
Suchmaschinen Optimierung

Prof. Dr. Clemens Cap
Universität Rostock
<http://www.internet-prof.de>

Die Fragestellung

Wie werde ich besser?

**Wenn das Ranking auf einer Seite so wichtig ist,
was kann ich tun, damit das Ranking meiner Seiten besser wird?**

Die Kurzantwort:

- Eine für andere inhaltlich wertvolle Website aufbauen

Die Langantwort:

- **Technologisch:** Tricks bei der Gestaltung der Webseite
- **Redaktionell:** Fokus auf gesuchte Bezeichnungen
- **Strategisch:** Fokus auf gesuchte Trends und Inhalte

Vorsicht:

- Aktuelle Geheimtipps wechseln
- Suchmaschinen entwickeln sich weiter

Der Wert

Was ist mir das wert?

Frage der Traffic-Evaluation ist problematisch

- Stark abhängig vom Anliegen der Web Site

Wert einer Optimierung = Kosten einer gleichwirksamen Keyword-Werbung

- Suchmaschinen-Optimierung ist eine Methode, den Traffic zu erhöhen
- Traffic kann man auch über Keyword-Werbung bekommen
Da sind die Kosten leicht zu eruieren

Beispiel

- Eine Maßnahme die 100.000 Impressions bringt
- Angenommener Cost Per Click bei Keyword Werbung 0,30 €
- Wert der Maßnahme 30.000 €

Die Bedeutung

Es zählen

- die ersten 5 Positionen der Suchmaschine
- allenfalls noch die zweite Seite der Suchmaschine

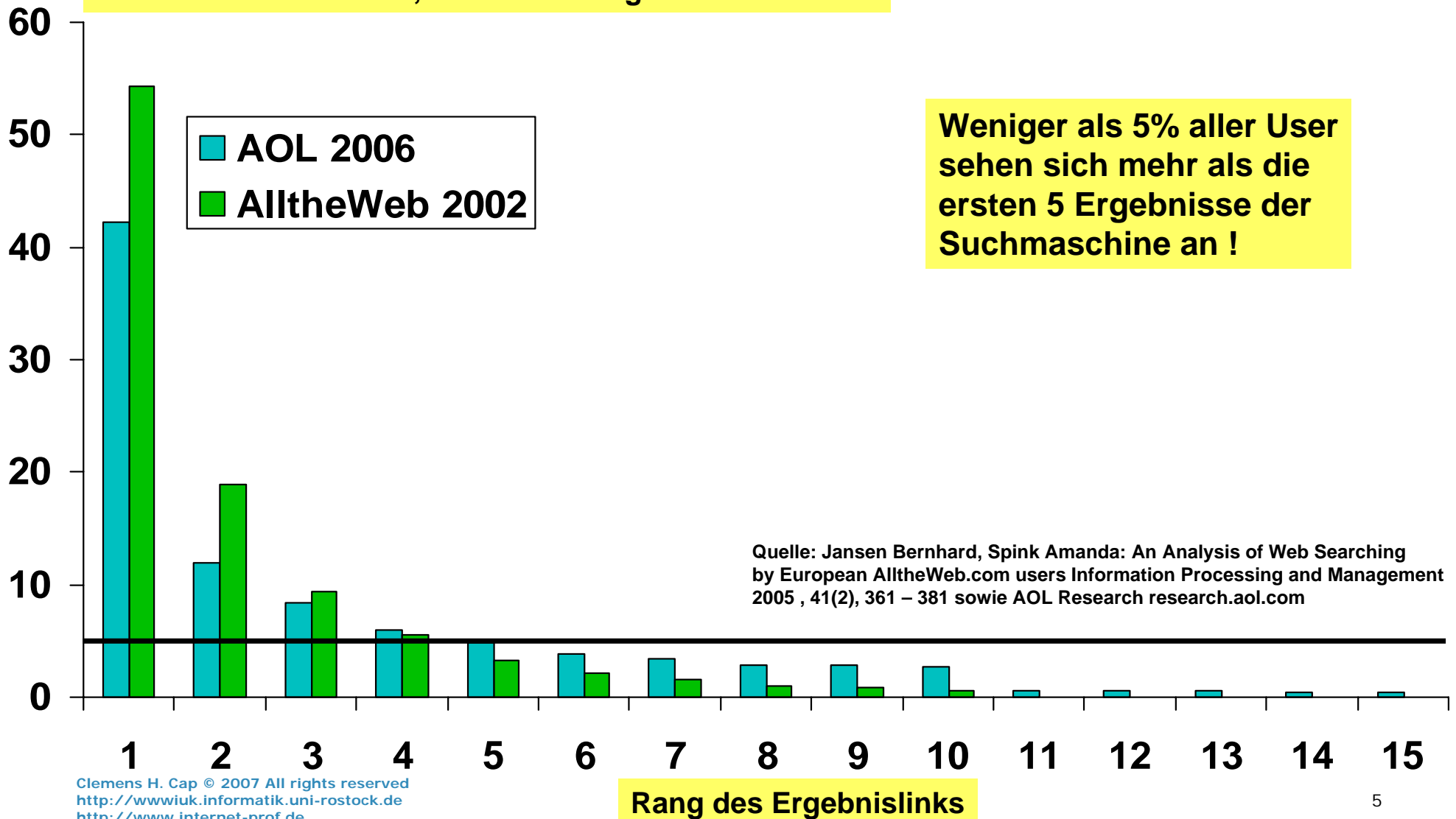
Der Rest der Treffer ist unterhalb der Relevanzgrenze für den Massentraffic !

