

Deutsche Suchmaschinen im Vergleich: AltaVista.de, Fireball.de, Google.de und Lycos.de

Joachim Griesbaum¹ / Marc Rittberger² / Bernard Bekavac¹

¹Universität Konstanz
Informationswissenschaft
Fach D 87
D-78457 Konstanz

²Heinrich-Heine-Universität Düsseldorf
Informationswissenschaft
Institut für Sprache und Information
D-40225 Düsseldorf

Inhalt

- ▲ **Motivation & Zielsetzung**
- ▲ Vorgehensweise & Konzeption
- ▲ Testanordnung
- ▲ Pretest & Test
- ▲ Ergebnisse & Analyse
- ▲ Einschätzung & Ausblick

Motivation & Zielsetzung

- ▲ Bedeutung der Suchmaschinen als Instrumente der informationellen Selbstbestimmung & Absicherung
Wissenherrschaft der Suchmaschinen?
- ▲ Goedart Palm, Die Welt ist fast alles, was Google ist.
<http://www.heise.de/tp/deutsch/special/auf/12187/1.html>, 28.03.02
- ▲ Zugriff auf vorhandenes Wissen zunehmend mit Hilfe von Suchmaschinen
- ▲ Entscheidende Frage: Wie hilfreich sind die Maschinen tatsächlich?
- ▲ Ansatz: Ermittlung der Qualität der Suchergebnisse - Retrievaleffektivität
- ▲ Aufbauend auf den Ergebnissen der Vorgängeruntersuchung „Evaluierung hybrider Suchsysteme im WWW“ die Aussagekraft (Validität, Reliabilität, Effizienz) von Retrievaltests bei Suchmaschinen sicherzustellen bzw. zu erhöhen
- ▲ Hilfestellung für Suchdienstnutzer

Inhalt

- ▲ Motivation & Zielsetzung
- ▲ **Vorgehensweise & Konzeption**
- ▲ Testanordnung
- ▲ Pretest & Test
- ▲ Ergebnisse & Analyse
- ▲ Einschätzung & Ausblick

Vorgehensweise

Methodische Grundlage: Tague-Sutcliffs Vorgehensweise zur Entwicklung von Retrievaltests
The Pragmatics of Information Retrieval Experimentation, Revisited (1992)



Ziel: gültige, verallgemeinerungsfähige Ergebnisse



Einhaltung von „Standards“
Berücksichtigung webspezifischer Eigenheiten und Kriterien
(Datenbestand, Hypertextstrukturen, Nutzerverhalten)

1. Testen oder nicht testen - Motivation
2. Art des Tests - Bestimmung des grundsätzlichen Testverfahrens
3. Variablendefinition und -zuordnung
4. Verwendetes Informationssystem - ausgewählte Suchmaschinen
5. Erschließung der Informationsbedürfnisse und Suchanfragen
6. Durchführung der Suchanfragen
7. Aufbau der Testanordnung
8. Erfassung der ermittelten Daten
9. Datenauswertung
10. Ergebnispräsentation

Konzeption - Ablauf

Ausarbeitung der Testanordnung

↓ Remodifikation ↑

Pretest
Ermittlung der Tauglichkeit der Testanordnung



Testdurchführung



Datenauswertung & Ergebnisanalyse

Konzeption - Testverfahren & Variablen

- ▲ Test anhand einer Testkollektionen (Kontrolle vs. Realitätsnähe)
- ▲ Relevanz der Treffer als Grundlage der Bewertungsmaße
- ▲ Relevanzeinstufung der Ergebnisseiten durch Testpersonen (Juroren)
- ▲ Unkenntlichmachung der Herkunft der Ergebnisse
- ▲ Effektivitätsmaß Top20 Precision
- ▲ Suchanfragen nichtkommerzielle, eher sozialwissenschaftliche politische Fragestellungen

Konzeption - Auswahl der Suchmaschinen

▲ Kriterien

- ▲ Reichweite und Nutzungsgrad
- ▲ primär roboterbasiert (Suchmaschinen)
- ▲ deutschsprachige Dienste
- ▲ hohe Qualitätseinschätzung (z.B. Stiftung Warentest)

