



Hochschule für Angewandte Wissenschaften Hamburg
Hamburg University of Applied Sciences

DEPARTMENT INFORMATION

Masterarbeit

Anforderungsanalyse für die Bereitstellung von Suchanfragedaten: Literaturanalyse und Empfehlungen für die Aufbereitung von Logfiledaten für weitere Studien

vorgelegt von
Sonja von Mach

Studiengang Informationswissenschaft und -management

erster Prüfer: Prof. Dr. Dirk Lewandowski
zweiter Prüfer: Prof. Dr. Franziskus Geeb

Hamburg, Februar 2011

Abstract

Die Intention hinter einer Suchanfrage wird in einigen Studien erforscht. Nicht immer geht aus den eingegebenen Stichwörtern das tatsächliche Informationsbedürfnis eines Nutzers hervor. Um die Zufriedenheit der Nutzer mit den Suchergebnissen und damit auch den Suchmaschinen zu verbessern wurde eine Vielzahl von Studien durchgeführt. Diese zeigen, dass der Forschungsbedarf auf diesem Gebiet noch groß ist. Deshalb wird hier eine Analyse der wichtigsten Studien durchgeführt und deren Methodik und Aufbau dargestellt, um aufzuzeigen, welche Forschungsvorhaben für die Erkennung der Suchintention hinter einer Suchanfrage und die zukünftige Verbesserung der Suchergebnisse durchgeführt werden sollten und welche Daten aus den Logfiles einer Suchmaschine für diese Studien benötigt werden.

Schlagwörter

Suchintention; Klassifikation von Suchanfragen; Logfileanalyse, Suchanfragen; Web Information Retrieval

Inhalt

Abbildungsverzeichnis	IV
Tabellenverzeichnis	IV
Abkürzungsverzeichnis	IV
1. Einleitung.....	1
2. Daten und Datengewinnung aus Logfiles.....	3
3. Evaluation des Informationsbedürfnisses eines Nutzers	4
3.1 Broder 2002	5
3.2 Kang und Kim 2003.....	6
3.3 Rose und Levinson 2004.....	9
3.4 Lee et al. 2005	11
3.5 Dai et al. 2006.....	14
3.6 Jansen et al. 2007	17
3.7 Calderon- Benavides et al. 2010	22
4. Klassifikation von Suchanfragen nach Themen.....	24
4.1 Spink et al. 2001	24
4.2 Jansen und Spink 2006	26
4.3 Klassifikation von Suchanfragen nach Anfragetyp und Thema	27
4.4 Lewandowski 2006.....	27
4.5 Baeza-Yates et al. 2006	28
5. Suchverhalten von Suchmaschinennutzern	33
6. Einsatz qualitativer Methoden zu Erfassung der Nutzerbedürfnisse.....	39
6.1 Hargittai 2002.....	39
6.2 Jansen und McNeese 2005.....	43
6.3 Otte und von Mach 2009	47
7. Zusammenfassung der Literaturanalyse	50
7.1 Evaluation des Informationsbedürfnisses eines Nutzers.....	50
7.2 Klassifikation der Suchanfragen nach Thema.....	53
7.3 Einsatz qualitativer Methoden zur Erfassung der Nutzerbedürfnisse.....	55
8. Grundlagen für die Entwicklung neuer Studien	56
8. 1 Einflüsse auf Probanden	57
8.2 Persönlichkeitsschutz der Probanden.....	58
8.2.1 Empfehlungen für Datensammlungen.....	64
8. 3 Probanden befragen.....	66

8. 4 Logfiledaten sammeln und analysieren	68
9. Entwicklung neuer Studien.....	72
9.1 Entwickelte Studien – Evaluation des Informationsbedürfnisses eines Nutzers	72
I. Laborexperiment zur Überprüfung der Einordnung von Suchanfragen durch Juroren	73
I.1 Methodik und Datengrundlage	74
I.2 Priorität und Durchführbarkeit.....	77
II. Klickdaten zu Identifikation des Informationsbedürfnisses	78
II.1 Methodik und Datengrundlage.....	78
II.2 Priorität und Durchführbarkeit.....	79
III. Überprüfung der Suchanfragen-Klassifikation nach Broder 2002..	80
III.1 Methodik und Datengrundlage.....	80
III.1.a Alternative Gruppendiskussion	82
III.2 Priorität und Durchführbarkeit.....	83
IV. Tagebuchstudie „Onlineaktivität“	83
IV.1 Methodik und Datengrundlage	84
IV.2 Priorität und Durchführbarkeit	86
9. 2 Entwickelte Studien – Klassifikation von Suchanfragen nach Thema	86
V. Überprüfung der Themenzugehörigkeit der Suchanfragen	87
V.1 Methodik und Datengrundlage	87
V.2 Priorität und Durchführbarkeit	88
9. 3 Entwickelte Studien – Suchverhalten von Suchmaschinennutzern/Einsatz qualitativer Methoden.....	88
VI Entwicklung optimaler Suchergebnisseiten.....	89
VI.2 Methodik und Datengrundlage	89
VI. 3 Priorität und Durchführbarkeit	92
10. Empfehlungen für die Bereitstellung von Logfiledaten.....	92
10.1 Datencharakteristika.....	92
10.2 Logfiledaten für entwickelte Studien	93
11. Fazit.....	95
Literatur	101
Anhang.....	105
1. Beispielhafte Kombination der entwickelten Studien.....	105

Abbildungsverzeichnis

Abbildung 1: Nutzerfoto mit Beschriftung, eigene Darst., 2010 nach RETTIG UND GOEL 2005, S. 31.....	90
--	----

Tabellenverzeichnis

Tabelle 1: Übersicht über die Vor- und Nachteile der verschiedenen Aufzeichnungsmethoden, eigene Darst. nach HAWKEY 2009, S. 89	63
--	----

Abkürzungsverzeichnis

OPAC - Online Public Access Catalogue
IP-Adresse – Internetprotokoll Adresse
PC – Personal Computer

1. Einleitung

Suchmaschinen sind die Hilfsmittel, die viele Menschen täglich nutzen, um Zugang zu Informationen, Internetseiten, Services oder anderen Internetressourcen zu erlangen (JANSEN ET AL. 2007, S. 1251).

Studien zeigen, dass die Intention von Suchenden in Suchmaschinen nicht immer aus der eingegebenen Suchanfrage hervorgeht und das Verständnis, für die von den Nutzern einer Suchmaschine gestellten Suchanfragen, von Seiten der Suchmaschinen verbesserungswürdig ist, um das Informationsbedürfnis des Suchenden zufriedenstellend zu erfüllen.

Bekannt ist dabei auch, dass die Nutzer einer Suchmaschine eine niedrige Toleranzschwelle haben, wenn es darum geht, Ergebnisse, die nicht auf der ersten Suchergebnisseite einer Suchmaschine stehen, anzusehen. Das Scrollen zu Ergebnissen unterhalb des Bildschirmknicks wird von den Nutzern noch akzeptiert, aber sie formulieren ihre Suchanfrage eher neu als sich durch die Ergebnislisten zu arbeiten (SPINK ET AL. 2006, S. 261).

Die Erforschung der Intention hinter einer Anfrage lässt sich in drei Unterkategorien einteilen:

1. Empirische Studien und Befragungen der Suchmaschinennutzer
2. Manuelle Analyse von Suchmaschinen-/Transaktions-Logs
3. Automatische Klassifikation von Suchanfragen

(vgl. JANSEN ET AL. 2007, S. 1252).

Zu diesen Forschungsfeldern wurden einige Untersuchungen durchgeführt, bei denen die Forscher am Ende feststellten, dass die gewählte Methode nicht die optimalen Ergebnisse lieferte oder die durchgeführte Studie eine Art Vorarbeit für folgende Forschungsvorhaben war und die Methodik verbessert werden kann (vgl. u.a. KANG UND KIM 2003, S. 70, CALDERON-BENAVIDES ET AL. 2010, S. 24, BAEZA-YATES ET AL. 2006, S. 108, JANSEN UND MCNEESE 2005, S. 1499, JANSEN ET AL. 2007, S. 1262).

Resultierend aus dem offensichtlichen Forschungsbedarf auf diesem Gebiet wird die Hochschule für angewandte Wissenschaften in Hamburg in Kooperation mit der Deutschen Telekom und der n-size GmbH ein Projekt zum Thema „Query Understanding in Web Search“ durchführen.

Durch dieses Projekt soll die Ergebnispräsentation der Suchmaschinen verbessert werden, indem neben den relevanten Suchergebnissen auch die Intention der Suchenden beachtet wird (vgl. LEWANDOWSKI 2010, S. 3). Ziel ist es, die Suchmaschine beim Abschicken der Anfrage erkennen zu lassen, welche Intention der Nutzer hat und mit passenden Ergebnissen und Hilfestellungen zu reagieren (vgl. LEWANDOWSKI 2010, S. 3).

Um dieses Ziel zu erreichen, ist eine Analyse der bisherigen Forschungsarbeiten sowie die Entwicklung von neuen Hypothesen und Forschungsvorhaben nötig.

Diese Arbeit stellt Studien vor, die sich mit der Erforschung der Logfiles von Suchmaschinen beschäftigen. Neben der Evaluierung des Informationsbedürfnisses des Nutzers, der die Suchanfrage eingibt, wird auch die Einteilung der Suchanfragen in verschiedene Themengebiete und die Erforschung von Logfiles mit Hilfe der Nutzung von ergänzenden Methoden wie bspw. Fragebögen betrachtet. Die gewählten Arbeiten warten dabei mit innovativen Methoden und Ansätzen auf, weshalb sie aus den Veröffentlichungen aus dem Bereich der Analyse von Suchanfragen und der dahinterliegenden Intention des Anfragestellers ausgewählt wurden. Es wurden sowohl verschiedene Forschungsansätze betrachtet als auch Studien, die mit ähnlichen Methoden arbeiteten um einen Vergleich herstellen zu können.

Die Studien wurden analysiert, um die Stärken und Schwächen der dort genutzten Methoden sowie das Vorgehen der Forscher zu eruieren. Eine Zusammenfassung der jeweiligen Studien führt zu einer Übersicht der Ergebnisse und genutzten Methoden, aus denen neue Forschungsvorhaben entwickelt wurden.

Diese neuen Studien zeigen, welche Logfiledaten zur Erhebung von Daten oder zur Durchführung von qualitativen Experimenten aus den Logfiles benötigt werden und führen zu einer Empfehlung zur Bereitstellung der Logfiledaten.

2. Daten und Datengewinnung aus Logfiles

Die Daten der durchgeführten Studien stammen meistens aus Logfiles.

Diese zeichnen die Interaktionen zwischen dem Nutzer und dem jeweiligen System auf. Die Aufzeichnungen können in verschiedenen Systemen vorgenommen werden. Zu diesen gehören Webseiten, OPACs, Computer von Privatpersonen, Blogs, Onlinezeitungen usw. (vgl. JANSEN ET AL. 2009, S. 2).

Um die gesammelten Daten auszuwerten, gibt es verschiedene Methoden, auf die im Laufe dieser Arbeit genauer eingegangen wird.

Logfiles enthalten die elektronischen Aufzeichnungen der Interaktionen zwischen einem System und dem jeweiligen Nutzer (vgl. JANSEN ET AL. 2009, S. 2).

Diese Aufzeichnungen starten in den Computersystemen automatisch (vgl. HÖCHSTÖTTER 2009, S. 183).

Ein Eintrag aus den Logfiles ist für gewöhnlich folgendermaßen aufgebaut:

```
99.99.999.999 - - [14/Dec/2010:22:59:49 +0100] "POST /listen/liste.php HTTP/1.1" 200
19314 "http://www.aaaaaaaaaaaaa.de/listen/liste.php" "Mozilla/5.0 (Windows; U; Windows
NT 6.0; de; rv:1.8.1.11) Gecko/20071127 Firefox/2.0.0.11"
```

- "99.99.999.999" – ist die IP-Adresse des Computer, der die Anfrage stellt
- "- -" - Falls ein Benutzername und ein Passwort verwendet wurde, werden diese hier angezeigt
- [14/Dec/2010:22:59:49 +0100] – Datum und Uhrzeit, zu der die Anfrage gestellt wurde
- "POST /listen/liste.php HTTP/1.1" – was angefordert wurde
- "200" – Status der Webseite
- "19314" – Menge der gesendeten Daten
- "http://www.aaaaaaaaaaaaa.de/listen/liste.php" – Anzeige, woher der Nutzer kam
- "Mozilla/5.0 (Windows; U; Windows NT 6.0; de; rv:1.8.1.11) Gecko/20071127 Firefox/2.0.0.11" – genutzte Software

(vgl. HÖCHSTÖTTER 2009, S. 180 und KAMMERER 2002).

In den Logfiles einer Suchmaschine werden zudem die Suchanfrage des Nutzers sowie die durchgeführten Aktionen wie bspw. die Klicks auf die Ergebnisse dargestellt (vgl. TANG ET AL. 2007, S. 2).

Bei der Analyse von Logdateien findet keine Kommunikation mit dem jeweiligen Nutzer statt. Sie erfordert keine Erwiderung von ihm und gilt daher als sogenannte unüberwachte Methode. Laborexperimente werden dagegen als aufdringlich bezeichnet, da die Teilnehmer hier aktiv partizipieren (vgl. JANSEN ET AL. 2009, S. 2).

Der Vorteil der Sammlung von Daten in Logfiles ist, dass die Daten in einer natürlichen Umgebung gesammelt wurden und der Nutzer dementsprechend unbeeinflusst handelt. Der Nachteil ist, dass der Forscher nicht nachfragen kann, warum der Nutzer bei einer Suche z.B. nur auf die erste Seite der Suchergebnisse geguckt hat. Er weiß so nicht, ob das gewünschte Ergebnis dort gefunden wurde oder ob der Nutzer die Suche abgebrochen hat (vgl. JANSEN ET AL. 2009, S. 6).

Die alleinige Nutzung der Logfiles eignet sich somit nur für quantitative Forschungsfragen, in denen es bspw. um die Verwendung bestimmter Terme oder die Beschaffenheit der Suchanfragen allgemein geht und nicht um die Intention der Suchenden hinter der jeweiligen Suchanfrage (vgl. TANG ET AL. 2007, S. 4).

3. Evaluation des Informationsbedürfnisses eines Nutzers

Um das Informationsbedürfnis hinter einer Suchanfrage zu eruieren, wurden verschiedene Studien durchgeführt, die sich der Informationen der Logfiles bedienten. Ausgewählte Studien aus diesem Bereich werden hier chronologisch aufgeführt.

3.1 Broder 2002

Im Jahr 2002 untersuchte Broder (vgl. BRODER 2002) 400 Suchanfragen und ordnete diese verschiedenen Kategorien zu. Die untersuchten Suchanfragen stammten von der Suchmaschine AltaVista. Zusätzlich wurden die zugehörigen Logdateien der Suchmaschine untersucht (vgl. BRODER 2002, S. 6). Ziel der Untersuchung war es, die Suchanfragen zu klassifizieren und damit den Suchmaschinen zu zeigen, wie sie mit den verschiedenen Kategorien von Informationsbedürfnissen umgehen sollten (vgl. BRODER 2002, S. 3).

Anhand einer Umfrage fand Broder das Informationsbedürfnis der Nutzer heraus und erarbeitete folgende Klassen (vgl. BRODER 2002, S. 5):

- Informationsorientiert
Die informationsorientierten Suchanfragen sollten nach Möglichkeit mehrere passende Treffer hervorbringen, da es dem Suchenden hier darum geht, sich über ein Thema zu informieren und dieses zu beleuchten (vgl. BRODER 2002, S. 5f.). In diesem Fall liegen die Informationen in einer statischen Form vor, das heißt, es sind keine weiteren Interaktionen durch den Nutzer erforderlich (außer dem Anklicken und Lesen der Ergebnisse) (vgl. BRODER 2002, S. 5).
- Navigationsorientiert
Die Intention der navigationsorientierten Suchanfragen ist es, eine bestimmte Webseite zu finden. In dem Fall hat der Suchende bereits eine Webseite im Kopf und sein Informationsbedürfnis kann folglich mit einem Treffer gestillt werden (vgl. BRODER 2002, S. 5).
- Transaktionsorientiert
Suchende, die eine transaktionsorientierte Suchanfrage abschicken, möchten eine webgestützte Transaktion durchführen. Dabei kann es sich bspw. um den Kauf einer Software oder den Download eines Dokuments/einer Datei handeln (vgl. BRODER 2002, S. 6).

An dieser Klassifikation orientieren sich viele der chronologisch nachfolgenden Studien. Entweder übernehmen sie die gesamte Klassifikation, Teile davon oder sie fügen eigene Klassen hinzu (vgl. ROSE UND LEVINSON 2004; KANG UND KIM 2003; JANSEN ET AL. 2007; OTTE UND VON MACH 2009).

In den Studien werden die untersuchten Suchanfragen oft nur einer oder zwei Kategorien zugeordnet. Sie können aber durchaus mehrere Bedeutungen haben und somit auch in mehr als zwei Kategorien passen (vgl. CALDERON- BENAVIDES ET AL. 2010, S. 21). Somit wäre nicht die gesamte Bandbreite der Intention hinter der Suchanfrage beleuchtet worden. Auch können die Suchanfragen unterschiedliches Vokabular gebrauchen, welches aber das Gleiche meint. Je nach verwendetem Suchwort, kann die Qualität der Ergebnisse unterschiedlich sein (vgl. KANG UND KIM 2003, S. 64), denn bei einigen Suchworten ist die Intention hinter der Suchanfrage leichter zu vermuten als bei anderen und somit ist es leichter, die Suchanfrage einem (oder mehreren) Anfragetyp(en) zuzuordnen.

3.2 Kang und Kim 2003

Kang und Kim gehen davon aus, dass unterschiedliche Informationsbedürfnisse der Suchenden mit unterschiedlichen Strategien gestillt werden können.

Um die unterschiedlichen Intentionen der Suchanfragen zu eruieren, werden diese in die folgenden unterschiedlichen Kategorien eingeteilt:

- Topic relevance task (informationsorientiert)
- Homepage finding task (navigationsorientiert)
- Service finding task (transaktionsorientiert)

(vgl. KANG UND KIM 2003, S. 64).

Die Definition der einzelnen Kategorien gleicht der von Broder 2002 aufgestellten Einteilung. Bei dieser sind informationsorientierte Anfragen jene, bei denen Informationen zu einem Thema gefunden werden sollen, navigationsorientierte Anfragen zielen auf eine bestimmte Webseite ab und transaktionsorientierte Anfragen verfolgen das Ziel, eine Transaktion durchzuführen (vgl. KANG UND KIM 2003, S. 64 f. und BRODER 2002, S. 5). Nutzer können aber mit der gleichen Anfrage unterschiedliche Bedürfnisse haben und die Anfragen können in mehrere Kategorien passen (vgl. KANG UND KIM 2003, S. 65). Kang und Kim berechnen in ihrer Studie die Wahrscheinlichkeit, mit der eine der Klassen auf eine Suchanfrage zutrifft. Dazu kombinieren sie mehrere Hinweise, die sich aus Inhaltsinformationen (z.B. Titel und (Anchor-)Texte der Homepage), Linkinformationen (Links die zur

Homepage führen oder dort enthalten sind) und URL-Informationen (Name einer Organisation, Anzahl der Slashes als Hinweis, ob es sich um eine Hauptseite handelt) zusammensetzen (vgl. KANG UND KIM 2003, S. 65 f.). Für ihre Untersuchung nutzen sie Anfragen aus einer Test-Collection aus dem Jahr 2001. Sie nutzen 49 Anfragen für die informationsorientierten Anfragen und 145 Anfragen für die navigationsorientierten Anfragen. Die transaktionsorientierten Anfragen werden hier nicht untersucht, da davon ausgegangen wird, dass sich die automatische Klassifikation nach der Untersuchung auf diesen Anfragetyp übertragen lässt (vgl. KANG UND KIM 2003, S. 65 f.).