Fazit:

- Entweder eine Top-Site
- Oder Keyword-Marketing

Welche Rang-Positionen im Ergebnis sehen sich die User an ?

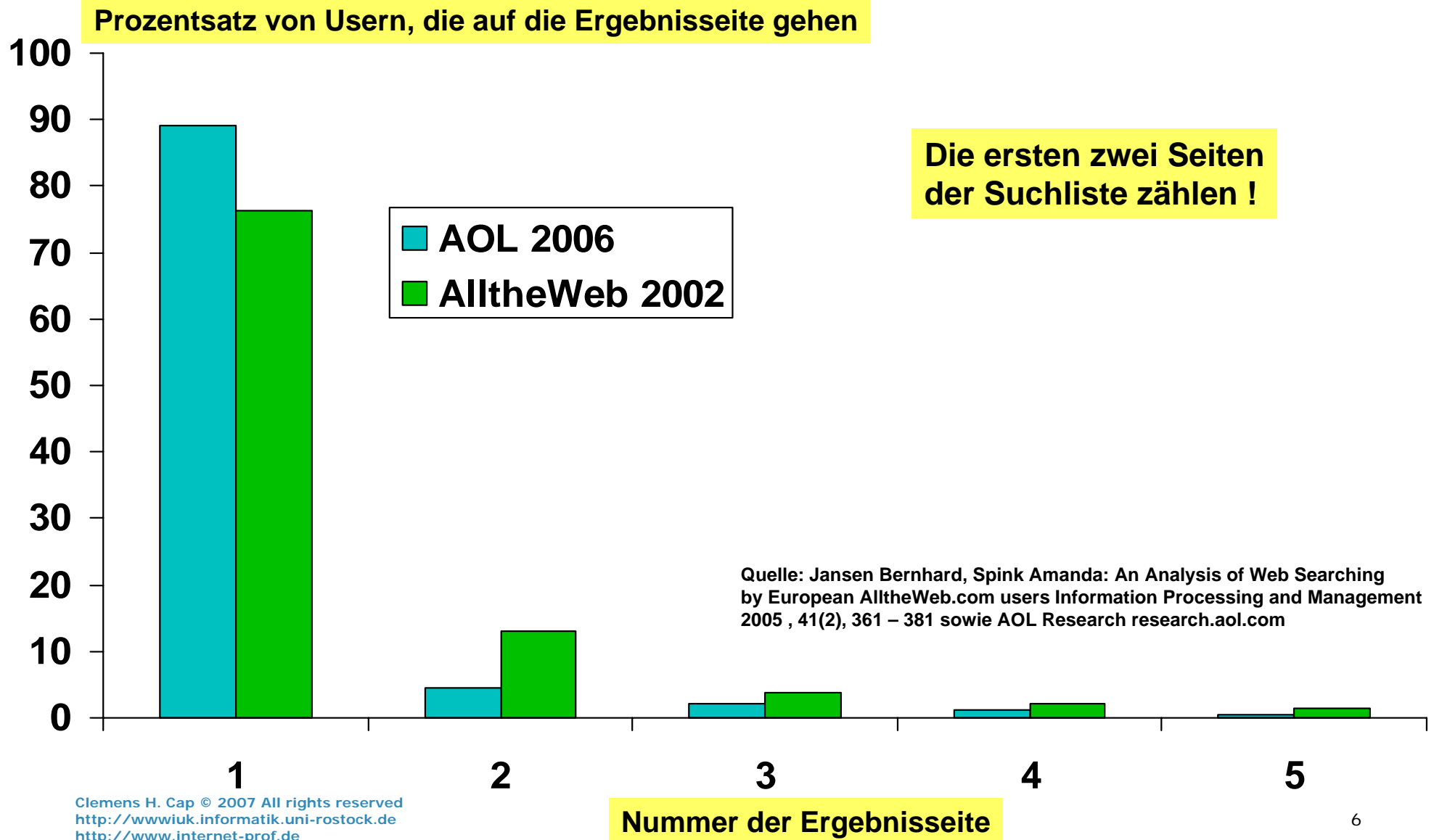
Prozentsatz von Usern, die auf den Ergebnislink klicken



