

Query Understanding

Dirk Lewandowski
Hochschule für Angewandte Wissenschaften Hamburg,
Department Information
Finkenau 35, 22081 Hamburg
dirk.lewandowski@haw-hamburg.de

Abstract. In diesem Kapitel wird beschrieben, wie Suchmaschinen Suchanfragen interpretieren können, um letztendlich den Nutzern besser auf ihren Kontext zugeschnittene Ergebnisse liefern zu können. Nach einer Diskussion der Notwendigkeit und der Einsatzmöglichkeiten des Query Understanding wird aufgezeigt, auf welcher Datenbasis und an welchen Ansatzpunkten Suchanfragen interpretiert werden können. Dann erfolgt eine Erläuterung der Interpretationsmöglichkeiten anhand der Suchanfragen-Facetten von Calderon-Benavides et al. (2010), welcher sich eine Diskussion der Verfahren zur Ermittlung der Facetten anschließt.

Keywords. Suchmaschinen, Suchanfragen, Klassifikation, Anfrageinterpretation, Anfragetypen

Einleitung

Suchmaschinen liefern für eine große Vielfalt von Suchanfragen gute oder zumindest zufriedenstellende Ergebnisse [1]. Während die mittlerweile bei den Suchmaschinen etablierten Rankingverfahren, die vor allem aus Textstatistik, Popularitätsmessung, Aktualitätsbestimmung und Ortsbestimmung bestehen [2], ausgereizt erscheinen und bahnbrechend neue Rankingverfahren nicht in Sicht sind¹, gibt es noch einen enormen Verbesserungsbedarf auf der Ebene der Interpretation der Suchanfragen. Die Hoffnung dabei ist, dass durch ein besseres Verständnis der Suchanfragen auch bessere Ergebnisse erreicht werden können, indem diese stärker als bislang beispielsweise auf das Thema und die Intention der Anfrage zugeschnitten werden.

Eine Modifikation der Suchanfragen fand im klassischen Information Retrieval vor allem für drei Zwecke statt: Erstens ging es darum, die Suchanfragen mittels Stemming so aufzubereiten, dass der Abgleich mit den Dokumenten möglichst optimal gestaltet wurde. Zweitens wurden Vorschläge zur Verbesserung der Suchanfrage gemacht, wenn in der Anfrage ein Rechtschreibfehler festgestellt wurde. Und drittens wurden Phrasen in der Suchanfrage erkannt, um Fehlinterpretationen zu vermeiden. So kann die Anfrage *New York Hotels* in „New York“ und „Hotel“ zerlegt werden. Dies erleichtert den Abgleich mit den Dokumenten, da bereits in der Vorverarbeitung die Anfrage auf weniger Elemente, die mit dem Index abgeglichen werden müssen, reduziert wird.

¹ Die oftmals als nächster Schritt des Rankings angesehenen Sozialen Rankingfaktoren sind im Sinne dieser Klassifizierung unter die Popularitätsfaktoren zu rechnen. Auch bislang schon wurden thematische Communities für das Ranking ausgenutzt; nun werden diese Communities sozial definiert. Ausführlich zu den Formen der Sozialen Suche s. den Beitrag von Burghardt, Heckner und Wolff in diesem Band.

In dem letztgenannten Fall wurde schon eine Interpretation im Sinne einer Entscheidung zwischen mehreren Möglichkeiten durchgeführt. Eine weitere, sicher weniger populäre Variante wäre, die Anfrage in die Bestandteile „New“ und „York Hotels“ zu zerlegen. Die Entscheidung für die populärere Variante wird aber sicher zu zufriedeneren Nutzern führen.

Ein weiteres sehr einfaches Beispiel für die Anfrageinterpretation ist die Berücksichtigung der Reihenfolge der eingegebenen Suchbegriffe: Die Suchanfragen *Paris Hilton* und *Hilton Paris* unterscheiden sich fundamental und sollten auch entsprechend mit verschiedenen Dokumenten beantwortet werden.

Einen Schub nun hat das Thema „Query Understanding“ in den letzten Jahren dadurch erfahren, dass vor allem bei den Web-Suchmaschinen große Mengen von Anfragedaten anfallen und durch ihre Analyse Nutzerwünsche und -bedürfnisse nachvollzogen werden können.

Croft et al. [3], die Organisatoren des „Query Representation and Understanding“ auf der ACM-SIGIR-Konferenz 2010², führen das zunehmende Interesse an der Anfrageinterpretation direkt auf die Verfügbarkeit entsprechender Datenmengen zurück: “With the recent availability of large amounts of data about user behavior and queries in web search logs, there has been an upsurge in interest in new approaches to query understanding and representing intent” [3].

Aber auch abseits der Suchmaschinen können Suchanfragen ausgewertet werden; als populäre Beispiele kann man Google Trends zur Analyse des Suchanfrageaufkommens und Google Analytics (nicht nur) zur Interpretation der externen wie internen Suchanfragen auf der eigenen Website nennen. Während mit Google Trends zeitliche Verläufe von Suchanfragen-Häufigkeiten nachvollzogen werden können, können mit Google Analytics Verlaufspfade nach der Eingabe von Suchanfragen analysiert werden. Dabei kann es sich sowohl um in einer externen Suchmaschine (durch Nutzer, die dann auf der eigenen Website gelandet sind) eingegebenen Suchanfrage handeln als auch um solche aus der Website-internen Suche.

In der Praxis werden Verfahren des Query Understandings bereits angewendet, auch wenn dies von den Nutzern kaum bemerkt wird, da die Anfrageinterpretation meist nicht offengelegt wird. Bemerkbar macht sie sich jedoch trotzdem: So werden etwa von Google für die Suchanfrage „Rechtsanwalt“ lokale Ergebnisse ausgegeben, obwohl eine lokale Beschränkung in der Anfrage nicht explizit gemacht wurde (Abb. 1).

Eine Suchmaschine, die ihre Anfrageinterpretationen auf den Ergebnisseiten explizit macht, ist Wolfram Alpha. Abb. 2 zeigt die Interpretation der Suchanfrage „London“ mit der Möglichkeit, eine andere als die aufgrund ihrer Popularität vorgeschlagene Interpretation auszuwählen.

In diesem Kapitel soll ein Überblick über die Möglichkeiten des Query Understanding gegeben werden. Zuerst wird nach der Notwendigkeit für diese Verfahren gefragt, dann werden die Einsatzmöglichkeiten gezeigt. Danach wird gezeigt, auf welchen Ebenen Suchanfrageninterpretation überhaupt möglich ist und welche Hindernisse jeweils zu beachten sind. Es folgt eine Darstellung der Suchanfrage im zeitlichen Verlauf, bevor dann ausführlich auf die zahlreichen Bestrebungen, Suchanfragen zur Interpretation der Nutzerintentionen zu klassifizieren, eingegangen wird. Anhand der von Calderon-Benavides et al. [4] gebildeten Facetten von Suchanfragen werden die Ansätze systematisch besprochen und auf ihre Tauglichkeit hin bewertet. Das Kapitel

² <http://ciir.cs.umass.edu/sigir2010/qru/>.

schließt mit einem Fazit, welches auch Hinweise für die Anwendung von Query Understanding in der Praxis bietet.

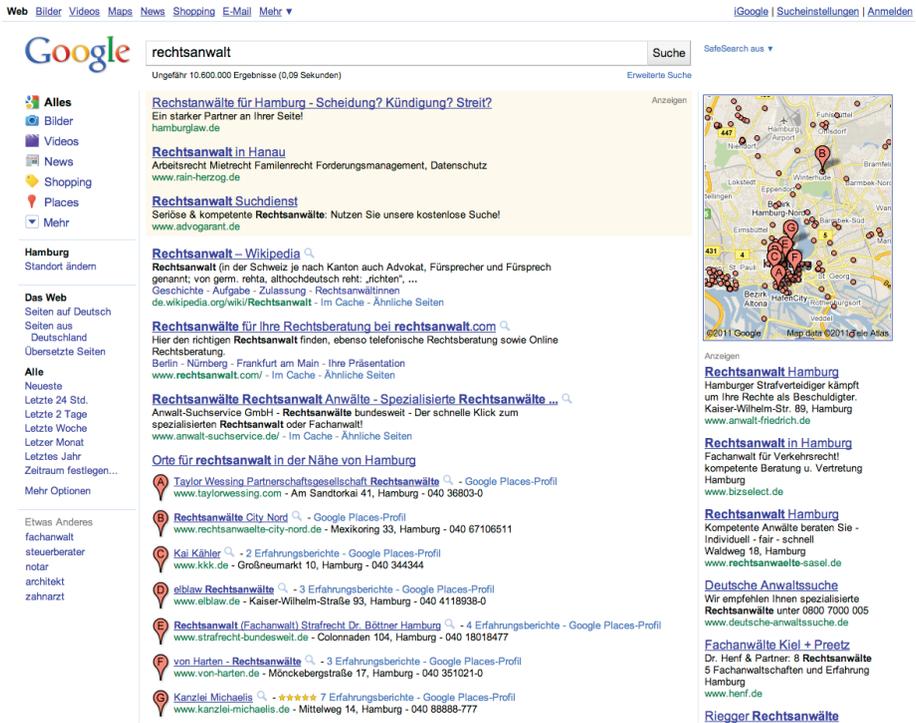


Abbildung 1: Einbindung lokaler Ergebnisse ohne explizite Benennung des lokalen Bezugs in der Suchanfrage (Beispiel Google)