Die Suchanfragen wurden zur Untersuchung nach der URL der Ergebnisse aufgeteilt und in zwei Datenbanken eingegeben. Eine Datenbank enthielt die Anfragen mit Hostnamen, die andere enthielt die Dokumente, die keinen Hostnamen in der URL hatten (vgl. KANG UND KIM 2003, S. 70 f.). Die Nutzung von Schlüsselworten wie Verben und Anchortexte zeigte wenig Abdeckung, allerdings wurden mit unterschiedlichen Verteilungen der Suchanfragen und Beziehungen (Nutzung bestimmter Verben in den Anfragetypen) eine gute Precision und eine gute Abdeckung erreicht. Die besten Recall- und Precision-Ergebnisse wurden durch die Kombination verschiedener Faktoren erreicht (vgl. KANG UND KIM 2003, S. 71).

Bei ihren Berechnungen hatten Kang und Kim bei den informationsorientierten Anfragen die besseren Ergebnisse als mit den Anchortexten (bei automatischer Zuordnung der Suchanfragen mit den Inhaltsinformationen der Homepage). Der Titel und die (Anchor-)Texte einer Homepage bieten nicht genug Informationen, um die Bedürfnisse einer informationsorientierten Anfrage zu befriedigen.

Bei den navigationsorientierten Suchanfragen wurden mit den Anchortexten das gleiche Ergebnis erzielt. URL-Informationen und Linkinformationen sind somit gut für die Suche nach einer Homepage, aber schlecht für die informationsorientierte Suche (vgl. KANG UND KIM 2003, S. 66).

Die informationsorientierten Anfragen enthalten gewöhnlich ein Schlagwort oder eine Erklärung darüber, was der Suchende finden möchte. Man kann aber nicht davon ausgehen, dass die anderen Nutzer, die Informationen zu

diesem Thema suchen, die gleichen Schlagworte verwenden (vgl. KANG UND KIM 2003, S. 66).

Kang und Kim kommen zu dem Ergebnis, dass für die unterschiedlichen Informationsbedürfnisse unterschiedliche Retrieval-Strategien gebraucht werden (vgl. KANG UND KIM 2003, S. 66).

Sie gehen davon aus, dass bestimmte Wörter wie bspw. „Fanclub“ auf eine navigationsorientierte Suche hindeuten, während andere wie bspw. „prime factor“ oder als Frage formuliert „What is a prime factor?“, eine informationsorientierte Suche nahelegen. Diese Annahmen beruhen auf Berechnungen während der Studie (vgl. KANG UND KIM 2003, S. 64 und 67 ff.). Desweiteren gehen Kang und Kim davon aus, dass navigationsorientierte Anfragen kein Verb enthalten. Sie weisen deshalb alle Anfragen, die kein Verb außer „to be/sein“ enthalten, der Kategorie Informationsorientiert zu (vgl. KANG UND KIM 2003, S. 69).

Kang und Kim stellen außerdem fest, dass die navigationsorientierten Anfragen (in ihrer Studie homepage finding tasks genannt) normalerweise einen Namen oder Eigennamen enthalten. Transaktionsorientierte Anfragen (hier service finding task) dagegen enthalten oft ein Verb, das die gewünschte Aktion näher beschreibt, wie „kaufen“ o.ä. Da diese sich oft in den Anchortexten befinden, schlagen Kang und Kim vor, diese genauer zu analysieren, um den Anfragetyp einer Suchanfrage zu evaluieren (vgl. KANG UND KIM 2003, S. 70).

Sie schlagen außerdem vor, die Klassifikation von Suchanfragen nicht mehr anhand der einzelnen Wörter vorzunehmen, sondern anhand des gesamten Suchverlaufs und der Handlungen der Nutzer (vgl. KANG UND KIM 2003, S. 70).

Sie empfehlen, den unterschiedlichen Bedürfnissen der Suchenden mit unterschiedlichen Strategien zu begegnen, denn ein für informationsorientierte Suchanfragen entwickelter Algorithmus ist nicht gleichzeitig gut für navigationsorientierte Suchanfragen (vgl. KANG UND KIM 2003, S. 71). So kann man zum Beispiel für navigationsorientierte Anfragen URL- und Link-Informationen nutzen. Diese Informationen sind aber schlecht für informationsorientierte Anfragen (vgl. KANG UND KIM 2003, S. 70).

3.3 Rose und Levinson 2004

Rose und Levinson nutzen bei ihrer Untersuchung die Klassifikation von BRODER 2002 als Grundlage. Sie versuchen, in ihrer Studie das Informationsbedürfnis hinter einer Suchanfrage zu eruieren, ohne die bisher verwendeten Faktoren wie z.B. die Anzahl der Worte einer Suchanfrage zu nutzen. Denn das Informationsbedürfnis eines Nutzers zu befriedigen, bedeutet, die Suche zufriedenstellend und schnell zu beantworten. Dazu gehört, die Intention hinter der Frage zu verstehen, denn ein Suchwort kann für mehrere Intentionen stehen (bspw. „Krebs“ für das Tier oder die Krankheit) (vgl. ROSE UND LEVINSON 2004, S.13).

Wenn die Suchmaschine das Ziel hinter der Suche wüsste, könnte sie bei Suchanfragen mit einem transaktionsorientiertem Hintergrund Werbung einblenden. Diese wäre bei einem Recherchehintergrund von Seiten des Nutzers eher nicht erwünscht (vgl. ROSE UND LEVINSON 2004, S.13).

Rose und Levinson orientieren sich an den von Broder aufgestellten Kategorien. Sie gliedern allerdings die transaktionsorientierten Anfragen in die Unterkategorien download, entertainment, interact und obtain und verwenden statt transaktionsorientiert die übergeordnete Bezeichnung „resource“, da sie bei der Einordnung von Beispielsuchanfragen feststellten, dass Nutzer die z.B. nach „Beatles Lyrics“ suchen, eine Seite suchen, auf der sie Texte einfach nur einsehen können (vgl. ROSE UND LEVINSON 2004, S.13f.). Navigationsorientiert wird als Suche nach einer Homepage, die der Suchende schon im Kopf hat, definiert und informationsorientierte Suchanfragen als Informationssuche zu einem bestimmten Thema. Diese Suche zielt auf keine bestimmte Webseite ab (vgl. ROSE UND LEVINSON 2004, S.14).

Um die Suchanfragen den einzelnen Kategorien zuzuordnen, nutzten Rose und Levinson die Suchanfrage, die Ergebnisse, die die Suchmaschine auf die Anfrage ausgegeben hat, das vom Nutzer geklickte Ergebnis und weitere Suchen oder andere Aktionen des Nutzers (vgl. ROSE UND LEVINSON 2004, S.16). So konnten sie feststellen, ob ein Nutzer seine Suchanfrage bei einer erneuten Eingabe verfeinert hat oder nicht zur Suchmaschine zurückkehrte, was den Schluss nahelegt, dass sein Informationsbedürfnis befriedigt wurde. Anhand der geklickten Webseiten konnte das Informationsbedürfnis zusätzlich abgelesen werden. Klickte der Nutzer eine offizielle

Webseite zu einem Computerspiel, war die Suche navigationsorientiert, klickte er eine (oder mehrere) Seite(n), auf der sich allgemeine Informationen, wie bspw. Veröffentlichungsdaten befanden, war sie informationsorientiert (vgl. ROSE UND LEVINSON 2004, S.17). Die verwendeten Daten stammen aus einem Alta Vista-Log und es wurden 1.500 Suchanfragen untersucht (vgl. ROSE UND LEVINSON 2004, S.14 und CALDERON- BENAVIDES ET AL. 2010, S. 21).

Die Ergebnisse zeigen, dass fast 40 Prozent der Anfragen nicht informationsorientiert waren und die informationsorientierten Suchanfragen hatten zu einem großen Teil die Intention, ein Produkt oder eine Dienstleistung zu finden. Mehr als 35 Prozent der Anfragen hatten einen Recherchehintergrund, für den die traditionellen Information Retrieval Systeme geschaffen wurden (also das Bedürfnis, Informationen über ein Thema zu erhalten) und wurden somit dem Anfragetypen Informationsorientiert zugeordnet (vgl. ROSE UND LEVINSON 2004, S.17).

Verglichen mit den von Broder im Jahr 2002 veröffentlichten Ergebnissen, hatten die Suchanfragen von Rose und Levinson einen höheren Anteil von informationsorientierten Anfragen und einen kleineren Anteil von navigationsorientierten und transaktionsorientierten bzw. ressourcenorientierten Anfragen (vgl. ROSE UND LEVINSON 2004, S.17 f.). Die Unterschiede können in den unterschiedlichen Definitionen der Klassen oder in den unterschiedlichen Methoden, die genutzt wurden, liegen (vgl. ROSE UND LEVINSON 2004, S.18). Auch erwähnen die Autoren, dass die Nutzer der Suchmaschine AltaVista nicht repräsentativ auf die gesamte Nutzerschaft von Suchmaschinen übertragen werden können, da AltaVista mit dem Ruf, besonders gute Suchwerkzeuge kombiniert mit einem relativ kleinen Marktanteil, eher auf Nutzer mit schwierigen informationsorientierten Suchanfragen zugeschnitten ist. Dadurch ist diese Suchmaschine nicht unbedingt die erste Wahl für typische Nutzer mit allgemeinen Suchanfragen (vgl. ROSE UND LEVINSON 2004, S.19).

3.4 Lee et al. 2005

Lee et al. untersuchen in ihrer Studie ebenfalls, ob und wie man Suchanfragen automatisch klassifizieren kann (vgl. LEE ET AL. 2005, S. 391). Ihr erstes Ziel war herauszufinden, ob man durch bloßes Ansehen einer Suchanfrage das Ziel dahinter erkennen kann oder ob man zur Einordnung in die von Broder definierten Klassen mehr Informationen von dem Suchenden braucht. Sie versuchen, Charakteristika zu identifizieren, die auf die Intention hinter der Suchanfrage hindeuten (vgl. LEE ET AL. 2005, S. 391). Die Vorhersagbarkeit eines Ziels hinter der Suchanfrage wurde anhand einer Befragung zu den ausgewählten Suchanfragen untersucht. Anschließend wurden Klicks auf die Ergebnisse der Suchergebnisseite aus der Vergangenheit und die Anchor-Link-Verteilung (also die Links, die in ihrem Anchortext die gleichen Worte haben, die auch in der Suchanfrage enthalten sind) als potenzielle Charakteristika untersucht (vgl. LEE ET AL. 2005, S. 391). Diese wurden dann anhand der automatischen Einordnung einer großen Anzahl von Suchanfragen überprüft (vgl. LEE ET AL. 2005, S. 391 f.). Um die Vorhersagbarkeit eines Ziels hinter einer Suchanfrage zu untersuchen, wurden 28 Teilnehmer (Studenten des Departments Computerwissenschaften der Universität Los Angeles) gebeten, die möglichen Ziele hinter den 50 populärsten Suchanfragen, die bei Google über die Computer des Departments eingegeben wurden, abzuleiten (vgl. LEE ET AL. 2005, S. 391 f.). Es wurden nur Studenten des Departments gewählt, weil davon ausgegangen wurde, dass diese das Ziel der Suchanfragen am ehesten beschreiben könnten (vgl. LEE ET AL. 2005, S. 392). Die Probanden füllten einen Fragebogen aus, in welchem sie die Suchanfragen den Kategorien navigationsorientiert oder informationsorientiert zuordneten. Sie wurden auch gefragt, ob ihnen die jeweilige Suchanfrage geläufig war (vgl. LEE ET AL. 2005, S. 393). Transaktionsorientierte Suchanfragen wurden hier wegen der unterschiedlichen Benennungen und der dadurch fehlenden Möglichkeit zum Vergleich nicht untersucht (vgl. LEE ET AL. 2005, S. 392). Die meisten Probleme in der Befragung der Probanden gab es mit Software- und Personensuchen, diese untersuchten die Forscher noch einmal gesondert. Dabei stellte sich heraus, dass die Softwareanfragen entweder den Besuch der Homepage nahelegten oder Download, Kommentare oder Bewertungen zur Software als Ziel vermutet wurde. Bei den Personen-

suchen war die Person den Probanden entweder sehr geläufig und sie wollten die Homepage der Person besuchen, oder sie kannten die Person nicht und wollten Informationen über sie suchen. Diese Anfragen wurden aus der weiteren Untersuchung entfernt, da ihr Ziel nicht oder nur schwer vorhersagbar ist (vgl. LEE ET AL. 2005, S. 394).

Die Ergebnisse der Befragung zeigen, dass die meisten Suchanfragen ein vorhersagbares Ziel haben (entweder navigationsorientiert oder informationsorientiert) (vgl. LEE ET AL. 2005, S. 392). Einige Anfragen aus bestimmten Themengebieten sind schwer vorhersagbar und sollten bei einer automatischen Klassifizierung anders behandelt werden als die Anfragen, die eindeutig navigationsorientiert oder informationsorientiert zugeordnet werden können (vgl. LEE ET AL. 2005, S. 394).

Die Klicks aus der Vergangenheit und die Anchor-Link Verteilung werden als mögliche Charakteristika genutzt, um die Intention hinter der Suchanfrage vorherzusagen. Lee et al. gehen davon aus, dass Nutzer mit einem navigationsorientierten Informationsbedürfnis hauptsächlich auf das Ergebnis klicken werden, das sie im Kopf haben, wenn sie die Suche beginnen. Wenn die Klicks aus der Vergangenheit für diese Suchanfrage untersucht werden, kann man das dahinter liegende Ziel herausfinden. Es wird davon ausgegangen, dass wenn „ebay“ gesucht wird, die Homepage von Ebay gemeint ist und der Suchende das Ergebnis mit der Hauptseite von Ebay anklicken wird. Dafür wurden die Zieladressen der Links zu den einzelnen Schlagworten der Suchanfragen beobachtet und so die potenziellen Ziele der Suchenden identifiziert (vgl. LEE ET AL. 2005, S. 392).

Um herauszufinden, welche Ergebnisse die Nutzer in der Vergangenheit geklickt haben, wurde die Klickverteilung für die einzelnen Anfragen angesehen. Wenn das erste Ergebnis zu der entsprechenden Anfrage sehr viel öfter geklickt wurde als die weiteren Ergebnisse, dann sollte das Ziel hinter der dazugehörenden Anfrage navigationsorientiert sein. Sollte es keinen Ausschlag der Kurve beim ersten Ergebnis geben und die Kurve der Verteilung der Klicks eher flach ausfallen, sollte das Ziel hinter der Suchanfrage informationsorientiert sein. Lee et al. untersuchten außerdem die Anzahl der Ergebnisse, die zu einer Suchanfrage geklickt wurden. Bei einer navigationsorientierten Anfrage sollte nur ein Ergebnis geklickt werden, nämlich das der Website, die der Suchende im Kopf hatte. Bei einer informations-

orientierten Suchanfrage klickt der Suchende mehrere Ergebnisse an, um die gewünschten Informationen zu sammeln (vgl. LEE ET AL. 2005, S. 394). Für die Nutzung einer automatischen Klassifizierung müssen der Suchmaschine genügend Daten aus der Vergangenheit vorliegen. Da viele Suchanfragen öfter eingegeben werden, ist dies möglich (vgl. LEE ET AL. 2005, S. 395).

Für die Suchanfragen, die nicht oft eingegeben werden, kann der Anchortext der Webseite zur Bewertung hinzugezogen werden, denn in diesem tauchen die gesuchten Schlagworte eher auf als in den Webadressen der Homepages. Um dieses Kriterium zu untersuchen, wurden die Anchortexte mit den entsprechenden Schlagwörtern im Web lokalisiert und deren Zieladressen extrahiert. Danach wurde das Vorkommen der Zieladresse gezählt und nach der absteigenden Reihenfolge ihres Auftauchens in der Suchergebnisliste sortiert. Es wurden Histogramme entwickelt, die zeigen, wie viele Links mit dem jeweiligen Schlagwort der Suchanfrage auf welche Position in der Suchergebnisliste verweisen (vgl. LEE ET AL. 2005, S. 395).

Für navigationsorientierte Anfragen galt dabei wieder, dass die Verteilung der verweisenden Links bei dem ersten Treffer ausschlagen sollte, da es hier nur eine richtige Antwort bzw. Webseite als Antwort gibt. Bei den informationsorientierten Ergebnissen sollte es eine flachere Verteilung der verweisenden Links geben, da die Links nicht nur auf eine Webseite verweisen.

Ein Problem bei diesem Vorgehen können Spamseiten sein, die eine große Anzahl der Stichworte in ihren Anchortexten verwenden und so die Messung verfälschen können (vgl. LEE ET AL. 2005, S. 395).

Die Effektivität der möglichen Charakteristika wurde mit Hilfe des Vergleichssets von Suchanfragen aus der Befragung evaluiert. Jedes einzelnes Charakteristikum war zu 80 Prozent korrekt, in Kombination waren die Ergebnisse zu 90 Prozent korrekt (vgl. LEE ET AL. 2005, S. 392).

Lee et al. stellten ihre genutzten Charakteristika denen von Kim und Kang (Anchor Nutzungsrate, Schlagwortverteilung und Zuverlässigkeit des Ausdrucks) gegenüber und stellten fest, dass die Charakteristika von Kim und Kang nicht so effektiv sind, da sie mit dem gleichen Datensatz nur 60 Prozent richtig einordnen, die von Lee et al. genutzten Merkmale jedoch in

Kombination 90 Prozent richtig zugeordnete Suchanfragen hervorbringen (vgl. LEE ET AL. 2005, S. 399).

3.5 Dai et al. 2006

Dai et al. beschäftigen sich mit Suchanfragen, die eine kommerzielle Intention haben. Sie gehen davon aus, dass eine kommerzielle Intention hinter jeder Suchanfrage stecken kann, egal ob sie navigationsorientiert, transaktionsorientiert oder informationsorientiert ist (vgl. DAI ET AL. 2006, S. 2). Die Intention hinter einer Suchanfrage zu kennen, kann einer Suchmaschine und E-Commerce-Webseiten helfen, Inhalte zu personalisieren und die Zufriedenheit der Nutzer zu erhöhen (vgl. DAI ET AL. 2006, S. 1). Dai et al. gehen davon aus, dass die Suchenden, bevor sie eine Transaktion abschließen, Recherche betreiben und bspw. Preise vergleichen oder sich das Produkt anschauen. Danach werden sie eine Seite suchen, um das Produkt/den Service zu erwerben. Deshalb wird die kommerzielle Suche in zwei Phasen geteilt: „Recherche“ und „Zweck bezogen“ (der Zweck wäre hier einen Kauf/eine Transaktion durchzuführen) (vgl. DAI ET AL. 2006, S. 2).

Eine Anfrage ist kommerziell, wenn sie eine Auktion, einen Verkauf oder einen bezahlten Service verlangt, also z.B. den zahlungspflichtigen Download einer Musikdatei oder den Preisvergleich eines Produkts. Nicht kommerziell ist sie, wenn sie keinen Bezug zu einer kommerziellen Aktivität oder Service hat, wie z.B. die Suche nach einem Liedertext oder einem E-Mail-Dienst (vgl. DAI ET AL. 2006, S. 2). Da die wenigsten Suchanfragen wegen ihrer Länge Schlagworte für die Einordnung in kommerziell/nicht kommerziell enthalten, nutzen die Autoren die manuelle Form der Einordnung der Suchanfragen und den Inhalt der besuchten Webseiten (vgl. DAI ET AL. 2006, S. 2 und S. 5).

Einige Anfragen sind auch nicht eindeutig der kommerziellen oder nicht kommerziellen Intention zuzuordnen. Wenn man z.B. den Begriff „Kunst“ betrachtet, ist eine Webseite über die Geschichte der Kunst nicht kommerziell, während eine Auktionsseite für Gemälde kommerziell ist. Aus diesem Grund konnte keine existierende Systematik für die Einordnung der Suchanfragen herangezogen werden (vgl. DAI ET AL. 2006, S. 3).

Die Autoren griffen auf Menschen als Einordnungshilfe zurück. Sie sollten die Suchanfragen als kommerziell oder nicht kommerziell einordnen. Wenn die Absicht hinter der Suchanfrage nicht eindeutig zu erkennen war, hatten die Juroren die Möglichkeit, die Suchanfrage als „confusing“ einzuordnen. Die Einordnung erfolgte erst, wenn sich die Mehrheit der drei Juroren bei der Einordnung einig war. Jede Suchanfrage wurde von jedem Juror einzeln eingeordnet (vgl. DAI ET AL. 2006, S. 3 und 6).

