

Suchverhalten im Web – Erhebung, Analyse und Möglichkeiten

Nadine Höchstötter, Karlsruhe

1 Einleitung

Im Januar 2005 nutzten 13 Millionen Deutsche Suchmaschinen. Im Schnitt werden Suchmaschinen von jedem Nutzer 17 Mal pro Monat genutzt. Google hält seine Quasi-Monopolstellung in Deutschland mit 87 Prozent aller gestellten Suchanfragen. Darauf folgen Yahoo mit 3,3 Prozent und MSN mit 2,4 Prozent¹. Deutschland war zudem im Jahr 2004 der zweitgrößte Markt für Suchmaschinen-Werbung. Es wurden ungefähr 165 Millionen Euro für Anzeigen in Suchmaschinen und Programme für Suchmaschinen-Marketing ausgegeben. Für das Jahr 2010 werden von Forrester 400 Millionen Euro prognostiziert (Lake 2005). Das zeigt, dass Suchmaschinen eine Dienstleistung darstellen, die sowohl im Alltag als auch im Berufsleben eingesetzt wird, um Informationen oder neue Webseiten beziehungsweise Produkte zu finden. Damit entwickelte sich das Online-Marketing verstärkt zu einem Suchmaschinen-Marketing, bei dem es vor allem darum geht, die obersten Plätze in den Ergebnislisten einzunehmen, wenn eine Suchanfrage passend zu einer Seite eingegeben wird. Um Informationen oder Webseiten tatsächlich effizient im Web anzubieten, reicht jedoch ein reines Suchmaschinen-Marketing nicht. Anbieter von Web-Inhalten sollten sich auch überlegen, wie sie anhand des Suchverhaltens von Internetnutzern ihre Online-Sichtbarkeit verbessern können, die nicht nur durch eingehende Links und Buchung von Anzeigen in Suchmaschinen gefördert werden kann, sondern auch durch eine geeignete Wortwahl auf den Webseiten und eine Ausrichtung an der Bedürfnisstruktur der Suchenden (Schmidt-Mänz/Gaul (2005), Schmidt-Mänz/Koch (2006)).

Der vorliegende Beitrag geht im nächsten Kapitel auf die gängigsten Methoden der Datenerhebung zur Untersuchung des Suchverhaltens ein und legt deren Vor- und Nachteile dar. Im dritten Kapitel werden die Auswertungsmöglichkeiten, die gerade bei der Erhebung von Suchanfragen in Frage kommen, erklärt und an-

hand der Literatur der letzten Jahre zusammengefasst. Im letzten Kapitel wird darauf eingegangen, wie die Kenntnis des Verhaltens dafür verwendet werden kann, Suchmaschinen und Webseiten zu verbessern.

2 Erhebungsmethoden

Im folgenden Kapitel werden Methoden vorgestellt, mit denen das Suchverhalten und die Interaktion von Menschen mit Suchmaschinen erhoben werden können. Es wird dabei jeweils die Vorgehensweise kurz erklärt und anschließend die Vor- und Nachteile genannt.

2.1 Logfile-Analysen

Die häufigste Methode, die eingesetzt wird, ist die Analyse von Logdateien, die bei Suchmaschinen gespeichert werden. Eine Logdatei beinhaltet das elektronische Protokoll der Interaktionen aller Suchmaschinennutzer mit einer Suchmaschine, welches automatisch erstellt wurde. Informationen, die dabei gespeichert werden, sind beispielsweise die IP-Adresse, die Suchanfrage des Nutzers und die Zeit, zu der die Suchanfrage einging (Jansen 2006). Informationen wie die Angabe, ob Operatoren oder andere Einschränkungen wie die Bildersuche genutzt wurden, sind ebenfalls verfügbar. In vielen Fällen werden auch Cookie-Angaben gespeichert, die wie die IP-Adresse Rückschlüsse auf den Rechner geben, von dem aus die Suche gestartet wurde. Außerdem sind noch die Position der aufgerufenen Dokumente innerhalb der Ergebnisseiten und deren URL verfügbar (siehe Tabelle 1 auf der nächsten Seite). Eine der bekanntesten und ersten Auswertungen einer solchen Logdatei wurde von Silverstein et al. (1999) durchgeführt. Diese Logdatei wurde 1997 von der Suchmaschine Altavista zur Verfügung gestellt.

¹ www.webhits.de/deutsch/index.shtml?deutsch/webstats.html [1.2.2007]

Tabelle 1: Übersicht über die gängigsten Einträge einer Logdatei (In Anlehnung an Silverstein et al. (1999) und Jansen (2006))

Information in der Logdatei	Beschreibung
Identifizierung des Nutzers	IP-Adresse des Computers. Oft werden diese als Hash-Zahlen ausgegeben, damit Rückschlüsse auf einzelne Nutzer, aber nicht auf die IP möglich sind.
Zeitstempel (timestamp)	Gibt an, zu welcher Uhrzeit eine Suchanfrage abgesetzt wurde. Der Zeitstempel wird seit dem 1. Januar 1970 in Millisekunden gemessen.
Cookie	Gibt ebenfalls Auskunft über den Nutzer, der eine Suchanfrage abgesetzt hat; es handelt sich hierbei ebenfalls um eine eindeutige Kennung. Wenn ein Nutzer keine Cookies zulässt, bleibt dieses Feld leer.
Suchterme	Einzelne Wörter, Zahlen und Begriffe, genau so, wie sie vom Nutzer eingegeben wurden.
Ergebnisfenster	Die Seitennummer des Fensters mit den Suchergebnissen, die von einer Suchmaschine nach der gestarteten Suche zurückgegeben wird. Auf der Seite sind auch Anzeigentexte sichtbar, die passend zu einem Suchterm in der Suchanfrage ausgegeben werden.
Zusätzliche Information	Angabe von speziellen Suchbereichen oder sonstigen Einschränkungen, die von einem Nutzer vorgenommen wurden.
Browser	Angabe des Browsers, der genutzt wird.

