

PsychSpider – Erfahrungen aus dem Betrieb einer spezialisierten Suchmaschine

Christiane Baier & Peter Weiland
Zentrum für psychologische Information und
Dokumentation (ZPID)

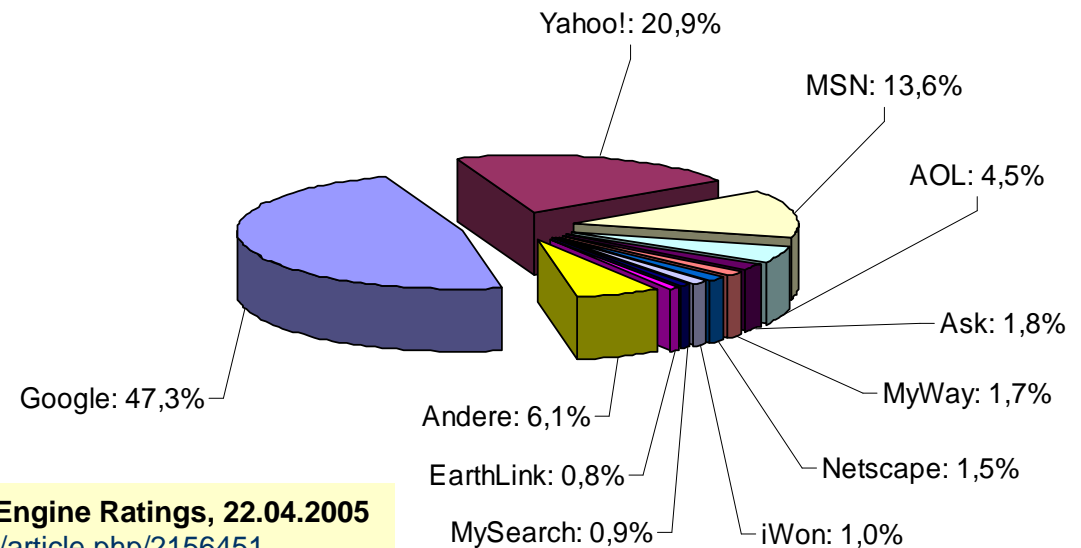
luK Jahrestagung 2005, Bonn, 10.05.2005

Agenda

- 1) Einführung
- 2) Universalsuchmaschinen vs. spezialisierte Suchmaschinen
- 3) Erfahrungen aus dem Betrieb
- 4) Auswertung des Nutzerverhaltens des PsychSpider
- 5) Fazit / Diskussion

Der Suchmaschinenmarkt

➤ Marktanteil Suchmaschinen in den USA (März 2005)



Quelle: Nielsen NetRatings Search Engine Ratings, 22.04.2005
<http://searchenginewatch.com/reports/article.php/2156451>

➤ Indexgrösse beliebter Suchmaschinen

- Google: 8.058.044.651 Seiten
- MSN: ~5.000.000.000 Seiten

Defizite von Universalsuchmaschinen

- **Einbindung von spezialisierten, so genannten "Deep Web"-Angeboten nur selten möglich**
(Beispiel: Yahoo! → 99% "Surface Web")
- **Qualitätskontrolle der Inhalte / Suchmaschinen-Spam**
- **Kommerzielle Abhängigkeiten / Neutralität**
- **Kontext bei der Suche**
Beispiel: Suche nach "Adler" in Google und PsychSpider

Suche im Kontext - Beispielsuche "Adler"

Google-Suche: Adler - Mozilla Firefox

File Edit View Go Bookmarks Tools Help

http://www.google.de

Google Google Scholar my del.icio.us gso-bookmarklet.js post to del.icio.us

Web Bilder Groups Verzeichnis News Froogle ^{Neu!}

Adler Suche Erweiterte Suche Einstellungen

Suche: Das Web Seiten auf Deutsch Seiten aus Deutschland

Web Ergebnisse 1 - 10 von ungefähr 5.340.000 für Adler. (0,04 Sekunden)

News-Ergebnisse für Adler - Meldungen des Tages

DGAP-Ad hoc: TA Triumph-Adler AG - sueddeutsche.de - 26. Apr. 2005
Adler Mannheim verlängern Vertrag mit Trainer Richer - Berliner Morgenpost - 25. Apr. 2005

ADLER Mannheim
Aufgeführt ist das Team des Eishockey-Clubs, die Spieltermine mit Berichten und Tabellen der Ergebnisse sowie Informationen zum Stadion und den Preisen.
www.adler-mannheim.de/ - 2k - Im Cache - Ähnliche Seiten

adler.de - ihre modemarke
Die Adler Modemärkte bieten aktuelle Mode für die ganze Familie zu attraktiven Preisen. Sie finden eine tolle Größenauswahl. Mit der Adler-Kundenkarte ...
www.adler.de/ - 7k - Im Cache - Ähnliche Seiten

heise online - Triumph-Adler streicht 100 Stellen
... Triumph-Adler streicht 100 Stellen. Mit einem weiteren Stellenabbau und Kosteneinsparungen ... Triumph-Adler in diesem Jahr aus den roten Zahlen kommen. ...
www.heise.de/newsticker/meldung/59016 - 30k - 26. Apr. 2005 - Im Cache - Ähnliche Seiten

Adler Schiffe - Insel- und Halligreederei

Find: Find Next Find Previous Highlight Match case

Done

ZPID PsychSpider - Psychologie Suchmaschine - Trefferliste, Hits - Mozilla Firefox

File Edit View Go Bookmarks Tools Help

http://www.zpid.de/F

Google Google Scholar my del.icio.us gso-bookmarklet.js post to del.icio.us

English | Ir

info@zpid.de | © 1996-2005 ZPID

PsychSpider ^{ZPID} - Psychologie Suchmaschine - Suchergebnisse

Hinweis der Woche: Verwenden Sie Bindestriche in Ihrer Suche, dann achten Sie darauf, kein Leerzeichen unmittelbar vor dem Bindestrich stehen zu haben. Der Bindestrich wird sonst als NOT interpretiert (siehe auch Tipp 11).