Abbildung 2: Darstellung der Anfrageinterpretation bei Wolfram Alpha

1. Notwendigkeit von Query Understanding

Die Frage nach der Notwendigkeit einer Anfrageinterpretation stellt sich erst, wenn man davon ausgehen muss, dass ein Nutzer nicht oder nur unzureichend in der Lage ist, sein Informationsbedürfnis in eine Anfrage zu übertragen. Daher ist es auch nicht verwunderlich, dass dieses Thema vor allem von den Betreibern der Web-Suchmaschinen forciert wird: Bei der Recherche in Fachdatenbanken wird stillschweigend davon ausgegangen, dass die Recherche von einem erfahrenen Nutzer durchgeführt wird, der zielgenaue Anfragen stellen kann. Ähnlich verhält es sich in dem traditionellen Paradigma der vermittelten Recherche: Hier ist der Informationsvermittler der Garant für sorgfältig formulierte Suchanfragen. Allerdings hat sich inzwischen deutlich gezeigt, dass auch in professionellen Kontexten die Fachkompetenz der Nutzer nicht notwendigerweise mit Recherchekompetenz einhergeht. Daher schwenken auch viele Anbieter von Fachinformationssystemen auf endnutzerorientierte Recherchesysteme um, wobei der Schwerpunkt allerdings vor allem auf der Interface-Gestaltung liegt.

Die Hoffnung, dass Endnutzer die kompetente Recherche in Suchmaschinen (wie auch in anderen Suchsystemen) mit der Zeit lernen würden, hat sich nicht erfüllt. Auch Schulungen haben, auch wenn sie im Einzelfall natürlich Erfolg gehabt haben können, in der Masse keine Veränderung hin zu elaborierteren Suchanfragen gebracht. Studien zeigen deutlich, dass Nutzer unabhängig vom Suchsystem einfache Suchanfragen bevorzugen und nicht bereit sind, viel Zeit in die Sichtung der Ergebnisse zu investieren (s. u.a. [5][6]). Aus diesem Grund ist auf der einen Seite eine Verbesserung des Rankings (also auf der Ergebnisseite), auf der anderen Seite eine Verbesserung der Anfrageinterpretation (also auf der Eingabeseite) vonnöten.

Suchanfragen bestehen zum überwiegenden Teil aus nur wenigen Wörtern; am häufigsten sind aus nur einem Wort bestehende Suchanfragen, gefolgt von Zweiwortanfragen (die aktuellsten für deutsche Suchanfragen stammenden Zahlen kommen aus Höchstötter & Koch [7]). Die durchschnittliche Länge der Suchanfragen liegt bei deutschsprachigen Anfragen bei etwa 1,7 Wörtern [7], S. 55). Eine Entwicklung hin zu längeren Suchanfragen ist nicht festzustellen.³

Es ist davon auszugehen, dass kurze Suchanfragen zum überwiegenden Teil wenig ausdrucksstark sind. Sie enthalten allein die Suchwörter und keine Informationen über den Kontext bzw. das hinter der Suchanfrage stehende Informationsbedürfnis. Um jedoch ein optimales Ergebnis, welches vom Nutzer als relevant eingeschätzt wird, zu erreichen, bedarf es mehr als des reinen Abgleichs zwischen Suchanfrage und Dokumenten. Mizzaro [8] verdeutlicht, dass Relevanz eben nicht nur als ein solch einfacher Abgleich gesehen werden kann, sondern beispielsweise auch den Abgleich zwischen dem (nicht explizit genannten) Informationsbedürfnis und dem Informationsgehalt der ausgegebenen Dokumente betreffen kann. Ein solches Verständnis wiederum führt zu der Notwendigkeit, entweder die Nutzer dazu zu bringen, ihre Informationsbedürfnisse besser in Suchanfragen umzusetzen oder aber Systeme zu schaffen, die die Suchanfragen besser verstehen.

³ Zu den Suchanfragenlängen werden keine regelmäßigen, vergleichbaren Untersuchungen durchgeführt bzw. veröffentlicht. Aussagen, die durchschnittliche Länge der Suchanfragen an Web-Suchmaschinen läge inzwischen über zwei Wörtern, beziehen sich auf englischsprachige Anfragen, die aufgrund der Komposition im Deutschen naturgemäß länger sind.

Bislang nicht untersucht wurde allerdings die Frage, ob die automatisch schon während der Eingabe generierten Suchvorschläge, die meist zumindest aus zwei Wörtern bestehende Anfragen generieren, zu einem Anstieg der Suchanfragenlänge geführt haben.

Da Nutzer aber in der Regel mit den Ergebnissen ihrer Recherchen zufrieden sind, ist es schwierig, sie überhaupt für das Training von Recherchekompetenz zu sensibilisieren. Auch, weil Suchmaschinen ihren Nutzern mit der Interpretation der Suchanfragen entgegenkommen, ergibt sich für diese Nutzer immer weniger die Notwendigkeit, über eine Spezifizierung ihrer Suchanfragen überhaupt nachzudenken. Außerdem muss konstatiert werden, dass vor allem bei navigationsorientierten Anfragen ohne allzu große Interpretation eine hohe Erfolgsquote erreicht werden kann, was wiederum, wenn man den hohen Anteil dieses Anfragetyps an den an die Suchmaschinen gestellten Anfragen bedenkt, zu einer höheren Nutzerzufriedenheit führt.

Eine Chance für ein solch besseres Verständnis ergibt sich aus den bei den Suchmaschinen anfallenden Daten: Sie verfügen über viele Millionen von Suchanfragen, die sehr breit gestreut sind. Ausnutzen lassen sich allerdings Informationen, die aus der Masse generiert werden können; von Bedeutung sind hier vor allem die Häufigkeitsverteilungen und ihre Veränderungen im Lauf der Zeit. Auf beide wird weiter unten ausführlicher eingegangen werden; an dieser Stelle sei nur angemerkt, dass sich die Interpretation einer Suchanfrage im Lauf der Zeit – und mitunter sehr schnell – verändern kann. Daraus folgt, dass in der Anfrageinterpretation nicht mit statischen Daten gearbeitet werden kann.⁴

Zusammengefasst lässt sich sagen, dass die Suchmaschinen versuchen, mittels Anfrageinterpretation die von den Nutzern nicht explizit gemachten Kontextinformationen zu rekonstruieren. Diese Kontextinformationen mögen für verschiedene Nutzer unterschiedlich sein und damit stellt sich die Frage, inwieweit die Anfrageinterpretation auf den einzelnen Nutzer, auf Nutzergruppen oder aber auf die Gesamtmenge der Nutzer einer Suchmaschine bezogen sein sollte. Allerdings sollte hier vor allem unter Gesichtspunkten des Datenschutzes eine wohlüberlegte Entscheidung getroffen werden.

2. Einsatzmöglichkeiten

Die Einsatzmöglichkeiten der Anfrageinterpretation ergeben sich vor allem auf vier Ebenen: Anfrageerweiterung, Generierung von Suchvorschlägen, Steuerung der organischen Ergebnisse, Steuerung der Werbung und Steuerung der Universal Search. Diese Einsatzmöglichkeiten werden im Folgenden kurz erläutert.

2.1. Anfrageerweiterung

Bei der Anfrageerweiterung (*query expansion*) werden einer ursprünglichen Suchanfrage weitere Begriffe hinzugefügt. Das Ziel ist, durch eine genauere Beschreibung einerseits zu *mehr* Dokumenten zu gelangen, andererseits zu *relevanteren*. Ersteres gelingt dadurch, dass der Suchanfrage Begriffe hinzugefügt werden, die in der ursprünglichen Version nicht enthalten waren, letzteres dadurch, dass der Abgleich zwischen Suchanfrage und Dokumenten auf der Basis von mehr Begriffen erfolgt, also eine genauere Gewichtung möglich ist. Problematisch bei der Anfrageerweiterung ist jedoch, wirklich

⁴ Dies gilt neben dem Web-Kontext für alle Bereiche, in denen sich beständige Veränderungen ergeben. In weniger dynamischen Bereichen (zum Beispiel Literaturdatenbanken) ist eine Arbeit mit einer statischen, nur in größeren Intervallen aktualisierten, Suchanfragedatenbank denkbar.

passende Begriffe zu finden, so dass es nicht zu einer Themenverschiebung (*topic drift*) kommt. Allerdings muss eine Anfrageerweiterung gar nicht so komplex sein, sondern es können der Anfrage schlicht Synonyme hinzugefügt werden oder alternative Schreibweisen. Die Begriffe in der erweiterten Suchanfrage werden dann entweder gleich behandelt (also die Verbindung der Begriffe mit dem Boole'schen OR) oder aber gewichtet (wobei Begriffe aus der ursprünglichen Suchanfrage höher gewichtet werden als die hinzugefügten).

2.2. Generierung von Suchvorschlägen

Die Generierung von Suchvorschlägen unterscheidet sich von der Anfrageerweiterung schlicht dadurch, dass dem Nutzer die Entscheidung überlassen wird, ob bzw. wodurch seine Suchanfrage erweitert bzw. in diesem Fall verändert werden soll. Die Auswahl von Suchvorschlägen kann einerseits dazu führen, dass eine vollständig neue Suchanfrage abgeschickt wird (wie dies bei der Implementierung bei den gängigen Suchmaschinen der Fall ist) oder aber, um die Anfrage zu erweitern [9]. Die Anfrageinterpretation ist sowohl bei der Anfrageerweiterung als auch bei der Generierung von Suchvorschlägen eine Methode, um die Qualität der Vorschläge zu verbessern.