Mit Hilfe einer Formel wurden verschiedene Faktoren wie bspw. die genutzten Schlagworte der Webseiten oder Buttons/Bilder mit der Aufschrift „order“ extrahiert, um Anhaltspunkte für die automatische Einordnung in kommerziell/nicht kommerziell zu finden (vgl. DAI ET AL. 2006, S. 4).

Für die Untersuchung wurden verschiedene Datenquellen genutzt:

- die Suchanfrage
- der Inhalt der Top-Landingpages, die von der Suchmaschine (auf der ersten Suchergebnisseite) empfohlen wurde
- die Suchergebnisseite der Suchmaschine
- die Anzahl der Klicks auf die Webseiten, die die Suchmaschine empfohlen hat

(vgl. DAI ET AL. 2006, S. 4).

Die Suchanfrage kann Worte enthalten, die helfen, die Intention des Nutzers zu erkennen. Die Suchergebnisseite enthält die URLs und die Kurzbeschreibungen der empfohlenen Seiten, aus denen sich Schlagworte ableiten lassen. Die Webseiten, die von den Nutzern ausgewählt werden, zeigen, wie die Nutzer diese auswählen und in welche Seiten sie besonderes Vertrauen haben (vgl. DAI ET AL. 2006, S. 5).

Für die Einordnung kommerziell/nicht kommerziell wurden die gerankten Webseiten als Untersuchungsfaktoren herangezogen und auf Schlagworte, die auf eine kommerzielle Intention schließen lassen, untersucht. Wenn die meisten Webseiten der Ergebnisseite Hinweise auf eine kommerzielle Aktion enthielten, wurden die Anfrage als kommerziell eingeordnet (vgl. DAI ET AL. 2006, S. 5).

Die Beschreibungen der Webseiten auf der Suchergebnisseite enthalten strukturierte Informationen, die man für eine Untersuchung nutzen kann. Die Webseiten sind von der Suchmaschine bzw. dem Algorithmus bereits gerankt worden und sind somit für die jeweilige Suchanfrage relevant. Die

Kurzbeschreibungen der Webseite enthalten oft nur den Teil mit den gesuchten Schlagworten, der relevanter sein kann, als die gesamte Homepage und sie haben eine bestimmte Länge. Dai et al. nutzen die ersten zehn Treffer der Suchergebnisseite um die Einordnung in kommerziell/nicht kommerziell vorzunehmen (vgl. DAI ET AL. 2006, S. 6).

Für das eigentliche Experiment wurden 1.408 englische Suchanfragen eines Tages aus dem MSN Suchlog untersucht, und die dazugehörigen ersten zehn Suchergebnisseiten der Suchmaschine sowie die ersten zehn Webseiten, auf die auf der Ergebnisseite verwiesen wurden, wurden gesammelt. Außerdem wurden 26.186 kommerzielle Webseiten zufällig ausgewählt, um mehr Webseiten mit kommerzieller Intention auf ihre Charakteristika untersuchen zu können (vgl. DAI ET AL. 2006, S. 6).

Die Suchanfragen, Suchergebnisse und Webseiten wurden von drei Juroren separat als kommerziell, nicht kommerziell oder confusing (nicht erkennbar) eingeordnet. Diejenigen Suchanfragen und Webseiten, bei denen sich mindestens zwei der Juroren einig waren und die nicht als confusing eingeordnet worden waren, wurden in die Untersuchung einbezogen (1.392 Suchanfragen und 25.897 englische Webseiten). Um ein Ungleichgewicht in den Trainingsdaten zu vermeiden, wurden aus den bereits eingeordneten Daten die gleiche Anzahl kommerzieller und nicht kommerzieller Webseiten herangezogen und zusätzlich wurden die Juroren gebeten, die jeweils gleiche Anzahl von kommerziellen und nicht kommerziellen Webseiten aus dem Internet zu extrahieren. Diese Methode spiegelt zwar nicht das Internet wieder, verhindert aber, dass mehr kommerzielle Webseiten in den Untersuchungsdaten vorhanden sind, so dass eine Beispieluntersuchung durchgeführt werden kann (vgl. DAI ET AL. 2006, S. 6).

Für die Evaluierung wurden die Daten zu gleichen Teilen in eine Trainings- und eine Testgruppe eingeteilt (vgl. DAI ET AL. 2006, S. 6). Aus den Trainingsdaten wurden Keywords ausgewählt und extrahiert, um diese dann an den Testdaten mit den gefundenen Faktoren zu testen. Dabei wurde auch untersucht, welcher Faktor (die zehn ersten Suchergebnisse oder die Inhalte der ersten zehn Webseiten) am besten die kommerzielle Intention voraussagen kann (vgl. DAI ET AL. 2006, S. 7). Die Resultate zeigen, dass die Vorhersage auf der Basis der ersten zehn Treffer auf der

Suchergebnisseite besser funktioniert als die Einordnung mit Hilfe der eigentlichen Webseiten (Landingpages), die für die jeweilige Suchanfrage als Topergebnis ausgegeben wurden. Am schlechtesten funktionieren beide Faktoren kombiniert (vgl. DAI ET AL. 2006, S. 7).

Um ein besseres Verständnis für die Intentionen der Nutzer zu bekommen, teilten die Forscher die Häufigkeit der Suchanfragen in fünf Bereiche ein: einfach, sehr niedrig, niedrig, mittel und hoch. Eine einfache Anfrage wird einmal im Monat gestellt, Suchanfragen aus dem Bereich hoch sind die beliebtesten Anfragen. Aus jedem Bereich wurden 10.000 zufällige Anfragen ausgewählt und mit diesen insgesamt 50.000 Suchanfragen wurde die Einordnung in kommerziell und nicht kommerziell wiederholt. Die Autoren stellten fest, dass 41 Prozent der beliebtesten Anfragen und 32 Prozent der einfachen Suchanfragen eine kommerzielle Intention haben (vgl. DAI ET AL. 2006, S. 8).

Folgende Faktoren wollen die Autoren in zukünftigen Studien verbessern. Sie wollen die Einteilung durch Menschen vermeiden, um subjektive Einschätzungen zu verhindern. Die Forschung sollte durch die Klickdaten der einzelnen Suchanfragen ergänzt werden. Es soll nach weiteren evtl. besseren Faktoren gesucht werden, um die kommerziellen Suchanfragen besser zu identifizieren. Außerdem soll die kommerzielle Intention des einzelnen Nutzers untersucht werden, da verschiedene Nutzer, die die gleiche Suchanfrage eingeben, damit verschiedene Ziele verfolgen können (vgl. DAI ET AL. 2006, S. 8).

3.6 Jansen et al. 2007

Jansen et al. nutzen in ihrer Studie die von Broder definierten Klassen transaktionsorientiert, navigationsorientiert und informationsorientiert, allerdings wurden hier noch weitere Unterklassen gebildet. Sie untersuchen 1,5 Millionen Suchanfragen,¹ die an einem Tag bei der Suchmaschine Dogpille.com aufgezeichnet wurden, um sie anschließend automatisch Kategorien zuzuordnen. Die Suchanfragen aus dem untersuchten Log enthielten folgende Daten: Eine Nutzer-ID, die automatisch von dem Webserver zugeordnet wurde, um den einzelnen Computer zu identifizieren;

¹ Anfragen ohne Worte (Null-Anfragen) und solche die durch die Rückkehr des Nutzers auf die Suchergebnisseite generiert wurden, wurden entfernt.

anonyme Cookies, um einzelne Nutzer zu identifizieren; Zeit in Stunden, Minuten und Sekunden; die Suchanfrage genau so, wie sie vom Nutzer eingegeben wurde; Quelle, den Inhalt, den der Nutzer zum Suchen gewählt hat (z.B.: Web, Bilder, Audio oder Video) (vgl. JANSEN ET AL. 2007, S. 1257).

Die Untersuchung der Suchanfragen fand mit Hilfe einer relationalen Datenbank statt. Um die einzelnen Suchanfragen zu identifizieren und chronologische Abläufe der Suche zu rekonstruieren, wurden Zeit, Nutzer-ID, die Suchanfrage und die Cookies als Identifikationsfelder gewählt. Suchanfragen, die von Agents in die Suchmaschine eingegeben wurden, wurden entfernt, damit sichergestellt war, dass die eingebenden Suchwörter von Menschen stammen (vgl. JANSEN ET AL. 2007, S. 1257). Die herausgefundenen Merkmale der unterschiedlichen Suchanfragen wurden in einen Algorithmus eingebunden, der die automatische Klassifikation der Suchanfragen übernahm (vgl. JANSEN ET AL. 2007, S. 1257).

Um die Merkmale der einzelnen Klassen zu erheben, isolieren die Autoren die Suchanfragen nach Klassen und untersuchen diese getrennt. Die gefundenen Merkmale wurden dann genutzt, um die Suchanfragen automatisch kategorisieren zu lassen (vgl. JANSEN ET AL. 2007, S. 1255).

Dabei nutzen sie zum Vergleich auch Juroren (insgesamt 3 Juroren, die sich auf eine Klasse einigten) (vgl. JANSEN ET AL. 2007, S. 1258), die 400 Suchanfragen den verschiedenen Anfragetypen informationsorientiert, navigationsorientiert und transaktionsorientiert zuordneten. Diese Anfragen werden auch maschinell zugeordnet. Der Vergleich der manuell und der automatisch klassifizierten Anfragen zeigt, dass 74 Prozent der Anfragen korrekt zugeordnet wurden, 25 Prozent der Anfragen sind allerdings zu vage oder zu facettenreich, um sie zuzuordnen.

Um die Suchanfragen zu untersuchen, wurden verschiedene aufgezeichnete Suchvorgänge, die nicht nur die Anfragen, sondern auch die Reihenfolge der Anfragen in einer Session, die Anfragenlänge und die Ergebnisseite zur Analyse enthalten, herangezogen. Diese weiteren Faktoren zeigten Attribute hinter der Anfrage und waren hilfreich für die Klassifikation (vgl. JANSEN ET AL. 2007, S. 1256).

Das Ergebnis der Suchcharakteristika der Nutzer stimmt bei Jansen et al. mit anderen Forschern überein und nutzt die von Broder und Rose und

Levinson genutzten Kategorien. Allerdings fügen sie diesen noch Unterkategorien hinzu.

Die Kategorie Informationsorientiert erhält die Unterklassen:

- Gezielt (→ weitere Unterklassen: geschlossen und offen)
- Ungezielt
- Finden
- Liste
- Auskunft

(vgl. JANSEN ET AL. 2007, S. 1259).

Die Unterklassen bedeuten, dass eine Frage gezielt gestellt wurde oder ungezielt, der Suchende also alle Informationen zu einem Thema haben wollte oder eine spezifische Frage gestellt hat. Unterteilt wird hier noch einmal in geschlossen, was eine Frage mit genau einer richtigen Antwort entspricht, und offen, also eine Frage, die zwei oder mehr Themen umschließt.

Finden beschreibt die Intention, einen „echten“ Dienstleister oder ein Produkt zu finden. Liste gibt eine Liste von Kandidaten aus, etwa alle Dinge die man in einer Stadt (z.B. an einem Urlaubsziel) unternehmen kann. Auskunft soll Auskünfte, Ideen, Vermutungen oder Anleitungen beschreiben (vgl. JANSEN ET AL. 2007, S. 1260).

Für navigationsorientierte Suchanfragen wurden folgende Unterklassen erschaffen:

- Navigation zu Transaktionsorientiert
- Navigation zu Informationsorientiert (→ weitere Unterklasse: online)

(vgl. JANSEN ET AL. 2007, S. 1259).

Navigationsorientiert beschreibt die Suche nach einer Homepage. Ist der Inhalt der gesuchten URL transaktionsorientiert, wurde sie der Unterklasse Navigation zu Transaktionsorientiert zugeordnet. Wollte der Suchende eine informationsorientierte URL, bekam die Suchanfrage die Unterklasse Navigation zu Informationsorientiert. Die Unterklasse „Online“ sagt dabei aus, dass die Quelle online zu finden sein soll (vgl. JANSEN ET AL. 2007, S. 1260).

Transaktionsorientiert erhält folgende Unterklassen:

- Beschaffen (→ weitere Unterklassen: offline und frei)
- Download (→ weitere Unterklassen: nicht frei und Links)

- Suchmaschinenergebnisseite (→ weitere Unterklasse: andere)
- Interaktion

(vgl. JANSEN ET AL. 2007, S. 1259).

Beschaffen beschreibt die Absicht, eine bestimmte Ressource oder ein Objekt beschaffen zu wollen. Die Unterklasse Offline sagt aus, dass die Ressource offline zu finden sein soll und evtl. weitere Aktionen des Nutzers erfordert. Frei bedeutet, dass der Download einer Datei kostenlos und frei sein soll.

Mit Download wird die Intention bezeichnet, eine Datei zum Downloaden zu finden. Die Unterklasse beschreibt, dass der Download nicht unbedingt frei sein muss. Link bedeutet, dass die Ressource im Titel, in der Zusammenfassung oder in der URL von einem oder mehreren Ergebnissen der Ergebnisseite erscheint. Suchmaschinenergebnisseite soll ein Ergebnis liefern, das gedruckt, gespeichert oder auf der Ergebnisseite gelesen werden kann. Die Unterklasse Andere bedeutet, dass die Ressource nicht in der Ergebnisliste auftaucht, aber irgendwo anders auf der Suchergebnisseite (z.B. gibt der Nutzer eine Anfrage ein, um die Rechtschreibung zu prüfen, ohne an den Suchergebnissen interessiert zu sein). Wenn eine Suchanfrage Interaktion zugeordnet wurde, möchte der Nutzer mit anderen Programmen oder Ressourcen auf anderen Webseiten interagieren (vgl. JANSEN ET AL. 2007, S. 1260).

Jansen et al. legten für die unterschiedlichen Suchanfragekategorien folgende Charakteristika fest:

Navigationsorientierte Suchen:

- Anfragen, die Firmen, Geschäfte, Organisationen, Personennamen enthalten
- Anfragen, die Domainendungen enthalten
- Anfragen, deren Quelle im Web zu finden ist
- Länge der Suchanfrage (weniger als drei Worte)
- Suchende sieht sich nur die erste Seite der Suchergebnisse an

Transaktionsorientierte Suchen:

- Anfragen, die Worte zu Filmen, Liedern, Texten, Rezepten, Bildern, Humor und Porno enthalten
- Anfragen mit „Beschaffungs“- Termen (z.B.: Texte, Rezepte usw.)
- Anfragen mit „Download“- Termen (z.B. Download, Software usw.)

- Anfragen bezogen auf Bilder-, Audio- oder Videosammlungen
- Anfragen die „Audio“, „Bilder“ oder „Video“ als Quelle angeben
- Anfragen mit „Entertainment“- Termen (z.B. Bilder, Spiele usw.)
- Anfragen mit „Interaktions“- Termen (z.B. kaufen, Chat usw.)
- Anfragen mit Endungen, die auf Filme, Songs, Texte, Bilder und Multimedia oder komprimierte Dateien hindeuten (jpeg, zip usw.)

Informationsorientierte Suchen:

- Nutzung von Frageworten (z.B. Wege nach, Wie, Was ist usw.)
- Anfragen mit natürlichen Formulierungen
- Anfragen, die informationsorientierte Terme enthalten (z.B.: Liste, Playlist usw.)
- Anfragen, die nach der ersten Anfrage gestellt wurden (Verfeinerung/Änderung)
- Anfragen, bei denen der Suchende mehrere Ergebnisseiten angesehen hat
- Anfragenlänge mehr als zwei Worte
- Anfragen, die die Kriterien von navigationsorientierten und transaktionsorientierten Anfragen nicht erfüllen

(vgl. JANSEN ET AL. 2007, S. 1261).

Mit den herausgefundenen Merkmalen wurden die Suchanfragen mittels eines Programms automatisch klassifiziert.

Die Ergebnisse zeigten, dass mehr als 80 Prozent der Suchanfragen informationsorientiert waren, während navigationsorientiert und transaktionsorientiert jeweils mit zehn Prozent vertreten waren.

Die Anzahl der navigationsorientierten Anfragen ist verglichen mit anderen Studien hier sehr niedrig (BRODER 2002: 20 bis 24%). Die niedrige Anzahl der transaktionsorientierten Anfragen war ebenfalls ungewöhnlich (BRODER 2002: 30 - 36%; Jansen/Spink 2005; 12 bis 24% bei 2.500 untersuchten Anfragen aus verschiedenen Logs). Die Unterschiede im Vorkommen der verschiedenen Anfragetypen können laut Jansen et al. an der Größe der untersuchten Stichproben liegen. Die zum Vergleich herangezogenen Studien haben viel kleinere Datenmengen untersucht. Es kann aber auch daran liegen, dass z.B. Rose und Levinson nur die erste Anfrage einer Session untersucht haben und daher evtl. mehr Anfragen der Klasse transaktionsorientiert zugeordnet haben (vgl. JANSEN ET AL. 2007, S. 1262).

Um die Effektivität der automatischen Klassifizierung zu prüfen, wurden 400 zufällig ausgewählte Anfragen von Juroren zugeordnet und die Ergebnisse mit denen der automatischen Klassifizierung verglichen (vgl. JANSEN ET AL. 2007, S. 1262).

26 Prozent der automatisch klassifizierten Suchanfragen wurden nicht korrekt zugeordnet. Der Algorithmus ordnete dem Anfragetyp Informationsorientiert zu viele und den Anfragetypen Transaktionsorientiert und Navigationsorientiert zu wenige Anfragen zu. Davon ausgehend, dass dieser Anteil im ganzen Datensatz so bleibt, würden informationsorientierte Anfragen zu 65 Prozent, navigationsorientierte Anfragen zu 15 Prozent und transaktionsorientierte zu 20 Prozent auftauchen. Diese Zahlen gehen aber davon aus, dass die Zuordnung von Menschen korrekt ist und jede Anfrage nur eine dahinterliegende Intention hat. Normalerweise benutzen verschiedene Nutzer die gleichen Suchworte mit unterschiedlichen Informationsbedürfnissen. Mehrere Intentionen können hinter 20–30 Prozent der Anfragen stecken. Diese Anfragen sind es auch, bei denen es zu Fehlzuordnungen kam (vgl. JANSEN ET AL. 2007, S. 1262).

Jansen et al. gehen davon aus, dass 75 Prozent richtig automatisch zugeordnete Anfragen zeigen, dass die automatische Klassifizierung funktionieren kann (vgl. JANSEN ET AL. 2007, S. 1263).

3.7 Calderon- Benavides et al. 2010

Eine andere Art der Einordnung von Suchanfragen nutzen Calderon-Benavides et al. und ordnen die untersuchten Suchanfragen in Facetten ein, um herauszufinden, welche für die Verbesserung der Suchergebnisse am geeignetsten sind. Die von den Autoren als repräsentativ eingestuften Facetten sind:

- das Genre, das Ziel,
- die Spezifität/Genauigkeit,
- der Bereich/die Sparte,
- das Thema/Topic,
- die Aufgabe/der Auftrag,
- Behördensensitivität (authority sensitivity),
- räumliche Sensitivität,

- Zeitsensitivität.

In dieser Studie werden allerdings nur die ersten vier Facetten betrachtet (vgl. CALDERON- BENAVIDES ET AL. 2010, S. 22).

Zur Einordnung in die verschiedenen Facetten wurde ein Suchanfragenset von 5.249 Suchanfragen aus einem Weblog genutzt. Die Suchanfragen wurden von Personen klassifiziert. Um die Klassifikationen zu prüfen, wurden zehn Prozent der Anfragen von zwei Menschen zugeordnet. Die Übereinstimmung der Juroren für die Facetten Zeit Sensitivität, Bereich und Ziel waren nahezu perfekt, für Behördensensitivität, Thema, Aufgabe und räumliche Sensitivität war die Übereinstimmung erheblich, die Facette Genre schnitt mäßig ab und die Facette Spezifität schnitt ausreichend ab. Diese Ergebnisse der Übereinstimmungen der Juroren und statistischer Berechnungen zeigen, dass die Zuordnung durch Menschen konsistent und vertrauenswürdig ist, denn gerade bei der Bereitstellung von neuen Klassen hätten die Überschneidungen geringer ausfallen können (vgl. CALDERON- BENAVIDES ET AL. 2010, S. 22).