Der Vorteil bei der Auswertung solcher Logdateien ist die verfügbare Menge an Informationen zu einem Nutzer. Durch die IP-Adresse oder die Cookie-Information kann der Suchprozess von der ersten bis zur letzten Anfrage ausgewertet werden. Ein Vorteil bei der Auswertung von solchen Massendaten ist auch, dass man einen Eindruck von deskriptiven Statistiken von Suchanfragen bekommt. Die Verteilung von Suchanfragen und deren Längen kann beispielsweise erhoben werden. Ebenso kann berechnet werden, wie groß der Anteil von Operatoren in der Menge aller erhobenen Suchanfragen ist. Damit sind zahlreiche Möglichkeiten gegeben, den Suchprozess zu beobachten. Bei Spink et al. (2000) wird beispielsweise auf die Veränderung einer Suchanfrage während des Suchprozesses eingegangen. Ein Problem bei der Betrachtung von Suchsitzungen (*sessions*) ist allerdings, dass nicht genau festgestellt werden kann, wann diese beendet ist. Ein Nutzer ruft einen Link in der Ergebnisliste auf und „verlässt“ dadurch die Suchmaschine. Zur Definition, wann genau eine Suchsitzung eines Nutzers beendet ist, werden deshalb verschiedene Möglichkeiten gegeben. Die erste Suchanfrage eines Nutzers eröffnet dabei immer die Sitzung, fraglich ist hingegen, wann diese Sitzung als beendet gilt. Eine Möglichkeit ist die Einführung einer zeitlichen Begrenzung. Ruft der Nutzer einen Link innerhalb der Suche auf und kehrt von dieser aufgerufenen Seite innerhalb von beispielsweise zehn Minuten nicht zurück,

dann gilt die Sitzung als beendet. Die nächste Suchanfrage des Nutzers gilt dann als die erste Anfrage einer neuen Suchsitzung.

Ein weiterer Vorteil bei der Betrachtung von Logdateien ist auch, dass sich die Nutzer bei dieser Erhebungsmethode unbeeobachtet fühlen, da sie nicht wissen, dass ihre Daten erhoben werden. Die Analyse von Logdateien von Suchmaschinen ist damit als Feldexperiment zu sehen. Die Ergebnisse sind objektiver und realistischer beziehungsweise als repräsentativer einzuschätzen, als dies bei anderen Methoden der Fall ist.

Nachteilig bei der reinen Auswertung ist das fehlende Wissen über die Sozio-Demographie des Nutzers. Außer einer Identität in Form einer Nummer sind keine Informationen wie Alter, Geschlecht, Familienstand, Gehalt oder Angaben zur Bildung bekannt. Ein weiterer Nachteil ist die gefährdete Privatsphäre von Nutzern, wie das Beispiel der Logdatei der Suchmaschine AOL zeigt, die 2006 zum allgemeinen Download bereitgestellt wurde. Die Logdatei enthielt ungefähr 20 Millionen Suchanfragen von rund 500.000 Nutzern über den Zeitraum März bis Mai 2006². Durch die Suchanfragen zu einem Nutzer, der durch eine gesonderte Identitätsnummer klar nachverfolgbar ist, sind Rückschlüsse auf dessen Gewohnheiten und zum Teil auch auf seinen Namen oder den von Bekannten möglich. Oft werden von Personen, die ihre eigene Website besitzen, ihre eigenen Namen eingegeben, um nachzuprüfen, wie ihre Seite gegenüber anderen Suchergebnissen auf den Ergebnisseiten abschneidet. Das Gleiche gilt bei der Suche nach Geschäftspartnern oder Bekannten, die gesucht werden.

Problematisch ist auch die Bereitstellung solcher Dateien, weil Suchmaschinen den Zugriff auf ihre Logdateien erlauben müssen. Dies wird wiederum von nahezu allen Suchmaschinen kategorisch abgelehnt. Im Hinblick auf das Beispiel der Veröffentlichung der AOL-Logdateien ist dies durchaus verständlich. Es werden meist, wenn überhaupt, auch zu Forschungszwecken nur sehr kleine Datensätze von einem Tag verfügbar gemacht. Beispiele zu der Analyse von Logdateien und ein Überblick solcher Ergebnisse sind in Jansen/Spink (2006) und Spink/Jansen (2004) zu finden.

2.2 Liveticker

Logdateien von Suchmaschinen stehen in den meisten Fällen nicht zur Verfügung. Es ist jedoch durch die Beobachtung sogenannter Liveticker möglich, auf eine ähnlich umfangreiche Datenbasis zuzugreifen. Manche Suchmaschinen wie Lycos oder Fireball stellen derartige Liveticker³ zur Verfügung. Die Webpage wird beispielsweise alle zehn Sekunden aktualisiert und zeigt die gerade eingegangenen letzten zwanzig Suchanfragen. Mit Hilfe dafür konstruierter Programme ist es möglich, diese Ticker abzufragen. Um eine vollständige Liste mit allen Suchanfragen zu erhalten, muss der Liveticker öfter als im Zehntel Sekundenabstand abgefragt werden, denn innerhalb von zehn Sekunden gehen in der Regel mehr Suchanfragen ein als im Liveticker angezeigt werden. Die automatische Abfrage des Livetickers erfolgt deshalb ungefähr alle halbe Sekunde. Dadurch sind aber einige Suchanfragen doppelt in der Datenbasis. Die gespeicherten Listen (immer ein Stapel von zwanzig Suchanfragen in unserem Beispiel) werden jeweils miteinander verglichen und so kombiniert, dass alle redundanten Suchanfragen entfernt werden und eine zeitlich geordnete Reihe von Suchanfragen entsteht. So ist es möglich bei Stoßzeiten alle Suchanfragen zu erhalten. In ruhigeren Zeiten, in denen weniger Suchanfragen an die Suchmaschine geschickt werden, werden sehr viele redundante Listen erzeugt. Insgesamt lassen sich auf diese Weise jedoch nahezu vollständige Datenbestände zur Auswertung erzeugen. Von eigens dafür vorgesehenen Suchanfragen (100.000 fingierte eindeutige Suchanfragen) waren ungefähr 99,5 Prozent in der Erhebung enthalten. Anschließend wurden diese wieder gelöscht, um die Ergebnisse nicht zu verfälschen (Schmidt-Mänz 2007).