▲ Auswahl

- ▲ Altavista.de
- ▲ Fireball.de
- ▲ Google.de
- ▲ Lycos.de

Konzeption - Erschließung der Anfragen/Infobedürfnisse

- ▲ Realitätsnahe Erschließung & Ausgestaltung der Suchanfragen
z.T. Rückerschließung der Informationsbedürfnisse & Relevanzkriterien

Quelle	Anzahl Suchanfragen
Cross Language Track (Trec 7)	7
Cross Language Track (Trec 8)	5
GIRT (Trec 8)	17
Web Track (Trec 8)	4
Ask Jeeves	27

Anzahl Suchbegriffe	Anzahl Suchfragen
1	2
2	30
3	18
4	8
5	1
6	1

Anzahl geschlossene Fragestellungen	Anzahl offene Fragestellungen
11	49

Konzeption - Durchführung der Suchanfragen

- ▲ Explizites Ziel: Validität und Effizienz im Vergleich zur „händischen Durchführung“ der Suchanfragen bei der Vorgängeruntersuchung zu erhöhen
- ▲ Automatisierung der Suchanfragendurchführung durch ein Abfrageskript (Perl)
 - ▲ minimiert zeitliche Verzerrungen
 - ▲ erspart lokale Sicherung der Ergebnisseiten

Inhalt

- ▲ Motivation & Zielsetzung
- ▲ Vorgehensweise & Konzeption
- ▲ **Testanordnung**
- ▲ Pretest & Test
- ▲ Ergebnisse & Analyse
- ▲ Einschätzung & Ausblick

Testanordnung - Beispiel einer Trefferliste

Datum:

Mon Jan 14 17:00:17 2002

Suchanfrage:

importbeschränkungen eu

Thema:

Welche Berichte und Analysen gibt es zu den Importbeschränkungen der EU?

Relevanzkriterien:

Relevante Dokumente befassen sich mit den Importbeschränkungen der EU bzw. EG. Dazu gehören auch Berichte Verhandlungen mit anderen Staaten.

Treffer Suchmaschine 1:

[1.http://ticker.fireball.de/ticker/BUSINESS/2001-06-23T124007Z_01_B350361_RTRIDST_0_ODEBS-USA](http://ticker.fireball.de/ticker/BUSINESS/2001-06-23T124007Z_01_B350361_RTRIDST_0_ODEBS-USA)

[2.http://www.bse-info.de/](http://www.bse-info.de/)

[3.http://www.harald-etil.at/bericht03.htm](http://www.harald-etil.at/bericht03.htm)

.....

.....

.....

[18.http://www.gb-f.de/lastmin.html](http://www.gb-f.de/lastmin.html)

[19.http://www.zerp.uni-bremen.de/deutsch/publikationen/monographien.php3](http://www.zerp.uni-bremen.de/deutsch/publikationen/monographien.php3)

[20.http://www.betrifft.de/aktuelles/wirtschaft/149147.html](http://www.betrifft.de/aktuelles/wirtschaft/149147.html)

Treffer Suchmaschine 2:

[1.http://www.konsument.at/seiten/p1764.htm](http://www.konsument.at/seiten/p1764.htm)

Testanordnung - Relevanzbeurteilung

- ▲ Juroren beurteilen nacheinander die Treffer 1-20 der vier Suchmaschinen zu einer Suchanfrage (80)
- ▲ Bewertungsmaße
 - ▲ relevant
 - ▲ verweist auf relevante Seite
 - ▲ nicht relevant
- ▲ Lern- und Ermüdungseffekte werden durch Variation der Reihenfolge der Trefferlisten der Suchmaschinen kompensiert

Inhalt

- ▲ Motivation & Zielsetzung
- ▲ Vorgehensweise & Konzeption
- ▲ Testanordnung
- ▲ **Pretest & Test**
- ▲ Ergebnisse & Analyse
- ▲ Einschätzung & Ausblick