Ihre Suche "adler" ergab 1977 Dokumente von 698268.
20 Dokumente werden angezeigt.
Ergebnis-Seite: 1 2 3 4 5 [>] [>>]

0.6466 Individualpsychologie Alfred Adler (1870-1937)
http://www.stanql-taller.at/ARBEITSBLAETTER/WISSENSCHAFTPSYCHOLOGIE/PSYCHOLOGIE/individualpsychologie/AlfredAdler(1870-1937).shhtml
Individualpsychologie: Alfred Adler (1870-1937)

0.5942 PersonalityTheories: Adler, Quiz1
http://www.tulsa.oklahoma.net/~inichols/PersonalityTheories-Adler-Quiz1.html
8. To Adler, a brief therapy approach was a. a matter of shortening the length of session clients would not have to be in an emotionally charged atmosphere for long b. technique solve clients problems in a specifiable and small number of sessions c. a means of shortening the relationship between the therapist and patient, regardless of the number of sessions d. a ridiculous idea that was not in the best interest of the client 10. Early in the history of the profession of psychology, the idea of a brief therapy approach was a. a matter of shortening the length of session clients would not have to be in an emotionally charged atmosphere for long b. technique solve clients problems in a specifiable and small number of sessions c. a means of shortening the relationship between the therapist and patient, regardless of the number of sessions d. a ridiculous idea that was not in the best interest of the client 10. Early in the history of the profession of psychology, the idea of a brief therapy approach was

0.5812 APSaA
http://www.apsa.org/cfp/pubinfo/about/bios/adler.html
There is a biography by Phyllis Bottome, Alfred Adler: A Portrait from Life (1957), another Edward Hoffman, The Drive for Self: Alfred Adler and the Founding of Individual Psychology (1994), and a study by Heinz and Rowena Ansbacher, The Individual Psychology of Alfred Adler (1964). There are Alfred Adler Institutes and a journal The Journal of Individual Psychology. Adler's work includes A Study of Organ Inferiority (1917), The Practice and Theory of Individual Psychology (1927), and The Practice and Theory of Individual Psychology (1927).

