

# Untersuchung des Online Suchverhaltens im Web

## Interaktion von Internetnutzern mit Suchmaschinen\*

*Nadine Höchstötter*

geb. Schmidt-Mänz, ETU, Universität Karlsruhe (TH), *nsh@topicflux.de*

### **Zusammenfassung**

Es wurden zwei empirische Studien durchgeführt. In der einen Studie konnten Internetnutzer ihr Suchverhalten aus ihrer Sicht in einer Online Umfrage beschreiben. Die Fragen orientieren sich an den Ergebnissen früherer Studien und sind durch neue Fragestellungen bei der Online Suche erweitert worden. In der anderen Studie wurden vier weltweit Datensätze von Suchanfragen generiert. Die Datensätze decken einen Zeitraum von mehr als einem Jahr ab. Erstmals sind hier auch Datensätze von Metasuchern gesondert in Betracht gezogen worden. Auf Basis dieser Daten werden Auswertungsmethoden aus der Literatur aufgegriffen und weiterentwickelt. Fehlerhafte und unvollständige Aussagen in anderen Studien können somit formalisiert und korrigiert werden. Die existierende heterogene Begriffswahl wird zu einer neuen, konsistenten Terminologie zusammengefasst. Anhand von diesen vier unabhängigen Datensätzen werden vollständige Statistiken von Suchanfragen erstellt. Die Formalisierungen können als Grundlage und Vergleichswerte für zukünftige Auswertungen von Suchanfragen herangezogen werden.

### **I Beobachtung von Suchmaschinen-Livetickern**

Es werden bekannte Auswertungsmethoden für Suchanfragen aggregiert und daraus Formalisierungen abgeleitet, welche die Vergleichbarkeit empirischer Datensätze von Suchanfragen gewährleisten. Im Anschluss wird die erhobene Datenbasis vorgestellt und die entsprechenden Kennzahlen berechnet. Die Suchanfragen der vier Suchmaschinen Fireball, Lycos, Metaspinner und Metager sind über einen Zeitraum von mehr als einem Jahr gespeichert worden. Die Suchmaschinen Fire-

---

\* Veröffentlicht in: OSSWALD, Achim; STEMPFHUBER, Maximilian; WOLFF, Christian (Hrsg.) (2007). Open Innovation. Proc. 10. Internationales Symposium für Informatikwissenschaft. Konstanz: UVK, 367-370.

ball, Lycos und Metaspinner wurden untersucht, da diese einen Liveticker zur Verfügung stellen, in dem die laufend eingehenden Suchanfragen angezeigt werden. Durch ein Programm wurden diese Suchanfragen abgefragt, um sie anschließend zur weiteren Verarbeitung zu speichern. Von Metager wurden die Top-4000 Suchanfragen für diese Erhebung bereitgestellt. Bei der Suchmaschine Lycos, die von den beobachteten Suchmaschinen am stärksten frequentiert ist, umfasst dieser Datensatz rund 190 Millionen Suchanfragen bzw. knapp 340 Millionen Terme.

In der Arbeit wird dargestellt, wie sich das Volumen von Suchtermen im Zeitverlauf mit der Zu- und Abnahme des Interesses der Suchenden für ein bestimmtes Thema verändert. Diese unterschiedlichen charakteristischen Kurven des Volumens von Termen pro Tag werden in grundsätzliche Klassen eingeteilt. Ein Evergreen-Term liegt bspw. in 90 Prozent aller Zeitintervalle über einem definierten Schwellwert (siehe Abbildung 1).