Weniger als 5% aller User sehen sich mehr als die ersten 5 Ergebnisse der Suchmaschine an !

Quelle: Jansen Bernhard, Spink Amanda: An Analysis of Web Searching by European AlltheWeb.com users Information Processing and Management 2005 , 41(2), 361 – 381 sowie AOL Research research.aol.com

Welche Such-Seiten im Ergebnis sehen sich die User an ?



Spamdexing

Spamdexing: Spamming einer Suchmaschine

- Hochtreiben des Rankings mit unfairen Methoden
- Suchmaschine versucht Spamdexer zu detektieren und dann abzuwerten

Negativ werten die meisten Suchmaschinen daher

- **Stuffing:** Worte in Meta Tags, Keywords, Titel, Text massiv wiederholen
- **Reciproke Link Stuffing:** Unreal hoher Prozentsatz wechselseitig gesetzter Links
- Verlinkung von Sites mit **identischem Domäneninhaber**
- Verlinkung von Sites mit **identischer Adresse** im Impressum
- Offenkundig unzutreffende oder irreführende Worte in Keywords, Meta Tags, Titel

Genereller Tipp

- Alles vermeiden, was als Spamdexing interpretierbar wäre

Zeitentwicklung

Beobachtung 1: Eine gute Site wird regelmäßig gepflegt

Beobachtung 2: Spammer wollen den raschen Erfolg (Hit and run offences)

Positiv werten die meisten Suchmaschinen daher

- Hohes Alter der Seite
- Lange Registrationsdauer der Domäne
- Hohes Alter eines Links auf die Seite
- Sinnvolle Rate des Hinzufügens neuer Inhalte auf der Site / Seite
- Hohe Rate, mit der Fehler (Bsp: broken links) behoben werden

Negativ werten die meisten Suchmaschinen daher

- Rate, mit der Linkzahl auf die Seite hochgeht, ist unrealistisch hoch
Verdacht: Link Spamming

Alleinstellungsmerkmale

- Beobachtung 1:** Gute Seiten verlinken bevorzugt auf gute Seiten
Daraus ergaben sich HITS / Pagerank Algorithmen
- Beobachtung 2:** Vereinzelt lassen sich inhaltlich gute Seiten auch
an anderen Kriterien messen

Positiv werten die meisten Suchmaschinen daher

- Text ist einzigartig, es gibt keine Kopien im Web
und der Text ist gut verlinkt (also kein "seltsames" Artefakt)
- Seite verlinkt mit wissenschaftlich hochstehenden Seiten
Bsp: Universitäten – Top Level Domäne .edu oder Liste von Unis auf dmoz.org
- Seite verlinkt mit von Menschen als gut bewerteten Seiten
Bsp: Seite ist bei www.dmoz.org gelistet

Site-Struktur

Siehe auch Vortrag
"Aufbau einer Suchmaschine"
Abschnitt Crawler, Deep Web

Beobachtung: Gute Sites sind gut durchdacht

Negativ werten die meisten Suchmaschinen

- **(Zu) (viele) Redirections**
- **(Zu) hohe Pfadlänge im Verzeichnis**
Bsp: <http://www.bsp.de/web/site/version1/de/marketing/kunden/produkt1/p.html>
- **Identische Kopien** von Inhalten auf einer Site
- **Verkehrslenkungssites** ohne echten eigenen Inhalt

