

Erkenntnisse aus dem Suchverhalten im Web „Muster in Suchanfragen“

Nadine Schmidt-Mänz, Institut für Entscheidungstheorie und Unternehmensforschung (ETU),
Universität Karlsruhe (TH), nadine.maenz@wiwi.uni-karlsruhe.de

Einleitung und Motivation

Die hohe Relevanz von Suchmaschinen bei der online Recherche erfordert es, das Suchverhalten von Internetnutzern zu erforschen. Ohne die Kenntnis der Suchstrategie potentieller Besucher von Webseiten können Marketing-Maßnahmen nicht zielgerichtet eingesetzt werden. Die Erforschung des Online Suchverhaltens erhält damit einen sehr wichtigen Bezug zum (Online) Marketing.

Für Betreiber von Suchmaschinen ist es wichtig, Erfahrungen über die Struktur von Suchanfragen zu sammeln, um dem Nutzer mit Hilfe einer optimierten, benutzerorientierten Suchoberfläche Empfehlungen auszusprechen oder Hilfestellungen bei der Formulierung von Suchanfragen zu geben. Weiterhin können Störfaktoren für Internetnutzer bei der Suche im Internet eliminiert werden. Web-Angebote können direkt am Verhalten der Suchenden ausgerichtet werden, indem die Inhalte an die Interessen der Suchenden und deren Suchbegriffe angepasst werden.

Ein Ziel dieser Arbeit lautet daher, nicht nur beschreibende Tatsachen und Auszüge der Ergebnisse darzulegen, sondern ebenfalls Implikationen für die beteiligten Parteien (Suchende, Suchmaschinen und Webseiten) bei der Online Suche abzuleiten.

Im Folgenden werden Kernergebnisse einer 2004 durchgeführten Online Umfrage angesprochen und dazu die Ergebnisse der Analyse von Suchanfragen verschiedener Suchmaschinen präsentiert. Im Anschluss werden Dauerbrenner und Eintagsfliegen bei Suchanfragen vorgestellt. Aus den Ergebnissen werden dann Handlungsempfehlungen abgeleitet und die weiterführende Forschung auf Basis der erhobenen Daten angesprochen.

Kernergebnisse der Umfrage

Die Startseite der Umfrage „Suchmaschinen und Suchverhalten im Internet“ wurde von 6773 Personen aufgerufen, von denen 5925 nahezu alle Fragen beantworteten. In der Umfrage wurden Fragen zum Suchverhalten, zu Suchstrategien und zum Personalisierungsbedarf der Internetnutzer gestellt.

Durch den selbstselektiven Charakter der Online Umfrage nahmen insgesamt versiertere Internetnutzer teil. Durch die Abfrage des Wissens über die Funktionsweise von Suchmaschinen kam aber heraus, dass nur ein Drittel der Teilnehmer tatsächlich über Grundkenntnisse verfügen. Weiterhin war der Altersunterschied bei den antwortenden Personen geringer und der Anteil der männlichen Teilnehmer höher, als dies bei klassisch durchgeführten Umfragen der Fall ist. Dieser Effekt wird bereits bei Theobald (2000) angesprochen.

Insgesamt werden nur wenige Suchbegriffe eingegeben und Ergebnisseiten angeschaut. Die Recherche nach Personen und Produkten spielt eine große Rolle bei der Internetrecherche. Die Probanden lehnen Werbung im Internet ab und führen keine Personalisierung „ihrer“ Suchmaschine durch, indem die Suchmaschine den Bedürfnissen angepasst wird. Die Teilnehmer störte vor allem die Angabe von Ebay-Webseiten und Trefferduplikaten in den Ergebnislisten.

Bei der Durchführung von Gruppenvergleichen zeigte sich, dass Angestellte gegenüber der Gruppe der Studenten eine zeitoptimiertere Herangehensweise an die Internetrecherche haben. Frauen stöbern gegenüber Männern mehr in Nachrichten. Nutzer, die mehr Grundkenntnisse über die Funktionsweise von Suchmaschinen besitzen, lehnen Werbung stärker ab und gehen bei der Suche zielgerichteter vor.