2.3. Steuerung der organischen Ergebnisse

Werden zu einer Suchanfrage zusätzliche Informationen gewonnen, so lassen sich die Suchergebnisse entsprechend anpassen. Wird etwa festgestellt, dass eine Suchanfrage einen aktuellen Bezug hat (was beispielsweise dadurch erkannt werden kann, dass diese Suchanfrage in den letzten Tagen oder Stunden im Vergleich zu vergangenen Zeiträumen besonders häufig gestellt wurde), können bevorzugt aktuelle Dokumente angezeigt werden. Bei einem einfachen Abgleich mit konventionellen Rankingfaktoren würden hingegen ältere Dokumente bevorzugt, da der in den linktopologischen Verfahren gemessene Linkaufbau erst erfolgen muss, was entsprechend Zeit beansprucht (s. [10], S. 183).

Aber nicht nur für die Steuerung der konventionellen Trefferlisten ist das bessere Verstehen der Suchanfragen von Bedeutung, sondern auch, um zu entscheiden, ob überhaupt eine Trefferliste angezeigt werden sollte oder ob sich die Suchanfrage direkt beantworten lässt.⁵ So kann die Anfrageinterpretation dazu dienen, die Anzahl der überhaupt zur genauen Suchanfrage angezeigten Treffer und ggf. die Erweiterung der Trefferdarstellung um Ergebnisse für verwandte Suchanfragen zu steuern. Weiterhin kann entschieden werden, ob einer Trefferliste oder eine alternative Trefferdarstellung ausgegeben werden soll. Zu denken ist hier bspw. an die direkte Beantwortung von Faktenfragen (wobei im Vorfeld die grundlegende Frage zu beantworten ist, ob der Nutzer für seine Suchanfrage eine Liste von Dokumenten oder eine direkte Antwort auf eine – nicht unbedingt als solche formulierte – Frage wünscht).

Die Steuerung der organischen Ergebnisse bzw. die Verbesserung von deren Relevanz ist als der zentrale Einsatzbereich des Query Understanding anzusehen.

⁵ Eine alleinige Beantwortung von Suchanfragen ohne die Darstellung einer Trefferliste ist heute bei den Web-Suchmaschinen nicht üblich. Allerdings werden zunehmend im Rahmen der Universal Search (s. Abschnitt 2.4) direkte Antworten gegeben.

2.4. Steuerung der Werbung

Von Suchmaschinen auf den Trefferseiten eingeblendete Textanzeigen basieren auf der aktuell eingegebenen Suchanfrage. Zwar können jeweils möglichst viele der im Anzeigenbestand des Suchmaschinenanbieters vorhandenen Anzeigen eingeblendet werden (bis zu einem gewissen Grenzwert), jedoch ist dies vor allem in den Fällen sinnvoll, in denen Nutzer bereits eine kommerzielle Intention haben, beispielsweise auf der Suche nach einem Produkt sind. Ist die Anfrage eher rechnerorientiert, so können weniger oder auch keine Anzeigen eingeblendet werden. Durch eine intelligente Steuerung der Anzeigen lässt sich erreichen, dass diese nicht störend auf den Nutzer wirken müssen, sondern sogar als nützlich empfunden werden können.

2.5. Steuerung der Universal Search

In der sog. Universal Search [11] werden Ergebnisse aus unterschiedlichen Datenbeständen kombiniert und innerhalb einer einzigen Ergebnispräsentation zusammengeführt. So ergeben sich Suchergebnisseiten, die weit mehr zu bieten haben als einfache, nach Relevanz sortierte Listen, sondern beispielsweise gesondert hervorgehobene Ergebnisse aus dem Nachrichten- oder dem Video-Index enthalten.

Bei der Steuerung der Universal Search ergeben sich zwei Fragen: Wann sollen Ergebnisse aus einer bestimmten Kollektion eingeblendet werden? Und wie sollen diese Ergebnisse ggf. platziert werden.

Wenn man bei dem oben beschriebenen Beispiel einer aktuell besonders häufig gestellten Suchanfrage bleibt, so ist es nicht nur sinnvoll, aktuelle Ergebnisse in den organischen Treffern anzuzeigen, sondern es sollten auch Ergebnisse aus dem Nachrichtenbestand bzw. anderer besonders aktueller Bestände (von Blogbeiträgen bis hin zur Real-Time-Suche [12]) eingeblendet werden.

Dass diese Ergebnisse gesondert in die Trefferpräsentation eingebunden werden und nicht einfach in der organischen Trefferliste angezeigt werden, hat zwei Gründe: Zum einen wird damit die Möglichkeit geboten, Nutzer, die besonders an bestimmten Treffertypen interessiert sind, direkt auf die passenden Kollektionen zu lenken. Zum anderen ist eine Aktualisierung des allgemeinen Web-Index oft schlicht nicht in der erforderlichen Geschwindigkeit möglich.

3. Ebenen der Analyse von Suchanfragen

Betrachtet man die Analyse von Suchanfragen auf Basis des Aggregationsgrades der vorhandenen Datenmenge, so zeigen sich vier Ebenen (s. Abb. 3). Die Basis der Pyramide bilden die Suchanfragen selbst, welche ohne Zusatzinformationen verwendet werden können.

Die zweite Ebene bilden die zu jeder Suchanfrage vorhandenen Klickdaten, aus welchen sich ersehen lässt, welche Treffer die Nutzer bevorzugt angeklickt haben. Bei diesen Klicks wiederum lassen sich die Häufigkeiten ausnutzen; etwa um festzustellen, ob es sich um eine informationsorientierte oder eine navigationsorientierte Suchanfrage handelt (s. Abschnitt **Error! Reference source not found.**).

Auf der dritten Ebene befinden sich die Such-Sessions, d.h. die Abfolgen von Suchanfragen, welche durch einen Nutzer innerhalb einer vorab bestimmten Zeitspanne ausgeführt werden (vgl. [13]). Zu beachten ist aber, dass die Sessionlängen stark variieren

können. So gibt es in der Websuche viele sehr kurze Sessions, die gerade einmal auf der Eingabe einer Suchanfrage und dem Anklicken eines einzigen Treffers bestehen [14].

Auf der obersten Ebene schließlich stehen die Nutzerprofile, welche aus den aggregierten Suchanfragen eines Nutzers bestehen. Während Daten auf dieser Ebene vor allem zum Zweck der Personalisierung [15] interessant sind, birgt eine solche Sammlung gleichzeitig die Gefahr des „gläsernen Nutzers“ und widerspricht dem Grundgedanken der Datensparsamkeit. Nichtsdestotrotz werden dem Nutzer zugeordnete Suchanfragen vor allem von Google im sog. Webprotokoll⁶ gesammelt und für die Optimierung der Ergebnisse eingesetzt.

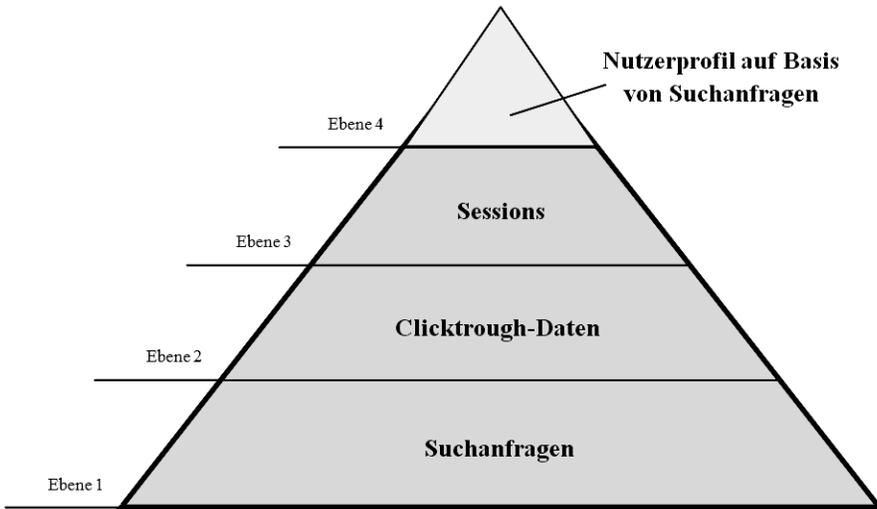


Abbildung 3: Ebenen der Suchanfragenanalyse

4. Zeitliche Stellung einer Suchanfrage

Um eine Suchanfrage möglichst genau zu interpretieren, ist es nötig, sie auch in ihrer zeitlichen Stellung zu betrachten. Dabei können sowohl Daten, die vor der Eingabe dieser Suchanfrage angefallen sind (entweder durch den aktuellen Nutzer oder aber durch andere Nutzer, die in der Vergangenheit dieselbe Suchanfrage eingegeben haben) als auch Daten, die nach der Eingabe dieser Suchanfrage angefallen sind, verwendet werden. In Abbildung 4 sind die verwendbaren Daten in ihrer zeitlichen Stellung dargestellt.

⁶ www.google.com/history/?hl=de.