Calderon- Benavides et al. fanden heraus, dass die Facetten Thema und Genre in Verbindung mit anderen Werten, wie bspw. der Aktion in der Facette Objekt, die besten Grundlagen für die Verbesserung der Websuche lieferten. Die Genre-Facette unterstützt die Thema-Facette und so bekommt man einen Eindruck von der Intention hinter der ausgeführten Suche. Wenn man also weiß, dass ein Suchender „HP Drucker“ sucht und diese Suche ebenso zu dem Thema Technologie und Elektronik gehört wie zu dem Genre Business, kann man die Facette Ziel dazu nehmen. Wenn letztere Facette eine Aktion angibt, kann man mit hoher Wahrscheinlichkeit davon ausgehen, dass der Suchende eine Transaktion im Bereich Technik und Elektronik durchführen möchte (vgl. CALDERON- BENAVIDES ET AL. 2010, S. 22 u. 24).

Die Autoren stellten in ihrer Untersuchung fest, dass die Suche in Suchmaschinen nicht als einfache Aufgabe, sondern in ihrer gesamten Komplexität gesehen werden sollte. Nur so kann auch die hinter der Anfrage stehende Komplexität herausgefunden werden.

Mit der Nutzung der Facette Genre als Meta-Klasse lassen sich die Suchanfragen gut nach Themen ordnen, genauso wie mit der Verbindung Genre und Ziel. Beispielsweise sind Anfragen, die dem Genre Business zuge-

ordnet werden, auch mit einer auszuführenden Aktion verbunden (diese Einordnung fällt in die Facette Ziel), was auf eine transaktionsorientierte Anfrage hindeutet und unabhängig vom Thema betrachtet werden kann (vgl. CALDERON- BENAVIDES ET AL. 2010, S. 24).

Die Autoren merken selber an, dass die Forschung in diesem Bereich am Anfang steht und weder komplett noch definitiv ist (vgl. CALDERON- BENAVIDES ET AL. 2010, S. 24).

4. Klassifikation von Suchanfragen nach Themen

Einige Studien beschäftigen sich nur mit der thematischen Zuordnung von Suchanfragen, um herauszufinden, ob es Unterschiede bei den Suchanfragen je nach Thema gibt.

4.1 Spink et al. 2001

Spink et al. haben 2001 über eine Millionen Suchanfragen untersucht und Themen zugeordnet. Die Suchanfragen stammten aus Logfiles der Suchmaschine Excite und wurden innerhalb eines Tages eingegeben (vgl. SPINK ET AL. 2001, S. 3).

Untersucht wurden

- diejenigen Begriffe, die Worte, Abkürzungen, Nummern und logische Operatoren (AND, OR, NOT) enthalten konnten. URLs und E-Mail-Adressen wurden als einzelnes Wort behandelt.
- Die Suchanfragen, die aus einem oder mehreren Wörtern bestehen konnten und erweiterte Suchoperatoren wie logische Operatoren enthalten konnten.
- Unique Suchanfragen: alle unterschiedlichen Anfragen, die von einem Nutzer während einer Session eingegeben wurden. Diese Anfragen können Modifikationen der vorherigen Anfrage sein oder komplett neue.
- Wiederholte Suchanfragen; die mehrfach auftretenden, gleichen Suchanfragen. Sie zeigen das Bedürfnis, mehrere Webseiten zu einer Suchanfrage ansehen zu wollen (wenn ein Nutzer auf die Sucher-

gebnisseite zurückkehrt, um eine weitere Webseite der Suchergebnisliste anzusehen, generiert Excite die gleiche Anfrage noch einmal).

- Null-Begriffe-Suchanfragen; Suchanfragen ohne Begriffe. Sie werden generiert, wenn ein Nutzer „More like this“ klickt, um mehr ähnliche Ergebnisse zu sehen oder eine Anfrage ohne Begriffe oder Symbole abschickt.
- Sessions, also das gesamte Set der Anfragen eines Nutzers über die aufgenommene Zeit. Dies kann eine einzelne Anfrage sein oder mehrere Unique oder sich wiederholende Anfragen enthalten.
- Ergebnisseiten; diese zeigen dem Nutzer die Suchergebnisse zu einer Suchanfrage. Eine Seite enthält zehn Ergebnisse, die nach Relevanz sortiert sind. Der Nutzer kann entscheiden, ob er sich eine oder mehrere Ergebnisseiten ansieht

(vgl. SPINK ET AL. 2001, S. 3).

Neben den Charakteristika wie Länge der Anfragen, Anzahl der Wörter und Nutzung von Operatoren wurde auch die Einordnung in Themen vorgenommen (vgl. SPINK ET AL. 2001, S. 8 und S.11). Für diese Einordnung wurden die Daten qualitativ und quantitativ untersucht.

Bei der qualitativen Untersuchung der Suchanfragen wurden die Themen der Wortpaare evaluiert, die in einer Suchanfrage vorkamen (vgl. SPINK ET AL. 2001, S. 12).

Die qualitative Untersuchung wurde anhand von 2.414 Suchanfragen durchgeführt (vgl. SPINK ET AL. 2001, S. 13). Die Klassifizierung wurde von Studenten, die Library and Information Science studierten, durchgeführt.

Diese entwickelten die folgenden elf Hauptkategorien:

1. Entertainment, recreation
2. Sex, pornography, preferences
3. Commerce, travel, employment, economy
4. Computers, the Internet
5. Health, the sciences
6. People, places, things
7. Society, culture, ethnicity, religion
8. Education, the humanities
9. The performing and fine arts
10. Government

11. Unknown, incomprehensible

Jede dieser Klassen hat weitere Unterklassen, die in dem Artikel allerdings nicht genannt werden. Bei der Einordnung der Suchanfragen fielen die meisten Anfragen in den Bereich „Entertainment, recreation“, dicht gefolgt von „Sex, pornography, preferences“. Die wenigsten Suchanfragen entfielen auf die Themenbereiche „Government“ und „Unknown“ (vgl. SPINK ET AL. 2001, S. 13).

Durch die quantitativen Untersuchungen haben Spink et al. herausgefunden, dass die meisten der Suchanfragen, die an Suchmaschinen gestellt werden kurz und von simpler Struktur sind. Sehr wenige Suchanfragen enthalten „erweiterte Suche“-Elemente wie Boolesche Operatoren oder „“ als Kennzeichnung für eine Phrasensuche.

Trotz der großen Ergebnismenge, die die Nutzer auf ihre Anfragen erhalten, schauen sie nur sehr wenige Ergebnisse an, meistens nur die erste und zweite Ergebnisseite.

Einige Suchanfragen werden oft gestellt, die Mehrheit wird aber in einem bestimmten Zeitraum nur einmal gestellt (Long Tail) (vgl. SPINK ET AL. 2001, S. 14).

4.2 Jansen und Spink 2006

Im Jahr 2006 analysieren und vergleichen Spink und Jansen neun Studien aus den USA und Europa auf Grundlage von Log-Dateien verschiedener Suchmaschinen (AlltheWeb.com, Excite, AltaVista), um Trends und Unterschiede im Suchverhalten zu erkennen (vgl. JANSEN UND SPINK 2006, S. 252). In dieser vergleichenden Studie fließen auch die Ergebnisse aus der zuletzt beschriebenen Studie ein (vgl. JANSEN UND SPINK 2006, S. 252). Datengrundlage sind dabei über eine Milliarde Suchanfragen. Sie untersuchen u.a. die Länge der Suchanfragen, wie viele Webseiten sich die Nutzer ansehen, die Nutzung von Suchoperatoren und die thematische Einordnung der Suchanfragen. Für den letztgenannten Punkt bezogen sie allerdings nur 2.600 Suchanfragen in die Stichprobe ein. Die Suchanfragen wurden von zwei unabhängigen Juroren den Themen (Menschen, Plätze oder Dinge; Computer oder Internet; Handel, Reise, Arbeit oder Wirtschaft; Sex oder Pornografie; Entertainment oder Erholung; Gesundheit oder Wissenschaft; Gesellschaft, Kultur, Volkszugehörigkeit oder Religion; dar-

stellende oder bildende Kunst; Bildung oder Geisteswissenschaften; Regierung; Unbekannt oder Andere) zugeordnet. Anschließend trafen sich die Juroren, um Unstimmigkeiten zu lösen (vgl. JANSEN UND SPINK 2006, S. 258 f.).

Die meistgenutzten Themenfelder in der Studie waren Menschen, Plätze oder Dinge mit fast der Hälfte aller Suchanfragen (AlltheWeb und AltaVista) sowie Handel, Reise, Arbeit oder Wirtschaft mit einem Anteil von 25 Prozent (Excite) (vgl. JANSEN UND SPINK 2006, S. 258 f.). Die Autoren stellten in Vergleichen verschiedener Jahre (insgesamt Daten von 1997 bis 2002) fest, dass es einen Anstieg bei den Suchen nach Personen (Menschen, Plätze oder Dinge); Handel, Reise, Arbeit oder Wirtschaft und Computer oder Internet gab, dafür aber einen Rückgang bei den Suchen nach Sex und Pornografie sowie Entertainment oder Erholung (vgl. JANSEN UND SPINK 2006, S. 259).

Daraus leiten Spink und Jansen den Trend ab, dass die Websuche stärker als Werkzeug für die Suche nach Informationen und Handel genutzt wird und weniger zur Unterhaltung (vgl. JANSEN UND SPINK 2006, S. 260).

4.3 Klassifikation von Suchanfragen nach Anfragetyp und Thema

Weitere Forschungen zu den Anfragetypen gibt es in Studien, die die Kategorisierung der Suchanfragen in Verbindung mit unterschiedlichen Themen untersuchen.

4.4 Lewandowski 2006

Lewandowski erhob 1.500 Suchanfragen verschiedener deutscher Suchmaschinen und analysierte diese ebenfalls nach Themenzuordnung und Anfragetyp. Diese beiden Klassifikationen wurden kombiniert, um herauszufinden, ob bestimmte Anfragetypen häufiger in einem Themenbereich vorkommen (vgl. LEWANDOWSKI 2006, S. 33). Lewandowski nutzte dabei die von Broder definierten Anfragetypen informationsorientiert, navigationsorientiert und transaktionsorientiert (vgl. LEWANDOWSKI 2006, S. 36). Die Themen wurden in dieser Studie von Jansen und Spink übernommen (vgl. LEWANDOWSKI 2006, S. 35 und JANSEN UND SPINK 2006, S. 259).

Um die Suchanfragen zu klassifizieren und den Themen zuzuordnen, wurden sie von zwei Juroren eingeordnet. Wenn die Juroren nicht einer Meinung waren, wurde die Einordnung diskutiert, um dann zu einer gemeinsamen Entscheidung zu kommen (vgl. LEWANDOWSKI 2006, S. 35). Bei den informationsorientierten Anfragen wurden die meisten Anfragen aus dem Themenbereich Gesundheit und Wissenschaft sowie aus dem Thema Gesellschaft, Kultur, Ethnizität und Religion gestellt. Die wenigsten Anfragen dieses Typs kamen aus den Bereichen Personen und Orte sowie Computer und Internet. In den Themen Personen, Wirtschaft und Arbeit, Unterhaltung und Freizeit, Regierung und Verwaltung sowie Kunst sind die meisten Anfragen navigationsorientiert.

Suchanfragen des Anfragetyps transaktionsorientiert finden sich häufig in den Themen Computer und Internet, Wirtschaft und Arbeit und ebenfalls in Sex und Pornografie (vgl. LEWANDOWSKI 2006, S. 40).

Bei den bisherigen Studien wurde festgestellt, dass jede Suchmaschine die von Broder identifizierten Anfragetypen bedienen können muss, da diese in allen Suchmaschinen einen nennenswerten Anteil der Suchanfragen ausmachen (vgl. BRODER 2002, S. 9). Auch Lewandowski kommt zu diesem Schluss (vgl. LEWANDOWSKI 2006, S. 39). Allerdings könnte es für einige Themenblöcke von Vorteil sein, wenn die Suchmaschine hier mit automatisierten Hinweisen arbeiten würde, die bspw. auf Downloadverzeichnisse für Software o.ä. im Bereich Computer verweisen oder im Themenbereich Unterhaltung eine Audio- oder Videosuche anbieten würden (vgl. LEWANDOWSKI 2006, S. 43).

4.5 Baeza-Yates et al. 2006

Baeza-Yates et al. nutzen die gespeicherten Interaktionen und Informationen in Logfiles, um die Intention hinter den Suchanfragen an eine Suchmaschine herauszufinden (vgl. BAEZA-YATES ET AL. 2006, S. 99).

Die Autoren ordnen die Suchanfragen nach Kategorien und nach Themen. Sie nutzen die Kategorien

- Informationsorientiert: Nutzer möchte eine Information im Web erhalten, unabhängig von dem Wissen der Quelle (er hat keine bestimmte Homepage im Kopf wie bspw. bei BRODER 2002)

- Nicht informationsorientiert: Dies sind Anfragen, die zu einem anderen Ziel oder zu einer bestimmten Transaktion führen (z.B.: Kauf, Download, Reservierung etc.)
- Mehrdeutig: Dies sind Anfragen, bei denen das Ziel nicht direkt von der Anfrage ablesbar ist (manchmal, weil das Ziel des Nutzers mehrdeutig ist)

(vgl. BAEZA-YATES ET AL. 2006, S. 99 f).

Um die Belange der Suchenden zu erforschen, ordneten die Autoren die Suchanfragen Themen zu, da diese das Informationsbedürfnis hinter der Suchanfrage angeben (vgl. BAEZA-YATES ET AL. 2006, S. 100).

Die Themen wurden vom Open Directory Project ODP² übernommen und um drei zusätzliche Themen ergänzt. Damit ergab sich folgende Liste:

- Arts
- Games
- Kids and Teens
- Reference
- Shopping
- World
- Business
- Health
- News
- Society
- Technology/Computer
- Home
- Recreation
- Science
- Sport
- Various (Anfragen, die in mehr als ein Thema passen)
- Other (Anfragen, die in keins der Themen passen)
- Sex

(vgl. BAEZA-YATES ET AL. 2006, S. 100).

Um die Intention der Suchanfragen zu eruieren, wurden zwei Methoden gewählt. Die Support Vector Machines Model-Methode wurde schon in anderen Studien genutzt, um bspw. Texte zu klassifizieren. Die Autoren

² Open Directory Project abrufbar unter: [http:// www.dmoz.org/](http://www.dmoz.org/)

gehen davon aus, dass die Wörter der Homepage, die die Nutzer auswählen (anklicken), die Suchanfrage widerspiegeln. Daher schien diese Methode geeignet, um Suchanfragen zu klassifizieren (vgl. BAEZA-YATES ET AL. 2006, S. 100).

Da Baeza-Yates et al. herausfinden wollten, was Nutzer dazu motiviert, ihre Suche im Web durchzuführen, wurde die Probabilistic Latent Semantic Analysis (PLSA) als zweite Methode gewählt. Sie wurde in der Vergangenheit genutzt, um Nutzerinteressen zu erforschen (vgl. BAEZA-YATES ET AL. 2006, S. 101).

Untersucht wurden 6.042 Suchanfragen und die getätigten Klicks der Nutzer (insgesamt 22.190 Klicks auf 18.527 unterschiedliche Webseiten). Diese Daten wurden anhand der Wörter, die in den Suchanfragen und in den geklickten Webseiten auftauchten, zu Gruppen zusammengefasst. Eine Gruppe bestand aus Wörtern, die das gleiche Ziel (Thema, Subjekt) unterschiedlich ausdrücken. Um die automatische Klassifikation überprüfen zu können, wurde ein Teil der gruppierten Suchanfragen manuell von einem Team von Juroren klassifiziert. Durch die gegebene Zuordnung zu einer Gruppe, wussten die Juroren etwas über den Hintergrund der Suchanfrage und konnten sie so einfacher den vorgegebenen Klassen und Themen zuordnen (vgl. BAEZA-YATES ET AL. 2006, S. 101).

Die eingeordneten Suchanfragen wurden als Trainingsdaten für die automatische Klassifikation genutzt, damit die Programme „lernen“, wie die unterschiedlichen Suchanfragen den Klassen und Themen zugeordnet werden sollen. Unterschiede tauchten bei den mehrdeutigen Anfragen und im Thema Various auf. In diesen Feldern kann die Suchintention allerdings mehrere Bedeutungen haben (vgl. BAEZA-YATES ET AL. 2006, S. 104).

Für die klassifizierten Suchanfragen wurden Recall und Precision gemessen. Bei den informationsorientierten Suchanfragen war die Precision hoch und der Recall nahezu perfekt. Dies rührt daher, dass in dieser Klasse eine Homepage zutreffend ist, welche von der Suchanfrage abgelesen werden kann (im Gegensatz zu den nicht informationsorientierten und vielseitigen Anfragen). Die Precision war aber auch bei den anderen beiden Klassen bei über 50 Prozent (vgl. BAEZA-YATES ET AL. 2006, S. 104 f.).

Bei der Themenzuordnung sind die Ergebnisse generell gut. Suchanfragen aus dem Bereich Computer lassen sich leichter identifizieren als bei anderen Themen, weil es hier ein relativ spezielles Vokabular gibt (vgl. BAEZA-YATES ET AL. 2006, S. 105).

Untersucht wurden auch die Beziehungen zwischen den Zielen der Nutzer und den Themen der Suchanfragen. Informationsorientierte Anfragen sind in den Themen Business, Recreation, Society, Education und News häufiger vertreten, dafür seltener in den Themenfeldern Shopping, Home, Sex und Others. Letztere Themen zeigen, dass die Nutzer hier nicht an Informationen interessiert sind, sondern zum Beispiel eher ein Spiel downloaden wollen (vgl. BAEZA-YATES ET AL. 2006, S. 105).

Nicht informationsorientierte Anfragen fallen eher in die Themengebiete Sex, Recreation und Technology/Computer, sind dafür aber seltener in den Themen Society, Health, Science und News zu finden (vgl. BAEZA-YATES ET AL. 2006, S. 105 f).

Die vielseitigen Anfragen haben kein einzelnes Ziel und sind mit keinem Thema verbunden. Das größte Vorkommen gibt es daher im Themenfeld Various (vgl. BAEZA-YATES ET AL. 2006, S. 106).

Um die Verbindung verschiedener Suchanfragen auf dem Themenlevel zu untersuchen, wurden die Themen, zu denen eine Anfrage gehören konnte, als Variable genommen, die einen Nutzer die Suchanfrage abschicken lässt.

Die Autoren sind der Auffassung, dass es Mängel bei der Transformation des Informationsbedürfnisses gibt, welches der Nutzer zu der Anfrage, die er abschickt oder der Schlagwortsammlung, die dieses Bedürfnis beschreibt, im Kopf hat. Außerdem beantwortet eine Suchmaschine eine Frage des Nutzers nicht kontextgebunden oder subjektiv.

Um die wahrscheinliche Zugehörigkeit einer Suchanfrage zu einem Themengebiet zu berechnen, wurde der Inhalt der einzelnen Themengruppen (die zusammengefassten Worte, die das gleiche Ziel beschreiben) analysiert. So wurden direkte und indirekte Verbindungen zwischen Suchanfragen und Themen entdeckt. Durch diese Informationen konnte das Interesse des Nutzers ermittelt werden (vgl. BAEZA-YATES ET AL. 2006, S. 106).

Die übernommenen Themen konnten nicht alle mit genügend Suchanfragen gefüllt werden. Dafür fiel hier auf, dass Themen wie bspw. Cars oder Law fehlten.

Die Autoren fanden heraus, dass viele Suchanfragen zu den Themen Sex und Entertainment gehören (z.B.: Madonna erotic). Dies umschließt auch annähernd 73 Prozent der als mehrdeutig eingeordneten Suchanfragen (2.168 Suchanfragen).

76 Prozent der 2.719 als informationsorientiert identifizierten Suchanfragen kamen aus den Themen References, Education, Health, Computers, Society und Home.

Die nicht informationsorientierten Anfragen (1.155) gehörten größtenteils (fast 70 Prozent) in die Themen Computer, Entertainment, Society und Sex (vgl. BAEZA-YATES ET AL. 2006, S. 107).

Der größte Unterschied zwischen Suchanfragen, die zu dem mehrdeutigen Anfragecluster gehören und jenen im nicht informationsorientierten Cluster ist, dass die nicht informationsorientierten Anfragen direkt auf Fotos von Prominenten, Teile eines Computers oder Software-Downloads Bezug nehmen (vgl. BAEZA-YATES ET AL. 2006, S. 107).