Vorteil bei dieser Erhebungsmethode ist, dass man nicht von der Bereitstellung von Logdateien durch Suchmaschinen abhängig ist. Zudem kann eine solche Erhebung über einen sehr langen Zeitraum fortgeführt werden. Durch die Beobachtung verschiedener Liveticker ist der Vergleich von verschiedenen Suchmaschinen und deren Nutzung möglich. Vorteilhaft

2 www.infoworld.com/article/06/08/07/HNaol_searchdata_1.html [1.2.2007]

3 www.lycos.de/suche/livesuche.html [1.2.2007]

ist auch, dass die Anonymität der Nutzer auf jeden Fall gewahrt bleibt und die erhobenen Daten eine objektive Sicht auf die Nutzungsweise geben, da sich die Nutzer unbeobachtet fühlen.

Nachteilig sind die fehlenden Informationen zu der genauen Interaktion der Nutzer während einer Suchsitzung. Es können also keinerlei Auswertungen auf Basis verschiedener Nutzer oder Suchsitzungen gemacht werden. Die einzige Information, die bei Livetickern gegebenenfalls angezeigt wird, ist die Nutzung von speziellen Suchbereichen. Zudem wird bei dem Aufruf der nächsten Ergebnisseite die Suchanfrage erneut angezeigt, wodurch die Anzahl der gestellten Suchanfragen verfälscht werden kann. Es ist auch nicht möglich, nachzuvollziehen, inwieweit Suchanfragen automatisiert eingegeben wurden.

Ergebnisse aus der Erhebung von Livetickern sind bei Schmidt-Mänz (2007) und Zien et al. (2000) beschrieben.

2.3 Umfragen

Eine weitere Möglichkeit ist die Durchführung einer Umfrage. Bei einer Umfrage zum Thema Suchmaschinen bietet es sich an, diese online durchzuführen; ein Vorteil sind die geringeren Kosten und die elektronische Verfügbarkeit der Daten. Schwierig ist die Wahl der Fragen, mit deren Hilfe das Verhalten von Personen bei der Suche abgefragt werden soll. Beispiele solcher Umfragen sind in Schmidt-Mänz (2007, S. 29-59) ausführlich zusammengefasst. Bei der Durchführung einer Online-Umfrage ist darauf zu achten, dass mit leichten Fragen begonnen wird, um dann systematisch das Abstraktionsniveau zu erhöhen. So kann am Anfang mit der Frage nach der Häufigkeit der Internetnutzung und der Nutzung von Suchmaschinen begonnen werden, um später Fragen nach der Nutzung von Operatoren oder speziellen Suchbereichen zu stellen. Es wird davon abgeraten, gleich zu Beginn nach demographischen Angaben zu fragen. Der Teilnehmer sollte zuerst den Sinn und Zweck der Umfrage und die gestellten Fragen kennen, bevor er um persönliche Angaben gebeten wird. Eine weitere Erleichterung ist die Aufteilung der Umfrage in mehrere Seiten, um der Ermüdung von Teilnehmern durch Scrollen vorzubeugen. Eine weitere Hilfe ist die Darstellung des Fortschritts, indem Angaben zur Länge der Umfrage gemacht werden (zum Beispiel „Seite 2 von 5“).

Der große Vorteil von Umfragen besteht darin, dass Angaben zur Person erfasst werden können. Es ist dadurch möglich, verschiedene Gruppen einander gegenüberzustellen. Hölscher/Strube (2000) vergleichen so Anfänger mit Experten, die durch die Dauer der Internetnutzung bestimmt wurden, was über die Angabe des Jahres des Internetbeginns be-

stimmt werden konnte. Bei Schmidt-Mänz (2007) werden weitere Gruppen wie Männern mit Frauen oder Angestellte mit Studenten verglichen. Es können auch Fragen nach der Funktionsweise von Suchmaschinen gestellt werden, um herauszufinden, inwiefern Suchmaschinennutzer eine Vorstellung darüber haben, wie Suchmaschinen Suchanfragen bearbeiten oder ob es unterschiedliche Suchmaschinen-Nutzertypen gibt (Körber 2000). Ein weiterer Vorteil ist die Formulierung so genannter „offener“ Fragen, wodurch Störfaktoren in Suchmaschinen oder Verbesserungsvorschläge erhoben werden können, wenn dies gewünscht ist.

Der Nachteil von Umfragen ist, dass sich die Teilnehmer der Befragung bewusst sind und sich deswegen bei „kritischen“ Fragen in ein besseres Licht rücken wollen. Die Häufigkeit der Nutzung von Operatoren oder der Phrasensuche wird beispielsweise überdurchschnittlich oft angegeben werden. Bei den persönlichen Angaben ist ebenfalls mit zum Teil falschen oder gar keinen Antworten zu rechnen. Eine Möglichkeit ist die Einführung von Kontrollfragen beziehungsweise die Durchführung einer Konsistenzprüfung (Jahr 1950 als Interneteinstieg nicht möglich). Es zeigte sich, dass bei einer Umfrage eine Anhebung des Antwortniveaus stattfindet, aber bei dem Vergleich von verschiedenen Gruppen durchaus wieder Unterschiede deutlich werden (Schillewaert et al. 1998; Schmidt-Mänz 2007, S. 96).

Eine weitere Problematik ist die Repräsentativität der Umfrageergebnisse. Da nicht von „dem“ Internetnutzer ausgegangen werden kann, muss auf eine breite Teilnehmerschaft geachtet werden. Eine weitere Möglichkeit ist die Befragung einer festen Zielgruppe, wie die Nutzer einer bestimmten Suchmaschine (Spink et al. 1999). Eine ausführliche Diskussion zur Repräsentativität und Beispiele zu Umfragen sind bei Schmidt-Mänz (2007, S. 64) zu finden.