Analyse von Suchanfragen

Seit 2004 werden drei Suchmaschinen (Fireball, Lycos, Metaspinner) beobachtet, indem die Suchanfragen aus der LiveSuche gespeichert und geeignet ausgewertet werden. Die LiveSuche wird nicht von allen Suchmaschinen angeboten. Es werden bei dieser Funktion zeitgleich durchgeführte Suchanfragen anderer Suchmaschinennutzer der gleichen Suchmaschine in einer Liste angezeigt. Von Metager wurden die top 4000 Suchbegriffe des Tages per eMail zugesendet.

Suchmaschine	Erster Tag	Letzter Tag	Zeitraum	Anzahl Tage
Fireball	18.08.2004	20.09.2005	13 Monate	399
Lycos	14.08.2004	20.09.2005	13 Monate	403
Metager	11.11.2004	20.09.2005	10 Monate	314
Metaspinner	28.09.2004	20.09.2005	12 Monate	358

Tabelle 1: Erhebungszeiträume der Suchmaschinen

Die Erhebung der Suchanfragen wird für ausgewählte Suchmaschinen fortgeführt, um die aktuelle Datenbasis von Suchanfragen weiter auszubauen. Die Ergebnisse, die an dieser Stelle vorgestellt werden, beziehen sich auf einen Zeitraum bis zum September 2005. Die genauen Erhebungszeiträume sind in Tabelle 1 dargestellt.

Als Eckdaten zur Beschreibung des Datensatzes seien die gesamte Anzahl der Suchanfragen (SA) und der Terme genannt. Die Summe aller Suchanfragen, die während eines bestimmten Zeitraumes eingegangen sind, wird als Bruttoanzahl der Suchanfragen bezeichnet. Das gleiche gilt für die Terme. Die Anzahl der eindeutigen Suchanfragen bzw. Terme wird als Nettoanzahl angeführt. Hierbei wird nicht die Anzahl des Auftauchens summiert, sondern die Anzahl Suchanfragen oder Terme, die überhaupt aufgetaucht sind. Die Zahlen sind in Tabelle 2 gegenübergestellt. Von Spink/Jansen (2004) oder Silverstein et al. (1999) wurden ähnliche Erhebungen durchgeführt. Die Datenbasis war jedoch jeweils eine andere und wurde zudem zu unterschiedlichen Zeitpunkten erhoben. Es mangelt dadurch an Vergleichbarkeit der einzelnen Studien.

Suchmaschine	Brutto SA	Netto SA	Brutto Terme	Netto Terme
Fireball	132.833.007	17.992.069	241.833.877	6.296.833
Lycos	189.930.859	29.322.366	344.242.099	11.232.710
Metager (Top-4000)	4.407.566	678.655	7.333.343	430.338
Metaspinner	4.089.731	1.287.417	7.853.501	627.507

Tabelle 2: Brutto- und Nettoanzahlen der Suchanfragen und Terme

Bei der Analyse der Suchanfragen stellte sich beispielsweise heraus, dass die Anfragelänge in allen Suchmaschinen ähnlich kurz ist. Die durchschnittliche Anzahl der Terme pro Suchanfrage beträgt 1,6 (Metager), 1,7 (Lycos) und 1,8 (Fireball und Metaspinner). Diese Zahlen basieren auf der Bruttomenge der Suchanfragen.

Der prozentuale Anteil von allen Suchanfragen, die eine Phrasensuche (z.B. "Münchner Kreis") enthielten, war sehr gering. Bei Fireball waren nur 2,1%, bei Lycos nur 1,7% und bei Metaspinner nur 2,1% vorhanden. Für Metager wurde diese Zahl nicht bestimmt, da dies auf Grund der vorliegenden Datenbasis nicht möglich war.

Bei Fireball und Metaspinner konnten zusätzlich die Suchbereiche ausgewertet werden. Es stellte sich hier heraus, dass bei beiden Suchmaschinen die „Deutsche Suche“ bevorzugt wird. Bei Fireball lag hier der Anteil bei 66,4% und bei Metaspinner bei 88,1%. Bei Fireball spielt danach die Bildersuche mit 24,5% die nächstwichtigste Rolle. Diese Ergebnisse sind nicht

überraschend, da die Suche nach deutschen Inhalten der Voreinstellung entspricht. Passend zu den Antworten der Umfrage wird also keine personalisierte Suche durchgeführt.