Die weitere Analyse zeigt, dass sich das Interesse eines Nutzers identifizieren lässt, wenn man eine bestimmte Präsentation der Suchanfragen (die Juroren wussten, welche Dokumente die Suchanfrage beantwortet hatte und das Themenfeld war bekannt) mit der automatischen Klassifikation kombiniert.

In der automatischen Klassifikation waren jene Informationsbedürfnisse der Nutzer, die mit Themen wie Entertainment, Sex oder Business verbunden werden, leicht zu erkennen und zwischen diesen Themen wurden wichtige Verbindungen entdeckt, denn die meisten der Suchanfragen aus einem dieser Themenfelder können auch zu einem anderen Thema gehören. Aber nicht bei allen Themenfeldern ist es so einfach, das Informationsbedürfnis hinter den Suchanfragen zu erkennen. Dies liegt daran, dass die in den Suchanfragen und auf den Webseiten genutzten Wörter nicht die gleichen sind und oft mehrere unterschiedliche Themen beschreiben können. Nur elf der 18 übernommenen Themen wurden genutzt. Dafür fiel den Autoren auf, dass zwei Themenfelder (Cars und Law) fehlten. Aus

diesem Grund empfehlen die Autoren, die Themen neu zu wählen, auch um Überlappungen zu vermeiden (vgl. BAEZA-YATES ET AL. 2006, S. 108).

5. Suchverhalten von Suchmaschinennutzern

Schmidt-Mänz analysierte das Verhalten der Nutzer bei einer Web-Suche. Die Autorin möchte in ihrer Untersuchung das Suchverhalten generell darstellen, um einen Überblick zu bieten. Sie möchte Erkenntnisse über das Suchverhalten sammeln und Unterschiede bei Personen aus verschiedenen Berufssparten, bei Personen mit unterschiedlichen Wissensleveln, zwischen der Internetgeneration und anderen Gruppierungen und Unterschiede zwischen Männern und Frauen im Suchverhalten darstellen (vgl. SCHMIDT-MÄNZ 2007, S. 61 f.).

Sie führte zwei empirische Studien durch, um das tatsächliche Suchverhalten von Nutzern zu eruieren. In den Studien wurden Fragen nach der Orientierung der Suchenden und ihrem Informationsbedürfnis beantwortet (vgl. SCHMIDT-MÄNZ 2007, S. 5). Ziel der Untersuchungen war es, kundenfreundlichere Suchmaschinen zu entwickeln und den Betreibern von Webseiten durch die Aufdeckung des Informationsbedürfnisses die Optimierung ihrer Webseite zu ermöglichen, z.B. durch bessere/passendere Schlagwortvergabe. Außerdem sollte der Suchende selbst sein Verhalten besser kennenlernen, um es evtl. verbessern zu können (vgl. SCHMIDT-MÄNZ 2007, S. 6).

Zusätzlich zu den eigenen Studien stellte Schmidt-Mänz die Ergebnisse bereits durchgeführter Studien einander gegenüber. Zusammengefasst lässt sich aus den verschiedenen Studien der Schluss ziehen, dass Suchanfragen meistens sehr kurz sind (meistens aus zwei Wörtern bestehend). Modifikationen, Operatoren sowie natürliche Fragen tauchen in den Suchanfragen selten auf. Die Suchmaschinennutzer schauen sich nur sehr wenige Suchergebnisseiten an. In mehreren Studien wurde die Einordnung in Themengebiete vorgenommen, eine Veränderung dieser lässt sich allerdings nicht feststellen (vgl. SCHMIDT-MÄNZ 2007, S. 28).

Um diese Faktoren zu ermitteln, wird ein Fragebogen genutzt, der sich in fünf Blöcke gliedert:

1. Standardnutzung von Suchmaschinen (Häufigkeit, Suchverhalten)

2. Nutzungsweise spezieller Suchfunktionen und Suchstrategien (Komplexität der Suchanfragen (Nutzung von Operatoren), Suchstrategien)
3. Einschätzung von Suchdiensten (Wissen über Suchmaschinen, falsch/wahr Antworten, Störfaktoren in Suchmaschinen)
4. Navigation im Internet (beliebteste Einstiegspunkte ins Netz; Beliebtheit/Akzeptanz verschiedener Werbeformen)
5. Allgemeine Fragen (demografische Daten, Internetzugang, Vertrauen in verschiedene Medien)

(vgl. SCHMIDT-MÄNZ 2007, S. 63 f).

An dieser Umfrage nahmen Nutzer teil, die mit dem Internet vertraut waren (Online-Umfrage). Dementsprechend war die Nutzungshäufigkeit der Suchmaschinen sehr hoch. Das Durchschnittsalter war relativ niedrig und die weibliche Bevölkerung unterrepräsentiert. In dieser Umfrage ist eine hohe Nutzungshäufigkeit als positiv zu vermerken, da Vielnutzern Problematiken eher auffallen als Menschen, die nicht so oft online sind (vgl. SCHMIDT-MÄNZ 2007, S. 94).

Um die verschiedenen Nutzergruppen (Anfänger und Fortgeschrittene) zu untersuchen, wurden 45 Merkmale ausgewählt und einander gegenübergestellt (vgl. SCHMIDT-MÄNZ 2007, S. 97). Um herauszufinden, ob die Nutzer Anfänger oder Fortgeschrittene sind, wurden in der Umfrage Wissensfragen gestellt (vgl. SCHMIDT-MÄNZ 2007, S. 98).

Die Ergebnisse der Umfrage stimmen bei der Wortanzahl, die in Suchanfragen genutzt wird (kurze Anfragen) und den Erkenntnissen über den Suchmaschinenmarkt (eine Suchmaschine (google) dominiert) mit vorherigen Studien überein. In der Umfrage wurde von den Probanden oft nach Personen, Produkten oder fachlichen Inhalten gesucht. Die entsprechenden Quellen waren den Suchenden bekannt. Die Probanden empfanden (zu sehr) optimierte Seiten und doppelt auftauchende Links in den Suchergebnissen als störend. Außerdem stuften sich die Probanden selbst als weiter fortgeschrittene Internetnutzer ein als die Auswertung der Wissensfragen (s.o.) es zeigt (vgl. SCHMIDT-MÄNZ 2007, S. 95).

Die befragten Nutzer lehnen AdWords, gesponserte Links, Bannerwerbung oder Pop-ups stark ab (vgl. SCHMIDT-MÄNZ 2007, S. 96).

Die Ergebnisse zeigen, dass die Anzahl der Suchanfragen in einer Session (auch Neuformulierungen) nicht vom Wissenstand der Probanden abhängt.

Allerdings bewerten fortgeschrittene Nutzer gefundene Webseiten schneller und kehren somit auch schneller zur Suchergebnisseite zurück. Auch die Phrasensuche wird von fortgeschrittenen Nutzern häufiger verwendet (vgl. SCHMIDT-MÄNZ 2007, S. 99 f.). Suchexperten geben eine Suche nicht so schnell auf wie die anderen Nutzer und suchen zielgerichteter (vgl. SCHMIDT-MÄNZ 2007, S. 101). Experten kennen sich im Internet besser aus und nutzen bspw. Onlinenachrichtendienste oder setzen Bookmarks. Insgesamt nutzen sie das Internet effizienter für ihre Suche (vgl. SCHMIDT-MÄNZ 2007, S. 102 f.).

Der Vergleich zwischen Studenten und Angestellten zeigt, dass Studenten weniger Suchanfragen mit mehr Wörtern stellen und schneller ans Ziel kommen als Angestellte. Studenten setzen oft die Phrasensuche ein, was auf die Suche nach sehr speziellen Dokumenten wie bspw. Literatur hinweisen kann (vgl. SCHMIDT-MÄNZ 2007, S. 106 f.). Das Informationsbedürfnis von Studenten ist nicht so spezifisch wie das der Angestellten. Sie durchstöbern auch mal Nachrichten, die nichts mit der ursprünglichen Suche zu tun haben. Portale werden von Studenten eher abgelehnt, sie wissen aber, wo sie die gewünschten Informationen am ehesten finden. Die verschiedenen Werbeformen werden von beiden Gruppen abgelehnt (vgl. SCHMIDT-MÄNZ 2007, S. 108). Allgemein sind Studenten sicherer im Umgang mit dem Internet. Sie nutzen Online-Nachrichten und soziale Netzwerke, um auf neue Webseiten aufmerksam zu werden und fühlen sich von Werbung weniger gestört. Im Vergleich zu den Angestellten haben sie später begonnen, einen Onlinezugang zu nutzen (vgl. SCHMIDT-MÄNZ 2007, S. 108).

Alle Personen unter 29 Jahre wurden in dieser Untersuchung der Internetgeneration zugeordnet. Im Vergleich mit der Nicht-Internetgeneration haben sie weniger Interesse, Zeit in die Recherche zu investieren, kehren schneller zu den Ergebnisseiten zurück und brechen die Suche eher ab (vgl. SCHMIDT-MÄNZ 2007, S. 110 f.). Es lässt sich sagen, dass die junge Generation eher nach allgemeinen Themen wie Nachrichten usw. sucht, während die ältere Generation eher spezielle Suchanfragen wie bspw. Reiseziele stellt. Bei der Suche haben die Jungen häufig eine Strategie und nutzen die Phrasensuche öfter (vgl. SCHMIDT-MÄNZ 2007, S. 112). Um neue Webseiten zu finden, nutzt die Internetgeneration Communities, Links oder

Mundpropaganda von Bekannten in/aus sozialen Netzwerken. Diese sind bei der älteren Generation nicht verbreitet und kaum bekannt (vgl. SCHMIDT-MÄNZ 2007, S. 114).

Der Geschlechtervergleich zeigt, dass Männer das Internet und somit auch die Suchmaschinen häufiger nutzen als Frauen. Männer geben mehr Suchbegriffe in die Suchmaschine ein und kehren häufiger zu der Suchergebnisseite zurück, wenn sie das Gesuchte auf einer Webseite nicht finden. Die Suchanfragen von Männern sind komplexer und sie nutzen die Phrasensuche häufiger (vgl. SCHMIDT-MÄNZ 2007, S. 116 f.). Pop-ups und Werbemails werden von Frauen häufiger akzeptiert als von Männern (vgl. SCHMIDT-MÄNZ 2007, S. 117 f.).

Zusammenfassend lässt sich für alle Gruppen sagen, dass sich die Teilnehmer in solchen Umfragen oft als kompetenter in der Websuche darstellen als sie sind. Gestört fühlen sich die Suchenden vor allem von toten Links und Links, die auf Ebay verweisen. Optimierungsmaßnahmen von Webseiten(betreibern) sollten nicht nur auf die Suchmaschine ausgerichtet sein, sondern auch den Besucher zufriedenstellen (vgl. SCHMIDT-MÄNZ 2007, S. 118 f.).

Für die Analyse der Suchanfragen wurden diese über ein Jahr bei der Live-Suche von Fireball und Metaspinner abgefangen und in einer Datenbank gespeichert (vgl. SCHMIDT-MÄNZ 2007, S. 121).

Für die Untersuchung bekamen die Suchanfragen einen Zeitstempel, so dass sie in eine chronologische Reihenfolge gebracht werden konnten. Bevor sie in der Datenbank gespeichert wurden, wurden sie auf benutzte Operatoren (AND, OR und „“, nicht kleingeschrieben oder auf Deutsch) untersucht (vgl. SCHMIDT-MÄNZ 2007, S. 122). Die Identifikation eines Users durch Cookies oder IP-Adressen fehlt hier, da es sich um beobachtete Daten handelt, nicht um Logfiles (vgl. SCHMIDT-MÄNZ 2007, S. 123). Insgesamt wurden Suchanfragen aus vier Suchmaschinen (Fireball, Lycos, Metager, Metaspinner) untersucht (vgl. SCHMIDT-MÄNZ 2007, S. 127). Es wurde das Vorkommen einzelner Suchanfragen und Suchterme in einer bestimmten Zeit berechnet und innerhalb der vier Suchmaschinen verglichen. Bei den Metasuchmaschinen tauchen Terme dabei öfter auf als einzelne Suchanfragen (vgl. SCHMIDT-MÄNZ 2007, S. 128 f.).

Die Länge der Suchanfragen war ähnlich lang (1,6 Wörter in Metager; 1,7 Wörter in Lycos; 1,8 Worte in Fireball und Metaspinner). Im Vergleich zu internationalen Studien sind deutsche Suchanfragen kürzer, dies könnte aber an aneinandergereihten Substantiven liegen, bspw. Tannenbaum auf Deutsch und christmas tree auf Englisch (vgl. SCHMIDT-MÄNZ 2007, S. 129). Die Top-Ten-Suchanfragen in den Suchmaschinen wurden berechnet und es wurde markiert, welche Anfragen bei allen vier Suchmaschinen in den Top-Ten vertreten waren (vgl. SCHMIDT-MÄNZ 2007, S. 138). Gleiches geschah auch mit den Top-Ten-Suchtermen (vgl. SCHMIDT-MÄNZ 2007, S. 139), außerdem wurden natürliche Fragen in den Suchanfragen ausfindig gemacht. Dazu wurden Fragewörter (warum, was, wer usw.) in den Suchanfragen identifiziert und als Indikator für eine natürliche Frage genutzt. Der Anteil liegt bei Fireball, Lycos und Metager bei 0,1 Prozent, bei Metaspinner bei 0,2 Prozent (vgl. SCHMIDT-MÄNZ 2007, S. 141 f). Phrasensuchen wurde zwischen 2,13 und 2,46 Prozent durchgeführt (vgl. SCHMIDT-MÄNZ 2007, S. 142).

Schmidt-Mänz betrachtete außerdem, ob es Unterschiede zwischen den Wochentagen und den Suchanfragen gibt und zu welcher Tageszeit besonders viele Suchanfragen generiert werden (vgl. SCHMIDT-MÄNZ 2007, S. 145). Die meisten Suchanfragen werden zwischen 13 und 15 Uhr generiert, die wenigsten zwischen fünf und sechs Uhr (vgl. SCHMIDT-MÄNZ 2007, S. 146). Montags werden mehr Suchanfragen gestellt als an anderen Wochentagen. Vermutlich werden die am Wochenende gesammelten Informationen montags gesucht und verarbeitet (vgl. SCHMIDT-MÄNZ 2007, S. 147).

Über die Zeit betrachtet gibt es in den Top-Ten-Suchlisten keine Veränderungen, in kurzen Abständen (Woche/Tag) betrachtet aber schon (vgl. SCHMIDT-MÄNZ 2007, S. 148). Populäre oder „heiße“ Terme sind Terme, die nur in einer bestimmten vorher festgelegten Zeit (Stunden, Tage, Wochen) nachgefragt werden und danach aus den Suchlisten verschwinden (vgl. SCHMIDT-MÄNZ 2007, S. 148 f.). Für Suchmaschinenbetreiber ist es wichtig, wiederkehrende Terme zu kennen, um mit Empfehlungen oder angepassten Caches reagieren zu können (vgl. SCHMIDT-MÄNZ 2007, S. 149.). Eintagsfliegen sind Terme, die in nur ein oder zwei Zeitintervallen vorkommen, Dauerbrenner sind Terme, die in allen Zeitintervallen auftauchen.

Dazwischen befindet sich das Mittelfeld (vgl. SCHMIDT-MÄNZ 2007, S. 151). Die hier genutzten Ergebnisse stammen von Lycos und sind im Tagesintervall gewählt (vgl. SCHMIDT-MÄNZ 2007, S. 154). Bei 61 gewählten Termen entfallen die meisten Dauerbrenner (19 Terme) in den Bereich Erotik, elf sind für die Suche irrelevante Füllwörter (vgl. SCHMIDT-MÄNZ 2007, S. 155). Die Dauerbrenner-Terme aller untersuchten Suchmaschinen wurden verglichen. Aufgrund der unterschiedlichen Frequentierung der Suchmaschinen jedoch mit unterschiedlichen Parametern. Die Ergebnisse gleichen sich, wenn auch in Metasuchmaschinen differenzierter gesucht wird als in Generellen (vgl. SCHMIDT-MÄNZ 2007, S. 158).

Events sind Ereignisse mit festem Datum (bspw. Weihnachten) oder in regelmäßigen oder unregelmäßigen Abständen wiederkehrende Ereignisse. Impulse sind unvorhergesehene Ereignisse (bspw. Naturkatastrophen) (vgl. SCHMIDT-MÄNZ 2007, S. 160).

Der Vergleich zwischen den unterschiedlichen Gruppen zeigt, dass zwar alle Suchanfragen sehr kurz sind, es aber Unterschiede im Auffinden von neuen Webinhalten gibt (vgl. SCHMIDT-MÄNZ 2007, S. 181).

Durch die Kennzeichnung der Suchanfragen mit einem Zeitstempel wurde es möglich, Dauerbrenner zu identifizieren. Die Suchmaschinen könnten nun Caches mit Dauerbrennern aufbauen und so mehr Kapazität für die Berechnung der Relevanz neuer Webseiten/Suchanfragen zur Verfügung stellen. Desweiteren üben Nachrichten einen großen Einfluss auf die Suche aus. Mit der Identifikation von Events können Nachrichtenseiten besser und schneller auf unvorhergesehene Ereignisse reagieren und den Suchenden die Navigation erleichtern (vgl. SCHMIDT-MÄNZ 2007, S. 182).

Für die weitere Forschung empfiehlt Schmidt-Mänz, dass Suchterme unter Beachtung ihrer Eigenschaften untersucht werden sollten (Dauerbrenner, Event?) und sie sollten zeitreihenanalytisch unterschiedlich betrachtet werden (vgl. SCHMIDT-MÄNZ 2007, S. 183).

Die Suchmaschinen sollten, den Ergebnissen dieser Studie folgend, versuchen, unseriöse Ergebnisse aus der Ergebnisseite zu entfernen. Sie sollten Strategien entwickeln, um solche Seiten nicht mehr in den Trefferlisten zu haben und das Verfahren des Rankings sollte überwacht werden, so dass z.B. auch tote Links nicht mehr auftauchen (vgl. SCHMIDT-MÄNZ 2007, S. 183).

Webseitenbetreiber können Fachjargon verwenden, denn die Suchenden besitzen oft Hintergrundwissen zu den Themen, die sie suchen. Da die Suchenden oft Links folgen, sollte bei externen Links darauf geachtet werden, dass die Webseiten thematisch im selben Kontext stehen (vgl. SCHMIDT-MÄNZ 2007, S. 185 f.).

Mit Hilfe der Dauerbrenner, Impulse und Events können geeignete Vokabeln für die Gestaltung der Webseite gefunden werden. Damit kann die Sichtbarkeit der Webseite erhöht werden. Trends können früh erkannt werden und die Webseite entsprechend gestaltet werden. Auch können die Wörter bei den Suchmaschinen besser gebucht werden, um Werbung zu schalten (vgl. SCHMIDT-MÄNZ 2007, S. 186).

6. Einsatz qualitativer Methoden zu Erfassung der Nutzerbedürfnisse

Die Analysen von Weblogs haben den Nachteil, dass die affektiven, emotionalen und kognitiven Wahrnehmungsaspekte des Nutzers weder erfasst noch erkannt werden. Diese Mängel können durch den ergänzenden Einsatz von qualitativen Methoden wie bspw. Interviews oder Face-to-Face Befragungen ausgeglichen werden. Hier werden Informationen wie persönliche Angaben durch direktes Befragen der Teilnehmer gewonnen (vgl. RAINIE UND JANSEN 2009, S.39).

Zusätzlich können die Befragungen auch genutzt werden, um die Beziehungen zwischen unterschiedlichen statistischen Variablen zu analysieren (vgl. RAINIE UND JANSEN 2009, S. 40).

Nutzt man die Möglichkeit der elektronischen Befragung, kann man die Kosten relativ gering halten (vgl. RAINIE UND JANSEN 2009, S. 40).

6.1 Hargittai 2002

Viele Studien fokussieren sich auf einen Aspekt der Websuche und beziehen das Verhalten des Nutzers nicht mit ein. Um den gesamten Suchprozess und das Verhalten eines Nutzers während einer Suche zu untersuchen, führt Hargittai eine Studie durch, die die sonstige Mediennutzung,

die generelle Internetnutzung und demografische Daten der Nutzer einbezieht (vgl. HARGITTAI 2002, S. 1239).