Positiv werten die meisten Suchmaschinen

- **Site Map** (Tipp: Google Sitemap Tool benutzen)
- **Statische URLs**

Seiten-Struktur

Beobachtung: Gute Sites produzieren gutes HTML

Positiv werten die meisten Suchmaschinen daher

- **Titel** vorhanden und enthält Suchwort
- **Metatags** vorhanden und enthalten Suchwort
- **Syntaktisch korrektes HTML**

Werkzeug: HTML Validator (Bsp: <http://validator.w3.org>)

- **Keine broken links**

Tip: Auf **geringe Anzahl** achten und bei Fehlern **rasch** beheben

Werkzeug: wusage, Log-Analyser, Google Webmaster Tools

Negativ werten die meisten Suchmaschinen daher

- Seiten mit "problematischem" Inhalt (Viren, Trojaner, Porno)

Text Limits

Suchmaschinen indizieren nicht immer den ganzen Text

Experiment:

- Seiten unterschiedlicher Länge generieren
- Kunstwörter unterbringen
- Detektieren, ob von Suchmaschine gefunden

Wie weit hinein in den Text arbeitet der Indexer

Yahoo	210 kB
Google	520 kB
MSN	1030 kB

Fazit

- Der wirklich wichtige Text sollte eher am Anfang stehen
- Keine extrem langen Seiten einstellen

Worte (1)

Worte nutzen in

- Meta Tags
- Text von Titel, Seite, Incoming & Outgoing Links, alt-Attribut von Bildern
- Text in der Nähe von Incoming & Outgoing Links

Erweiterung um Co-Suchbegriffe

- Welche Suchworte wurden von Usern in Abfragen mit einem gegebenen Begriff auch noch benutzt?
- **Bsp:** "Auto" erweitern auf "Auto Autokauf" oder "Auto Mietwagen"
- **Tipp:** keyDB nutzen <http://www.kso.co.uk/de/tools.html>

Vom Substantiv zum Verb erweitern

- Oftmals wird nur an Substantiv gedacht
- **Tipp:** Um geeignetes Verb erweitern (**Was kann man damit tun ?**)
- **Bsp:** Blumen – pflegen, pflücken, pflanzen, riechen, gießen, schenken, stecken, ...

Worte (2)

Erweiterung auf Stamm-Verwandte

- Suchmaschine testet auch auf Vorhandensein von Stamm-Verwandten
Insbesondere im Text des Artikels
- **Bsp:** spielen, Spiel, Spieler, verspielt
- **Tipp:** Sicherstellen, daß genutzt werden

Erweiterung auf Thesaurus-Verwandte

- Suchmaschine teste auch auf Vorhandensein von Thesaurus-Verwandten
- **Tipp:** Thesaurus nutzen, um Wortliste zu erweitern

