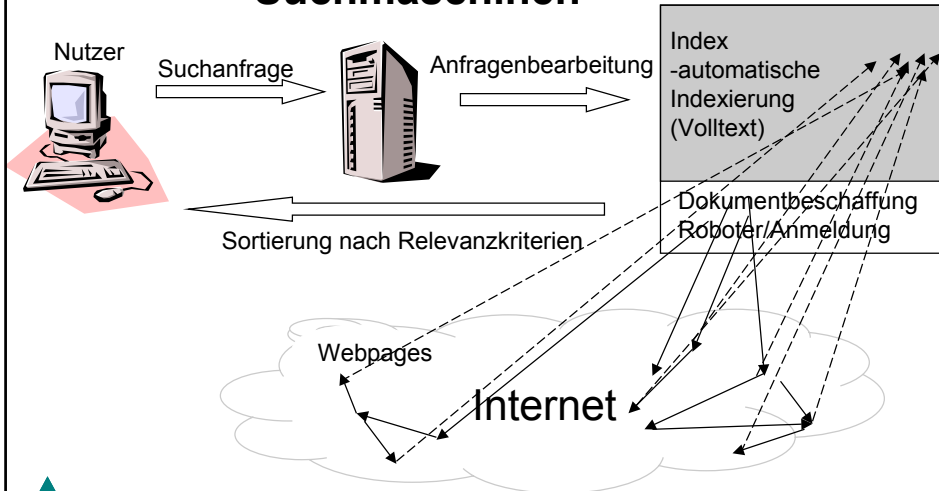
### Sortierverfahren, Rankingkriterien

- ▶ Bei Browsingzugriff hierarchische Zuordnung entscheidend
- ▶ Alphabetisch
- ▶ Wortkookurenzen in den erfassten Metadaten (Position, Häufigkeit, Funktion)
- ▶ Ergebnisse eher von hoher inhaltlicher Qualität



## Suchdienstetypen - Suchmaschinen

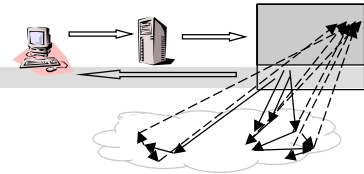
### Suchmaschinen



25. DGI-Online-Tagung: Unbeschränkter Zugang zu Wissen? Leistungsfähigkeit und Grenzen von Suchdiensten im Web. Zwischen informationeller Absicherung und manipulierter Information

11

## Suchdienstetypen - Suchmaschinen



### Erschließung der Wissensbestände

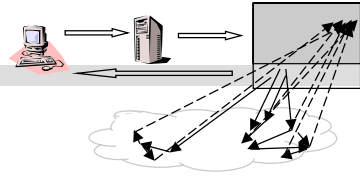
- ▶ Automatische Erfassung durch Roboterprogramme
- ▶ Volltextinvertierung
- ▶ ca 30-40% des Surface Web
- ▶ Problembereiche: Deep Web - anbieterspezifische Datenbanken (professionelle Informationsanbieter) und/oder Zugangsbeschränkungen
- ▶ Paid-Inclusion, Trusted Feed, Indexierung gegen Bezahlung



25. DGI-Online-Tagung: Unbeschränkter Zugang zu Wissen? Leistungsfähigkeit und Grenzen von Suchdiensten im Web. Zwischen informationeller Absicherung und manipulierter Information

12

## Suchdienstetypen - Suchmaschinen



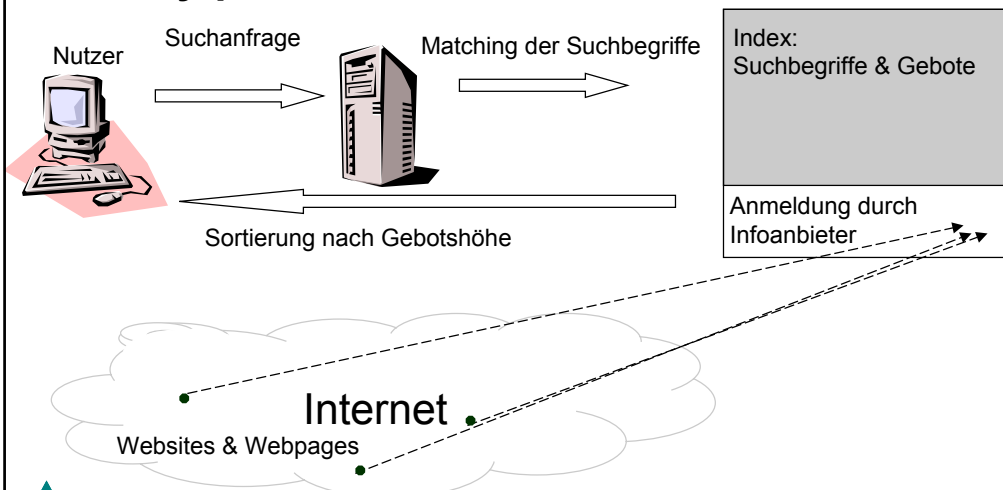
### Sortierverfahren, Rankingkriterien

- ▶ Wortkookurrenzen in Bezug auf Position, Häufigkeit, Funktion, Formatierung >>> Manipulationspotential
- ▶ Analyse von Referenzstrukturen (Pagerank)
  - ▶ populäre Seiten werden bevorzugt
  - ▶ Anpassung im Linkverhalten (Linkfarmen, Linkverkauf)
- ▶ Weiterhin Manipulationspotential vorhanden