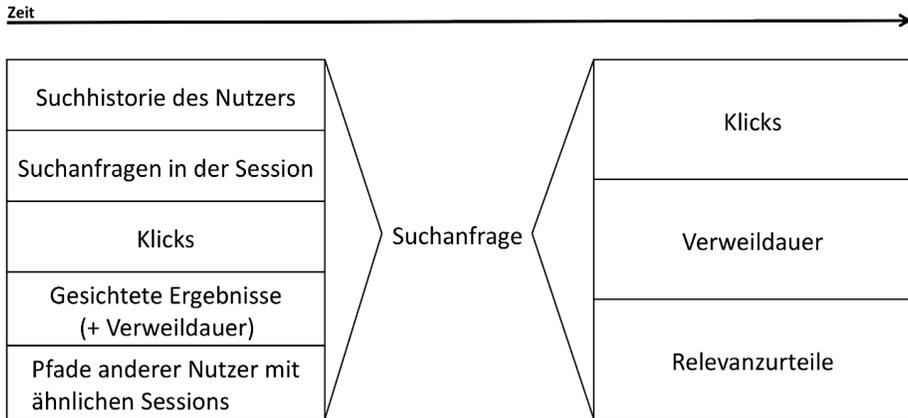


Abbildung 4: Zeitliche Stellung einer Suchanfrage

4.1. Daten, die vor der Eingabe einer Suchanfrage anfallen

Vor der Eingabe einer bestimmten Suchanfrage hat ein Nutzer bereits eine ganze Historie früher eingegebener Suchanfragen vorzuweisen. Werden diese Daten in ihrer Gesamtheit ausgewertet, hat dies zur Folge, dass ganze Nutzerprofile erstellt werden. Man spricht dann von Personalisierung, da das Ziel verfolgt wird, die Ergebnisse auf Basis dieses Nutzerprofils für den individuellen Nutzer zu optimieren. Der Preis dafür ist allerdings die Speicherung aller Suchanfragedaten und das Zuordnen dieser Daten zu einem individuellen Nutzer.

Es können aber vor der Eingabe einer Suchanfrage auch bereits in der aktuellen Session Suchanfragen angefallen sein, welche für das Verständnis der eingegebenen Suchanfrage nützlich sind. So kann es sich bei der aktuellen Suchanfrage bspw. um eine Reformulierung einer bereits eingegebenen Anfrage handeln. Doch nicht nur in diesem Fall sind die vorher eingegebenen Suchanfragen nützlich, sondern auch, wenn es darum geht, das Thema der aktuellen Suchanfrage zu bestimmen. Wie Rose und Levinson zeigen konnten [16], können diese Informationen bei der Zuordnung von Suchanfragen – nicht nur auf die Themen beschränkt – helfen.

Auch die bereits vor der Eingabe der aktuellen Suchanfrage getätigten Klicks auf Dokumente können zum Verständnis der Suchanfrage herangezogen werden. Einerseits kann allein ein Klick auf ein Dokument als Indikator für das Ziel und/oder das Thema einer Suchanfrage genutzt werden, andererseits können auch zusätzliche Kennwerte, wie die *dwell time* oder die *bounce rate* verwendet werden:

Unter *dwell time* versteht man die Verweildauer auf einem bestimmten Dokument, welche bei einer bestimmten Länge als Qualitätsindikator für das Dokument angesehen werden kann, bei sehr kurzer Dauer (und einem Zurückkehren auf die Trefferliste der Suchmaschine) aber als Indikator gegen die Qualität des Dokuments verwendet wird (sog. *bounce rate*).

Zuletzt lassen sich die vor der aktuellen Suchanfrage angefallenen Daten mit vergleichbaren Sessions anderer Nutzer aus der Vergangenheit vergleichen. Aus den bekannten Sessions lassen sich dann Annahmen über die Intention und das weitere Verhalten des nun suchenden Nutzers gewinnen.

4.2. Daten, die nach der Eingabe einer Suchanfrage anfallen

Auch auf der anderen Seite, also nach der Eingabe der Suchanfrage, fallen Daten an, die für das Verständnis der aktuellen Suchanfrage genutzt werden können. Hierbei muss es sich natürlich um historische Daten anderer Nutzer handeln, da die Daten ja bereits zum Zeitpunkt, an dem eine Suchanfrage eingegeben wird, benötigt werden.

Einerseits können auch hier wieder die Dokumente, die Nutzer in der Vergangenheit nach der Eingabe der fraglichen Suchanfrage besucht haben, verwendet werden, wobei wiederum nicht nur die Klicks selber, sondern auch die Verweildauer analysiert werden kann. Neben dieser impliziten Bewertung der Relevanz der Dokumente können explizite Relevanzbewertungen, die aus Befragungen gewonnen werden können, hinzukommen.

Während es bei der Analyse von Suchanfragen stets wünschenswert wäre, sowohl auf allen Aggregationsebenen als auch auf den beschriebenen zeitlichen Ebenen⁷ zu arbeiten, sind die Analysen sehr stark von den tatsächlich vorhandenen Daten abhängig. Wie Lewandowski & Höchstötter [17] zeigen, können je nach Erhebungsmethode Daten in unterschiedlichem Umfang anfallen, so etwa bei der Ticker-Analyse nur Daten auf der Ebene der Suchanfragen, wobei außer dem Anfragezeitpunkt keine den Kontext erhellenden Daten gewonnen werden.

Klar wird also, dass die Möglichkeiten der Anfrageinterpretation auch immer mit der Verfügbarkeit der Daten zusammenhängen. Auf den ersten Blick mag man annehmen, dass den Betreibern von Suchsystemen alle genannten Daten auch zur Verfügung stehen. Dabei ist aber zu bedenken, dass es nicht alleine um das Ansammeln der Daten geht, sondern Werkzeuge geschaffen werden müssen, die eine entsprechende Aggregation der Daten und ihre Auswertung ermöglichen.

5. Klassifikation von Suchanfragen

Wenn nun festgestellt wurde, auf welchen Ebenen Suchanfragen interpretiert werden können, soll im Folgenden beschrieben werden, in welchen Facetten eine Klassifikation dieser Suchanfragen erfolgen kann. Das Ziel der Anfrageinterpretation ist schließlich, möglichst umfangreiche ergänzende Daten zu einer Suchanfrage zu gewinnen. Die Diskussion soll anhand der von Calderon-Benavides et al. [4] definierten Facetten erfolgen.

Die Autoren beschreiben folgende neun Aspekte samt entsprechender Ausprägungen, die eine Suchanfrage umfassend charakterisieren: Genre, topic, task, objective, specificity, authority sensitivity, spatial sensitivity und time sensitivity. Die Facetten mit Erläuterungen sind in Tabelle 1 dargestellt. In den folgenden Abschnitten werden diese Erläuterungen jeweils zum Ausgangspunkt der Diskussion genommen, auch wenn sich zeigen wird, dass vor allem die zu den Facetten benannten Werte nicht notwendigerweise gewählt werden müssen.

Die bisherige Forschung konzentriert sich vor allem auf die Klassifikation von Suchanfragen nach Anfragetypen und nach Themen sowie auf die Analyse der Klicks auf Trefferdokumente. Methodisch lassen sich dabei nach Jansen, Booth & Spink [18] drei Forschungsansätze unterscheiden: “(1) empirical studies and surveys of search engine use, (2) manual analysis of search engine transaction logs, and (3) automatic classification of Web searches” [18].

⁷ Hinzu kommt ein möglicher Wandel der Suchintentionen zu einer Suchanfrage im zeitlichen Verlauf (vgl. [40]).

Tabelle 1: Facetten von Suchanfragen (aus [4])

Facet	Values	Description	Use
Genre	News, Business, Reference, Community	It is consider a meta-category that provides a generic context to the user query intent.	This facet clearly delimits the search engines area of search, and is expected to make the search faster and more accurate. For example if the genre = Community, the pages that have to be looked at are blogs, forums, chats, and communities, such as Facebook.
Topic	Adult & Sex, Arts & Culture, Beauty & Style, Cars & Transportation, Computers & Internet, Education, Entertainment, Music & Games, Finance, Food & Drink, Health, Home & Garden, Industrial Goods & Services, Politics & Government, Religion & belief systems, Science & Mathematics, Social Science, Sports, Technology & Electronic, Travel, Undefined, Work	This list of topics was created from the first level of categories offered by ODP, Yahoo!, and Wikipedia, but leaving out categories that are actually a genre such a News and Reference.	Although this is one of the first features that are considered by the search engines, they should ideally present the answers to the user organized by topic
Task	Informational, Not Informational, Both	Considered as the primary need reflected by a query, could be also the intent itself.	If the query is Informational, then Web resources to be recommended should avoid those with transactional content, or the ones where the user has to assume an active role of interaction, such as fulfilling a form (i.e., Not-Informational resources).
Objective	Resource, Action	Represents the aim of a query, without considering the format of the information to retrieve.	If the user wants to perform an Action, then he/she may also be interested in commercial sites that offer a product or service, hence appropriate ads may be presented to the user.
Specificity	Specific, Medium, Broad	This facet describes how specialized is a query.	The search engines can help the user by presenting many similar pages, for example on the same topic, or a broad variety of pages.
Scope	Yes, No	Shows whether the query contains polysemic words or not.	The answers presented to the user by the search engine should ideally be presented by topic, and trying to cover most of the topics related to the polysemic words, allowing the user to select the right one.
Authority Sensitivity	Yes, No	Is the query designed to retrieve authoritative and trusted answers?	Through this dimension, search engines have the opportunity to be very selective regarding the authority of the web pages presented to the user. For example, it is critical to trust the pages where an antivirus software would be download from. Moreover, it is essential to reach the right page for all the navigational queries.