Pretest & Test

▲ Pretest

- ▲ Unterschiedliche Trefferzahlen bei der Durchführung der Suchanfragen
 - ▲ mehrfache Anfragewiederholung nötig
 - ▲ Modifikation: Suchanfragen werden vorab durchgeführt
 - ▲ zeitliche Verzerrungen werden in Kauf genommen
- ▲ Relevanzbeurteilung unproblematisch; Trefferlistenzuordnung gelingt nicht

▲ Test

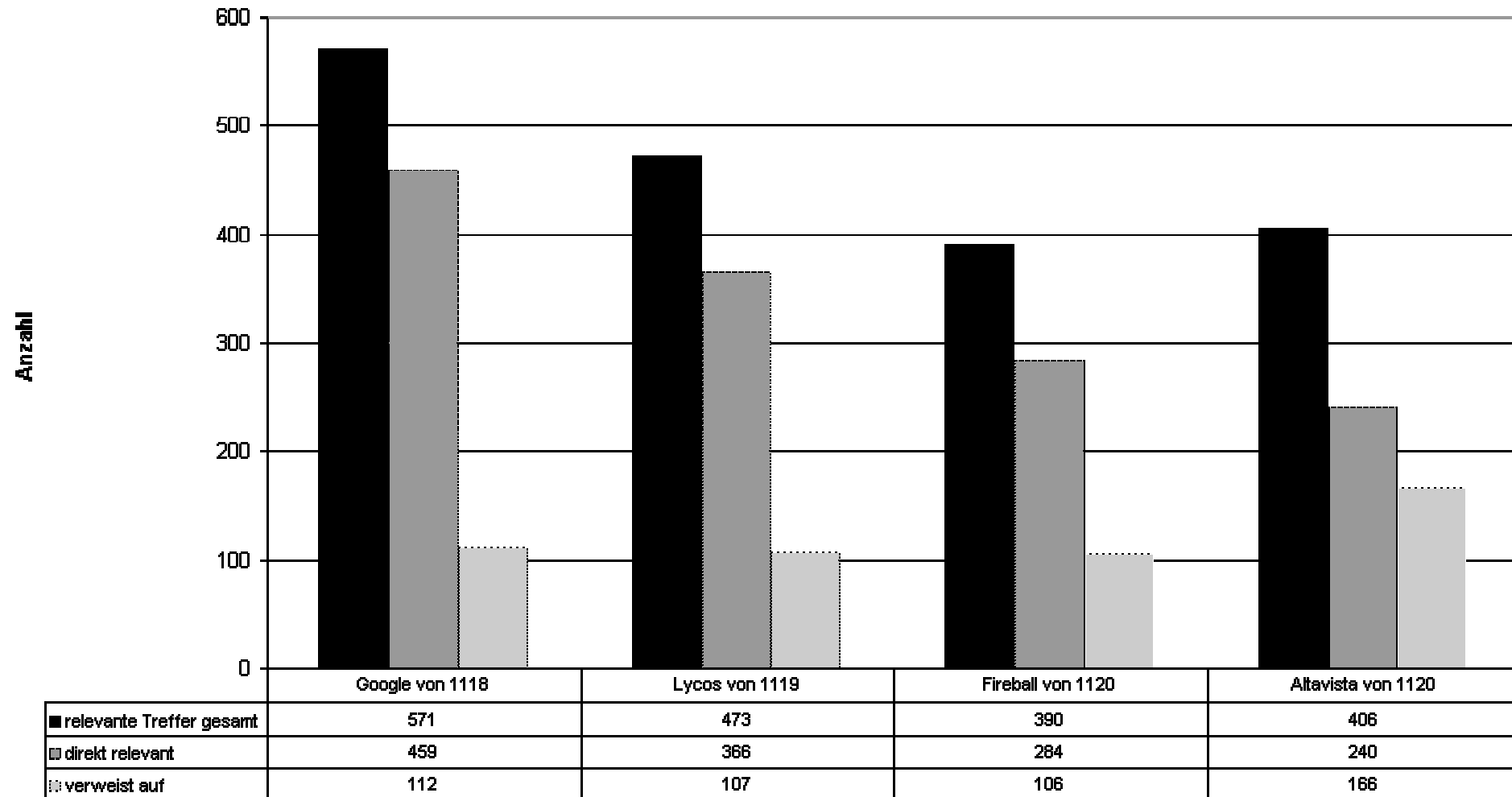
- ▲ 28 Juroren (Teilnehmer der Retrievalkurse & Mitarbeiter im Fachbereich) bewerten insgesamt 56 Suchanfragen