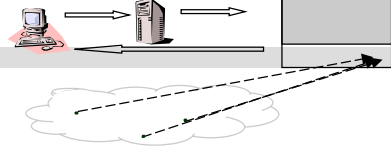


## Suchdienstetypen - Pay-per-Click

### Pay-per-Click Suchdienste



## Suchdienstetypen - Pay-per-Click

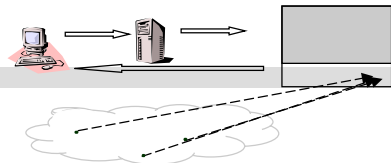


### Erschließung der Wissensbestände

- ▶ redaktionelle Begutachtung angemeldeter Seiten
- ▶ Suchbegriffe
- ▶ Zahlungsbereitschaft & Qualität
- ▶ Erschließen nur einen sehr kleinen Teil des Internets



## Suchdienstetypen - Pay-per-Click



### Sortierverfahren, Rankingkriterien

- ▶ Nach Gebotshöhe (& Klickhäufigkeit - Google)
- ▶ Redaktionelle Kontrolle sichert eine grundlegende Einschlägigkeit, bzw. Relevanz
- ▶ Qualität abhängig davon, inwieweit die Zahlungsbereitschaft der Anbieter mit der Relevanz ihrer Angebote korrespondiert





## Inhalt

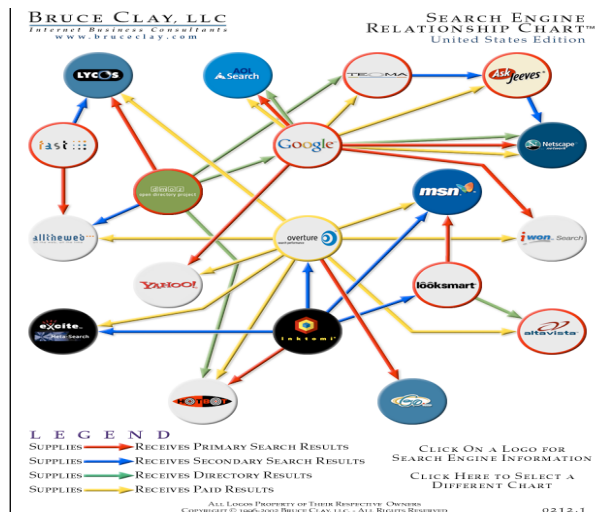
- ▶ Google & Co.: Pfadfinder oder Flaschenhalse im Datenschwung des Internet?
- ▶ Suchdienstetypen - Erschließung der Wissensbestände, Sortierverfahren, Rankingkriterien
- ▶ **Ausprägung des Suchdienstemarktes**
- ▶ Ergebnis & Schlussfolgerungen



25. DGI-Online-Tagung: Unbeschränkter Zugang zu Wissen? Leistungsfähigkeit und Grenzen von Suchdiensten im Web. Zwischen informationeller Absicherung und manipulierter Information

17

## Ausprägung des Suchdienstemarktes



[Quelle Stand Dezember 2002  
www.bruceclay.com/searchenginechart.pdf]



25. DGI-Online-Tagung: Unbeschränkter Zugang zu Wissen? Leistungsfähigkeit und Grenzen von Suchdiensten im Web. Zwischen informationeller Absicherung und manipulierter Information

18

## Inhalt

- ▶ Google & Co.: Pfadfinder oder Flaschenhalse im Datenschlingel des Internet?
- ▶ Suchdienstetypen - Erschließung der Wissensbestände, Sortierverfahren, Rankingkriterien
- ▶ Ausprägung des Suchdienstemarktes
- ▶ **Ergebnis & Schlussfolgerungen**



## Ergebnis & Schlussfolgerungen

- ▶ Anbieterseitig kann weniger von einem Google-Monopol als eher von einem Yahoo, Overture, Google-Oligopol gesprochen werden.
- ▶ Nutzer können auf große **Teilbereiche des Surface Web** zugreifen. Datenbankbasierte Inhalte größerer Wissensbasen werden jedoch nur zu einem geringen Teil erfasst.
- ▶ Zusammensetzung der Indizes & Sortierung der Ergebnisse wird neben technischen und inhaltlichen Kriterien zunehmend durch **Marketingstrategien und ökonomische Potenz** der Informationsanbieter sowie der **Vermarktungsbereitschaft** der Suchdienste beeinflusst.



## Ergebnis & Schlussfolgerungen

Auswirkungen auf die Qualität der Ergebnisse?

- ▶ Möglicherweise wird durch die Kommerzialisierung in vielen Suchkontexten eine Qualitätssteigerung bewirkt. (Minimierung von Spam/Suche nach Produkten & Dienstleistungen)
- ▶ Kommerzialisierung der Suchtreffer Voraussetzung für das Weiterbestehen der populären kostenfreien Dienste.



## Ergebnis & Schlussfolgerungen

- ▶ Wichtig scheint es ...
  - ▶ Nicht kommerziell ausgerichtete Inhalte hinreichend zu berücksichtigen.
  - ▶ **Dem Nutzer die Einflussfaktoren auf die Ergebnisse transparent zu machen. Damit dieser fähig ist, die Vollständigkeit, Genauigkeit und Glaubwürdigkeit der Ergebnisse überhaupt einzuschätzen.**
- ▶ Systematische Einbindung geeigneter Deep Web-Suchdienste & Informationsanbietern könnte dazu beitragen, die informationelle Absicherungsfähigkeit der Nutzer zu erhöhen [<http://turbo10.com/>].



Vielen Dank für Ihre Aufmerksamkeit!