Hargittai führt ihre Studie an einer Stichprobe der normalen Internetpopulation durch. Sie bemängelt, dass bei vorhergegangenen Studien nur eine bestimmte Nutzergruppe befragt wurde (meistens Studenten) (vgl. HARGITTAI 2002, S. 1239 f.). Um verschiedene Nutzergruppen für ihre Studie zu erreichen, schrieb sie zufällig ausgewählte Adressen aus einer Postdatenbank an. Hargittai erklärte ihr Vorhaben und die potenziellen Probanden bekamen eine Webseite genannt, auf der sie weitere Informationen finden konnten. Ein paar Tage nachdem die Briefe versandt wurden, wurden die Haushalte per Telefon kontaktiert. Aus diesen Haushalten wurde derjenige Internetnutzer (Person über 18 Jahre, die mindestens einmal im Monat online ist) mit dem nächsten Geburtstag ausgewählt (vgl. HARGITTAI 2002, S. 1240).

Den Probanden wurden 40 Dollar für ihre Teilnahme angeboten, die sie am Ende erhalten sollten. Die Untersuchung wurde auf dem Universitätsgelände durchgeführt. Ihre E-Mail Adresse wurde aufgenommen und vereinbart, wie der weitere Kontakt gewünscht wurde (E-Mail oder Post). Einen Tag vor der Studie wurde ein Erinnerungsanruf getätigt (vgl. HARGITTAI 2002, S. 1241).

Für die Untersuchung wurden Windows-PCs und Macintoshs benutzt, außerdem die drei populärsten Browser der USA (Internet Explorer, Netscape Communicator und America Online). Um die beobachtete Session aufzuzeichnen, wurde das Programm HyperCam auf den Windows-PCs und äquivalent SnapZPro auf den Macintoshs genutzt. Durch diese Programme wurden audiovisuelle Dateien erstellt, die die Aktivitäten auf dem Bildschirm und die verbalen Kommentare der Probanden zeigten. Zusätzlich wurde ein Programm eingesetzt, das die Browser-History löscht, so dass jeder Proband mit einem „sauberen“ Browser startete und nicht durch Aktionen anderer beeinflusst wurde. Die erstellten Dateien wurden passwortgeschützt auf den Servern der Universität gespeichert und jede Nacht gesichert, um einen Datenverlust zu vermeiden (vgl. HARGITTAI 2002, S. 1241).

Die Studie begann für die Probanden mit einem 20-minütigen Interview über die allgemeinen Internetgewohnheiten. Nach der beobachteten

Session füllten die Probanden einen Fragebogen aus (vgl. HARGITTAI 2002, S. 1241).

Die Daten, die auf diese Weise gesammelt wurden, enthielten Informationen über die Häufigkeit und den Ort der Internetnutzung, die Art der Seiten, die die Probanden besuchen, die Aktivitäten, die sie online durchführen, die Nutzung von anderen Medien und ihre Aktivitäten in sozialen Netzwerken (vgl. HARGITTAI 2002, S. 1241).

Den Probanden wurde eine Liste mit computer- und internetbezogenen Wörtern vorgelegt und sie wurden gebeten, diese zu erklären. Eine Liste von Multiple Choice-Fragen maß ihr Wissen über die computer- und internetbezogenen Wörter, diese wurden später in der Studie genutzt. Das Ziel war hier zu sehen, ob das Wissen mit dem Verhalten und den Aktionen im Internet korreliert (vgl. HARGITTAI 2002, S. 1241).

Die hier gewählte Methode beinhaltet, dass die Probanden an einen ihnen unbekanntem Ort (Universität) kommen und dort Computer benutzen, die nicht so eingerichtet sind, wie ihre eigenen. Diese Tatsachen können das Verhalten der Probanden beeinflussen (vgl. HARGITTAI 2002, S. 1241 f.).

Die Probanden bekamen in der beobachteten Session 17 Aufgaben gestellt und hatten jeweils fünf Minuten Zeit, diese zu bearbeiten (vgl. HARGITTAI 2002, S. 1242).

Das pure Aufzeichnen der Webseiten, die die Nutzer besuchen, zeigt den Forschern nicht, welche Informationen auf der jeweiligen Seite gesucht werden. Deshalb wurden die Probanden während der Suche gebeten, Kommentare abzugeben und die Forscher stellten Fragen zu den jeweiligen Aktivitäten (vgl. HARGITTAI 2002, S. 1242).

Um die aufgezeichneten Aktivitäten zu codieren, wurde eine von der Autorin entwickelte Liste genutzt. Mit Hilfe dieser Liste konnten Informationen darüber gewonnen werden, ob der Nutzer

- die Webseite direkt betreten hat (durch eintippen der URL),
- eine Suche in einer Suchmaschine durchgeführt hat,
- den Suchergebnissen einer Suche gefolgt ist,
- auf ein Kategorienverzeichnis oder einen Web-Guide geklickt hat,
- auf eine Werbung geklickt hat,
- auf einen Link auf einer Webseite geklickt hat.

Als zusätzliche Information wurde notiert, wenn der Nutzer Browserfeatures wie die Autovervollständigung, den „Zurück“-Button oder die Browserhistory nutzte oder die URL verkürzte, falls die gewünschte Webseite nicht zu finden war. Alle Suchterme und deren Ergebnisse wurden einzeln gezählt, egal ob sie bspw. gesponserte Links oder empfohlene Seiten waren. Die URL jeder besuchten Webseite wurde aufgezeichnet, genauso wie die exakte Zeit des Besuchs (vgl. HARGITTAI 2002, S. 1242). Zu jeder Aufgabe und jedem Probanden wurde außerdem notiert, ob die Aufgabe erfolgreich abgeschlossen wurde und wie lange der Proband dafür gebraucht hat (vgl. HARGITTAI 2002, S. 1243).

Die während der Suche aufgezeichneten Kommentare der Probanden wurden als zusätzliche Informationen über die gewöhnlichen (Such-)Strategien der Probanden genutzt.

Die geführten Interviews enthielten folgende Informationen über die Nutzer:

- ob die Nutzer soziale Netzwerke wie Foren nutzen, um dort Hilfe bei einem bestimmten Problem zu erhalten und Webseitenempfehlungen von anderen Nutzern zu bekommen
- wie die Nutzer eine Webseite kennenlernen (z.B. Werbung in anderen Medien oder Artikeln)
- wie sich Onlineerfahrungen für Nutzer anfühlen (z. B. Verwirrung, weil sie bestimmte Features nicht kennen oder Frustration bei Problemen mit Webseiten)
- was genau sich Nutzer anschauen, wenn sie die Suchergebnisse einer Suche sehen (z.B.: Wissen sie, wie man eine URL liest, benutzen sie diese als Hinweis auf das Suchergebnis?)
- ob die Nutzer wissen, wer eine Webseite eingestellt hat und ob sie vertrauenswürdig ist

(vgl. HARGITTAI 2002, S. 1243).

Die Ergebnisse fallen sehr unterschiedlich aus. Einige der Probanden benötigten 20 Minuten, um die 17 gestellten Aufgaben zu erledigen, andere 100 Minuten. Ein paar Probanden wussten kaum, dass es einen „Zurück“-Button gibt oder kannten keine anderen Features des Webrowsers.

Die Probanden, die bei einer Aufgabe mehrere Wörter in die Suchmaschine eingaben, kamen leichter zu einem Ergebnis (vgl. HARGITTAI 2002, S. 1243).

In dieser Studie hat sich gezeigt, dass sich nicht alle Internetnutzer gut mit dem Internet und mit Suchmaschinen auskennen. Es sollte ein Mix aus verschiedenen Methoden (z.B. Befragungsinstrumente mit Beobachtungen) genutzt werden, um nicht nur die Suchstrategien, sondern auch die sozialen Aspekte der Suche und deren Unterschiede betrachten zu können (vgl. HARGITTAI 2002, S. 1243).

6.2 Jansen und McNeese 2005

Jansen und McNeese untersuchen in ihrer Studie die Effektivität von automatischen Unterstützungen bei der Suche mit Suchmaschinen. Die kommerziellen Suchmaschinen verwenden bereits eine automatisierte Hilfestellung in Form der „Meinten Sie...“-Vorschläge auf der Suchergebnisseite (JANSEN UND MCNEESE 2005, S. 1481).

Jansen und McNeese nutzen zur Evaluation der automatischen Unterstützung ein System mit besagter Funktion und zum Vergleich eines ohne (vgl. JANSEN UND MCNEESE 2005, S. 1480).

In der Vergangenheit wurde versucht verschiedene Faktoren, wie die Verweildauer auf einer Homepage, Browsing-Muster oder Daten wie die Clickthrough- Daten als Indikator der Suchzufriedenheit zu nutzen (vgl. JANSEN UND MCNEESE 2005, S. 1481).

Jansen und McNeese konzentrierten sich in ihrer Arbeit auf die Fragen, ob automatisierte Assistenz den Suchprozess verbessert und wie die Nutzer mit der automatisierten Suchassistenz während der Suche agieren (vgl. JANSEN UND MCNEESE 2005, S. 1483).

Dazu wurde die Nutzung aus drei Blickwinkeln betrachtet:

1. Wie oft brauchen und implementieren die Nutzer die automatisierte Suchassistenz?
2. Wo im Suchprozess brauchen die Nutzer die automatisierte Suchassistenz?
3. Wo im Suchprozess implementieren die Nutzer die automatisierte Suchassistenz?

Diese Fragen wurden anhand der Aktionen (Nutzung der Suchassistentz) und dem Feedback der Nutzer beantwortet (vgl. JANSEN UND MCNEESE 2005, S. 1484). 40 Teilnehmer interagierten mit einem System und ihre Interaktionen wurden aufgezeichnet. Außerdem wurden die Teilnehmer während der Session gefilmt und dazu animiert, laut zu denken und ihre Absichten mitzuteilen. Des Weiteren schrieben die Forscher ein Protokoll der Suchsession (vgl. JANSEN UND MCNEESE 2005, S. 1484).

Für jede Suchanfrage wurden die Dokumente notiert, die der Nutzer für relevant gehalten hat. Dies geschah sowohl für die Suche mit der Suchassistentz als auch für die ohne. Die Dokumente, die doppelt auftauchten, wurden entfernt und die Leistung der beiden Systeme verglichen (vgl. JANSEN UND MCNEESE 2005, S. 1484).

Es gibt in dieser Studie Aktion-Objekt-Paare, an denen das System die durchzuführende Hilfestellung erkennt. Beispiele für dieses implizierte Feedback sind Paarungen wie copy-paste, save-document, print-document oder submit-query. Das System erstellt aus diesen Aktionen und den Objekten (z.B.: Webseiten, auf denen diese ausgeführt wurden) eine automatische Assistentz. Die Aktionen zeigen aber nicht, warum der Nutzer sie durchgeführt hat (vgl. JANSEN UND MCNEESE 2005, S. 1485).

Das automatische System wird bei fünf verschiedene Nutzeraktionen aktiviert (vgl. JANSEN UND MCNEESE 2005, S. 1485 f.).

Suchanfragen-Neuformulierung:

Die Applikation versucht hier, den Nutzern zu helfen, die Probleme mit den Ergebnissen haben. Deshalb werden die Ergebnisse mit Hilfe der durchgeführten Interaktionen und der Anzahl der Ergebnisse entweder reduziert oder aufgestockt. Dies geschieht, indem der Nutzer den AND oder PHRASE Operator wählen kann, um die Ergebnisse aufzustocken oder den OR Operator, um die Ergebnisse einzuschränken.

Suchanfragen-Verfeinerung:

Viele Suchende verfeinern ihre Anfragen nicht, auch nicht, wenn es andere Wörter gibt, die sich direkt auf ihr Informationsbedürfnis beziehen. Mit der Nutzung eines Thesaurus versucht das System, Synonyme für die Suchanfrage zu finden.

Relevanz-Feedback:

Das Feedback wird erstellt, wenn der Nutzer eine Aktion wie bspw. das Kopieren einer Passage durchführt. Das System nutzt Wörter aus dieser Passage, um dem Nutzer neue Suchwörter für weitere Anfragen vorzuschlagen.

Gleiche Anfragen:

Anhand früherer Suchanfragen mit den gleichen Wörtern, sucht das System aus einer Datenbank die Suchanfrage und die Anfragen, die dieser folgten und stellt dem Nutzer drei ähnliche Anfragen zur Auswahl für die weitere Recherche.

Rechtschreibung:

Falsch geschriebene Suchanfragen generieren häufig weniger Ergebnisse. Das System teilt die Suchanfrage in einzelne Wörter und kontrolliert deren Schreibweise, in dem es auf ein Online-Wörterbuch zurückgreift (vgl. JANSEN UND MCNEESE 2005, S. 1486).

Für die empirische Studie wurden Suchanfragen aus sechs Themenbereichen ausgewählt. Die 40 Probanden (Studenten, Library and Information Science) arbeiteten mit zwei identischen Systemen, wobei das eine die automatische Suchassistenten anbot und das andere nicht. Für die Studie wurden der Microsoft Thesaurus sowie die Wörterbücher in den Internet Explorer integriert (vgl. JANSEN UND MCNEESE 2005, S. 1486).

Die Probanden wurden nach ihrer Interneterfahrung, ihrem Alter, Geschlecht und weiteren demografischen Daten befragt (vgl. JANSEN UND MCNEESE 2005, S. 1487 f.). Die Studenten füllten während der Studie einen Fragebogen aus, der ihre Zufriedenheit mit der Problemlösung bzw. dem Suchergebnis darstellte. Die Fragen in diesem Fragebogen enthielten eine Sechs-Punkte-Skala, die von „Ich stimme voll zu“ bis zu „ich stimme gar nicht zu“ reichte. Damit wollten die Autoren herausfinden, wie die Probanden sich während der Tests fühlten und zwischen der Suche mit automatischer Suchassistenten und der ohne vergleichen können (vgl. JANSEN UND MCNEESE 2005, S. 1488).

Die Probanden durften 15 Minuten mit jedem System suchen. Dazu bekamen sie eines der sechs gewählten Themen sowie schriftliche Anweisungen zum Ausfüllen des Fragebogens. Sie hatten die Aufgabe, so

viele relevante Dokumente wie möglich zu finden. Waren sie mit dieser Aufgabe fertig, wechselten sie zu dem anderen System (mit Suchassistenten-Suchende wechselten zu dem System ohne und umgekehrt). Bei dem System mit der automatischen Suchassistenten wurden Hilfestellungen in Form von Hinweisen auf die verschiedenen Hilfestellungen (z.B. Buttons oder Vorschläge für neue Suchanfragen) gegeben (vgl. JANSEN UND MCNEESE 2005, S. 1489).

50 Prozent der Probanden fanden mit dem System mit automatischer Suchassistenten mehr relevante Dokumente als mit dem System ohne Assistenten. Allerdings waren die Dokumente nicht relevanter als die des Systems ohne Assistenten (vgl. JANSEN UND MCNEESE 2005, S. 1490).

Die Interaktionen der Nutzer mit dem System wurden anhand der Videoaufzeichnungen und der Transaktionslogs ausgewertet. Anhand verschiedener Kategorien (View Results, Selection, View Document etc.) wurden die Aktionen eruiert, die am häufigsten auftraten. Auch die Nutzung der Suchassistenten wurde in die Kategorien aufgenommen und so konnten Jansen und McNeese untersuchen, wie oft diese Unterstützung von den Teilnehmern genutzt wurde. Die am häufigsten getätigte Interaktion war das Ansehen der Ergebnisse, gefolgt von der Auswahl einer Webseite (vgl. JANSEN UND MCNEESE 2005, S. 1493). Die Anzahl der Aktionen mit der automatischen Suchassistenten stehen aber nicht im Zusammenhang mit der Anzahl der relevanten Ergebnisse, die der Suchende gefunden hat (vgl. JANSEN UND MCNEESE 2005, S. 1494).

Aus den getätigten Aktionen lassen sich Verhaltensmuster der Probanden ableiten. Offensichtlich leiten die Suchenden aus den Metadaten des Dokuments, die auf der Ergebnisseite angezeigt werden, ab, ob es relevant ist. Fast ein Viertel der Probanden (24,2 Prozent) prüfen die automatische Suchassistenten direkt nach der Eingabe der Suchanfrage. Dies legt nahe, dass die Probanden entweder unsicher bezüglich der eingegebenen Suchwörter oder interessiert an den Vorschlägen der Assistenten sind. Viele Probanden (24,2 Prozent) nutzten die Assistenten erst, nachdem sie sich ein (nicht relevantes) Dokument angesehen hatten (vgl. JANSEN UND MCNEESE 2005, S. 1495).

Die Nutzer, die unzufrieden mit dem Ergebnis waren, das sie sich angesehen haben, gingen häufig zurück auf die Ergebnisseite, schauten sich

weitere Ergebnisse an (scrollten auf der Ergebnisseite), bearbeiteten die Suchanfrage und schauten sich dann die Suchassistent an (vgl. JANSEN UND MCNEESE 2005, S. 1496).

Die Ergebnisse der Studie lassen darauf schließen, dass man die Interaktionen der Nutzer mit einem automatischen System vorhersagen und somit durch Forschung die Systeme, die eine automatisierte Suchassistent anbieten, weiter verbessern kann.

Für die weitere Forschung auf diesem Feld empfehlen die Autoren auch Faktoren wie das Scrollen auf der Ergebnisseite, das Scrollen auf Dokumenten und die Inhalte aus den Abstracts der Webseiten (gewonnen aus den Klickdaten) einzubeziehen (vgl. JANSEN UND MCNEESE 2005, S. 1499).

6.3 Otte und von Mach 2009

Eine Kombination aus qualitativen und quantitativen Methoden setzten von Mach und Otte 2009 ein und untersuchten einen Datensatz der Deutschen Telekom mit dem Fokus auf die Klassifizierung der Suchanfragen durch Juroren. Dazu nutzten sie die Klassifikation von Suchanfragen, die Broder 2002 erarbeitete, um die Suchanfragen verschiedenen Anfragetypen zuzuordnen. Hier wurden die bestehenden Kategorien navigationsorientiert und transaktionsorientiert um die Kategorie kommerziell und lokal ergänzt. Für Suchanfragen, die keiner dieser Kategorien zugeordnet werden konnten, wurde angenommen, dass sie dem Anfragetyp informationsorientiert zuzuordnen sind. Kommerzielle Anfragen zielten auf eine kostenpflichtige Handlung wie kaufen oder mieten ab, die allerdings nicht im Internet getätigt werden musste. Lokale Anfragen hatten einen offensichtlichen örtlichen Bezug zu einem Land, einer Region oder einer Stadt (vgl. OTTE UND VON MACH 2009, S. 12 und 17). Die Zufallsstichprobe von circa 50.000 Suchanfragen wurde von je zwei Juroren in die unterschiedlichen Anfragetypen eingeordnet. Die Juroren erhielten dafür eine Anleitung mit der Erklärung der verschiedenen Anfragetypen und einer negativ/positiv Beispielliste für die jeweiligen Anfragetypen (vgl. OTTE UND VON MACH 2009, S. 12).

In dieser Arbeit wurde untersucht, ob und wie viele Suchanfragen verschiedenen Anfragetypen zugeordnet wurden. Die Kombination kommerziell, lokal, navigationsorientiert und transaktionsorientiert trat dabei am häufigsten auf. Dies erscheint schlüssig, wenn man die dahinter liegen-

de Intention betrachtet. Der Suchende hat vermutlich die Intention einen Kauf zu tätigen. Dies möchte er in seiner Nähe tun und sucht dazu eine Webseite, um sich zu informieren (vgl. OTTE UND VON MACH 2009, S. 30). Um die Einordnung durch die Juroren zu prüfen, wurde eine Klickdatenanalyse durchgeführt.

Um eine valide Prüfung durchzuführen, wurden nur Suchanfragen der vorliegenden Stichprobe genutzt, die mehr als 100 Klicks erhalten hatten (vgl. OTTE UND VON MACH 2009, S. 31).