Inhalt

- ▲ Motivation & Zielsetzung
- ▲ Vorgehensweise & Konzeption
- ▲ Testanordnung
- ▲ Pretest & Test
- ▲ **Ergebnisse & Analyse**
- ▲ Einschätzung & Ausblick

Ergebnisse - Anzahl Relevante Treffer

Übersicht relevante Treffer



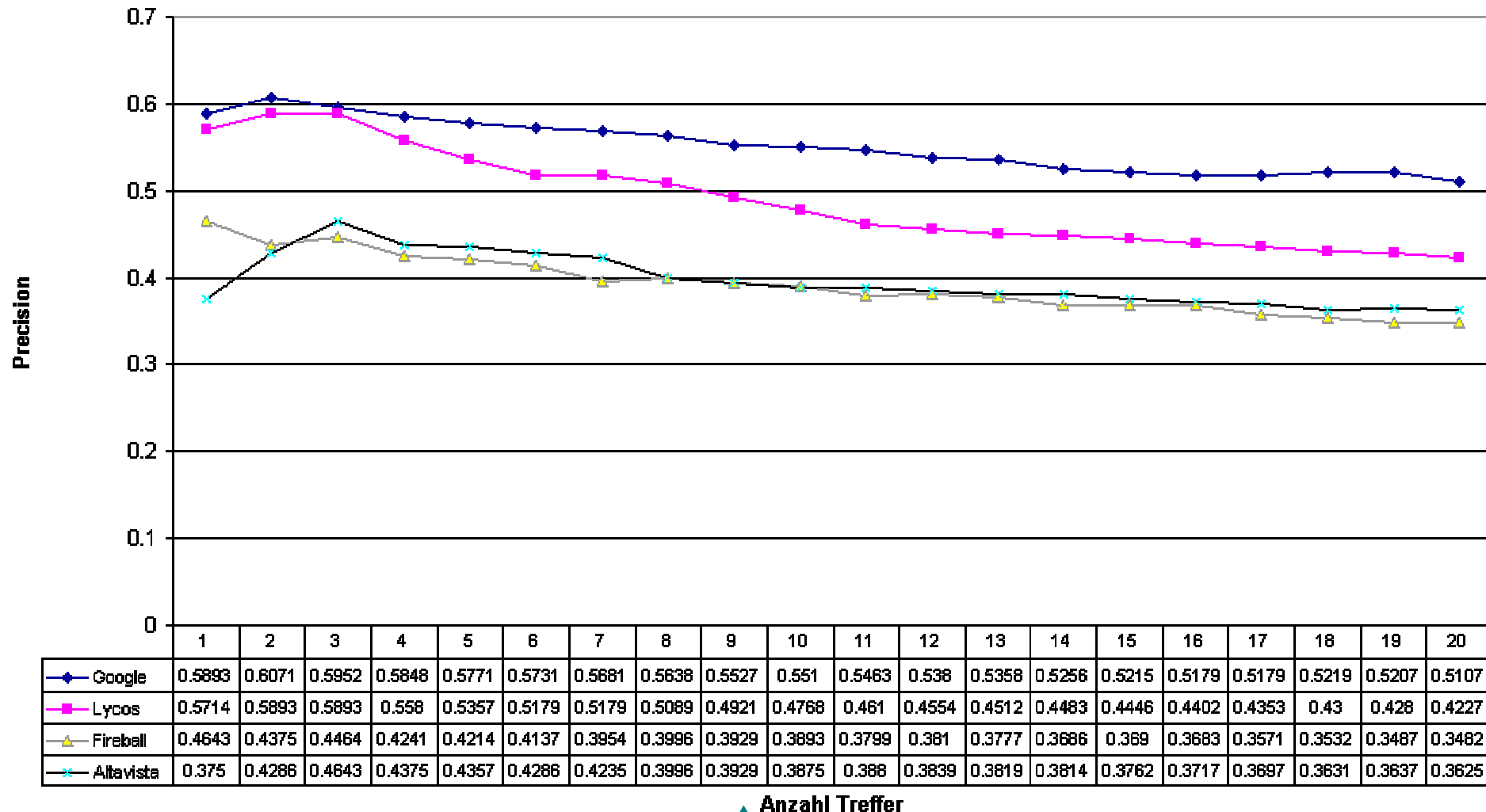
Suchmaschine

Ergebnisse - Beantwortete Suchanfragen

Suchmaschine	Beantwortete Suchanfragen	Nicht beantwortete Suchanfragen	Nicht beantwortete Suchanfragen und mehr als 20 Treffer
Google.de	53 (94.6%)	1	2
Lycos.de	50 (89.2%)	4	2
Fireball.de	51 (91.1%)	4	1
Altavista.de	55 (98.2%)	0	1

Ergebnisse - Top20 Precision

Top20 Precision - alle Suchanfragen (Anzahl 56)