2.4 Laborexperimente

Seltener als die Analyse von Logdateien und Umfragen werden Laborexperimente durchgeführt (Körber 2000; Machill et al. 2003). Weitere Beispiele solcher Laborexperimente sind bei Schmidt-Mänz (2007, S. 30) aufgeführt. Bei Laborexperimenten werden die Probanden bei der Nutzung von Suchmaschinen in einem dafür eingerichteten Computerlabor beobachtet. Es werden hierfür den Teilnehmern Suchaufgaben vorgelegt, die dann unter Beobachtung gelöst werden müssen (zum Beispiel „Wie heißt der fünfte Präsident der USA?“). Dabei sind verschiedene Fragestellungen möglich, es kann nach allgemeinen und ganz speziellen Themen gesucht werden. Damit lassen sich wiederum Rückschlüsse auf die Nutzung von

Operatoren und speziellen Suchfunktionen bei unterschiedlich schweren Aufgabenstellungen ziehen.

Ein Nachteil sind die Kosten der Durchführung eines solchen Experiments. Die Einrichtung und Bereitstellung eines Labors ist dabei kein unerheblicher Faktor. Weiterhin ist nachteilig, dass sich im Gegensatz zu Umfragen Teilnehmer nicht nur beobachtet fühlen, sondern den Beobachter direkt neben sich sitzen haben. Unter diesen Bedingungen wird sich kaum ein Proband wie bei der alltäglichen Suche verhalten. Es kann auch passieren, dass der Interviewer, der die Beobachtung durchführt, den Probanden aktiv beeinflusst. Zudem sind die Suchanfragen oft sehr konstruiert und entsprechen nicht unbedingt echten Informationsbedürfnissen, die oft spontan aufkommen und nach denen im Internet recherchiert wird. Ein weiterer Nachteil ist, dass solche Experimente sehr zeitaufwendig sind und deswegen auch nur wenige Probanden für die Teilnahme gewonnen werden können (meist zehn bis zwanzig Personen oder weniger). Daraus entsteht neben der möglichen Verminderung der Objektivität das weitere Problem der mangelnden Repräsentativität der Ergebnisse.

Ein Vorteil hingegen ist, dass der Beobachter nach einzelnen Schritten des Probanden bei der Suche beziehungsweise der Interaktion mit Suchmaschinen und nach deren Sinn und Zweck fragen kann (zum Beispiel „Wieso werden an einer bestimmten Stelle Operatoren eingesetzt?“). Ein weiterer Vorteil ist die Kenntnis der Probanden und deren soziodemographischen Merkmale. Diese werden oft durch einen ergänzenden Fragebogen erhoben.

2.5 Sonstige

Eine Kombination der verschiedenen Methoden ist durchaus denkbar. So gibt es beispielsweise Online-Panels, bei denen Internetnutzer bei verschiedenen Online-Aktivitäten beobachtet werden, worunter auch die Interaktion mit Suchmaschinen fällt. Der Vorteil besteht darin, dass detaillierte Angaben über Nutzer verfügbar sind, aber der Nutzer nicht unter der gleichen Beobachtung steht, wie dies bei Laborexperimenten der Fall ist. Auch die Analyse von Logdateien kann um einen Fragebogen erweitert werden, indem jeder n-te Nutzer einer bestimmten Suchmaschine diesen präsentiert bekommt, wodurch neben den Logdateien auch Hintergrundinformationen zu den Nutzern erhoben werden (Spink et al. 1999). Bei Laborexperimenten können auch Blickaufzeichnungsverfahren eingesetzt werden, die zusätzlich zeigen, wohin Nutzer tatsächlich schauen, wenn Ergebnisseiten von Suchmaschinen aufgerufen werden. Dies ist eine Erweiterung zu der Analyse, welche Ergebnisse aufgerufen

werden. So kann beurteilt werden, ob tatsächlich vor dem ersten Aufruf eines Links in den Ergebnisseiten die anderen Angaben in der Liste evaluiert wurden. Es ist ebenfalls interessant, inwiefern Links, die zu Werbezwecken eingesetzt werden (beispielsweise Adwords), betrachtet werden (Hotchkiss et al. 2005).

2.6 Zusammenfassung

Als die objektivste und aufschlussreichste Methode ist die Auswertung von Logdateien anzusehen. Hier sind die umfassendsten Analysen von Suchanfragen möglich. Zudem kann diese Erhebungsmethode durch eine Online-Umfrage erweitert werden, um auch soziodemographische Informationen zu den beobachteten Nutzern zu erheben. Durch den Charakter eines Feldexperiments fühlen sich die Nutzer unbeobachtet. Die Kosten und der Zeitaufwand der Erhebung sind dabei gering, weil die Logdateien automatisch angelegt werden und in elektronischer Form zur weiteren Analyse zur Verfügung stehen. Bei der Speicherung der Logdateien über einen langen Zeitraum wird lediglich der Speicherplatz wichtig, was auch einen Hauptfaktor der Kosten ausmacht. Bei der Analyse ist unbedingt auf die Wahrung der Privatsphäre der Suchmaschinennutzer zu achten.

Für tiefere Analysen, bei der die Vorgehensweise bei der Suche erforscht werden soll, eignen sich eher Laborexperimente und Umfragen, weil sich die Teilnehmer dabei aktiv zu ihrem Suchverhalten äußern. Logdateien geben nur das Resultat der Suche, also das letztendliche Verhalten wieder, während Laborexperimente und Umfragen das Verhalten während der Suche beziehungsweise Meinungen und Einstellungen von Suchmaschinennutzern aufgreifen können. Diese Erhebungsmethoden eignen sich besser, die Psychologie von Suchmaschinennutzern bei ihrer Recherche zu untersuchen.

3 Analyse und Auswertungen von Suchanfragen

Bei der Analyse von tatsächlich gestellten Suchanfragen, wie dies vor allem bei der Analyse von Logdateien und Livetickern von Suchmaschinen der Fall ist, gibt es drei grundsätzliche Auswertungsmöglichkeiten. In den folgenden Abschnitten werden unter Angabe der Literatur, in der die zugehörigen Ergebnisse detailliert zu finden sind, diese vorgestellt und die Ergebnisse zusammengefasst wiedergegeben. Auf Ergebnisse, die aus Umfragen und Laborexperimenten stammen, soll im Folgenden nicht eingegangen werden, da diese oft sehr unterschiedliche Aspekte des Nutzerverhal-

tens betrachten. Hierfür sei auf Schmidt-Mänz (2007, S. 29-59) verwiesen.