Ein weiteres Ergebnis aus der Analyse ist, dass ein starker Montagseffekt bei der Internetsuche vorhanden ist. Die Internetnutzer suchen verstärkt an Montagen. Das Volumen der Suchanfragen pro Tag nimmt dann während der Woche ab, um am Wochenende wieder langsam anzusteigen. Man kann hier die Hypothese unterstellen, dass die Suchenden die Informationen, die an Wochenenden gesammelt werden, an dem darauffolgenden Tag verarbeiten. Liegt auf einem Montag ein Feiertag (bspw. Ostermontag), dann verschiebt sich der Effekt auf den Dienstag darauf. Eine weitere Hypothese ist, dass die Internetrecherche nicht als Freizeitvergnügen angesehen wird. Sie ist durchaus zeitaufwendig und wird zu Gunsten anderer Aktivitäten am Wochenende zurückgestellt. Der Start in die Woche beginnt dann mit dem Abbau des „Informationsdefizits“.

Bei der Betrachtung der Top-10 der Suchanfragen und Terme wird deutlich, dass diese Listen nicht repräsentativ sind, beziehungsweise aus verschiedenen Gründen absichtlich verfälscht werden. In diesen Listen sind lediglich erotische Inhalte und Füllwörter zu finden. Zudem sind die Listen auf Seiten der Suchanfragen ebenfalls durch Terme besetzt: Die häufigsten Suchanfragen bestehen nur aus einem Term. Auf eine Darstellung wird in diesem Rahmen verzichtet.

Um die Relevanz von Termen oder Themengebieten zu erkennen, muss eine tiefere und zeitabhängige Auswertung vorgenommen werden.

Dauerbrenner und Eintagsfliegen

In diesem Zusammenhang werden nur noch Terme betrachtet, da Suchterme im Schnitt häufiger wiederauftauchen als Suchanfragen. Um einen Einblick in die Verteilung der Terme zu bekommen, wird betrachtet, wie groß die Nachfrage nach einem Term pro Tag ist. Taucht ein Term während der gesamten Periode nur an einem Tag auf, dann wird dieser als eine Eintagsfliege (Mayfly) bezeichnet. Kommt er an vielen Tagen der Erhebungsperiode vor, dann wird von einem Dauerbrenner (Evergreen) gesprochen. Man kann einen Dauerbrenner auch dahingehend definieren, dass er nur an 90% der erhobenen Tage auftauchen muss. Weiterhin kann wahlweise auch eine Mindestschranke eingeführt werden, über der das nachgefragte Volumen eines Terms an einem Tag liegen muss, damit er überhaupt in Betracht gezogen wird.

Die Dauerbrenner zeigen, dass über den erhobenen Observationszeitraum verstärkt nach Erotik, Reise, Multimedia und Füllwörtern gesucht wird. Der Einblick in die Dauerbrenner liefert bessere und detailliertere Ergebnisse als die Betrachtung der Top-10 Listen. Weiterhin ist erkennbar, dass die Suchenden keine Grundkenntnisse über die Funktionsweise von Suchmaschinen haben, da ansonsten nicht so viele Personen Füllwörter eingäben, die bei der Suche nach Webseiten keine verfeinernde Wirkung für die Suche haben.

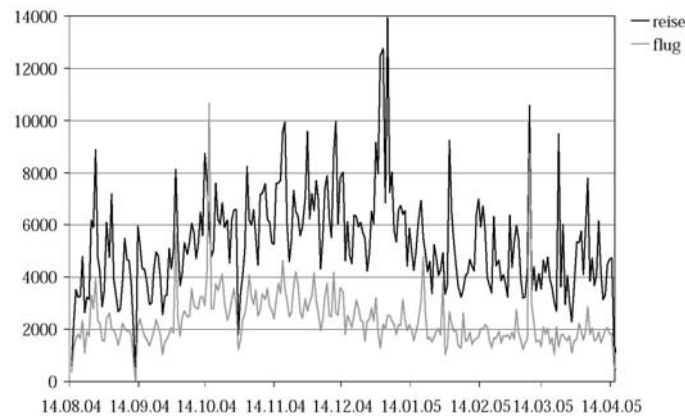


Abbildung 1: Nachfragevolumen von Dauerbrennern während der Erhebungsperiode

Sieht man sich den Verlauf des Nachfragevolumens über die Zeit der Erhebungsperiode in Abbildung 1 bei den Termen Reise und Flug an, dann ist erkennbar, dass sich Dauerbrenner auf einem sehr hohen Niveau befinden und kontinuierlich über einem imaginären Schwellwert liegen. Es treten aber trotzdem Schwankungen auf.