Find: Find Next Find Previous Highlight Match case

Done

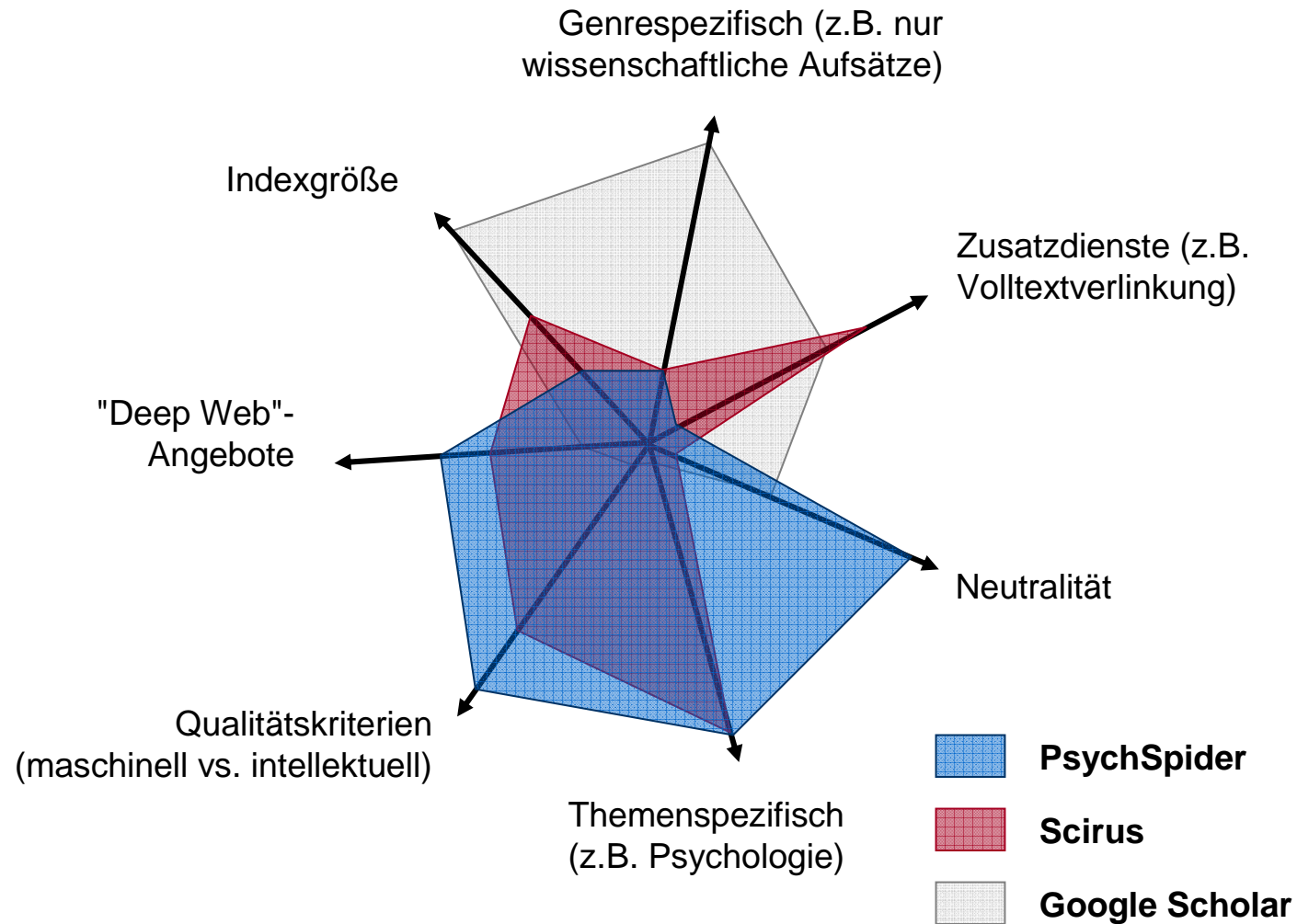
PsychSpider – Kurzbeschreibung

- Seit 2001 betreibt das ZPID den PsychSpider als spezialisierte Suchmaschine für Psychologie und angrenzende Themenbereiche

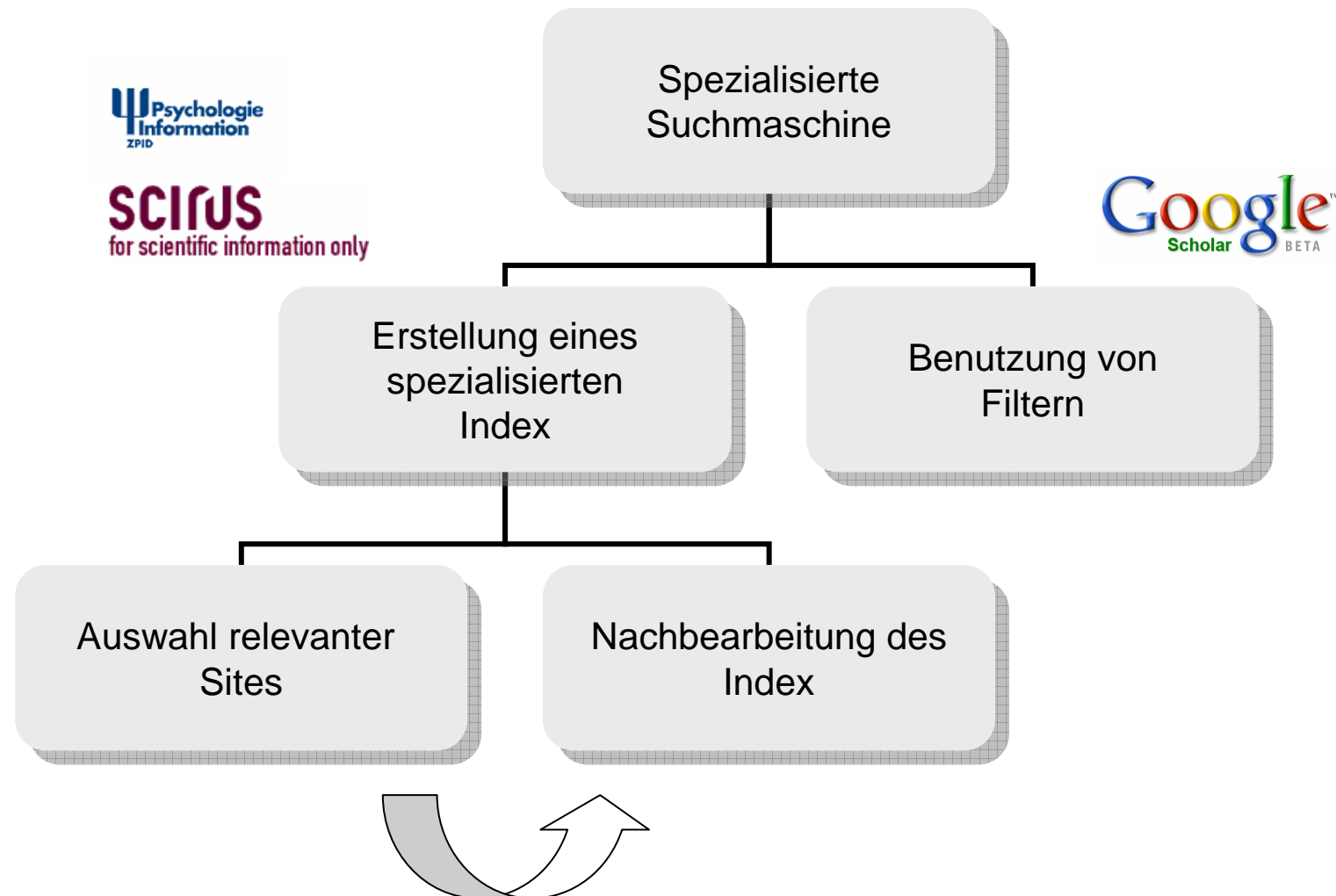
Der PsychSpider...