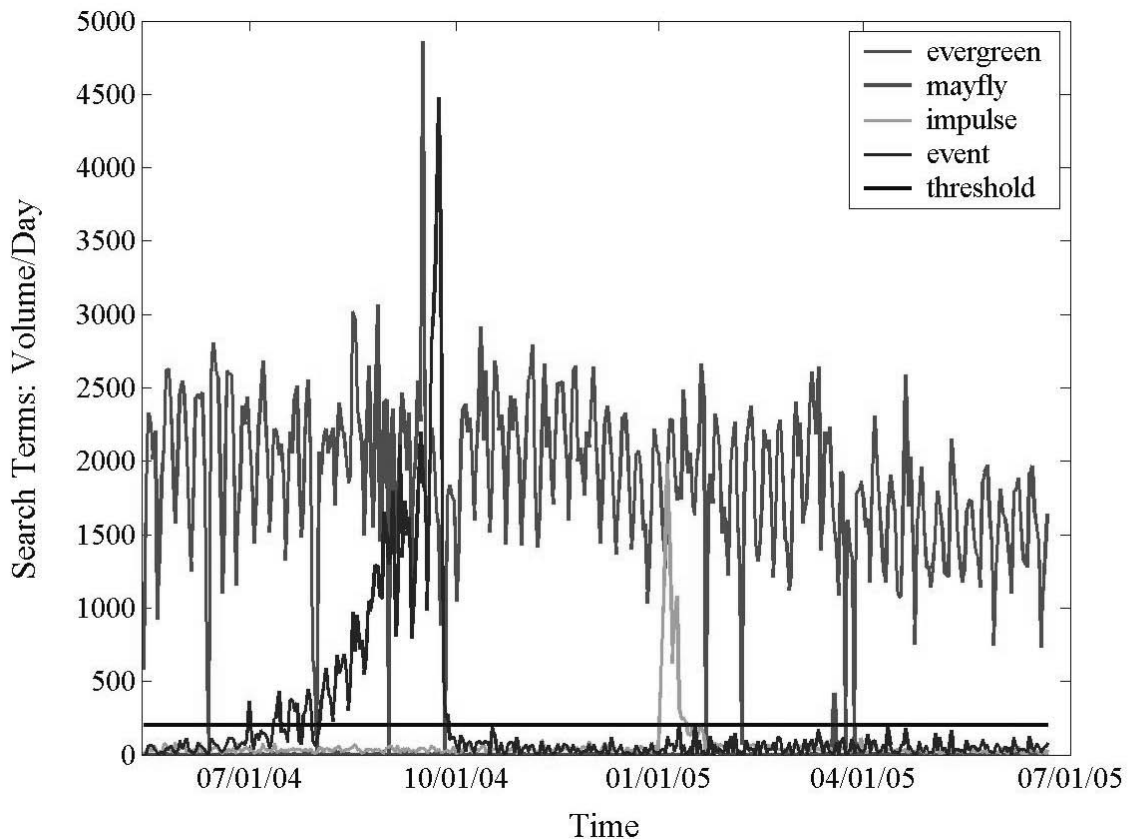


Abbildung 1: Generelle Klassifikation von Suchanfragen

Ein Mayfly-Term bzw. eine Eintagsfliege übersteigt den Schwellwert nur ein oder zwei Mal während der gesamten Erhebungsperiode. Ein Impuls zeigt nach einer Nachricht ein bestimmtes Muster. Direkt nach der Nachricht steigt das Volumen schlagartig an, um nach ein paar Tagen oder einer Woche wieder abzuklingen (Suchbegriffe: Tsunami, Papst). Ein Event ist mit einem wiederkehrenden Ereignis

verbunden und zeigt ein dazu gegenteiliges Muster. Das Volumen der Suchanfragen steigt sich bis zu dem Tag des Ereignisses, um danach abrupt abzufallen (Suchbegriffe: Weihnachten, Bundestagswahl).

Mit der Kenntnis dieser Verteilungen von Termen ist es möglich, diese auch inhaltlich zu klassifizieren. Bei dem vorgestellten Vorgehen werden die Terme aufgrund der Ähnlichkeit ihrer zeitlichen Struktur, mit der diese Terme nachgefragt werden, inhaltlich klassifiziert. Zeitlich ähnliche Terme können bei der Suche z. B. als Empfehlungen dienen.

## 2 Ausblick

Insgesamt stellt diese Arbeit durch den Umfang der Literaturstudie, die durchgeführte Umfrage und den Vergleich verschiedener Suchanfragen-Datensätze eine der vollständigsten Arbeiten auf dem Gebiet des Online Suchverhaltens dar. Durch die Formalisierung und Herausarbeitung der Statistiken von Suchanfragen wird eine Einheitlichkeit auf diesem Gebiet geschaffen, die vorher fehlte. Die Analyse von Suchanfragen und vor allem deren zeitabhängige Betrachtung, sind zukunftsweisend für weitere Forschungsansätze in dem Bereich der Trendforschung. Hierbei ist besonders die Betrachtung von Mikrotrends in der Konsumentenverhaltensforschung als ein wichtiger Punkt zu nennen. Die Suchanfragen spiegeln die Bedürfnisstruktur und das steigende und fallende Interesse von Internetnutzern an aktuellen Themen wider. Auch Phänomene wie Mund-zu-Mund-Propaganda werden bei Suchanfragen deutlich. Die weitere Erforschung von Veränderungen in den Volumina von Suchanfragen wird unter dem Begriff „topic flux“ zusammengefasst. Der Begriff „topic“ steht dabei für Thema oder Gesprächsgegenstand, „flux“ bezeichnet die Durchflussmenge (siehe [www.topicflux.de](http://www.topicflux.de)).

## 3 Literatur

Höchstötter, N. (2007): Suchverhalten im Web – Erhebung, Analyse und Möglichkeiten, wird erscheinen in: *Information: Wirtschaft & Praxis*, 58(2007)3.

Lewandowski, D./ Höchstötter, N. (2007): Qualitätsmessung bei Suchmaschinen, wird erscheinen in: *Informatik Spektrum*, Springer, Heidelberg.

Lewandowski, D./ Höchstötter, N. (2007): Web Searching: A Quality Measurement Perspective, to be published in: Spink, A./ Zimmer, M. (Eds): *Web Search: Interdisciplinary Perspectives*, Springer, Heidelberg.

Schmidt-Mänz, N. (2007). Untersuchung des Suchverhaltens im Web - Interaktion von Internetnutzern mit Suchmaschinen, Dissertation, Verlag Dr. Kovač, Hamburg.

- Schmidt-Mänz, N. (2006): Erkenntnisse aus dem Suchverhalten im Web – "Muster in Suchanfragen", in: Eberspächer, J./ Holtel, St. (Eds): Suchen und Finden im Internet, Springer, Heidelberg.
- Schmidt-Mänz, N. / Bomhardt, Ch. (2005): Wie suchen Onliner im Internet?, in: Science Factory 2/2005, Absatzwirtschaft.
- Schmidt-Mänz, N./ Gaul W. (2005). Web Mining and Online Visibility, in: Weihs, C.; Gaul, W. (Eds.) (2005): Classification – the Ubiquitous Challenge. Springer, Heidelberg, 418-425. Proceedings, Gesellschaft für Klassifikation, 2004.
- Schmidt-Mänz, N./ Koch, M. (2006). A General Classification of (Search) Queries and Terms, Proceedings of the Third International Conference on Information Technology: New Generations (ITNG'06), Las Vegas, Nevada, USA, 375-381.
- Schmidt-Mänz, N./ Koch, M. (2005): Patterns in Search Queries, in: Baier, D., Decker, R., and Schmidt-Thieme, L. (Eds.) (2005): Data Analysis and Decision Support. Springer, Heidelberg, 122-129.