3.1 Deskriptive Auswertungen

Unter deskriptive Auswertungen fallen die beschreibenden Angaben, mit denen verschiedene Datensätze von Suchanfragen verglichen werden können. Die Größe eines Datensatzes wird durch die Dauer der Erhebung, die Anzahl der Suchsitzungen, der Suchanfragen und -terme angegeben, diese können auch als Bruttozahlen von Suchanfragen oder -termen gesehen werden. Die Bruttozahl gibt somit die Menge aller Suchanfragen an. Eine weitere Angabe ist die Nettoanzahl von Suchanfragen oder -termen, wobei hier nur die eindeutigen Anfragen oder Terme gezählt werden. Darunter ist die Anzahl der Terme zu verstehen, die mindestens einmal gesucht wurden. Bei Schmidt-Mänz/Koch (2006) wurde ein Datensatz analysiert, der 403 Tage umfasste und durch die Beobachtung des Lycos-Livetickers erhoben wurde. Die Anzahl der Suchsitzungen konnte nicht erhoben werden, dies wurde im vorherigen Kapitel erläutert. Insgesamt wurden in diesem Zeitraum 189 Millionen Suchanfragen und 344 Millionen Terme erhoben, wobei davon 29 Millionen Suchanfragen und elf Millionen Terme eindeutig waren. Diese Zahlen sollen dazu dienen, eine Vorstellung von dem Umfang solcher Datensätze zu bekommen.

Beim Vergleich von Datensätzen von Suchanfragen verschiedener Suchmaschinen aus verschiedenen Jahren (Altavista, Alltheweb, Biwe, Excite, Fireball, Lycos, Metager, Metaspinner und Webcrawler) werden immer wieder ähnliche Analysen durchgeführt, die in ihrer Gesamtheit als Benchmarks von Suchanfragen bezeichnet werden können (Jansen/Spink 2006; Spink/Jansen 2004).

Auf die Menge aller Suchanfragen bezogen, sind Suchanfragen im deutschsprachigen Raum mit der Länge von 1,6 bis 1,8 Termen sehr kurz. Sie sind im Schnitt kürzer als englischsprachige Suchanfragen (Jansen/Spink 2006), weil in der englischen Sprache Substantive als Wortzusammensetzung aneinandergereiht werden und nicht wie im Deutschen als ein Wort geschrieben werden (zum Beispiel *christmas tree* vs. *Weihnachtsbaum*). Bezieht man die Länge nur auf die Menge der eindeutigen Suchanfragen, so variiert diese zwischen 2,0 und 2,7 Termen. Dies wird auch durch den sehr hohen Anteil von 48,4 bis 50,1 Prozent an ein-elementigen Suchanfragen deutlich. In Relation zur Menge aller Suchanfragen treten 7,9 Prozent aller Suchanfragen nur einmal auf. Sieht man diese Zahl relativ zu der Nettoanzahl der Terme, dann sind 58,3 Prozent der eindeutigen Suchanfragen auch einzigartig. Bei den Suchtermen ist diese Zahl bedeutend niedriger, 1,3 Prozent der Suchterme tauchen in Relation aller Such-

terme nur einmal auf und 49,1 Prozent der eindeutigen Suchterme sind auch einzigartig (Schmidt-Mänz 2007, S. 131/132).

Operatoren wie AND oder OR werden in den wenigsten Fällen genutzt; häufiger wird die Phrasensuche zur Verfeinerung der Suchergebnisse angewendet („Karl der Große“), um die Suchergebnisse einzuschränken. Insgesamt lag die Nutzungsweise von Operatoren und Phrasensuche unter drei Prozent, der Anteil der Phrasensuche schwankte dabei zwischen 2,1 und 2,5 Prozent. Nur in wenigen Fällen werden natürliche Fragen wie „Wer ist Karl der Große?“ gestellt. Suchanfragen werden als eine natürliche Frage definiert, wenn sie ein Fragewort enthalten. Nur zwischen 0,1 und 0,2 Prozent aller Suchanfragen enthielten solche Fragewörter. Spezielle Suchbereiche wie die Bildersuche werden auch eher selten benutzt, am häufigsten wird mit 65,8 Prozent (Fireball) oder 87,9 Prozent (Metaspinner) die Grundeinstellung der deutschen Suche beibehalten (Schmidt-Mänz 2007).

Als eine weitere deskriptive Auswertung werden oft die Top-10-Terme oder Suchanfragen angegeben. Wie sich zeigte, verändern sich diese selten und nur geringfügig (Schmidt-Mänz/Koch 2006), deswegen wird im nächsten Abschnitt auf die Auswertung von Suchanfragen im zeitlichen Verlauf eingegangen.

3.2 Zeitliche Auswertungen

Unter die zeitliche Auswertung von Suchanfragen fallen beispielsweise die durchschnittliche prozentuale Verteilung von Suchanfragen pro Wochentag oder im Ablauf eines Tages. Für diese Angabe werden jedoch größere Datensätze benötigt, die mehrere Tage oder Wochen abdecken. Oft werden solche zeitlichen Auswertungen schon auf der Basis eines Datensatzes durchgeführt, der lediglich einen Tag umfasst, was nicht sinnvoll ist. Das Tageshoch des durchschnittlichen Suchanfragenvolumens liegt zwischen 14.00 und 15.00 Uhr. Ungefähr sieben Prozent aller Suchanfragen werden um diese Tageszeit abgesetzt. Das Tagestief liegt mit gerade einem Prozent in den frühen Morgenstunden zwischen 5.00 und 6.00 Uhr. Montags werden ca. 14 Prozent aller Suchanfragen formuliert, das Wochentief liegt mit zehn Prozent aller Suchanfragen auf den Samstagen (Schmidt-Mänz 2007). Interessant bei der zeitlichen Auswertung ist vor allem die Betrachtung der Nutzung von Suchtermen über die Beobachtungsperiode hinweg. Die Häufigkeit, mit der Suchterme nachgefragt werden, gibt einen Anhaltspunkt, was Suchmaschinennutzer interessiert. Es wurde hier eine generelle Klassifikation von Suchtermen eingeführt (Schmidt-Mänz/Koch 2006, Schmidt-Mänz 2007, S. 148). Ein Suchterm, der in nahezu jedem Zeitintervall eines Erhebungszeitraumes häufig