- umfasst ca. 700.000 Dokumente
- indexiert über 370 Server
- beinhaltet die Datenbanken PSYINDEX Lit & AV und PSYINDEX Tests
- basiert auf Verity K2 unter SUN Solaris

Kriterien für die Einordnung von spezialisierten Suchmaschinen



Aufbau spezialisierter Suchmaschinen



PsychSpider – Quellen

- **Manuell überprüfte URLs**
alle URLs, die zum Index hinzugefügt werden sind intellektuell überprüft worden; laufende Kontrolle des Index auf nicht relevante Inhalte
- **Einbindung der Datenbanken PSYINDEX Lit & AV und PSYINDEX Tests**
- ***Geplant:*** Integration weiterer "Deep Web"-Angebote (beispielsweise Inhalte von weiteren Dokumentenservern über OAI)

PsychSpider - Herausforderungen

Die Gestaltung von Websites hat einen großen Einfluss auf die Güte der Indexierung:

- JavaScript in Menünavigationen
- Verwendung von Session Cookies
- Nichtvalide XML-Dokumente
- PDF-Dateien (Formate / Versionen)
- Inhalte, die nur über ein Suchinterface zugänglich sind (z.B. Dokumentenserver)
- Angebote, die explizit Suchmaschinen außer Googlebot aussperren
- Einsatz von Content Management Systemen

Maßnahmen zur Attraktivitätssteigerung

- **Weitere Einbindung von psychologierelevanten "Deep Web"-Angeboten, z.B. Volltextserver**
- **Kontinuierliche Erweiterung des Index**
(ist verbunden mit einem höheren Ressourceneinsatz)
- **Verbesserungen an der Suchoberfläche**
Einbindung erweiterter Ranking-Verfahren;
Berücksichtigung des Suchverhaltens der Benutzer

Nutzerverhalten

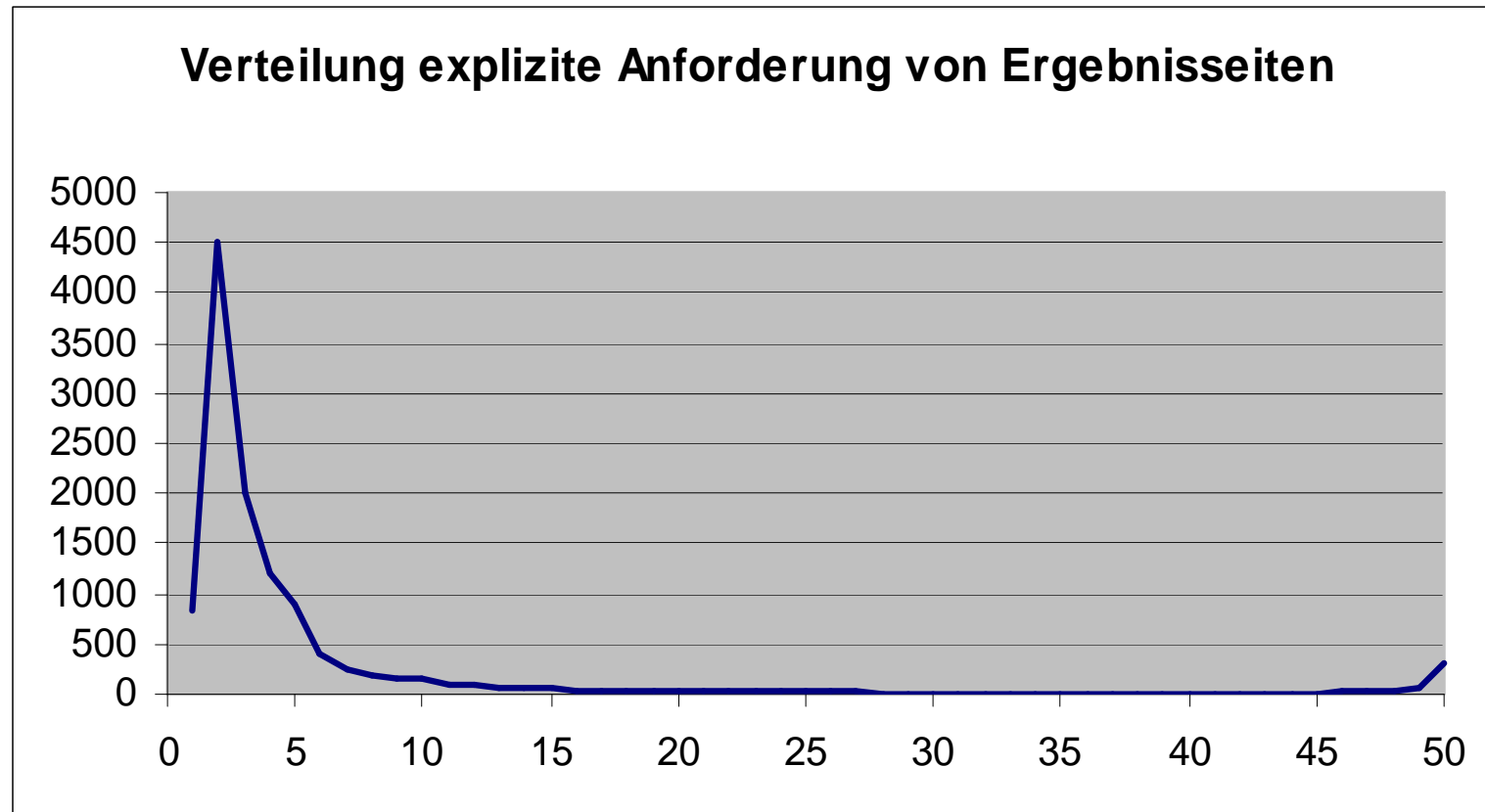
- Logfile-Analyse von PsychSpider
 - Zeitraum von 9 Monaten (Juli 2004 – April 2005)
 - 51403 Suchanfragen insgesamt
davon:
 - 37877 initiale Suchanfragen (per POST übergeben)
 - 12864 Aufrufe weiterer Ergebnisseiten (per GET)