Die Thesen, die dabei zu Grunde lagen, waren:

1. Für navigationsorientierte Suchanfragen erfolgt ein Hauptteil der Klicks auf den ersten Treffer der Suchergebnisseite.
Dies bedeutet, dass für einen steigenden Anteil von Klicks auf den ersten Treffer einer Suchergebnisseite auch der Anteil der als navigationsorientiert eingeordneten Anfragen steigen sollte. Im Rückschluss heißt dies, dass der Klickanteil auf die erste Position für navigationsorientierte Suchanfragen im Vergleich zum Durchschnitt erhöht sein müsste.
2. Für kommerzielle Anfragen entfällt ein, im Vergleich zu den übrigen Anfragetypen, größerer Anteil der Klicks auf die bezahlten Treffer/Anzeigen.
Ist die These korrekt, müsste bei steigendem Klickanteil auf die Anzeigen auch die Einordnung in die Kategorie „kommerziell“ ansteigen. Im Rückschluss müssten für kommerziell eingeordnete Suchanfragen im Vergleich zum Durchschnitt mehr Klicks auf die Anzeigen entfallen (vgl. OTTE UND VON MACH 2009, S. 32 f.).

Die Analyse der Klickdaten zeigte, dass die Einordnung der Suchanfragen in die Anfragetypen kommerziell und navigationsorientiert zum größten Teil richtig durchgeführt wurde, allerdings wurden durch die Untersuchung der Klickdaten einige Fehler aufgedeckt (vgl. OTTE UND VON MACH 2009, S. 45). Die Nutzung der Klickdaten zur Überprüfung der Einordnung durch Juroren eignet sich, um mit Hilfe der Klickmuster Fehleranteile und somit falsch zugeordneten Anfragetypen zu identifizieren. Ebenso eignet sie sich, um Zuordnungen durch Juroren zu bekräftigen.

Allerdings geben die Autoren zu bedenken, dass auch mit dieser Methode nicht eindeutig auf die tatsächliche Intention des Suchenden geschlossen

werden kann, da der Suchende nicht selbst nach seinem Informationsbedürfnis befragt wurde (vgl. OTTE UND VON MACH 2009, S. 45).

Um dem Problem der Fremdeinordnung zu begegnen, wurde eine Online-Umfrage nach dem Vorbild Broders durchgeführt (siehe Kapitel 3.1 Broder 2002). Diese hatte zum Ziel, die Informationsbedürfnisse der Suchenden mit den Einordnungen der Juroren zu vergleichen (vgl. OTTE UND VON MACH 2009, S. 46).

In der Umfrage wurden nur diejenigen Suchenden befragt, die eine Suchanfrage abschickten, die auch in der kategorisierten Stichprobe vorkam. Die Befragung startete, sobald die Suchanfrage abgeschickt wurde (vgl. OTTE UND VON MACH 2009, S. 50).

Die Auswertung der Onlineumfrage und eine anschließende Plausibilitätsprüfung lassen die Autoren die folgenden Schlüsse ziehen:

- die Juroren waren größtenteils nicht in der Lage, den korrekten Anfragetyp aus der Suchanfrage abzuleiten. Dies wird durch die in der Klickdatenanalyse aufgedeckten Fehler unterstützt. Diese Annahme könnte darauf hindeuten, dass Menschen nicht in der Lage sind, den Anfragetyp einer fremden Suchanfrage zu erfassen, oder dass die Juroren in dieser Studie nicht ausreichend geschult wurden (vgl. OTTE UND VON MACH 2009, S. 59).
- Die Auswertung der Ergebnisse zeigt außerdem, dass die Methode der Online-Befragung mit den gewählten Fragestellungen nicht geeignet war. Die Widersprüche der Probanden bei gleichen Suchanfragen und der im Vergleich zu der Stichprobe stark abweichende Anteil der Anfragetypen lassen diesen Schluss zu.

Die Autoren tendieren schlussendlich zu der Empfehlung, eine Online-Umfrage in zukünftigen Studien mit einer weiteren Methode zu triangulieren (vgl. OTTE UND VON MACH 2009, S. 59).

Da in verschiedenen Studien darauf hingewiesen wurde, dass die Einteilung von Broder 2002 in drei Anfragetypen (navigationsorientiert, transaktionsorientiert und informationsorientiert) ergänzungswürdig ist (vgl. ROSE UND LEVINSON 2004, S. 19 und BRODER 2002, S. 8 f.), führten Otte und von Mach einen Versuch mit Probanden durch, in dem untersucht

werden sollte, ob es einen weiteren Ansatz für die Einordnung von Suchanfragen gibt (vgl. OTTE UND VON MACH 2009, S. 60).

Die Aufgabe wurde mit 75 zufällig ausgewählten Suchanfragen der Stichprobe durchgeführt. Die Probanden waren Studenten im vierten Semester (Library and Information Science), die zwar Kenntnisse in der Klassifizierung von Daten, aber keine theoretischen Kenntnisse der Funktion von Suchmaschinen hatten (vgl. OTTE UND VON MACH 2009, S. 61).

Das Ergebnis der Sortierung durch die Studenten zeigte, dass diese dazu tendieren, die Suchanfragen thematisch zu ordnen, selbst dann, wenn sie gebeten werden, dies nicht zu tun. Die Ergebnisse dieser Aufgabe konnten die Autoren nicht verwerten, sehen aber die Methode als einen guten Ansatz, neue Kategorien aufzudecken (vgl. OTTE UND VON MACH 2009, S. 65).

7. Zusammenfassung der Literaturanalyse

Die vorgestellten Studien werden hier noch einmal zusammengefasst, um aus den Ergebnissen weiterführende Studien abzuleiten und deren Aufbau zu beschreiben.

7.1 Evaluation des Informationsbedürfnisses eines Nutzers

Broder erarbeitete die in vielen Studien genutzte Klassifikation von Suchanfragen anhand einer Online-Umfrage und bat die Anfragesteller, ihre Intention anhand eines Fragebogens wiederzugeben. Zusätzlich wurden 400 Suchanfragen aus Logfiles den erarbeiteten Klassen zugeordnet. In diesem Teil wird die Methodik nicht erwähnt, es ist aber zu vermuten, dass schon in dieser Studie die Suchanfragen nicht von dem Originalanfragesteller zugeordnet wurden (vgl. BRODER 2002, S. 6 ff).

Rose und Levinson, die anhand der Session untersucht haben, wann und ob das Informationsbedürfnis eines Nutzers befriedigt wurde, verfolgen damit den Ansatz, die Intention hinter einer Suchanfrage anhand eines glickten Ergebnisses zu ermitteln. Allerdings fehlt auch hier die Überprüfung, ob die Vermutungen der Forscher richtig sind und der Suchende seine Suche wirklich erfolgreich beendet hat oder ob er bspw. die Suchmaschine gewechselt hat (vgl. ROSE UND LEVINSON 2004, S.19).

Konform dazu sehen Kang und Kim, dass hinter einem verwendeten Suchwort unterschiedliche Informationsbedürfnisse stehen können (vgl. KANG UND KIM 2003, S. 70). Sie ermitteln Strategien für die unterschiedlichen Intentionen der Suchenden. Ihr Vorschlag ist es, die Klassifikation von Suchanfragen nicht mehr anhand der Wörter vorzunehmen, sondern anhand der Daten und Klicks (oder anderer Aktionen der Nutzer) (vgl. KANG UND KIM 2003, S. 70).

Auch Rose und Levinson empfehlen in ihrer Arbeit, dass die Untersuchung von Sessions mit qualitativen Studien kombiniert werden sollte (vgl. ROSE UND LEVINSON 2004, S.19). Dabei sollten die Faktoren (Suchanfrage, Ergebnisse der Suchmaschine, geklicktes Ergebnis, weitere Suchen oder Aktionen des Nutzers), die sie in ihrer Studie mit einbeziehen, unbedingt mit aufgezeichnet werden, da sie ein gutes Bild der Suche und damit auch des Informationsbedürfnisses des einzelnen Probanden ergeben (vgl. ROSE UND LEVINSON 2004, S.16).

Calderon- Benavides et al. teilen die Suchanfragen in Facetten ein und arbeiten Schritt für Schritt heraus, welche Intention hinter der Anfrage stehen könnte. Allerdings wurden hier die Suchanfragen nicht von den „Originalsuchenden“ klassifiziert und der Vergleich zur „echten“ Intention fehlt. Damit lässt sich nicht zuverlässig überprüfen, ob erstens die Zuordnung zu den verschiedenen Facetten und zweitens die dadurch gefundene Intention hinter der Suchanfrage stimmt (vgl. CALDERON-BENAVIDES ET AL. 2010, S. 22 ff.).

Dai et al. arbeiteten in ihrer Studie, in der sie die kommerzielle Intention einer Suchanfrage untersuchten, mit Faktoren der Suchergebnisseite und der Webseiten, um kommerzielle Suchanfragen automatisch zu ermitteln. Zu diesen Faktoren gehören Schlagwörter auf den dazugehörigen Webseiten und in den Kurzbeschreibungen der Treffer auf der Suchergebnisseite (z.B.: Kaufen, Download) sowie die Schlagworte der Suchanfrage, die auf eine kommerzielle Intention hinweisen können (vgl. DAI ET AL. 2006, S. 4).

Lee et al. untersuchen die Möglichkeiten der automatischen Klassifikation anhand der Merkmale „Klickverhalten in der Vergangenheit“ und des „Vorkommens der Stichworte im Anchortext“. Diese Merkmale in Kombination bieten eine zu 90 Prozent richtige Einordnung der Suchanfragen, wobei

hier zu beachten ist, dass die Einordnung der Suchanfragen nicht durch den Originalanfragensteller vorgenommen wurde, sondern von Juroren, die die Intention vermuteten.

Um die automatische Identifikation zu überprüfen, wurden Suchanfragen von Juroren zugeordnet (vgl. DAI ET AL. 2006, S. 3). Die Nutzung der manuellen Einordnung der Suchanfragen begründen die Forscher mit der fehlenden Aussagekraft von kurzen Suchanfragen, die selten Schlagworte enthalten die z.B. eine kommerzielle Intention vermuten lassen (vgl. DAI ET AL. 2006, S. 2 und S. 5 sowie LEWANDOWSKI 2006, S.38 und SPINK ET AL. 2001, S. 14).

Für weitere Studien empfehlen Dai et al. allerdings die Einordnung durch Juroren zu vermeiden, um deren subjektive Einschätzung zu umgehen. Außerdem empfehlen sie, die Klickdaten der einzelnen Suchanfragen in die Untersuchung einzubeziehen und die gefundenen Faktoren zu ergänzen, um kommerzielle Suchanfragen zu erkennen (vgl. DAI ET AL. 2006, S. 8). Auch Lee et al. ließen die Suchanfragen von Juroren in die Anfragetypen navigationsorientiert und informationsorientiert einordnen. Sie arbeiteten mit einem sehr eingeschränkten Datenbeispiel und überprüften zusätzlich anhand eines Fragebogens, ob die Suchanfrage dem jeweiligen Probanden geläufig war (vgl. LEE ET AL. 2005, S. 393).

Die gewählten Beispieldaten sind allerdings beschränkt auf die Suchen des Departments Computerwissenschaften der Universität Los Angeles und bieten daher nicht die gleiche Breite wie die Anfragen der gesamten (Internet-)Population. Außerdem wurden Personen- und Softwareanfragen aus dem Datenset entfernt, was zu einer Verfälschung der Ergebnisse führt (vgl. LEE ET AL. 2005, S. 395 - 399).

Jansen et al. entwickelten einen Algorithmus, um Suchanfragen automatisch den von Broder 2002 entwickelten Kategorien zuzuordnen. Dafür entwickelten sie für die einzelnen Kategorien Unterklassen und stellten eine Liste mit Charakteristika auf. Dabei widersprachen sich die Wortlängen von navigationsorientierten Anfragen, die weniger als drei Wörter haben sollten und die informationsorientierten Anfragen, die mehr als zwei Worte haben sollen. Die weiteren Merkmale waren die aufgegliederten Beschreibungen, die schon Broder verwendete. Die Autoren kamen bei der automatischen

Klassifizierung auf andere Ergebnisse als bspw. Kang und Kim 2003 und begründen dies mit der größeren Stichprobe.

Auch in dieser Studie wurden zur Überprüfung der automatischen Zuordnung Juroren genutzt, die die Suchanfragen in die bekannten Kategorien einordneten. Jansen et al. erwähnen das Problem, das bei einer solchen Aufgabe entsteht: Man kann nicht sicher sein, dass diese Einordnung korrekt ist und nicht jede Suchanfrage passt nur in eine Kategorie (vgl. JANSEN ET AL. 2007, S. 1262).

7.2 Klassifikation der Suchanfragen nach Thema

In den Studien, die sich mit der Zuordnung von Suchanfragen zu verschiedenen Themen beschäftigen, wurde herausgefunden, dass grundsätzlich alle Themen bedient werden müssen bzw. gesucht werden. Sowohl Lewandowski 2006 als auch Jansen und Spink 2006 ließen die Suchanfragen von zwei unabhängigen Juroren den Themen zuordnen. Waren diese unterschiedlicher Meinung, wurde durch Rücksprache ein gemeinsames Thema gefunden (vgl. JANSEN UND SPINK 2006, S. 258 f. und LEWANDOWSKI 2006, S. 35).

Spink et al. ließen bereits 2001 2.414 Suchanfragen von Studenten zuordnen. Diese entwickelten eine eigene Themenklassifikation, von der in dem Artikel nur die oberste Ebene erwähnt wird (vgl. SPINK ET AL. 2001, S. 13). In dieser Untersuchung wurden die Themengebiete „Entertainment, recreation“ und „Sex, pornography, preferences“ am häufigsten nachgefragt, wogegen „Government“ und „Unknown“ am wenigsten verlangt wurden (vgl. SPINK ET AL. 2001, S. 13).

Fünf Jahre später wurden die Themenfelder Menschen, Plätze oder Dinge (AlltheWeb und AltaVista) und Handel, Reise, Arbeit oder Wirtschaft (Excite) (vgl. JANSEN UND SPINK 2006, S. 258 f.) am häufigsten nachgefragt. Lewandowski 2006 und Baeza-Yates et al. 2006 ordneten die Suchanfragen sowohl Themen als auch Anfragetypen zu, um herauszufinden, ob ein Thema in einem Anfragetyp besonders häufig vorkommt. Lewandowski ließ die Suchanfragen von zwei Juroren zuordnen, die gleichzeitig auch den Anfragetyp bestimmten. Er kommt zu dem Schluss, dass alle von Broder definierten Anfragetypen bedient werden müssen und dass bei bestimmten Themen in den unterschiedlichen Anfragetypen Hilfestellungen in

Form von automatisch generierten Hinweisen bspw. auf Datenbanken oder Downloadverzeichnisse gegeben werden können (vgl. LEWANDOWSKI 2006, S.43).

Baeza-Yates et al. ordneten die Suchanfragen den Anfragetypen informationsorientiert, nicht informationsorientiert und mehrdeutig zu. Zusätzlich wurden die Anfragen insgesamt 18 Themengebieten zugewiesen (vgl. BAEZA-YATES ET AL. 2006, S. 99 f.). Zur Klassifikation der Suchanfragen wurde das Support Vector Machines Model genutzt, eine Methode, die sich bei der Klassifikation von Texten bewährt hatte. Desweiteren bedienten sich die Autoren der Probabilistic Latent Semantic Analysis (PLSA) Methode, die bereits in der Vergangenheit genutzt wurde, um Nutzerinteressen zu erforschen (vgl. BAEZA-YATES ET AL. 2006, S. 100 f.). Ein Teil der Suchanfragen wurde von Juroren den Themen und den Anfragetypen zugeordnet (vgl. BAEZA-YATES ET AL. 2006, S. 101). Diese manuell zugeordneten Anfragen dienten als Grundlage für die automatische Klassifikation (vgl. BAEZA-YATES ET AL. 2006, S. 104).

Die ermittelten Themen wurden als Grundlage für die Einordnung in Cluster genommen, aus denen die Interessen der Nutzer abgeleitet wurden. Die Autoren gingen davon aus, dass die Nutzer Probleme haben, das Ziel, das sie im Kopf haben, in einer Anfrage zu formulieren, weshalb sie sich für die Bildung von Themen-Clustern entschieden (vgl. BAEZA-YATES ET AL. 2006, S. 106).

Für die automatische Klassifikation stellten die Autoren fest, dass einige Themenfelder wiederkehrende Schlagwörter haben, an denen sich die Themen feststellen lassen. Allerdings stellten sich die übernommenen Themen als Defizit heraus, da nicht alle Themenfelder mit Suchanfragen gefüllt werden konnten, gleichzeitig den Autoren aber zwei Themen (Cars und Law) fehlten (vgl. BAEZA-YATES ET AL. 2006, S. 108).

7.3 Einsatz qualitativer Methoden zur Erfassung der Nutzerbedürfnisse

Die Studien, die ergänzende qualitative Untersuchungen beinhalten, arbeiten mit weniger Suchanfragen und weniger Probanden, beschäftigen sich dafür aber intensiver mit den einzelnen Faktoren.

Otte und von Mach nutzen in ihrer Studie sowohl quantitative als auch qualitative Methoden, um eine Datenstichprobe zu untersuchen. Durch die intensive Untersuchung der Daten, unter anderem durch die Nutzung von Klickdaten, konnten Defizite in der Einordnung durch Juroren herausgefunden werden (vgl. OTTE UND VON MACH 2009, S. 7 ff.). Neben der Überprüfung der Einordnung durch Juroren, also die Klassifizierung fremder Anfragen, war die Findung einer neuen oder anderen Klassifizierung der Suchanfragen ein Gegenstand der Untersuchung. Dieser Ansatz war hier noch nicht ausgereift und brachte nicht die erwarteten Ergebnisse (vgl. OTTE UND VON MACH 2009, S. 65).

Insgesamt stellen die Autoren aber fest, dass die Ausgangsfrage, ob Juroren fremde Suchanfragen bezüglich ihrer Intention bewerten können, unbeantwortet bleibt und empfehlen, die Methodik zu verfeinern, um zu schlüssigen Ergebnissen zu kommen (vgl. OTTE UND VON MACH 2009, S. 66).

Jansen und McNeese führten ihre Studie mit 40 Probanden durch. Diese sollten die Suchen mit einem System mit automatisierter Suchassistenz und einem System ohne automatisierte Suchassistenz durchführen und vergleichen. Neben den Auswertungen der relevanten Dokumente, die dabei in den beiden Systemen gefunden wurden, waren die Forscher auch daran interessiert, wie die Suchenden die Zufriedenheit mit den unterschiedlichen Systemen empfanden. Sie ließen die Probanden einen Fragebogen ausfüllen, der den Grad der Zufriedenheit aufzeigte (vgl. JANSEN UND MCNEESE 2005, S. 1488).

Um sich nicht auf subjektive Eindrücke verlassen zu müssen, wurden die Suchsessions der Probanden auf Video aufgenommen. Anschließend wurden die getätigten Aktionen der Probanden herausgearbeitet und analysiert (vgl. JANSEN UND MCNEESE 2005, S. 1493).

Durch diese Art der Forschung erfuhren die Autoren mehr über die Art, wie die Nutzer suchen und wie sie mit (neuen) Features umgehen.

Mit der Einbeziehung von qualitativen Methoden lassen sich Muster in der Interaktion von Nutzern ermitteln und die automatische Suchassistenten kann so verbessert werden. Allerdings sollten dafür weitere Faktoren wie das Scrollen auf der Ergebnisseite, auf Dokumenten und Inhalte aus den Abstracts der Webseiten (gewonnen aus Klickdaten) mit einbezogen werden (vgl. JANSEN UND MCNEESE 2005, S. 1499).

Hargittais Weg, Probanden für ihre Studie zu gewinnen, war aufwändig, aber sie bekam so Zugang zu zufällig ausgewählten Personen. Zumeist befragen Forscher wegen des einfacheren Zugangs Studenten und erhalten so nur Ergebnisse, die auf diese Gruppe übertragbar sind.

Sie nutzte eine andere Art der Aufzeichnung, die zwar die Emotionen in den Gesichtern der Probanden nicht zeigt, dafür aber die Bewegungen am Bildschirm und die zugehörigen verbalen Kommentare der Probanden aufzeichnet (vgl. HARGITTAI 2002, S. 1242).

Die Probanden wurden zusätzlich in einem Interview und über einen Fragebogen zu ihren Internetgewohnheiten befragt, was Rückschlüsse auf ihren Medienkonsum und die damit verbundene Versiertheit zuließ.

Hargittai stellte fest, dass sich die Kenntnisse der Probanden enorm unterscheiden. Sie empfiehlt, weiterhin verschiedene Methoden zu mixen und in weiteren Studien auch die Unterschiede in der Internetkompetenz der Probanden zu erfassen (vgl. HARGITTAI 2002, S. 1243).