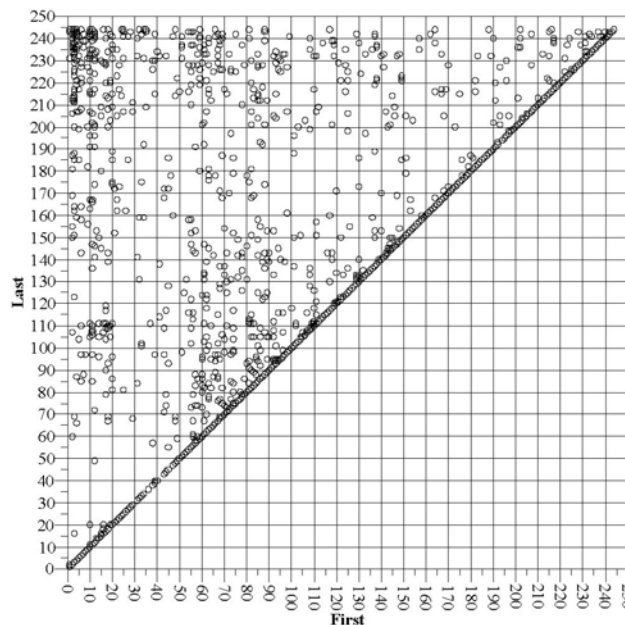


Abbildung 2: Erstes vs. letztes Auftreten eines Terms

In Abbildung 2 ist das erste Auftreten eines Terms (x-Achse) gegenüber dem letzten Auftreten (y-Achse) abgetragen. Wie man sieht sind sehr viele Eintagsfliegen vorhanden, welche die Winkelhalbierende bilden. Die Dauerbrenner sind als dunkler Pulk in der linken oberen Ecke zu finden.

Insgesamt konnten mit der Definition dieser Terme generelle zeitliche Klassen von Termen gefunden werden. Interessant ist die Kenntnis dieser Dauerbrenner, um beispielsweise Caching Strategien für Suchmaschinen abzuleiten. Terme und Suchanfragen, die immer nachgefragt werden, müssen nicht bei jeder neuen Anfrage auf ein Neues gesucht werden. Die Personen, die nach diesen Termen suchen, können einfach aus dem Cache bedient werden. Zudem müssen die Trefferlisten nicht jedes Mal neu berechnet werden, sondern können als statische Liste der empfohlenen Webseite ausgegeben werden, die nur von Zeit zu Zeit aktualisiert werden.

Handlungsempfehlungen und weitere Forschung

Da die Suchenden keine Operatoren bei der Suche einsetzen und auch die Phrasensuche selten genutzt wird, um Suchbegriffe einschränkender zu verwenden, wird empfohlen, eine Tabu-Suche einzuführen. In diesem Fenster kann der Suchmaschinennutzer Begriffe eingeben, die er auf keinen Fall in den darauf angezeigten Ergebnissen finden möchte. Dies hilft dem Konsumenten bei der Suche insofern, dass er meist nicht artikulieren kann, was er haben möchte, aber sehr wohl deutlich machen kann, was er nicht möchte.

Der Einsatz von Werbung sollte gezielt vorgenommen werden, da ein Ergebnis der Umfrage war, dass Werbung generell abgelehnt wird.

Die Caching Strategien für Dauerbrenner von Suchtermen wurden schon oben besprochen. Für die Gestaltung von Websites ist die Kenntnis dieser Terme interessant, um beispielsweise Themen auf den Seiten aufzunehmen, die einen Dauerbrenner-Charakter haben. Bei den Eintagsfliegen können keine Strategien abgeleitet werden, da diese Terme zu selten auftauchen, als dass sie für Marketing-Aktionen interessant werden.