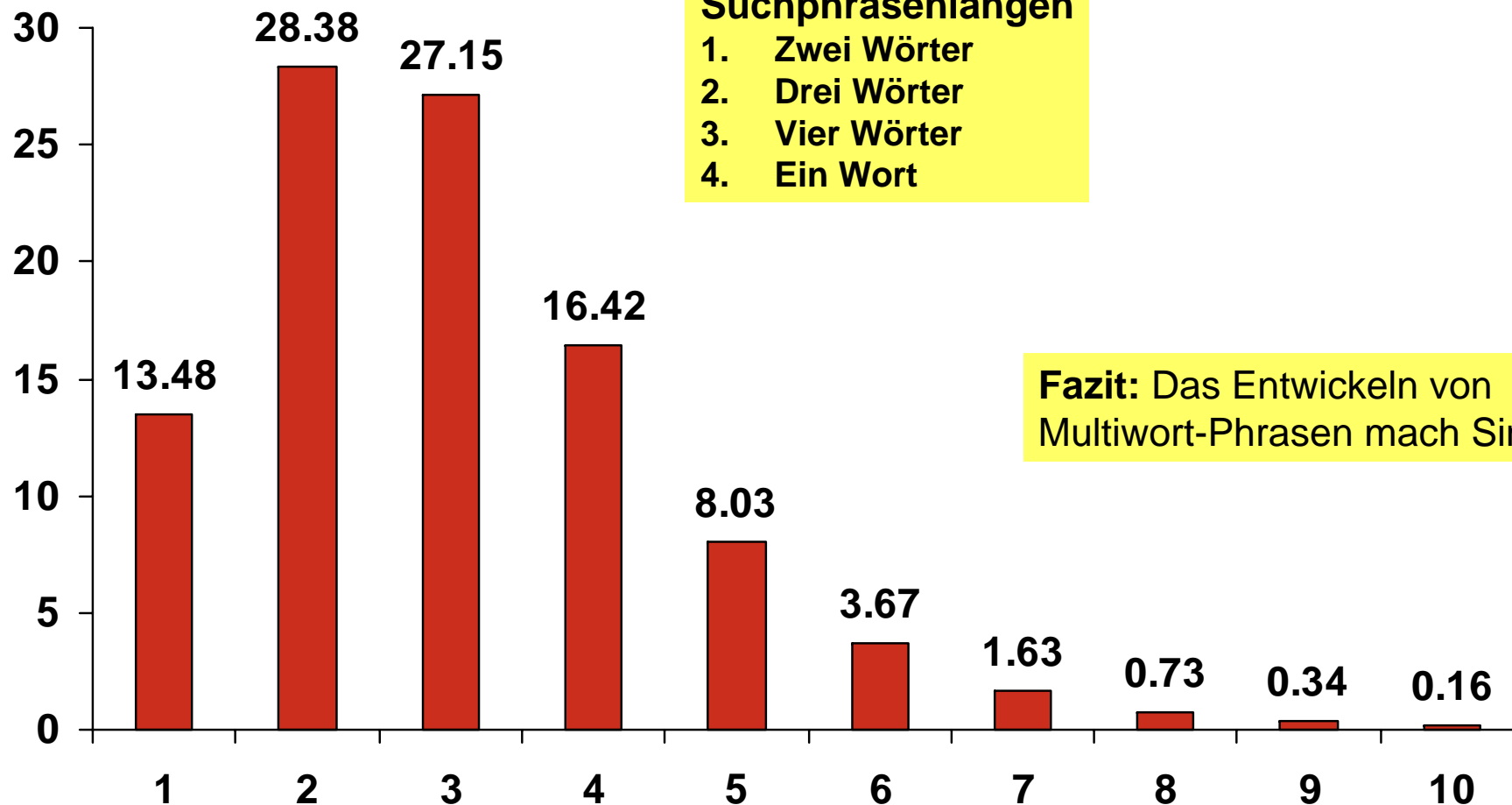
Nach wievielen Worten sucht der Nutzer?

Auswertung: 2 Mio Besucher
100 Länder, Jan 2007

Prozentsatz aller Queries

Ranking von Suchphrasenlängen

1. Zwei Wörter
2. Drei Wörter
3. Vier Wörter
4. Ein Wort



Fazit: Das Entwickeln von Multiwort-Phrasen mach Sinn !

Anzahl Worte in der Suche

Verbesserung externer Verlinkung

Technische Sichtbarkeit

- Jedes elektronisch abgegebene Dokument muß Link auf mich enthalten

Redaktionelle Sichtbarkeit

- Jedes erstellte Dokument muß Link auf mich enthalten

Strategische Sichtbarkeit

- Kooperationen und Zusammenarbeit offen legen
- Link-Sharing und Banner-Tausch

CMS, Wikis und dynamische Seiten

Problem:

- Viele Suchmaschinen ignorieren Seiten mit Query-encoded URLs
- Google ignoriert Seiten mit id Parameter (Idee: Ist Session Id)
- Viele Systeme generieren Seiten aufgrund Eingaben in Formulare

Lösung:

- **CMS System so einstellen, daß keine Query-encoded URLs genutzt werden**
Bsp: Mögliche Option in der Konfiguration von Typo 3
- **Zusätzlich alle möglichen Seiten auch statisch auskoppeln** (für Suchmaschine)
- **Sitemap** zur Verfügung stellen

**Siehe auch Vortrag
"Aufbau einer Suchmaschine"
Abschnitt Crawler, Deep Web**

Whistle Blower Page

Der exakte Suchbegriff, bei dem ein Problem auftritt (kopieren Sie diesen

Die zurückgegebene Google-Ergebnisseite, die das Problem aufweist (kc

Die spezifische verdächtige Webseite oder Website:

Problematische Vorgehensweise(n) (alle zutreffenden markieren):