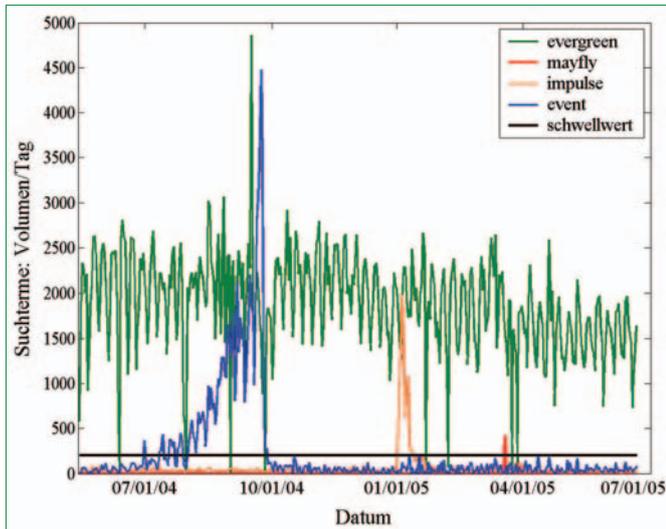


Abbildung 1: Beispielhafte Volumina-Verläufe für Klassen von Suchtermen

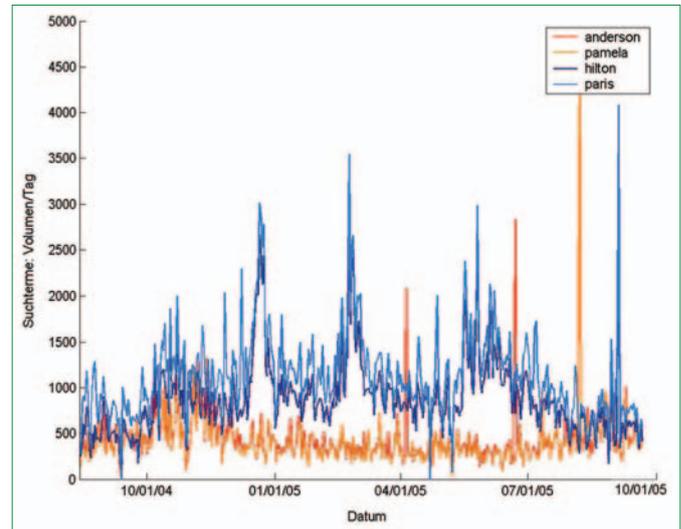


Abbildung 2: Darstellung zweier Cluster von Phrasensuchen

war, wird als Dauerbrenner (engl. *Evergreen*) bezeichnet. Die Zeitintervalle, in welche die Erhebungsperiode zerlegt wird, können beispielsweise als ein Tag oder als vier Stunden beziehungsweise als noch kleinere Taktungen gewählt werden. Ein Term wird dann als häufig eingestuft, wenn sein Auftreten über einem bestimmten Schwellwert (zum Beispiel 100mal pro definiertem Zeitintervall) lag. Es ist für einen Term ausreichend, wenn er nur in 80 bis 90 Prozent aller Zeitintervalle häufig genug war, um als Dauerbrenner bezeichnet zu werden. Unter den Dauerbrennern sind vor allem Terme aus den Begriffswelten der Erotik, der Reise, Multimedia und Musik zu finden. Viele dieser Dauerbrenner stellen auch falsch benutzte Operatoren dar, wie kleingeschriebenes AND oder OR. Füllwörter und Artikel sind ebenfalls unter den Dauerbrennern zu finden.

Eine weitere zeitliche Termklasse stellen die Eintagsfliegen (engl. *mayflies*) dar. Sie sind sozusagen das Gegenteil von Dauerbrennern, denn sie sind nur in ein oder zwei Zeitintervallen häufig genug, um in Betracht gezogen zu werden. Bei einem Schwellenwert von 200 Anfragen pro Tag wurden bei dem Lycos-Datensatz 4.837 Eintagsfliegen gezählt, bei einem Schwellenwert von 400 waren es nur noch 2.826. Die Zahlen wären ohne diesen Schwellenwert bedeutend größer.

Weitere Termklassen werden durch wiederkehrende Begebenheiten oder Nachrichten beeinflusst. Eine Termklasse ist die der Events, die immer vor dem Eintreten einer angekündigten Begebenheit eintritt (Neujahr, Wahlen, Weihnachten, etc.). Das Charakteristische an dem zeitlichen Verlauf dieser Terme ist, dass sie während einer langen Erhebungsphase nicht häufig auftreten und dass das Volumen zum Ereignis hin ansteigt. Das Maximum des Volumens liegt auf dem Tag des Events. Impulse werden hingegen durch Nachrichten hervorgerufen (zum Beispiel Tsunami, Tod des Papstes). Der charakte-

ristische Kurvenverlauf der Suchanfragen verhält sich gerade entgegengesetzt zu denen der Events. Am Tag des Ereignisses liegt das maximale Volumen an Suchanfragen vor, an den Tagen darauf fällt es schnell wieder ab und bleibt dann wieder unter dem Schwellwert, bis erneut eine Nachricht auftritt, bei der dieser Suchterm wieder wichtig wird.

In Abbildung 1 sind die charakteristischen Verläufe der jeweiligen zeitlichen Termklassen anhand der Lycos-Daten aus Schmidt-Mänz/Koch (2006) aufgezeigt.