- Vergleich mit anderen Studien
 - AltaVista, Excite, FAST

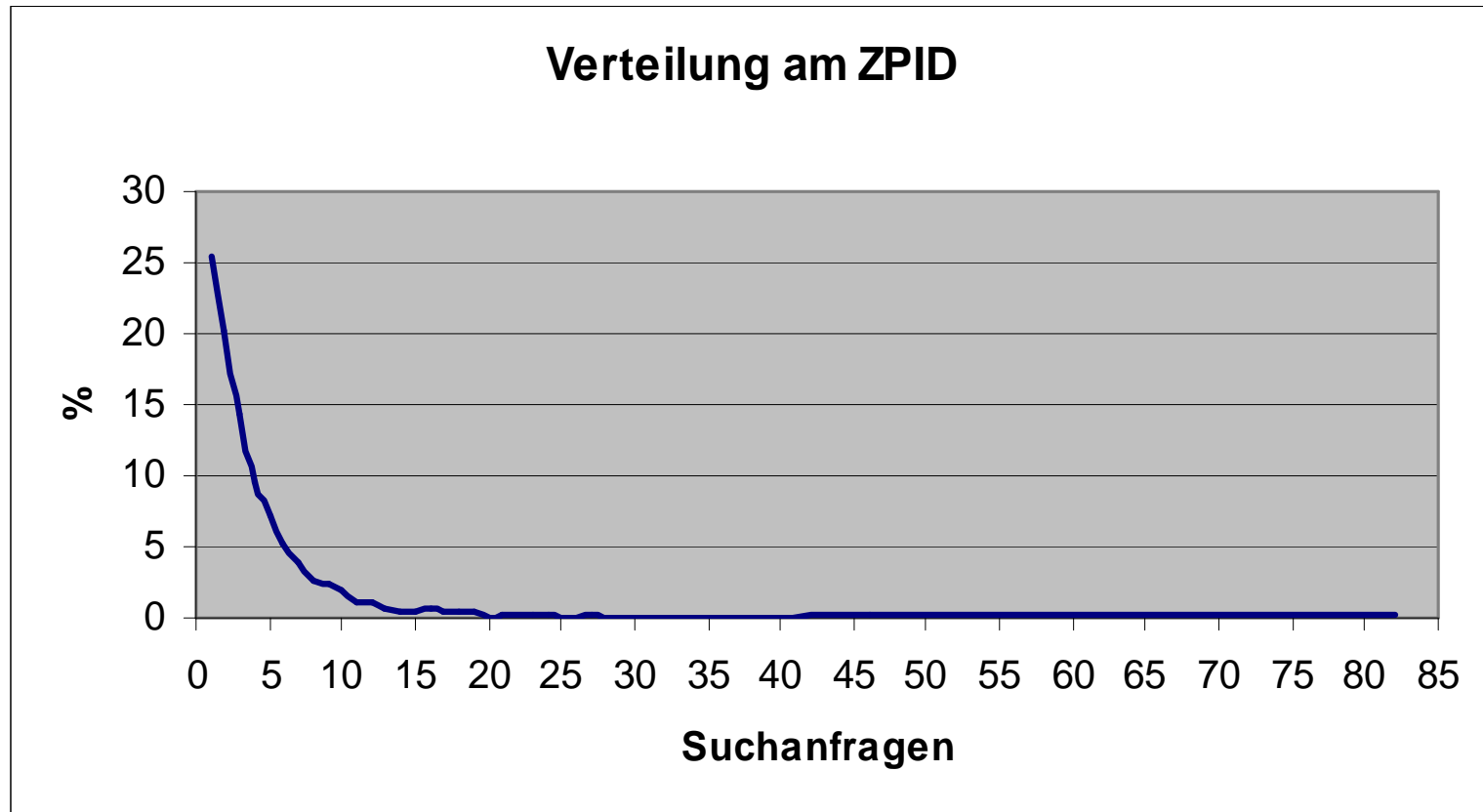
Analyse

- Navigation über Ergebnisseiten
- Suchanfragen per Session
- Terme pro Suchanfrage
- Verwendung von Operatoren
- Modifikationen der Suchanfrage
- Suchanfragen- / Termhäufigkeiten