- Verborgener Text oder Links
- Irreführende oder sich wiederholende Wörter
- Die Seite stimmt nicht mit der Beschreibung von Google überein
- Cloaking
- Irreführende Umleitungen
- Brückenseiten
- Doppelt vorhandene Sites oder Seiten
- Sonstige (bitte angeben)

Google fahndet nach
mißbräuchlichen Seiten

Weitere Informationen:

Trend-Analyse

Absolut Wenn ich weiß, wonach auf Suchmaschinen / Blogs gesucht wird dann kann ich abschätzen, ob auch nach mir gesucht wird

Vergleichend Wenn ich zwei Bezeichnungsalternative habe sollte ich die mehr gesuchte wählen

Idee Eine große Anzahl von Menschen irrt sich selten zumindest bei der Einschätzung von Trends

Werkzeuge

- Google Trends: Benutzte Suchwörter
- Google Zeitgeist
- Technorati: Durchsucht Blogs
- Hotsearches: Zu einzelnen Themenkreisen die aktuellen Suchwörter

Trend-Analyse

Zeitlicher Wechsel von Trends

Frage: Welche Themen liegen im Trend ?



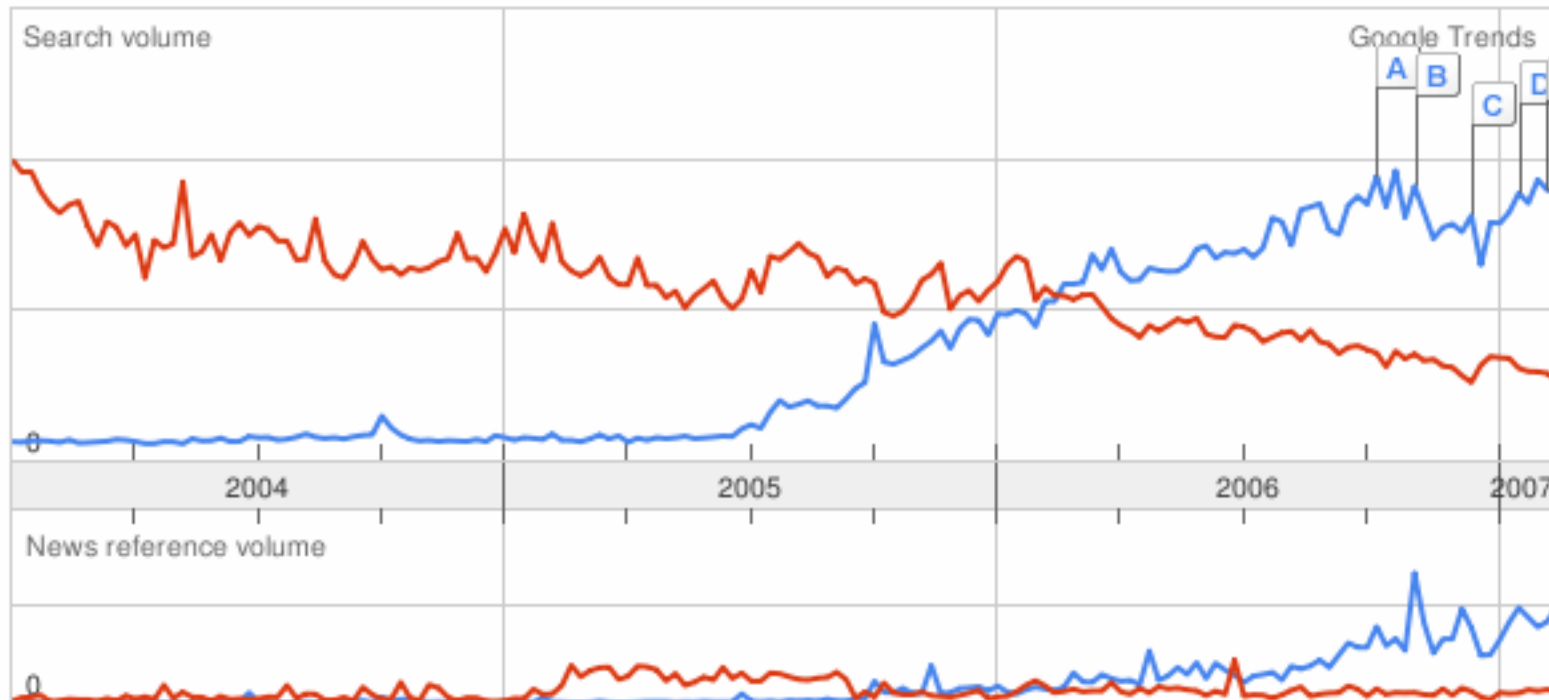
web 2.0, suchmaschinen

Search Trends

Tip: You can compare searches by separating with commas.