3.3 Inhaltliche Auswertungen

Bei der inhaltlichen Auswertung wird versucht, durch die Klassifikation von Termen in verschiedene Themengebiete (beispielsweise Erotik, Business, Reise) zu zeigen, mit welchen Themen sich Suchende auseinandersetzen. Bei Spink/Gunar (2001) und Spink et al. (2002) wurden solche Klassifikationen eingeführt. Lewandowski (2006) führte ebenfalls eine empirische Untersuchung durch, die sich mit Typen und Themen in Suchanfragen auseinandersetzt. Eine thematische Verlagerung von Suchanfragen von Erotik nach Business wird jedoch durch die Betrachtung der Dauerbrenner bei Schmidt-Mänz (2007) abgelehnt. Sehr kleine Datensätze können durchaus auch thematische Verzerrungen enthalten (Beitzel et al. 2004). Ein weiterer Nachteil bei der inhaltlichen Einteilung von Suchanfragen ist, dass oft nur sehr kleine Teildatensätze genommen werden, um diese wiederum durch eine Person in verschiedene Themengebiete einteilen zu lassen. Automatisierte Verfahren funktionieren ähnlich, indem eine Stichwortliste zu einem Themengebiet erstellt wird, anhand welcher ein Datensatz untersucht werden kann. Enthält eine Suchanfrage mindestens eines der Stichwörter, wird diese dem entsprechenden Themengebiet zugeordnet.

In anderen Veröffentlichungen werden Terme anhand der Ähnlichkeiten ihrer Volumina im zeitlichen Verlauf verschiede-

nen Klassen zugeordnet. Hier seien Vlachos et al. (2004) und Schmidt-Mänz (2007, S. 169) genannt. Der Vorteil dieser Methode liegt darin, dass verschiedene Terme nicht unbedingt einem Themenbereich zuzuordnen sind, aber trotzdem in ihrer Kombination dazu verwendet werden, nach einem bestimmten Thema zu suchen. Bei Vlachos et al. (2004) werden zu einem bestimmten Term, der von Interesse ist, zeitlich ähnliche Terme gesucht und diese ausgegeben. Bei Schmidt-Mänz (2007) werden alle Terme, die nicht zu den Klassen der Eintagsfliegen oder Dauerbrenner gehören, verglichen. Die untereinander zeitlich ähnlichsten Terme bilden dann jeweils eine eigene Klasse. Bei einer sehr feinen Einstellung des Algorithmus werden gerade noch die Phrasensuchen ausgegeben. In Abbildung 2 ist ein solcher Verlauf von zwei verschiedenen Clustern dargestellt. Durch die feine Einstellung des Algorithmus wurden nur zwei Cluster im gesamten Datensatz gefunden. Das eine Cluster bestand gerade aus den Termen „pamela“ und „anderson“, während das andere Cluster „paris“ und „hilton“ enthielt. Der Verlauf der Kurven zeigt, wie ähnlich sich diese Terme zueinander in ihren Volumina über die Zeit gesehen sind. Die jeweiligen Terme in den Klassen werden fast ausschließlich zusammen gesucht. Bei einer gröberen Einstellung werden mehr Cluster gefunden, die auch mehrere Terme enthalten (beispielsweise: last, minute, reise). Die Daten stammen aus dem Lycos-Datensatz (Schmidt-Mänz/Koch 2006).

3.4 Möglichkeiten der Umsetzung der Ergebnisse

Bei den Möglichkeiten der Umsetzung können verschiedene Bereiche wie Suchmaschinen, Portale oder Webseiten in Betracht gezogen werden. Deskriptive Angaben wie die Verteilung der Suchanfragen und die Nutzung von Operatoren beziehungsweise speziellen Suchbereichen zeigen, dass Suchmaschinen und ihre Nutz-

barkeit für Suchende verbessert werden müssen. Auf Basis der sehr kurzen Suchanfragen ist es verständlich, dass Rankingalgorithmen entwickelt werden müssen, die mit dieser geringen und oft allgemeinen Information umzugehen wissen. Auf der anderen Seite können Suchmaschinennutzer noch nicht mit Suchmaschinen geeignet umgehen, was die Auswertung der Dauerbrenner zeigt, unter denen viele falsch eingesetzte Operatoren und Füllwörter beziehungsweise Artikel zu finden sind, die keine Wirkung auf die Einschränkung einer Suchanfrage haben. Suchmaschinen müssen in ihrer Anwendbarkeit also noch intuitiver aufgesetzt werden. Eine weitere Möglichkeit ergibt sich auch noch durch die Kenntnis von Dauerbrennern. Suchanfragen, die in diese Termklasse fallen, müssen nicht gesondert bearbeitet werden. Für diese Anfragen können Themenlisten erstellt werden, die manuell auf ihre Qualität überprüft werden, da zu diesen Rubriken wiederholt die gleichen Suchanfragen eingehen. Dauerbrenner bieten aber auch Portalen eine Chance, ihre Inhalte zu strukturieren und geeignet darzustellen. Auf Portalen werden ebenfalls Suchanfragen generiert. Die oft durch Informationen überladenen Portale können durch Besucher gar nicht anders bewältigt werden. Die Dauerbrenner-Themengebiete könnten beispielsweise die Navigationsleiste bilden. Dies wird vor allem auch im Bereich von Websites wichtig, die tagesaktuelle Nachrichten oder Informationen anbieten. Oft sind solche Sites auf ähnliche Weise wie Portale strukturiert, es gibt verschiedene Rubriken, anhand der die einzelnen Artikel eingeordnet werden. Durch die Kenntnis von Suchanfragen können Anstiege in den Volumina oder neue Suchterme bemerkt werden, bevor andere Anbieter von Inhalten davon erfahren. Durch diese Kenntnis können Inhalte flexibel am Informationsbedürfnis ausgerichtet werden. Zum einen können Dauerbrenner-Themengebiete als Rubriken dienen, zum anderen kann ein spezielles, aktuelles Thema gesondert dargestellt werden. Aufkommende Trends können besonders kommuniziert werden, indem sie ebenfalls einen wichtigen Platz auf der Hauptseite eingeräumt bekommen. Anbieter von Websites können durch die Kenntnis von zeitlich ähnlichen Termklassen die Inhalte an den Suchbegriffen ausrichten, die den benutzten Termen auf der Webseite zeitlich ähnlich sind, aber häufiger gesucht werden. Dadurch erhöht sich die Chance, von potentiellen Besuchern gefunden zu werden. Durch die Kenntnis der Begriffswelt der Suchenden sind weitere Optimierungsmöglichkeiten denkbar.