Facet	Values	Description	Use
Spatial Sensitivity	Yes, No	Reflect the interest of the user to obtain a resource that is related to a particular spatial location (explicit or not).	Although a query does not mention a place, search engines should be able to identify spatial sensitivity to offer information related to the local area from which the query was submitted, emphasizing in queries that reflect a commercial need. For example: "school", "tai chi classes", "cars selling".
Time Sensitivity	Yes, No	Whether the information to retrieve, involves a date or period of time.	Search engines can take advantage of the time sensitivity of queries to suggest related information that occurred in the same period as what it is being sought for. Further more, this information can be used to recommend resources following a temporal order, as it is done with news.

5.1. Genre

Unter Genres werden hier breite Themengebiete verstanden, die vor allem dabei helfen können, die Suche zu beschleunigen. Wird beispielsweise erkannt, dass für eine bestimmte Suchanfrage nur Nachrichten gewünscht werden, kann die Suche auf den entsprechenden Datenbestand beschränkt werden, wodurch die Ergebnisse schneller generiert werden können. Darüber hinausgehend können Genres aber auch eingesetzt werden, um die Durchmischung der Trefferliste zu steuern. Wenn beispielsweise ermittelt wurde, dass eine Suchanfrage zu 80 Prozent auf einen Bedarf an Nachrichten hindeutet, zu 20 Prozent aber auf Ergebnisse aus Nachschlagewerken, so kann die Trefferpräsentation entsprechend durchmischt werden.

Die Einteilung nach Genres von Benavides et al. in News, Business, Reference und Community ist sicherlich zu grob. Es wurden einige weitere Vorschläge gemacht, so etwa in [19], wo eine Unterteilung in Help, Article, Discussion, Shop, Portrayal (non-private), Portrayal (private), Link Collection, Download vorgeschlagen wird. Allerdings sind die Genres hier auf die Dokumente und nicht auf die Suchanfragen bezogen. Natürlich lassen sie sich aber entsprechend adaptieren.

5.2. Thema

Unter Themen fassen Calderon-Benavides et al. grobe Themen, wie sie sich etwa auf der obersten Hierarchieebene von Universalklassifikationen finden lassen. Allerdings wird hier hinsichtlich der verwendeten Klassifikation pragmatisch vorgegangen und schlicht die oberste Ebene des Open Directory⁸ verwendet, wobei Klassen, die nach der Einteilung von Calderon-Benavides et al. als Genres zu bezeichnen sind, ausgeschlossen werden. Eine weitere pragmatische Vorgehensweise findet sich bei Spink, Wolfram, Jansen & Saracevic [20], die ihre Themenklassifikation aus den Daten heraus (Klassifikationsaufgaben im Rahmen eines Seminars) gewinnen.

Die Klassifikation von Suchanfragen nach Themen erfolgt zumeist auf manuelle Weise und bezieht sich entweder auf die grundsätzliche Verteilung von Themen inner-

⁸ <http://www.dmoz.org/>.

halb der untersuchten Anfragen, die Ausprägung einzelner Themenkomplexe oder Zusammenhänge zwischen Themen und Anfragetypen. Spink, Wolfram, Jansen & Saracevic [20] untersuchten bspw. die allgemeine Themenverteilung anhand einer Logfile-Analyse; Erkenntnisse zu bestimmten, häufig nachgefragten Themen, wie E-Commerce, Gesundheit oder Sexualität, finden sich u.a. bei Spink & Jansen [22]. Lewandowski [21] kombinierte die thematische Ausrichtung mit der Intention der Suche (differenziert nach informations-, navigations- oder transaktionsorientiert), um die Verteilung von Anfragetypen innerhalb bestimmter Themenklassen zu untersuchen, und konnte wie erwartet eine unterschiedliche Ausprägung in den einzelnen Themen-Klassen zeigen.

Manuell klassifizierte Suchanfragen eignen sich als Trainingsmenge für die maschinelle Klassifikation. Dazu ist allerdings eine eindeutige Zuordnung zu den Klassen nötig und das Ergebnis steht und fällt mit der Qualität der Ausgangsdaten.

5.3. Intention

Unter *task* verstehen Calderon-Benavides et al. das hauptsächliche hinter einer Suchanfrage stehende Informationsbedürfnis. Häufig spricht man jedoch eher von der *Intention* einer Suchanfrage bzw. dem Anfragetyp.

Schon vor dem Aufkommen der Web-Suche stellte die Untersuchung von Anfragetypen bspw. im Kontext der Bibliotheksrecherche ein Forschungsinteresse dar: Kantor [23] identifizierte in diesem Zusammenhang *known item searches* als Suchen nach einem bereits bekannten Element; Frants, Shapiro, & Voiskunskii [24] differenzierten zwischen einem konkreten und einem problemorientierten Informationsbedarf. Die Analyse der Anfragetypen im Web-Kontext basiert auf dem Ansatz von Andrei Broder [25], der Suchanfragen in informations-, navigations- und transaktionsorientiert unterteilt. Broders Definition folgend

- zeichnet sich eine *informationsorientierte Anfrage* durch das Bedürfnis, Informationen zu einem bestimmten Thema zu finden, aus. Als Ergebnis ist nicht ein bestimmtes einzelnes Dokument, sondern eine Auswahl an relevanten Dokumenten gewünscht.
- ist eine Suchanfrage als *navigationsorientiert* anzusehen, wenn der Nutzer eine bestimmte Webseite finden möchte, von deren Existenz er bereits weiß oder deren Existenz er zumindest vermutet. Typischerweise handelt es sich um die Suche nach der Homepage einer bestimmten Person, eines Unternehmens o.ä. Das Informationsbedürfnis ist mit der Anzeige eines (des korrekten) Treffers umfassend befriedigt.
- liefern *transaktionsorientierte Anfragen* als Ergebnis Webseiten, auf denen wiederum eine Interaktion erforderlich ist. Hierbei kann es sich u.a. um den Download einer Software, den Kauf eines Produktes oder die weitere Suche in einer Datenbank handeln.

Seit der Veröffentlichung dieses Aufsatzes wurde einige Forschung in diesem Bereich betrieben, vor allem hinsichtlich der automatischen Klassifikation von Suchanfragen in diese oder vergleichbare Anfragetypen gibt es intensive Bemühungen. In den meisten Fällen dient Broders Taxonomie als Ausgangsbasis, die dann modifiziert wird. So sind auch die von Caleron-Benavides et al. genannten (und auch schon in Baeza-Yates et al. [26] verwendeten) Werte zu sehen: Sie unterscheiden die Anfragetypen *informational* und *not informational*, wobei es auch eine dritte Kategorie (*both*) gibt, die Suchanfragen enthält, auf die beides zutrifft.

Kang & Kim [27] nutzen prinzipiell dieselben Anfragetypen wie Broder, bezeichnen diese aber abweichend als *topic relevance task*, *homepage finding task* und *service finding task*. Rose & Levinson [16] gliedern die transaktionsorientierten Anfragen in die Unterklassen *download*, *entertainment*, *interact* und *obtain* und verwenden statt *transaction* die übergeordnete Bezeichnung *resource*.

5.4. Ziel

Bei dem Ziel (*objective*) einer Suchanfrage unterscheiden Calderon-Benavides et al. zwischen Suchanfragen, die auf eine Aktion abzielen und solchen, die auf Ressourcen hin orientiert sind. Unter der Aktion ist hier analog zu Broder [25] die Transaktion zu verstehen; es geht hier vor allem darum, ob kommerzielle Ergebnisse gewünscht werden oder nicht.

Zahlreiche Studien haben sich bereits mit der automatischen Ermittlung kommerzieller Intentionen aus Suchanfragen beschäftigt. Für Suchmaschinen, die sich durch die Einblendung von kontextbasierten Werbeanzeigen auf den Trefferseiten finanzieren, ist die Ermittlung kommerzieller Intentionen von besonderer Bedeutung, um die Menge und Passgenauigkeit der einzublendenden Anzeigen möglichst exakt zu bestimmen. Allerdings ist auch für den Aufbau der organischen, d.h. auf regulären Suchergebnissen basierenden Trefferliste die Identifizierung kommerzieller Anfragen von Bedeutung, um kommerziell orientierte Ergebnisse bevorzugt zu platzieren bzw. (der weit schwierigere Fall) im Falle einer nicht-kommerziell orientierten Suchanfrage kommerziell orientierte Ergebnisse auf den vorderen Trefferplätzen zu unterdrücken.

Für die Ermittlung kommerzieller Anfragen wurden unterschiedliche Verfahren vorgeschlagen (u.a. [25][28][29]). Diesen ist gemeinsam, dass ihre Evaluierung auf einer relativ geringen Menge von Referenzdaten, welche zudem nur ungenau klassifiziert wurden (bspw. keine Berücksichtigung der Interrater-Reliabilität), beruht. Dies stellt die Ergebnisse zwar nicht grundsätzlich in Frage, zeigt aber, welches Potential für verbesserte Analysen hier noch besteht.