Trend history

● web 2.0 ● suchmaschinen



Trend-Analyse

Jahreszeitliche Schwankung von Trends

Frage: Fährt man im Sommer oder im Winter Ski?

Trends fragen !



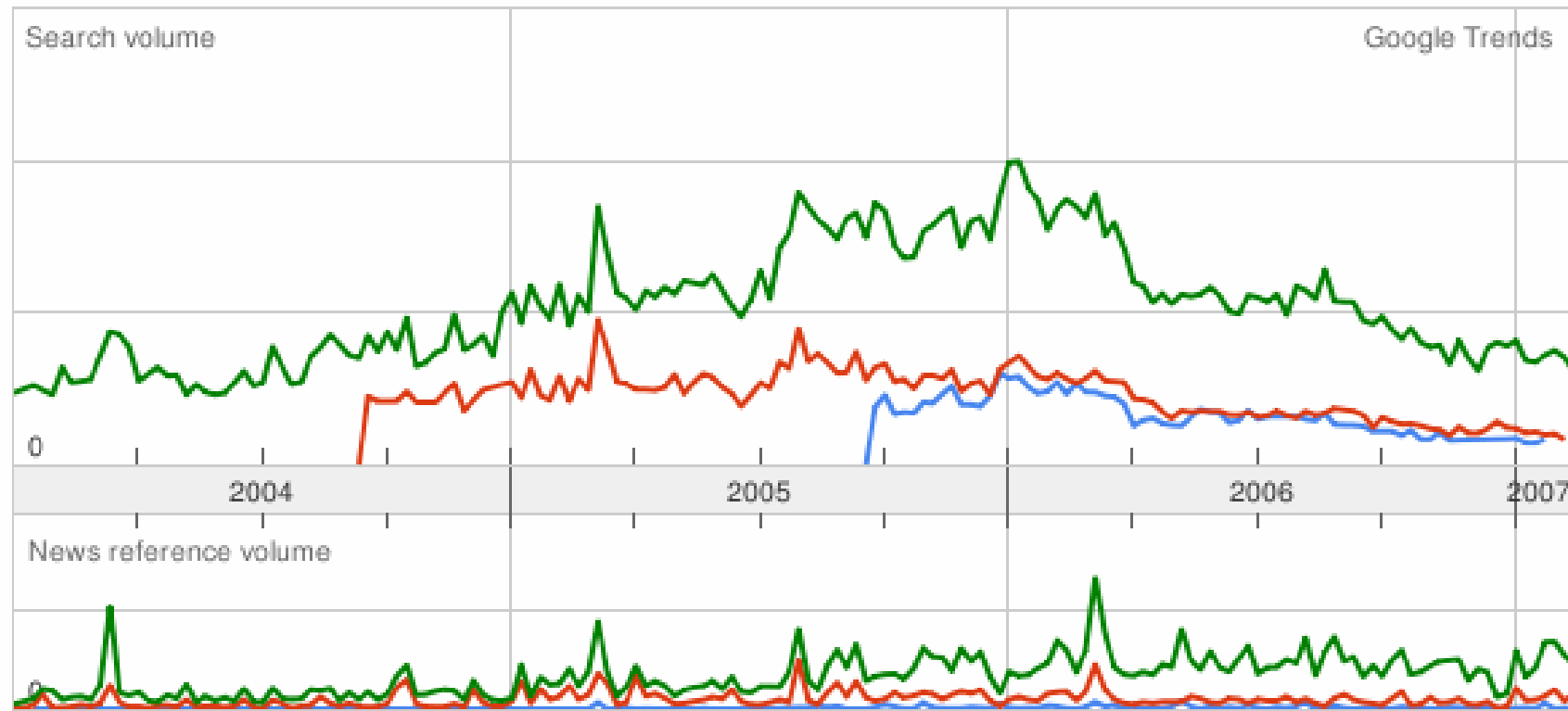
Trend-Analyse

Optimierung von Bezeichnungen

Frage: Welche Bezeichnungen sind besser für Produkte?