Navigation über Ergebnisseiten

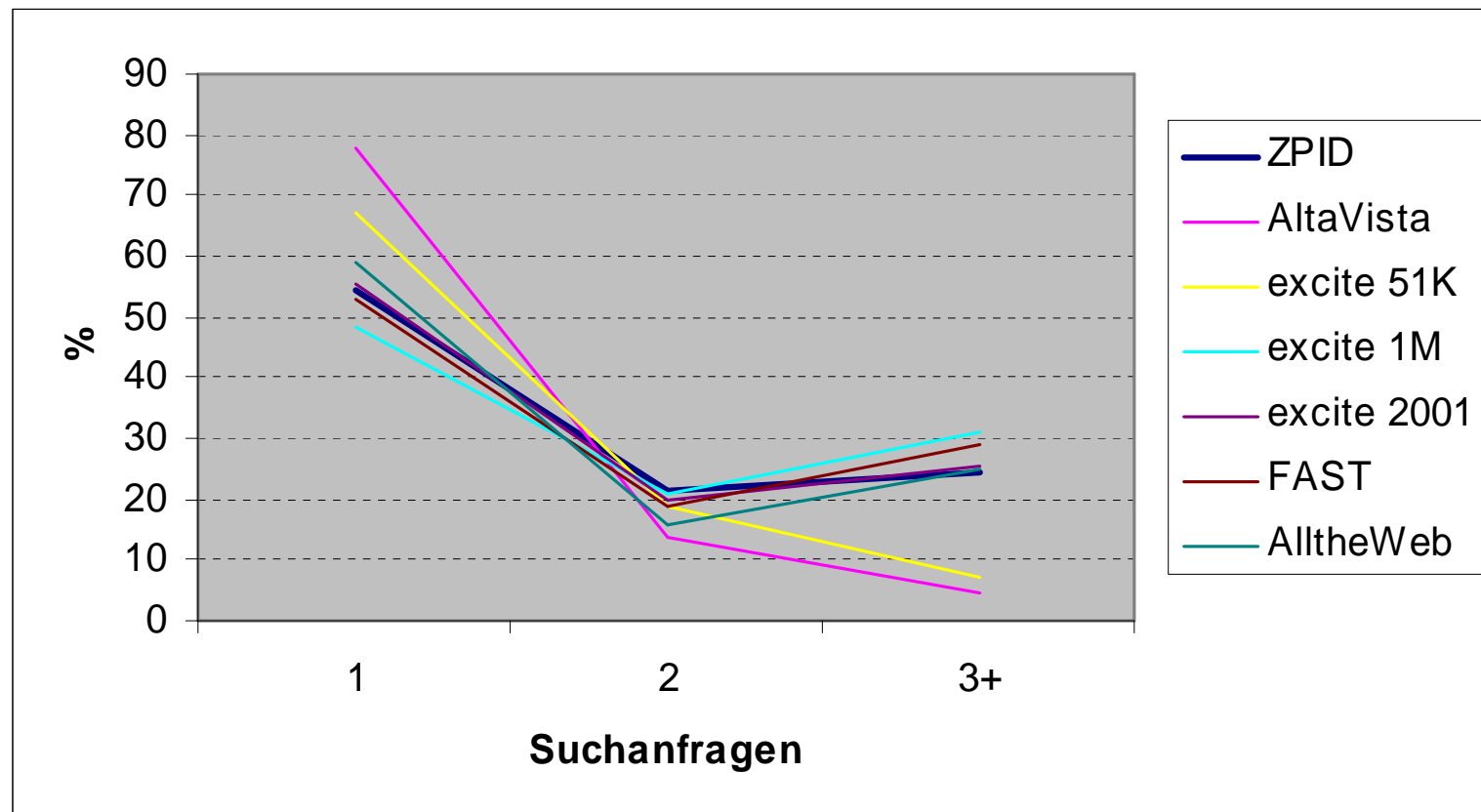


Suchanfragen per Session

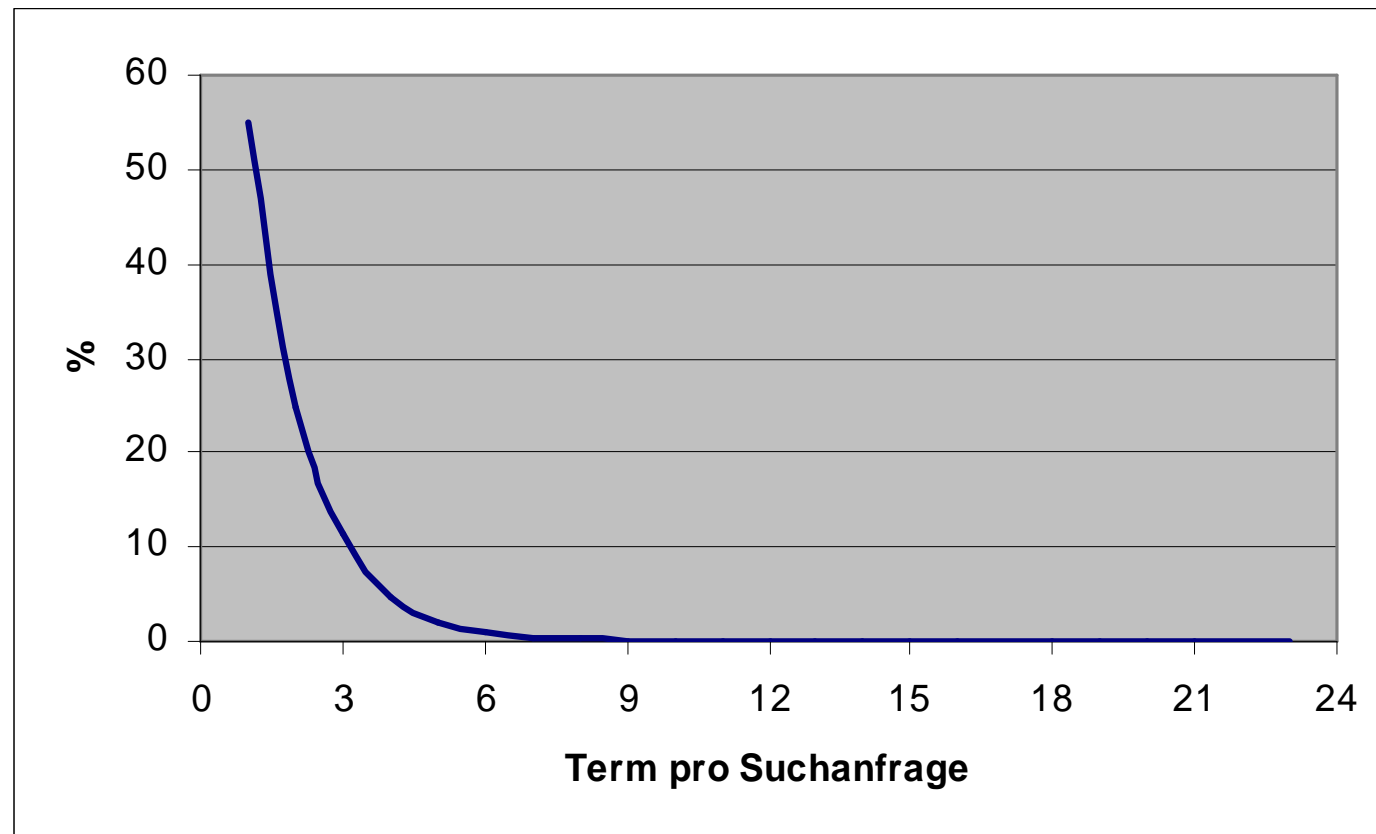


Suchanfragen per Session

➤ Vergleich mit anderen Studien

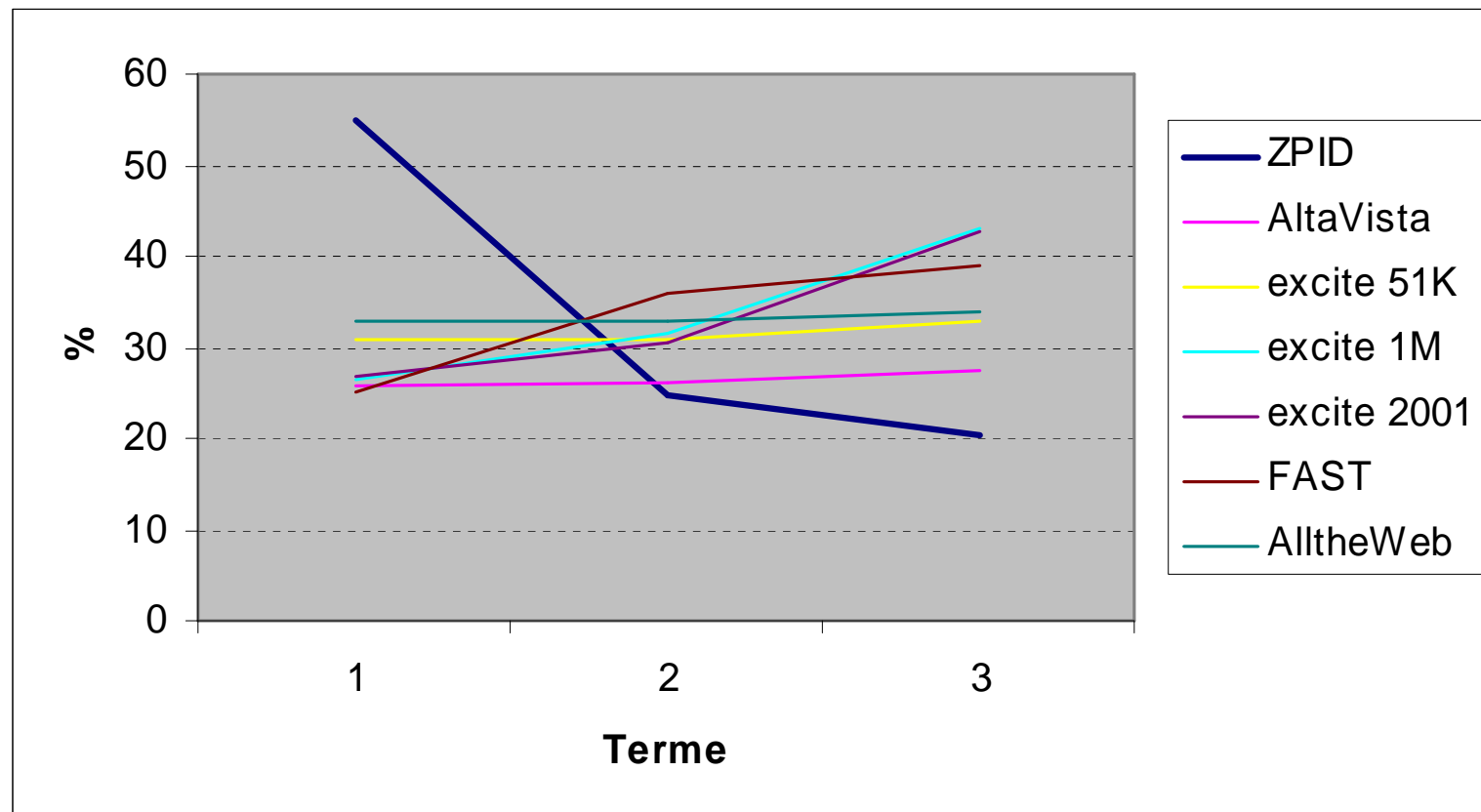


Terme pro Suchanfrage



Terme pro Suchanfrage

➤ Vergleich mit anderen Studien



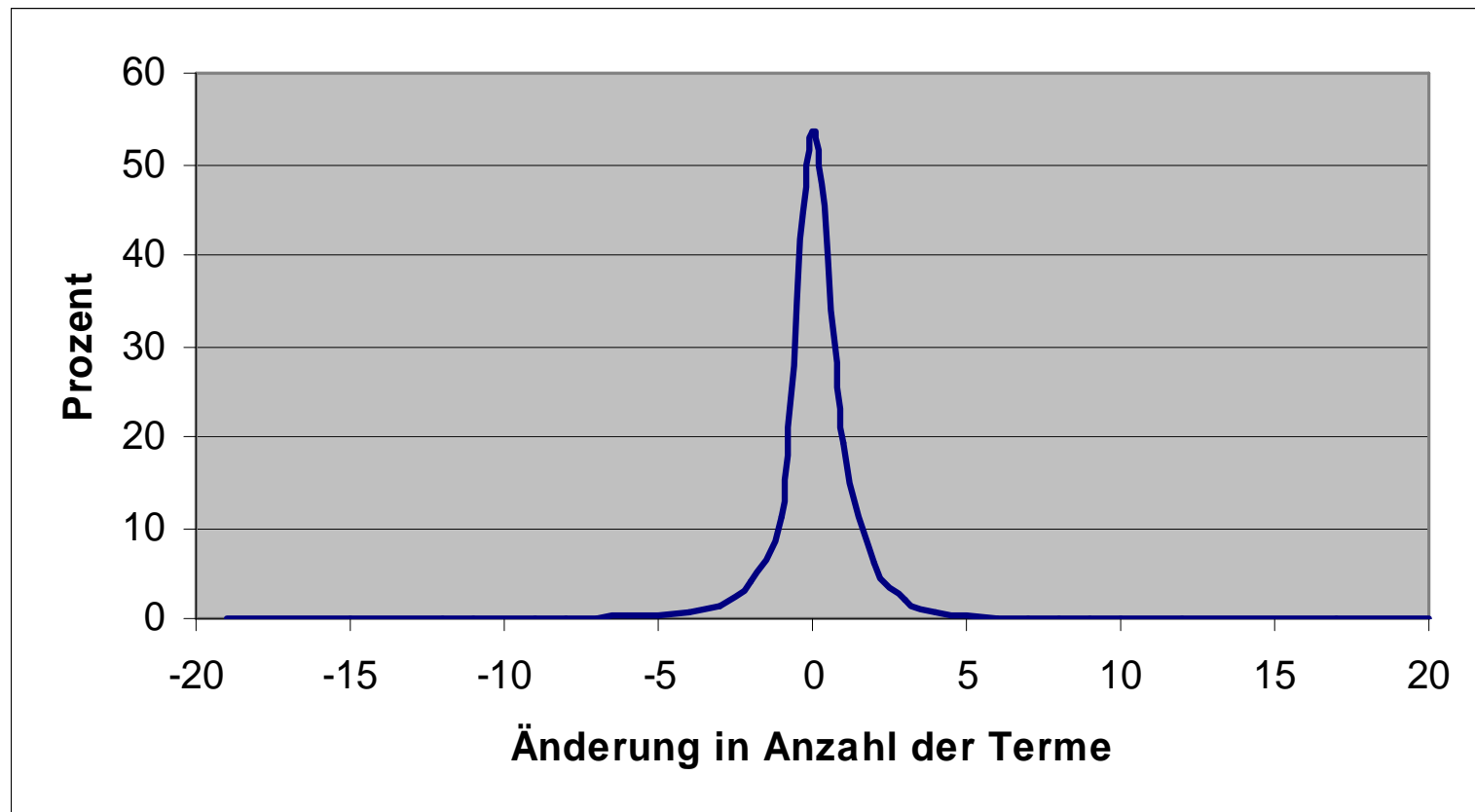
Verwendung von Operatoren

Operator	absolut	prozentual
Phrase	723	1,91
AND	4582	12,10
OR	271	0,72
NOT	61	0,16
+	392	1,03
-	79	0,21
Komma	751	1,98
Klammer	86	0,23
Wildcard *	286	0,76
Wildcard ?	16	0,04