5.5. Spezifität

Aus der Spezifität einer Suchanfrage können Schlüsse auf die gewünschte Trefferzusammenstellung gezogen werden. Bei einer sehr allgemein formulierten Suchanfrage ist davon auszugehen, dass der Nutzer sich einen Überblick über das Thema verschaffen möchte. Insofern ist es nicht nur sinnvoll, Überblicksartikel anzuzeigen (wobei der Überblick hier innerhalb des einzelnen Dokuments erfolgt), sondern auch die Trefferzusammenstellung so zu gestalten, dass insgesamt durch die ersten angezeigten Treffer ein Überblick gegeben wird. Dies kann beispielsweise dadurch geschehen, dass Dokumente zu unterschiedlichen Aspekten des Themas angezeigt werden oder Dokumente, die zwar das gleiche Thema behandeln, jedoch aus unterschiedlichen, miteinander nicht oder nur schwach vernetzten Communities stammen.

Ein solches Vorgehen lässt sich gut an der Suchanfrage „Mindestlohn“ erläutern. Die Anfrage ist unspezifisch, insofern kann man davon ausgehen, dass grundlegende Informationen gewünscht werden. Würde man nun einfach alle Überblicksartikel zum Thema anzeigen lassen, bestünde die Gefahr, dass nur einige Aspekte des Themas genannt werden. Wir haben es mit einem stark umkämpften Thema zu tun und die relevanten Informationsseiten stammen zu einem großen Teil von Interessensgruppen. Diese wiederum lassen sich in zwei Lager einteilen, nämlich Pro-Mindestlohn und

Contra-Mindestlohn. Suchmaschinen können über die Verlinkungsstruktur feststellen, dass die beiden Communities existieren, beide relevante Dokumente anzubieten haben, aber untereinander nicht vernetzt sind. Daraus kann geschlossen werden, dass es sinnvoll ist, jeweils einige Dokumente aus beiden Gruppen anzuzeigen.

5.6. Bandbreite

Unter Bandbreite (*scope*) einer Suchanfrage verstehen Calderon-Benavides et al. die Ausdifferenzierung der Suchanfrage durch Polysemie bzw. Homonymie. Mehrdeutige Suchanfragen sorgen für eine stärkere Ausdifferenzierung nach Themen durch die verschiedenen Bedeutungen, die abgedeckt werden müssen. Dabei wird davon ausgegangen, dass auf der Trefferseite zumindest die wichtigsten Bedeutungen abgedeckt werden müssen, um den Nutzer zufriedenzustellen. Es würde zur Enttäuschung des Nutzers führen, wenn nur Bedeutungen, die vom Nutzer nicht intendiert waren, abgedeckt werden.

Wie viele Bedeutungen einfache Wörter haben können, zeigt beispielsweise ein Blick auf die entsprechenden Disambiduirungsseiten in der Wikipedia. So werden für das Wort „Brücke“ unter anderem folgende Bedeutungen angegeben:

- ein Bauwerk zur Überquerung von Hindernissen
- ein Entwurfsmuster in der objektorientierten Programmierung
- eine Übung in der Gymnastik
- eine deutsche Künstlervereinigung von Expressionisten
- eine Verteidigungshaltung beim Ringen
- ein Zahnersatz in der Zahnmedizin⁹

Nutzer werden allerdings nur in seltenen Fällen ihre Anfrage entsprechend spezifizieren. Vielmehr machen sie sich keine Gedanken über eventuell bestehende, nicht intendierte Bedeutungen.

Natürlich können die ermittelten Bedeutungen auch dafür eingesetzt werden, um dem Nutzer Vorschläge zur Modifikation seiner Suchanfrage zu machen. Dies ist allerdings nur als zusätzliche Option zu empfehlen, da Nutzer die Ergebnisse von Suchmaschinen zuerst einmal aufgrund der direkt auf eine Suchanfrage hin ausgegebenen Treffer beurteilen. Insofern ist die Empfehlung zu geben, zuerst eine möglichst viele Aspekte abdeckende Trefferliste und ergänzend Vorschläge zur Auswahl einer Bedeutung der Suchanfrage zu machen.

5.7. Autoritätsbezug

Unter einer Autorität (*authority*) wird im Web-Kontext ein Dokument bzw. eine Website verstanden, die Inhalte besonderer Qualität bietet, für ein Thema von herausragender Bedeutung ist oder schlicht in der Ergebnismenge (bzw. sogar auf dem ersten Platz der Trefferliste) nicht fehlen darf. Die Bezeichnung *authority* wurde vor allem im Zusammenhang mit linktopologischen Rankingverfahren bekannt: Kleinberg [30] strebte in seiner Umsetzung eines linkbasierten Rankings an, auf der einen Seite *hubs*, also Zusammenstellungen relevanter Quellen zu einem Thema, und auf der anderen Seite *authorities* zu bestimmen.

⁹ [http://de.wikipedia.org/wiki/Brücke_\(Begriffsklärung\)](http://de.wikipedia.org/wiki/Brücke_(Begriffsklärung)).

Wenn nun nach dem Autoritätsbezug einer Suchanfrage gefragt wird, so geht es um die Frage, welche Bedeutung Autoritäten für diese spezielle Suchanfrage haben. Calderon-Benavides et al. illustrieren dies an einem Beispiel: Sucht ein Nutzer nach einer Antiviren-Software, so ist es essentiell, dass ihm Autoritäten ausgegeben werden. Nur diese können garantieren, dass es sich wirklich um eine Antiviren-Software handelt und gerade nicht um eine Schadsoftware. Im Gegensatz dazu ist bei navigationsorientierten Anfragen weniger auf die Autorität der Quelle zu achten, da Nutzer in diesem Fall ja schon wissen (oder zumindest eine Vorstellung davon haben), auf welche Seite sie gelangen wollen.

Schwieriger wird es bei informationsorientierten Anfragen. Hier ist je nach Dokumentenlage zu entscheiden: Sind genug autoritative *Dokumente* vorhanden, sollten diese natürlich bevorzugt gelistet werden. Wenn dies nun allerdings nicht der Fall ist, kann es sinnvoll sein, bevorzugt Dokumente aus autoritativen Quellen anzuzeigen. Die Annahme in diesem Fall ist, dass, falls es keine eindeutig bestimmbar hochwertigen Dokumente gibt, es doch wahrscheinlicher ist, dass ein Dokument aus einer bekannten Quelle relevant ist als eines aus irgendeiner beliebigen Quelle. Häufig ist diese Art der Dokumentbewertung zu beobachten, wenn auch zu abseitigen Themen Dokumente von den bekannten Nachrichtenwebsites angezeigt werden [31].

5.8. Ortsbezug

Ein Ortsbezug kann in einer Suchanfrage explizit oder implizit ausgedrückt sein. Typische explizite Ortsbenennungen sind Städte, aber auch größere geographische Regionen. Ein impliziter Ortsbezug ist beispielsweise bei den Namen von Sehenswürdigkeiten gegeben. Allerdings können auch Suchanfragen, die auf den ersten Blick keinen Ortsbezug enthalten, diesen doch implizieren. Sucht ein Nutzer beispielsweise nach „Supermarkt“, „Anwalt“ oder „Reinigung“, so ist anzunehmen, dass er Treffer mit lokalem Bezug bevorzugt wird.

Für die Anzeige relevanter Suchergebnisse ist in diesen Fällen ein Abgleich zwischen dem Standort des Nutzers und dem geographischen Bezug des Dokuments nötig. Während sich dies bei dem lokalen Bezug in den genannten Beispielen als recht einfach erweist (die Treffer sollten wohl im unmittelbaren Umfeld bzw. in der Stadt des Nutzers sein), kann der Suchradius bei anderen lokalen Anfragen stark variieren [32]. So ist „Freizeitpark“ zwar ebenso eine Suchanfrage mit Ortsbezug wie „Pizzeria“, im ersten Fall ist der Radius, in dem Suchergebnisse relevant sind, jedoch weit größer.

5.9. Zeitbezug

Auch der Zeitbezug einer Suchanfrage kann explizit oder implizit sein. Ein expliziter Zeitbezug wird durch die Eingabe eines Datums, einer Jahreszahl oder einer Zeitspanne deutlich. Ein impliziter Zeitbezug ist gegeben durch die Nennung von Ereignissen („Zweiter Weltkrieg“) oder Epochen („Renaissance“). Auf der Ebene der Dokumente ist zu unterscheiden zwischen einem Erstellungsdatum und dem behandelten Zeitraum.

Ein aktueller Bezug einer Suchanfrage lässt sich auch aus dem zeitlichen Verlauf der Suchhäufigkeiten ermitteln: Steigt das Anfragevolumen plötzlich stark an, kann auf einen aktuellen Bezug geschlossen werden. Dabei kann wiederum zwischen wiederkehrenden Ereignissen, Eintagsfliegen und Impulsen unterschieden werden [33]. Entsprechend dieser aktuellen Bezüge kann wiederum die Zusammenstellung der Ergebnisse angepasst werden.