Trend history

● dsl telefonie ● internet telefonie ● voip telefonie ● telefonie



Germany

All subregions

All years

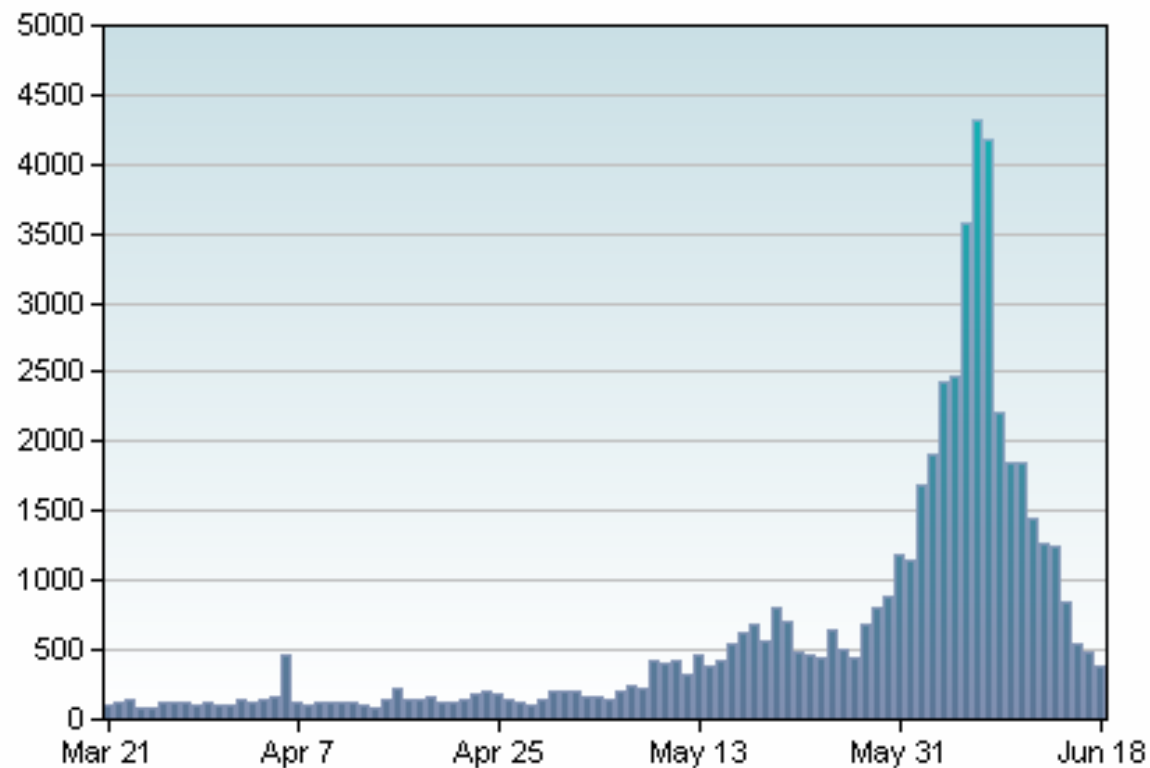
Trend-Analyse

Technorati: Analyse von Blogs

<http://technorati.com/chart/tag:g8>

Posts tagged **g8** in **all languages** per day for the last **90 days**

Chart It



Trend Analyse

hotsearches.aol.com

Spezifische Studien zu einzelnen Bereichen

Statistische Analyse plus menschliche Nachbearbeitung

Top searched cheap items on AOL

- 1) Cheap flights
- 2) Cheap hotels
- 3) Cheap cruises
- 4) Cheap car rental
- 5) Cheap gas
- 6) Cheap books
- 7) Cheap prescription drugs
- 8) Cheap vacations
- 9) Cheap cars
- 10) Cheap Broadway tickets

This month's top searched outdoor

- 1) Patio furniture
- 2) Hot tubs
- 3) Patio umbrellas
- 4) Outdoor kitchens
- 5) Outdoor lighting
- 6) Swimming pool
- 7) Outdoor rugs
- 8) Patio cushions
- 9) Patio doors
- 10) Grill

Top Searches Used Items

- 1) Used cars
- 2) Used trucks
- 3) Used boats
- 4) Used books
- 5) Used motorcycles
- 6) Used auto parts
- 7) Used golf clubs
- 8) Used textbooks
- 9) Used golf carts
- 10) Used furniture