Anzahl Treffer

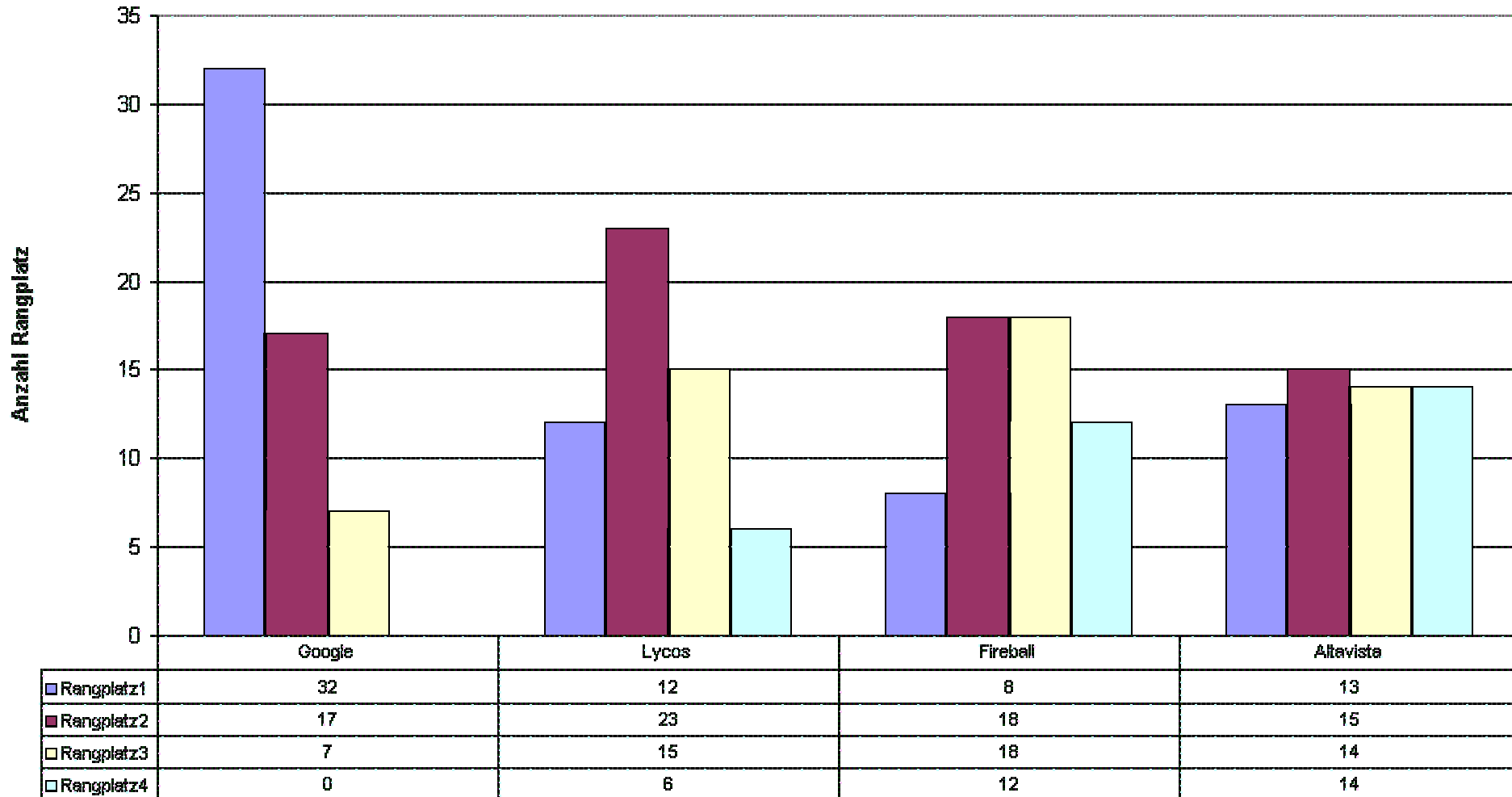


Ergebnisse - Top20 Mean Average Precision

Suchmaschine	Top20 Mean Average Precision