Modifikationen der Suchanfrage

- Welcher Art sind die Modifikationen?
 - 33,8 % Modifikation der Begriffe
(Altruismus → Altruismusskalen)
 - 25,4 % völlige Neuformulierung
(Angst → Ärger)
 - 6,4% Operatorenwechsel
 - 46,4 % Modifikation in der Termanzahl

Modifikationen der Suchanfrage



Modifikationen der Suchanfrage

- mit gleicher Termanzahl
 - 44,8 % Begriffsmodifikation
 - 29,4 % Neuformulierung
 - 3,1 % Operatorenwechsel

Suchanfragen- / Termhäufigkeiten

- Subset der 37877 initialen Anfragen
(gleiche IP, gleicher Tag und gleiche Anfrage eliminiert)

Top 10 Suchanfragen:

litzcke
stress
borderline
depression
angst
motivation
burnout
aggression
schizophrenie
mobbing

Top 10 Terme:

psychologie
test
kinder
fragebogen
soziale
stress
depression
therapie
motivation
angst

Zusammenfassung

- Spezialisierte Suchmaschinen sind eine Ergänzung zu Universalsuchmaschinen
 - Sie bieten Kontext für eine Suche, d.h. auch Ein-Wort-Suchbegriffe führen zum Ziel
 - Sie ermöglichen den Zugriff auf "Deep Web"-Angebote, die über andere Suchmaschinen nicht abgesucht werden können
- Beim Nutzerverhalten zeigen sich wenig signifikante Unterschiede zu Studien bei Universalsuchmaschinen