6. Methoden zur Gewinnung von Informationen aus Suchanfragen

6.1. Intellektuelle und maschinelle Klassifikation

Prinzipiell können alle genannten Facetten der Suchanfragen durch intellektuelle Klassifizierung bestimmt werden. Werden hier allerdings nur die Suchanfragen selbst vorgelegt, so kann es zu Missverständnissen kommen, da viele Suchanfragen nur mit Hilfe von Kontextinformationen verständlich werden. Ein Beispiel hierfür ist die häufig gestellte Suchanfrage „wow“. Was die Nutzer mit dieser Suchanfrage meinen, erschließt sich erst aus dem Kontext. Beispielsweise kann beobachtet werden, welche Dokumente von den Nutzern bevorzugt angeklickt werden. Dies sind in den meisten Fällen Dokumente zum Spiel „World of Warcraft“; die Interpretation der Suchanfrage wäre ohne diese Zusatzinformation kaum möglich.

Betrachtet man Studien, die sich methodisch der manuellen Zuordnung von Suchanfragen bedienen, lassen sich verhältnismäßig große Unterschiede in den Resultaten feststellen. Die Anteile der einzelnen Anfragetypen variieren und scheinen stark abhängig vom vorhandenen Datenbestand, dem verwendeten Untersuchungsdesign und den befragten Juroren zu sein.

Auch wenn eine direkte Vergleichbarkeit derartiger Studien nur schwer gegeben ist, stellt sich hier grundsätzlich die Frage, inwieweit die Menge an Suchanfragen, die intellektuell klassifiziert wird/werden kann, allein für ein aussagekräftiges Ergebnis ausreicht. Alle bisherigen Untersuchungen nutzen vor dem Hintergrund des erheblichen Zeit- und Arbeitsaufwandes, den eine manuelle Zuordnung mit sich bringt, lediglich eine überschaubare Anzahl an Suchanfragen [34].

Ein Beispiel für die Differenzen, die sich aus der Einschätzung der Intention durch den Suchenden selbst sowie durch einen externen Bewerter ergeben, findet sich in der Untersuchung von Broder [18]. Rose & Levinson [25] untersuchten Daten derselben Suchmaschine wie Broder, stellten ihren Juroren aber gleichzeitig ergänzende Informationen (über Ergebnisklicks und weitere Aktionen des Nutzers) zur Verfügung. Ihre Ergebnisse darüber, welchen Anteil jede Suchintention im Gesamt der Anfragen einnimmt, weichen stark von Broders Ergebnissen ab.¹⁰

Insgesamt ist die Aussagekraft von manuell klassifizierten Suchanfragen mit überschaubarer Datenmenge und ohne zusätzliche Informationen mit Vorsicht zu betrachten. Broder selbst geht so weit zu sagen, dass „Inferring the user intent from the query is at best an inexact science, but usually a wild guess“ [25]. Dies hält ihn aber nicht davon ab, in seiner Untersuchung die Suchanfragen genau auf diese Weise einordnen zu lassen.

Nun wird niemand annehmen, dass die intellektuelle Klassifikation von Suchanfragen ein Verfahren wäre, das bei Suchmaschinen im täglichen Betrieb zum Einsatz kommen kann. Jedoch werden die tatsächlich verwendeten maschinellen Lernverfahren auf der Basis von Trainingsmengen, welche manuell annotiert wurden, trainiert. Wenn jedoch diese Basismengen nicht zuverlässig klassifiziert werden (oder eben gar nicht zuverlässig klassifiziert werden können), so lässt auch die Zuverlässigkeit der maschinellen Verfahren zu wünschen übrig.

Gerade der Schwerpunkt der Forschung auf der maschinellen Klassifikation von Suchanfragen missachtet die Individualität der Informationsbedürfnisse und leidet daran, dass noch in weiten Teilen unklar ist, inwieweit diese Klassifikationen von Menschen

¹⁰ Dies kann natürlich auch andere Gründe haben. Zumindest ist es aber naheliegend, dass die Zusatzinformationen zu einer abweichenden Bewertung führen.

überhaupt geteilt werden. Zwar dürfte der Forschungsansatz der maschinellen Klassifikation von Suchanfragen weiterhin Fortschritte machen, ohne ein grundlegendes Verständnis der Suchanfragen jedoch auch künftig an enge Grenzen stoßen.

6.2. Analyse von Klickdaten

Analysen der Klickdaten verfolgen das Ziel, automatisch Nutzerintentionen zu identifizieren und legen wie bspw. bei Lee et al. [35] die Einteilung der Suchanfragen in navigations- oder informationsorientierte Anfragen zu Grunde. Basis ihrer Studie ist u.a. die Annahme, dass für navigationsorientierte Anfragen die Klicks auf einen einzigen Treffer fallen, während für informationsorientierte Anfragen eine breitere Streuung der Klickdaten zu beobachten ist.

Bereits im vorangegangenen Abschnitt wurde in dem World-of-Warcraft-Beispiel deutlich, dass Klickdaten eine wertvolle Ergänzung zur Interpretation der Suchanfragen sind. Noch einen Schritt weiter gehen Ansätze, die Suchmaschinenergebnisse automatisch durch die Nutzung von Klickdaten optimieren, was bereits 2002 von Joachims diskutiert wurde. Er nutzte die Klickdaten als Ersatz für Relevanzurteile und kam zu dem Schluss, dass Klickdaten relative Präferenzurteile darstellen und als Trainingsdaten zur Verbesserung des Rankings der Suchergebnisse erfolgreich eingesetzt werden können (vgl. [36], S. 141). Auch MacDonald und Ounis stellten fest, dass sich Klickdaten zur Beurteilung von Relevanz in Trainingsdatenbanken sehr gut eignen, in einigen Fällen sogar besser als die Relevanzurteile menschlicher Bewerter [37]. Weitere Arbeiten, wie z.B. von Chao et al. [38] oder Dou et al. [38] bestätigen dies.¹¹

6.3. Sessionanalyse

In die Sessionanalyse können sowohl Klickdaten, betrachtete Dokumente als auch vorher eingegebene Suchanfragen eingehen. Außerdem können vergleichbare Sessions anderer Nutzer aus der Vergangenheit herangezogen werden, um die Suchanfrage in den Kontext stellen zu können.

Aus der Analyse der bereits eingegebenen Suchanfragen kann das Thema der laufenden Session abgeleitet werden, was wiederum für die Interpretation der aktuellen Suchanfrage ausgenutzt werden kann. Die Verweildauer auf bereits angesehenen Dokumenten kann Aufschluss über in der Session aktuelle Präferenzen des Nutzers geben, beispielsweise welchem Genre die bislang betrachteten Dokumente angehörten. Hat ein Nutzer in der laufenden Session beispielsweise bevorzugt Nachrichtenbeiträge gelesen, so erscheint es plausibel, ihm bei weiteren Anfragen bevorzugt solche Beiträge anzuzeigen.

6.4. Analyse des Suchverlaufs/Webprotokolls

Aus der Analyse des kompletten Suchverlaufs können Informationen über die generellen Präferenzen eines Nutzers gewonnen werden. Dies betrifft sowohl die thematische Zuordnung der Suchanfragen als auch die Präferenz hinsichtlich der Quellen, der Verständlichkeit der ausgegebenen Dokumente, der Anzahl der ausgegebenen Dokumente, der Einbindung von Ergebnissen aus zusätzlichen Kollektionen sowie der Anzeige von Werbung.

¹¹ Zur Evaluierung der Trefferqualität von Suchmaschinen s. den Beitrag von Lewandowski in diesem Band.

Ebenso kann der lokale Bezug des Nutzers berücksichtigt werden, sogar wenn der Nutzer sich momentan nicht an seinem Heimatstandort befindet. Allerdings ist die Differenzierung zwischen Anfragen, die sich auf den aktuellen Aufenthaltsort beziehen und solchen, die sich auf den Heimatstandort beziehen, schwierig.

Besonders geeignet sind Daten aus dem Webprotokoll für Suchanfragen, bei denen bereits früher einmal gesichtete Dokumente wiedergefunden werden sollen. Wenn ein Nutzer dieselbe Suchanfrage in der Vergangenheit bereits einmal eingegeben hatte, ist es sinnvoll, ihm bevorzugt diejenigen Dokumente anzuzeigen, die er damals angeklickt hatte. Treffer hingegen, die in vorangegangenen Such-Sessions übergangen wurden, können durch neue Treffer ersetzt werden [39].

7. Fazit

Die in diesem Kapitel beschriebenen Möglichkeiten zur Interpretation von Suchanfragen zeigen, dass Suchmaschinen durch die Gewinnung von Kontextinformationen wesentlich besser auf die Bedürfnisse ihrer Nutzer eingehen können. Dies gilt sowohl für die Masse der Nutzer als auch für individuell auf den aktuellen Nutzer angepasste Ergebnisse.

Die Herausforderung liegt darin, die Facetten einer Suchanfrage zuverlässig automatisch zu bestimmen und aus ihrer zeitlichen Stellung Informationen zur Query Expansion zu gewinnen. Wenn dies gelingt, können aus vordergründig trivialen oder gar unverständlichen Suchanfragen Suchanfragen gemacht werden, wie sie ein Profi-Rechercheur stellen würde, der das Suchvokabular der von ihm verwendeten Suchdienste perfekt beherrscht und außerdem in der Lage ist, sein Informationsbedürfnis in eine Suchanfrage umzusetzen.