Eine weitere Anwendung der Datenbasis von Suchanfragen liegt in der geeigneten Buchung von Anzeigen in Suchmaschinen. Die Datenbasis gibt eine Antwort

darauf, wann und wie lange Anzeigen in Suchmaschinen gebucht werden sollen, um optimale Bietstrategien abzuleiten.

4 Fazit

Insgesamt zeigt die Darstellung der Analyse von Suchanfragen und von Suchmaschinennutzern, dass durch die Erhebung des Suchverhaltens im Web Möglichkeiten der Verbesserung von Suchmaschinen, Portalen und Websites möglich sind, die in vielen Fällen noch nicht (geeignet) umgesetzt wurden. Es sind jedoch noch weitere Untersuchungen vonnöten, um eine profundere Kenntnis der Bedürfnisstruktur zu bekommen. Es können so beispielsweise verschiedene Anordnungen von Ergebnisfenstern untersucht werden, um herauszufinden, an welcher Stelle Anzeigen am besten platziert werden sollten. Eine weitere Möglichkeit ist die Untersuchung verschiedener Eingabefelder in der Suche. In ein Feld werden beispielsweise alle gewünschten Wörter eingegeben, in ein anderes alle unerwünschten. Die Erforschung des Suchverhaltens ist bei weitem noch nicht abgeschlossen. Die Ergebnisse dieses Forschungsgebiets müssen aber mit anderen Forschungsrichtungen wie der Weiterentwicklung von Ranking-Algorithmen oder des Suchmaschinen-Marketings zusammengebracht werden.

Literatur

Beitzel, S.; Jensen, C.; Chowdhury, A.; Grossman, D.; Frieder, O. (2004): Hourly Analysis of a Very Large Topically Categorized Web Query Log. In: Proceedings of the 2004 ACM SIGIR Conference on Research and Development in Information Retrieval (ACM-SIGIR), Sheffield, UK, 321-328.

Hölscher, Ch.; Strube, G. (2000): Web Search Behavior of Internet Experts and Newbies. In: Proceedings of the 9th International World Wide Web Conference, 337-346.

Hotchkiss, G.; Alston, St.; Edwards, G. (2005): Eye Tracking Study – An in Depth Look at Interactions with Google Using Eyetracking Methodology. www.enquiro.com/Eyetracking2-Sample.pdf [01.02.2007].

Jansen, B. (2006): Search Log Analysis – What it is, What's been done, How to do it. In: Library & Information Science Research (28), 407-432.

Jansen, B.; Spink, A. (2006): How Are we Searching the World Wide Web? A Comparison of Nine Search Engine Transaction Logs. In: Information Processing and Management, 52(1), 248-263.

Körber, S. (2000): Suchmuster erfahrener und unerfahrener Suchmaschinennutzer im deutschsprachigen World Wide Web. Ein Experiment. Magisterarbeit, www.grin.com/de/preview/27.html [01.02.2007].

Lake, Ch. (2005): Die Online-Werblandschaft in Europa. DoubleClick, http://emea.doubleclick.com/uploadpdf/pdf/europe_online_de_05.pdf [01.02.2007].

Lewandowski, D. (2006): Query Types and Search Topics of German Web Search Engines, Information Services & Use, 26, 261-269.

Machill, M.; Welp, C. (Hrsg.) (2003): Wegweiser im Netz – Qualität und Nutzung von Suchmaschinen. Gütersloh: Verlag Bertelsmann Stiftung.

Schillewaert, N.; Langerak, F.; Duhamel, T. (1998): Non Probability Sampling for WWW Surveys: A Comparison of Methods. In: Journal of the Market Research Society, 40 (4), 307-322.

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

Schmidt-Mänz, N.; Gaul, W. (2005): Web Mining and Online Visibility. In: Weihs, C.; Gaul, W. (Eds.) (2005): Classification – the Ubiquitous Challenge. Heidelberg: Springer, 418-425.

Schmidt-Mänz, N.; Koch, M. (2006): A General Classification of (Search) Queries and Terms. In: Proceedings of the Third International Conference on Information Technology: New Generations (ITNG'06), Las Vegas, Nevada, USA, 375-381.

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

Spink, A.; Bateman, J.; Jansen, B. (1999): Searching the Web: Survey of Excite Users, Internet Research: Electronic Networking Applications and Policy, 9(2), 117-128.

Spink, A.; Gunar, O. (2001): E-Commerce Web Queries: Excite and Ask Jeeves Study. In: First Monday, Peer-Reviewed Journal on the Internet, 6(7), 4.

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

Spink, A.; Jansen, B.; Ozmutlu (2000): Use of query reformulation and Relevance Feedback by Excite Users. In: Internet Research: Electronic Networking Applications and Policy, 19(4), 317-328.

Spink, A.; Jansen, B.; Wolfram, D.; Saracevic, T. (2002): From E-Sex to E-Commerce: Web Search Changes. In: IEEE Computer, 35(3), 107-109.

Vlachos, M.; Meek, Ch.; Vagena, Z. (2004): Identifying Similarities, Periodicities and Bursts for Online Search Queries. In: Proceedings of the 2004 ACM SIGMOD International Conference on Management of Data. Paris: ACM Press, 131-142.

Zien, J.; Meyer, J.; Tomlin, J.; Liu, J. (2000): Web Query Characteristics and their Implications on Search Engines. Research Report, RJ 10199 (95073), Almaden Research Center.

Benutzerforschung, Suchmaschine, Marktforschung, Data Mining, Informationsverhalten, Empirische Untersuchung, Methode, Übersichtsbericht

DIE AUTORIN

Dr. Nadine S. Höchstötter



geb. Schmidt-Mänz arbeitet als Dozentin an der Universität (TH) und der Berufsakademie in Karlsruhe. In ihrer Dissertation befasste sie sich mit dem Suchverhalten in Suchmaschinen. Ihre weitere Forschung ist in der Vorhersage von Suchanfragevolumen und aufkommenden Trends angesiedelt.

Sophienstr. 167, 76185 Karlsruhe
nsh@topicflux.de
www.topicflux.de