Google.de	55.1%
Lycos.de	48.8%
Fireball.de	39.1%
Altavista.de	39.6%

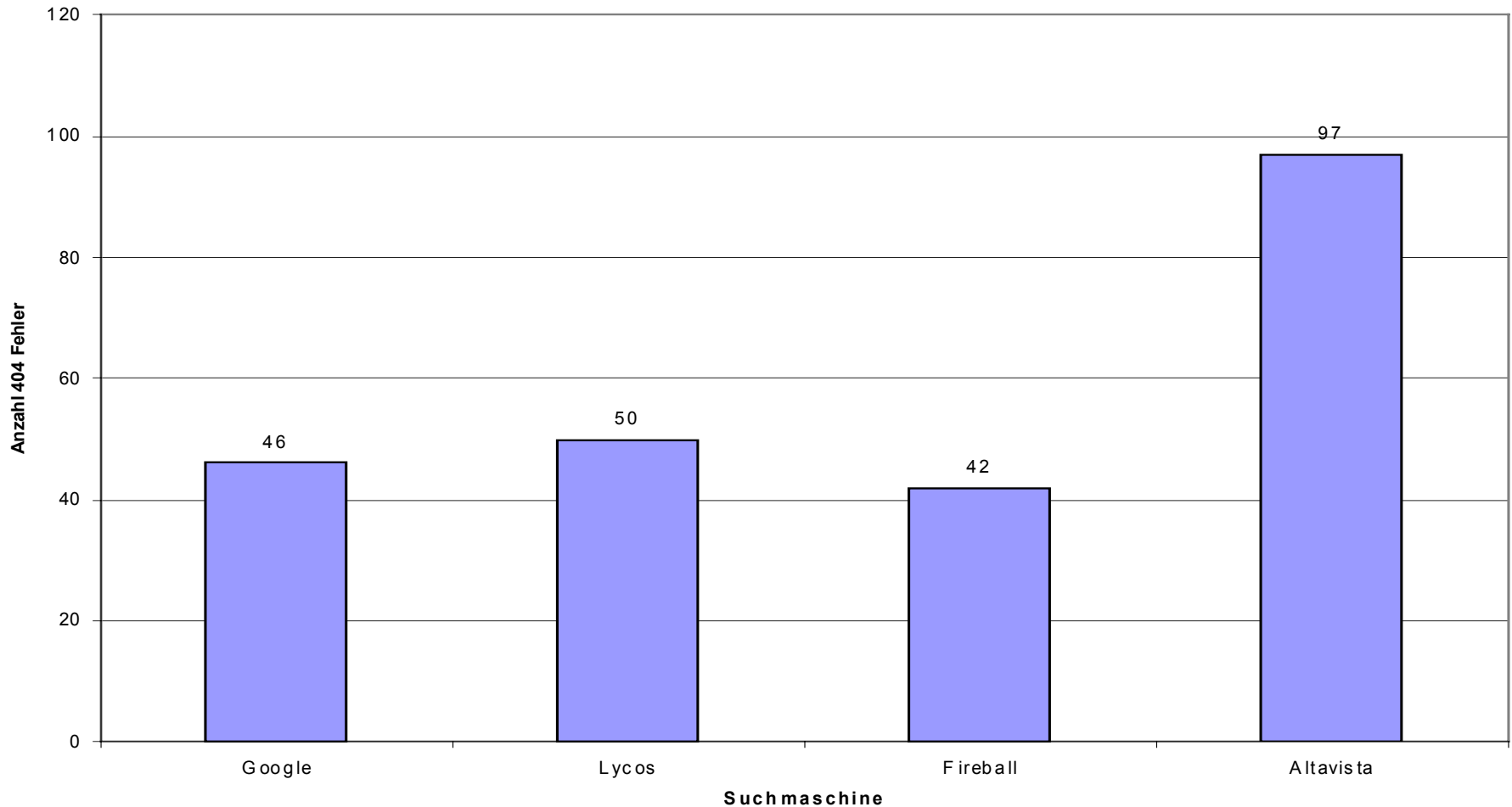
Ergebnisse - Beantwortung der Suchanfragen