Insgesamt wird die weitere Erforschung der online Informationsverarbeitung von Menschen angestrebt. Wenn es Dauerbrenner und Eintagsfliegen gibt, dann sind dazwischen ebenfalls Terme vorhanden, die andere Charakteristiken aufweisen. Beispielsweise sind periodisch auftretende Terme vorhanden und solche, die in Zusammenhang mit einem bestimmten Ereignis auftreten. Diese Charakteristiken sind interessant, um Empfehlungen auszusprechen,

wann und wie lange bezahlte Anzeigen in Trefferlisten von Suchmaschinen gebucht werden sollen. Diese Empfehlungen des Buchens geeigneter Schlagwörter können allgemein ausgesprochen werden, wie im Fall der Dauerbrenner, oder sie werden für eine Webseite angepasst, nachdem diese inhaltlich analysiert wurde. Man kann also noch zwischen der globalen und der spezialisierten Analyse solcher Muster unterscheiden.

Zudem ist es möglich, durch die Ähnlichkeiten im zeitlichen Verlauf der Nachfrage verschiedener Terme Rückschlüsse zu ziehen, dass diese ein gemeinsames Thema betreffen. Für weitere Schlüsse und Analysen ist die mathematische Formulierung verschiedener Termklassen und deren Charakteristika wichtig.

Die Kenntnis der Dauer des online Informationsbedürfnisses von Suchenden ist für die Zukunft wichtig, um mehr über die Psychologie des Menschen bei der Informationsaufnahme zu erfahren. Die changierenden Volumina von Suchanfragen pro rata temporis können als Gehirnströme gesehen werden, die das Informationsbedürfnis wiedergeben. Verschwindet dieser Strom, dann ist die Recherche online abgeschlossen, die Suchenden sind dann aber offline immer noch auf einem aktivierten Niveau der Informationsaufnahme. Mit dieser Kenntnis lassen sich auch offline Marketing-Aktionen verbessern. Hierbei ist die Länge der Informationsaufnahme und die Aktivität der Gehirnströme, die mit der Wichtigkeit bzw. Brisanz des Themas zusammenhängen, erhebenswert.

Weiterhin können durch das Erkennen von (Mikro)Trends in Suchanfragen neue Möglichkeiten von Produkten und Diensten aufgezeigt werden. Ein Anbieter kann so an einem sehr frühen Zeitpunkt auf Bedürfnisse des Konsumenten reagieren bzw. diese vorhersehen. Die Einbeziehung dieser psychologischen Aspekte bei der online Suche werden zukunftsweisend für die Weiterentwicklung von Suchmaschinen und anderer online Dienstleistungen sein.

Die Datenerhebung von Suchanfragen wird weitergeführt, um diese Forschungsansätze weiterzuverfolgen. Mit der wachsenden Länge der Observationsperiode werden zeitliche Analysen von Periodizitäten und Trends besser möglich. Die Muster auf globaler Ebene werden ebenfalls deutlicher und aussagekräftiger werden.

Die hier vorgestellten Ergebnisse stellen einen kurzen Auszug der durchgeführten Analysen und Ergebnisse dar. Detailliertere Auszüge sind bei Schmidt-Mänzl/ Koch (2005 und 2006) und bei Schmidt-Mänzl/Bomhardt (2005) zu finden.

Literatur

Schmidt-Mänz, N./Bomhardt, Ch. (2005): Wie Suchen Onliner im Internet?, in: Science Factory 2/2005, Absatzwirtschaft, <http://www.absatzwirtschaft.de/pdf/sf/Maenz.pdf>.

Schmidt-Mänz, N./Koch, M. (2006): A General Classification of (Search) Queries and Terms, angenommen, 3rd International Conference on Information Technologies: Next Generations, Las Vegas, Nevada, USA.

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, S. 122-129.

Silverstein, C./Henzinger, M./Marais, H./Moricz, M. (1999): Analysis of a Very Large Web Search Engine Query Log, ACM SIGIR Forum, 33(1), S. 6-12.

Spink, A./Jansen, B. (2004): Web Search: Public Searching of the Web, Kluwer Academic Publishers.

Theobald, A. (2000): Das World Wide Web als Befragungsinstrument, Gabler Edition Wissenschaft, Interaktives Marketing.