Nicht unerwähnt bleiben soll hier allerdings, dass die Interpretation einer Suchanfrage eben immer das ist – nämlich eine Interpretation aus mehreren möglichen. Dass die Suchmaschinen bei der Präsentation ihrer Ergebnisse auf die wahrscheinlichste bzw. wahrscheinlichsten Interpretation(en) setzen, ist verständlich, da sie damit doch die größte Menge von Nutzern befriedigen. Möchte man sich von der automatischen Interpretation der Suchmaschine lösen, so ist dies durch die Spezifizierung der Suchanfrage möglich. Für Laiennutzer bietet die Suchanfrageninterpretation aber den Vorteil, optimierte Ergebnisse zu erhalten, ohne sich mit der Funktionsweise und dem Vokabular der Suchmaschine auseinandersetzen zu müssen.

In dieser Hinsicht ist Query Understanding auch ein Thema für die Anbieter anderer Suchsysteme. Auch wenn man sich im fachlichen Kontext wünschen würde, dass die Nutzer in der Lage sind, ihre Informationsbedürfnisse auszudrücken und in elaborierte Suchanfragen umzusetzen, muss man doch eingestehen, dass auch in diesen Systemen der Anteil derjenigen Nutzer, denen eine Unterstützung durch Anfrageinterpretation helfen würde, sehr hoch ist. Von daher kann man nur dazu anregen, sich auch im fachlichen Kontext Gedanken über dieses Thema zu machen.

Literatur

- [1] D. Lewandowski, The retrieval effectiveness of web search engines: considering results descriptions, *Journal of Documentation*, 64, 2008, 915-937.
- [2] D. Lewandowski, Der OPAC als Suchmaschine, *Handbuch Bibliothek 2.0*, J. Bergmann and P. Danowski, eds., München: Saur / de Gruyter, 2010, 87-107.
- [3] B. Croft, M. Bendersky, H. Li, and G. Xu, eds., Query Representation and Understanding: Workshop of the 33rd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, 2010.
- [4] L. Calderon-Benavides, C. Gonzalez-Caro, and R. Baeza-Yates, Towards a Deeper Understanding of the User's Query Intent, *SIGIR 2010 Workshop on Query Representation and Understanding*, Geneva, Switzerland: ACM, 2010, 21-24.
- [5] M. Machill, Wegweiser im Netz: Qualität und Nutzung von Suchmaschinen, *Wegweiser im Netz*, M. Machill and C. Welp, eds., Gütersloh: Bertelsmann Stiftung, 2003, 13-490.
- [6] I. Rowlands, D. Nicholas, P. Williams, P. Huntington, M. Fieldhouse, B. Gunter, R. Withey, H.R. Jamali, T. Dobrowolski, and C. Tenopir, The Google generation: the information behaviour of the researcher of the future, *Aslib Proceedings*, 60, 2008, 290-310.
- [7] N. Höchstätter and M. Koch, Standard parameters for searching behaviour in search engines and their empirical evaluation, *Journal of Information Science*, 35, 2009, 45.
- [8] S. Mizzaro, Relevance: The whole history, *Journal of the American Society for Information Science*, 48, 1997, 810-832.
- [9] P. Anick, Using terminological feedback for web search refinement: a log-based study, *Annual ACM Conference on Research and Development in Information Retrieval*, peter.anick@overture.com: Association for Computing Machinery, 2003, 88-95.
- [10] D. Lewandowski, *Web Information Retrieval: Technologien zur Informationssuche im Internet*, Frankfurt am Main: Deutsche Gesellschaft f. Informationswissenschaft u. Informationspraxis, 2005.
- [11] S. Quirnbach, Universal Search - Kontextuelle Einbindung von unterschiedlicher Quellen und Auswirkungen auf das User Interface, *Handbuch Internet-Suchmaschinen*, D. Lewandowski, ed., Heidelberg: Akademische Verlagsgesellschaft Aka GmbH, 2009, 220-248.
- [12] D. Lewandowski, Real Time Suche, *Proceedings des 25. Oberhofer Kolloquiums*, M. Ockenfeld, ed., DGI, 2010.
- [13] B.J. Jansen, A. Spink, C. Blakely, and S. Koshman, Defining a session on Web search engines, *Journal of the American Society for Information Science and Technology*, 58, 2007, 862-871.
- [14] B.J. Jansen and A. Spink, How are we searching the World Wide Web? A comparison of nine search engine transaction logs, *Information Processing & Management*, 42, 2006, 248-263.
- [15] K. Riemer and F. Brüggemann, Personalisierung der Internetsuche - Lösungstechniken und Marktüberblick, *Handbuch Internet-Suchmaschinen*, D. Lewandowski, ed., Heidelberg: Akademische Verlagsgesellschaft Aka, 2009, 148-171.
- [16] D.E. Rose and D. Levinson, Understanding user goals in web search, *Proceedings of the 13th international conference on World Wide Web*, ACM, 2004, 13-19.
- [17] D. Lewandowski and N. Höchstätter, Web Searching: A Quality Measurement Perspective, *Web Search: Multidisciplinary Perspectives*, A. Spink and M. Zimmer, eds., Berlin, Heidelberg: Springer, 2008, 309-340.
- [18] B.J. Jansen, D.L. Booth, and A. Spink, Determining the informational, navigational, and transactional intent of Web queries, *Information Processing and Management*, 44, 2008, 1251-1266.
- [19] B. Stein and S. Meyer zu Eissen, Genre classification of Web pages user study and feasibility analysis, *KI 2004: Advances in Artificial Intelligence*, 2004.
- [20] A. Spink, D. Wolfram, B.J. Jansen, and T. Saracevic, Searching the web: The public and their queries, *Journal of the American Society for Information Science and Technology*, 52, 2001, 226-234.
- [21] D. Lewandowski, Query types and search topics of German Web search engine users, *Information Services & Use*, 26, 2006, 261-269.
- [22] A. Spink, *Web search: public searching on the Web*, Dordrecht: Kluwer Academic Publishers, 2004.
- [23] P.B. Kantor, Availability Analysis, *Journal of the American Society for Information Science*, 27, 1976, 311-319.
- [24] V.I. Frants, J. Shapiro, and V.G. Voiskunskii, *Automated Information Retrieval: Theory and Methods*, San Diego: Academic Press, 1997.
- [25] A. Broder, A taxonomy of web search, *ACM Sigir forum*, 36, 2002, 3-10.
- [26] R. Baeza-Yates, L. Calderón-Benavides, and C. González-Caro, The intention behind web queries, 4209 LNCS, 2006, 98-109.

- [27] I.H. Kang and G.C. Kim, Query type classification for web document retrieval, *Proceedings of the 26th annual international ACM SIGIR conference on Research and development in information retrieval*, Division of Computer Science, Department of EECS, KAIST: ACM, 2003, 64-71.
- [28] A. Ashkan and C.L.A. Clarke, Characterizing commercial intent, *Proceeding of the 18th ACM conference on Information and knowledge management*, Cheriton School of Computer Science, University of Waterloo, Waterloo, ON, Canada: ACM, 2009, 67-76.
- [29] H.K. Dai, L. Zhao, Z. Nie, J.R. Wen, L. Wang, and Y. Li, Detecting online commercial intention (oci), *Proceedings of the 15th international conference on World Wide Web*, Microsoft Corporation, One Microsoft Way, Redmond, WA 98052, United States Tsinghua University, Beijing, China Microsoft Research Asia, Beijing, China: ACM, 2006, 829-837.
- [30] J.M.D.O.-<http://doi.acm.org/10.1145/324133.324140> Kleinberg, Authoritative sources in a hyperlinked environment, *Journal of the ACM*, 46, 1999, 604-632.
- [31] N. Höchstätter and D. Lewandowski, What users see – Structures in search engine results pages, *Information Sciences*, 179, 2009, 1796-1812.
- [32] R. Jones, W. Zhang, B. Rey, Jhala, and E. Stipp, Geographic intention and modification in web search, *International Journal of Geographical Information Science*, 22, 2008, 1-20.
- [33] N. Schmidt-Mänz, *Untersuchung des Suchverhaltens im Web: Interaktion von Internetnutzern mit Suchmaschinen*, 2007.
- [34] Z. Dou, R. Song, X. Yuan, and J.R. Wen, Are click-through data adequate for learning web search rankings?, *Proceeding of the 17th ACM conference on Information and knowledge management*, Microsoft Research Asia, No. 49 Zhichun Road, Beijing, China Nankai University, No. 94 Weijin Road, Tianjin, China: ACM, 2008, 73-82.
- [35] U. Lee, Z. Liu, and J. Cho, Automatic Identification of User Goals in Web Search, *14th International World Wide Web Conference*, ACM, 2005, 391-400 ST - Automatic Identification of User Goal.
- [36] T. Joachims, Optimizing search engines using clickthrough data, *Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining*, ACM, 2002, 133-142.
- [37] C. Macdonald and I. Ounis, Usefulness of quality click-through data for training, *Proceedings of the 2009 workshop on Web Search Click Data*, Department of Computing Science, University of Glasgow, Scotland, United Kingdom: ACM, 2009, 75-79.
- [38] F. Guo, C. Liu, and Y.M. Wang, Efficient multiple-click models in web search, *Proceedings of the Second ACM International Conference on Web Search and Data Mining*, ACM, 2009, 124-131.
- [39] J. Teevan, How people recall, recognize, and reuse search results, *ACM Transactions on Information Systems*, 26, Sep. 2008, 1-27.
- [40] A.D. Halverson, K. Kenthapadi, N. Mishra, A. Slivkins, and U.A. Syed, Search queries with shifting intent, 2010.