Beantwortung der Suchanfragen



Ergebnisse - Festgehaltene 404 Fehler

Von Juroren festgehaltene 404 Fehler



Ergebnisanalyse

- ▲ Google.de ist hinsichtlich der Top20 Precision effektiver als die anderen Suchmaschinen
- ▲ Altavista.de beantwortet die größte Zahl der Suchanfragen
- ▲ Altavista.de deutlich mehr „404“ Fehler (veralteter Index?)
- ▲ Zeitliche Verzerrungen & Mehrwerte wie
 - ▲ passende Katalogrubriken (Google/Lycos)
 - ▲ Google Archiv
 - ▲ Zugriff auf PDF oder Postscript Dokumente (Google) bleiben unberücksichtigt

Inhalt

- ▲ Motivation & Zielsetzung
- ▲ Vorgehensweise & Konzeption
- ▲ Testanordnung
- ▲ Pretest & Test
- ▲ Ergebnisse & Analyse
- ▲ **Einschätzung & Ausblick**

Einschätzung & Ausblick

- ▲ These: Validität und Reliabilität der Untersuchung wird durch methodische Vorgehensweise, die Berücksichtigung, webspezifischer Eigenheiten und Einhaltung gängiger Standards zumindest näherungsweise sichergestellt
- ▲ Ziel der Optimierung durch Automatisierung der Durchführung der Suchanfragen wird nicht vollständig erreicht. Effizienzsteigerungen vs. zeitliche Verzerrungen

Einschätzung & Ausblick

- ▲ Problemfelder
 - ▲ Nullantworten
 - ▲ Nicht beantwortete Suchanfragen (Gefahr der Ergebnisverzerrungen)
 - ▲ Nicht beantwortete Suchanfragen und mehr als 20 Treffer (Top20 Grenze generalisierbar?)
Validität der Ergebnisse hängt auch von den Ergebniswerten selbst ab
 - ▲ Unkenntlichmachung der Herkunft der Treffer künftig durchzuhalten und sinnvoll?
Abbildung von Alleinstellungsmerkmalen wie Google Archiv oder Visualisierungsmehrwerten (www.kartoo.com)
- ▲ breites Optimierungspotential bei künftigen Untersuchungen