8. Grundlagen für die Entwicklung neuer Studien

Die bisher durchgeführten Studien zeigen, dass Bedarf an weiteren Forschungen im Bereich des Informationsbedürfnisses eines Nutzers hinter einer Suchanfrage besteht und die Erforschung der Suchmaschinennutzer, ihres Verhaltens und ihrer Intention hinter der eingegebenen Suchanfrage längst noch nicht abgeschlossen ist.

Bei der Interaktion der Nutzer mit einer Suchmaschine werden Daten produziert, die über die gewünschten Suchergebnisse der Nutzer zu einer bestimmten Suchanfrage Aufschluss geben können.

Dabei kann es sein, dass die Bewertung der Suchergebnisse je nach Nutzer unterschiedlich ausfällt.

Da ein Nutzer mit der Teilnahme an einer Studie einen Teil seiner persönlichen Verhaltensweisen preisgibt, wird im Folgenden beschrieben, wie sich das Verhalten von Probanden während einer Studie verändern kann und welchen Datenschutzbedenken es geben kann.

Anschließend werden Ideen für neue Forschungsvorhaben, die anhand der analysierten Literatur erarbeitet wurden, vorgestellt.

Aus diesen resultiert eine Anforderungsaufstellung, die darstellt, welche Daten für diese Art der Forschung aus den Logfiles der Suchmaschinen benötigt werden.

8.1 Einflüsse auf Probanden

Die Daten, die aus den Logfiles gewonnen werden, stammen von „normalen“ Menschen, die in alltäglichen Situationen Suchmaschinen nutzen.

Während der Durchführung einer Studie kann allerdings das Verhalten der Probanden beeinflusst werden. Die Beeinflussung kann durch interne Faktoren wie bspw. emotionale oder fachliche Einflüsse oder externe Faktoren wie Umgebung oder Situation verursacht werden (vgl. JANSEN ET AL. 2009, S. 3). Wissenschaftler versuchen daher, in den Studien die Umgebung zu berücksichtigen, in welcher die Daten gesammelt wurden (z.B. Aufzeichnung der Daten in privatem Umfeld oder in einem Labor?) und das Subjekt (Aktionen und Aspekte, z.B. Mimik des Subjekts) und das Verhalten (beobachtbare Ereignisse) zu erfassen (vgl. JANSEN ET AL. 2009, S. 4).

Egal in welcher Weise die Daten erhoben wurden, der Forscher sollte sie immer auf die Glaubwürdigkeit, die Validität und die Zuverlässigkeit überprüfen. Dies bedeutet eine Prüfung der Vertrauenswürdigkeit der Daten, wie aktuell und damit statistisch valide die Datensammlung ist und ob die gleichen Variablen auch mit den gleichen Messverfahren gemessen werden (können) (vgl. JANSEN ET AL. 2009, S. 6 f.).

8.2 Persönlichkeitsschutz der Probanden

Abgesehen von den gesetzlichen Regelungen, die den Datenschutz und die Privatsphäre von Teilnehmern an Studien regeln, kann es für die Forscher von enormem Vorteil sein, die Probanden über die Nutzung der gesammelten Daten aufzuklären. Besonders wichtig wird dies, wenn für eine Studie andere oder mehr Daten benötigt werden, als normalerweise in Logdateien aufgezeichnet werden (vgl. HAWKEY 2009, S. 80).

Die Informationen, die in Logdateien sichtbar sein können, sind fest mit der Interaktion des jeweiligen Nutzers am Webbrowser verbunden. Das Internet ist mehr und mehr zu einem Mittel geworden, sich als Nutzer mit seinen emotionalen Bedürfnissen zu beschäftigen und bspw. Hilfsforen zu besuchen, zu bloggen oder sich über Gesundheitsbelange zu informieren. So können die in Logdateien abgefangenen Daten neben neutralen, situationsgebundenen Begriffen wie Wetterinformationen auch sensible Daten wie bspw. die Suche nach sozial unerwünschten Inhalten (z.B. Pornografie), vertrauliche Geschäftsdaten oder private Aktivitäten während der Arbeitszeit enthalten (vgl. HAWKEY 2009, S. 81). Daher bestehen Bedenken bezüglich der Einhaltung von Privatsphäre und Datenschutz auf Seiten der Nutzer.

Bei der Aufzeichnung der Daten werden die Geräte-/Browsereigenschaften so wie „Bequemlichkeitseinstellungen“ des Browsers wie die Browser-history und Autovervollständigungen aufgezeichnet. Die meisten Teilnehmer berichten, dass sie versuchen, ihre Spuren im Internet mit Einstellungen im Browser unsichtbar zu machen, wenn sie darauf aufmerksam gemacht werden, dass ihre Daten aufgezeichnet werden. Sie nutzen z.B. Proxyserver oder andere Anonymisierungen (vgl. HAWKEY 2009, S. 81). Onlinedatenschutz beschäftigt sich generell mit den Belangen, die beim Transfer von persönlichen Daten zu geschäftlichen oder behördlichen Einrichtungen entstehen, also der (Online-)Beziehung zwischen Konsument und Institution. Dies sind andere Bedenken als jene, die mit den angesehenen Spuren der Webbrowseraktivitäten, wie im Fall der aufgezeichneten Daten im wissenschaftlichen Kontext, aufkommen (vgl. HAWKEY 2009, S. 81).

Die Feldforschung ermöglicht das Studium von realem Verhalten in einem realistischen Umfeld. Dass die Teilnehmer dabei beobachtet oder ihre Daten aufgezeichnet werden, kann sie dazu verleiten, ihr Verhalten zu ändern (vgl. HAWKEY 2009, S. 81 f.).

Sollte bei einer Studie Aufzeichnungssoftware auf den Rechnern der Probanden installiert werden, kann dies zusätzliche Bedenken verursachen, da die Teilnehmer persönliche Daten wie Nutzernamen und Passwörter nicht preisgeben möchten (vgl. HAWKEY 2009, S. 82).

Indem man versucht, die Daten der Probanden zu schützen, kann man die Bereitschaft der Teilnehmer, ihr natürliches Webbrowsingverhalten beizubehalten, erhöhen. Welche Methoden zur Verringerung der Bedenken seitens der Teilnehmer gewählt werden sollten, hängt von der Forschungsfrage und der Umgebung (privat oder Labor) ab (vgl. HAWKEY 2009, S. 82). Die Bedenken der Teilnehmer und wie viele Daten sie preisgeben möchten, ist bei jedem unterschiedlich und die Privatsphäre bei jedem individuell definiert. Ein Mensch kann den individuellen Schutz seiner Daten unterschiedlich definieren und sozusagen in verschiedenen Rollen auftreten. Als Privatperson können ihm seine Daten wichtiger und schützenswerter erscheinen als bspw. als Repräsentant einer Firma (vgl. HAWKEY 2009, S. 82).

Die Einstellung zum Datenschutz kann von Alter und Geschlecht beeinflusst werden. Die Lebenserfahrung, die technische Versiertheit oder das technische Verständnis wirken sich ebenfalls auf die Einstellung zum Datenschutz aus. Situationsbezogen können sich die Einstellung zur Privatsphäre und die Bereitschaft, Daten preiszugeben, ändern (vgl. HAWKEY 2009, S. 83).

Dies hängt auch damit zusammen, welche Beziehung die Probanden zu dem Empfänger der Informationen haben und welche Informationen mitgeteilt werden sollen. Bspw. ist das Wissen, dass während der Arbeitszeit nicht arbeitsbezogene Websites besucht werden oder dass der Proband Webseiten mit erotischen Inhalten besucht, unangenehmer für die Probanden als das Teilen von Kontakten oder anderen verfügbaren Daten. Je mehr Kontrolle der Proband über die Daten hat, die er teilen möchte, desto wohler fühlt er sich und ist dementsprechend eher bereit, die benötigten Daten aufnehmen zu lassen. Seine Bedenken können weiter reduziert

werden, wenn er weiß, wie und wofür seine Daten verwendet werden (vgl. HAWKEY 2009, S. 83).

Die staatlichen und organisatorischen Regelungen, die den Datenschutz betreffen, müssen bei jeder Studie eingehalten werden. Dazu gehören die Vorschriften für das Sammeln der Daten, für die Speicherung und die Sicherstellung der Teilnehmerdaten, ebenso wie die Nutzung, die Wiederverwendung und der Transfer der Daten (vgl. HAWKEY 2009, S. 85). Der Forscher muss sicherstellen, dass die gesammelten Daten keine Hinweise enthalten, die auf die Identität der Teilnehmer schließen lassen könnten, wie detaillierte demografische Daten oder IP-Adressen. Falls die Daten extern aufgezeichnet wurden und zum Forscher geschickt werden müssen, muss bei der Übermittlung der Daten sichergestellt sein, dass keine Verletzung der Datenschutzrichtlinien stattfindet (vgl. HAWKEY 2009, S. 85). Die Forscher sollten die Möglichkeit haben, die Daten zu anonymisieren bevor sie sie erhalten. Wenn die Daten clientseitig gesammelt werden, können diejenigen, die den Teilnehmer identifizieren könnten (z.B. die IP-Adresse) durch ein Skript gelöscht und willkürliche ID-Nummern vergeben werden. Wenn die Forschungsfrage es erfordert, können die IP-Adressen durch Bezeichnungen wie „zu Hause“, „Arbeit“ oder „Schule“ ersetzt werden (vgl. HAWKEY 2009, S. 86).

Während der Planung einer Studie kann sich die Frage ergeben, wo die Daten gespeichert werden sollen. Werden sie während eines Labor-experiments oder durch Software vom Proxyserver aufgezeichnet, werden die Daten normalerweise direkt auf einem Computer des Forschers gespeichert. Dabei sollte darauf geachtet werden, eine sichere Verbindung zu nutzen. Schwieriger wird es, wenn die Daten clientseitig gespeichert werden. Dann müssen sichere Wege für den Datentransfer und die Speicherung gefunden werden (vgl. HAWKEY 2009, S. 86).

Die Speicherung der Daten auf dem Computer der Teilnehmer wird durch eine delokalisierte Deinstallation der Aufzeichnungssoftware nach der Studie erleichtert. Die Gefahr, dass die Privatsphäre durch den Transfer verletzt wird, wird minimiert und der Teilnehmer muss nichts tun, um die Daten zu übermitteln. Aber die Forscher laufen Gefahr, die Daten zu ver-

lieren, wenn bspw. das Betriebssystem des Teilnehmers instabil wird. Sollten sich die Forscher entscheiden, die Daten auf dem PC des Teilnehmers öfter übermitteln zu lassen, könnte dies dem Probanden lästig werden. Außerdem müsste eine sichere Form der Übermittlung gefunden werden (vgl. HAWKEY 2009, S. 86).

Soll das natürliche Webbrowsingverhalten von Probanden aufgezeichnet werden, ist zu beachten, dass das Bewusstsein der Aufzeichnung das Verhalten der Probanden im Web verändern kann und sie sich selbst kontrollieren, indem sie bestimmte Webseiten während der Studie nicht besuchen. Für die Studien ist es jedoch wichtig, dass die Probanden ihren normalen Webaktivitäten folgen und sich ungeachtet von der sozialen Erwünschtheit oder der persönlichen Daten benehmen wie immer (vgl. HAWKEY 2009, S. 86).

Der Forscher muss versuchen, einen Kompromiss zwischen der Sammlung von ergiebigen und kontextabhängigen Daten und der Wahrung der Privatsphäre der Probanden zu finden. Er muss also versuchen, so viel Kontrolle zu behalten, dass er die benötigten Daten bekommt und dem Probanden gleichzeitig gewisse Freiheiten lassen, um nicht zu zudringlich zu werden (vgl. HAWKEY 2009, S. 86).

Ein Ansatz, um dieses Problem zu umgehen ist die Nutzung von archivierten Aufnahmen, die zu anderen Zwecken aufgezeichnet wurden oder die Nutzung von versehentlich hinterlassen Spuren, wie Daten von Nutzern, die durch serverseitige Aufzeichnungen erstellt wurden (vgl. HAWKEY 2009, S. 86 f.). Wenn erfasste Daten nachträglich genutzt werden, gibt es keine Verhaltensänderungen bei den Probanden, während sie beobachtet werden. Dieser Ansatz hat allerdings den Nachteil, dass die Daten für den aktuellen Forschungsansatz nicht geeignet sein können. Bei archivierten Daten besteht ebenfalls das Problem, dass der Proband sein Verhalten geändert hat, sobald ihm bewusst wurde, dass seine Daten aufgezeichnet wurden (vgl. HAWKEY 2009, S. 87). Außerdem wird bei den serverseitig gesammelten Daten die Breite der Daten begrenzt. Es wird nur der Eintritt auf eine bestimmte Seite oder über ein bestimmtes Portal aufgezeichnet. Meistens enthalten die Daten nicht mehr als die IP-Adresse des Nutzers,

einen Zeitstempel und die URL, die angewählt wurde. Webserverlogs können unvollständig sein. Wenn die vom Teilnehmer verlangte Seite im Cache liegt oder vom Proxyserver aufgezeichnet wurde, wird sie nicht empfangen und nicht aufgezeichnet (vgl. HAWKEY 2009, S. 87).

Die Datenschutzbedürfnisse der Nutzer und die Ansprüche der Webseiten, Daten über ihre Nutzer zu sammeln, widersprechen sich. Die Nutzung von Cookies verringert Probleme bei der Identifikation zurückkehrender Individuen. Nutzer können die Cookies aber auch ausschalten, um ihre Privatsphäre zu schützen. Zusätzlich gibt es dynamisch zugewiesene IP-Adressen, die verhindern, dass ein Nutzer immer die gleiche IP-Adresse hat. Die Internetnutzer können Proxyserver nutzen, die vielen Nutzern die gleiche IP-Adresse zuweisen und so die Identifikation von individuellen Nutzern schwierig macht (vgl. HAWKEY 2009, S. 87). Dafür können die Nutzer aber mit ihrem gewohnten Browser surfen, wobei die im Browser getätigten Aktionen nicht aufgezeichnet werden können. Durch die Nutzung von Javascript lassen sich zumindest Aktionen der Nutzer wie die Bewegungen der Maus, Scrollen und Tastenanschläge ermitteln (vgl. HAWKEY 2009, S. 88).

Der Vorteil dieser Methode gegenüber den Laborexperimenten ist, dass die Teilnehmer Zugang zu ihren üblichen Webtools sowie Browsern haben und sich in ihrer gewohnten physischen Umgebung befinden. Diese beinhaltet in diesem Fall auch die gewohnten Einstellungen in seinem Browser. Diese sollten auch bei clientseitigem Aufzeichnen nicht verändert werden. Da bei dieser Methode häufig Daten gesammelt werden, die zur Erstellung von Aktivitätsmustern dienen, sollte die datensammelnde Software den natürlichen Fluss der Aktivitäten während des Browsens nicht unterbrechen (vgl. HAWKEY 2009, S. 88).

Die Installation von Software auf den Computern der Teilnehmer erfordert mehr Beachtung des Datenschutzes. Es kann schwierig werden, die Software auf den Computern zu installieren und es sollte darauf geachtet werden, dass die Aufzeichnungssoftware robust ist und gut getestet wurde. Wenn der zur Aufzeichnung genutzte Computer nicht dem Teilnehmer gehört, sondern bspw. seinem Arbeitgeber, kann es schwierig sein, die Er-

laubnis für die Installation zu bekommen. Außerdem kann der beobachtete Arbeitnehmer seine Aktivitäten selbst regulieren, wenn er denkt, dass der Arbeitgeber Zugang zu den aufgezeichneten Daten hat (vgl. HAWKEY 2009, S. 88).

Eine Übersicht über die Vor- und Nachteile der verschiedenen Methoden zur Datensammlung bietet Tabelle 1.

	Serverseitiges Aufzeichnen	Proxy-Aufzeichnen	Clientseitiges Aufzeichnen
Reichhaltigkeit der Daten	Limitiert auf Navigation, Dateneingabe auf der Seite	Limitiert auf Navigation, einige Formulardaten – verbessert durch Scripts	Ergiebige Daten, inkl. Navigation, Tastenanschlag, Browser-Interaktion.
Vollständigkeit der Daten	Caching-Probleme, Webseiten abhängig	Caching-Probleme, die aber umgangen werden können	Können umgangen werden, indem ein anderer Browser oder PC genutzt wird
Möglichkeit, die individuellen Nutzer/Teilnehmer zu unterscheiden	Kann schwierig sein (Anonymisierungsservice)	Gut (muss sich einloggen)	Gut
Natürlichkeit der Teilnehmer und der Umgebung während des Browsens	Komplett natürlich	Teilnehmer weiß von Aufzeichnung, aber die Browserumgebung ist unverändert	Teilnehmer wissen von Installation. Umgebung hängt von der Software ab. (Hilfe (tools)generell transparent, aber der gewöhnliche Browser könnte nicht die üblichen Funktionen und Einstellungen haben)

Tabelle 1: Übersicht über die Vor- und Nachteile der verschiedenen Aufzeichnungsmethoden, eigene Darst. nach HAWKEY 2009, S. 89

Einige der Bedenken der Teilnehmer von Studien entstehen durch die Verknüpfung von Daten mit den spezifischen Nutzern sowie deren persönlichen Informationen. Allerdings sind die Daten primär auf Navigation mit den Webseiten und den Daten über den Eintritt auf eine Webseite (Verlinkung, Verweildauer usw.) beschränkt. Weblogs können die Daten von gecachten oder auf dem Server hinterlegten Webseiten aufzeichnen (vgl. HAWKEY 2009, S. 88).

Clientseitiges Aufzeichnen kann ergiebiger Daten liefern, die Datensammlung greift allerdings auch mehr in die Privatsphäre ein. Es muss Software auf dem persönlichen Computer installiert werden und der Teilnehmer sowie seine persönlichen Daten sind dem Forscher normalerweise bekannt. Je nach der Software, die zum Aufzeichnen verwendet wird, kann eine große Menge Daten aufgezeichnet werden inklusive der Interaktionen mit dem Webbrowser und den Tastenanschlägen (vgl. HAWKEY 2009, S. 89).

Wenn bei Proxyaufzeichnungen Javascript verwendet wird, können einige Nutzerinteraktionen abgefangen werden. So kann ein proxyseitiges Aufzeichnen, je nach Forschungsfrage, ein brauchbarer Kompromiss sein (vgl. HAWKEY 2009, S. 89).

8.2.1 Empfehlungen für Datensammlungen

Die wichtigsten Empfehlungen, die Forscher bei der Sammlung von Daten beachten sollten, werden hier zusammengefasst aufgezeigt.

Schule die Teilnehmer

Der Forscher sollte möglichst transparent mit den Daten umgehen, um die Bedenken der Teilnehmer zu mildern. Eventuell verwendete Einverständniserklärungen sollten den Datenschutz und die Wahrung der Privatsphäre explizit erklären, so dass die Teilnehmer ein klares Verständnis dafür haben, welche Daten gesammelt werden, wer die Möglichkeit hat, sie einzusehen, wie die Daten benutzt werden und wie sie kommuniziert werden. Detaillierte Informationen über die genutzten Mechanismen, die den Datenschutz gewährleisten, sollten generelle Bedenken ausräumen. Die Teilnehmer sollten die Möglichkeit haben, spezielle Bedenken zu äußern oder Unklarheiten zu erfragen. Einige der Teilnehmer werden die

angebotenen Schutzmechanismen nicht annehmen, andere bekommen aber durch die genutzten Mechanismen das Gefühl, Kontrolle über ihre Daten und die Aktivitäten im Web zu haben und werden so eher an der Studie teilnehmen (vgl. HAWKEY 2009, S. 90).

Nur so viele Daten aufzeichnen/empfangen wie gebraucht werden

Im Bereich des E-Commerce wird von Kunden vermutet, dass die dort gesammelten Daten benötigt werden, um die Seiten zu personalisieren bzw. zu optimieren. Hier die gesammelten Daten einzuschränken, kann die Bereitschaft der Nutzer, ihre Daten preiszugeben, erhöhen. Forscher sollten bei der Sammlung von Daten die gleiche Vorgehensweise nutzen. Den Teilnehmern sollte der Nutzen der Daten vermittelt werden und dementsprechend sollten nur die Daten gesammelt werden, die zur Beantwortung der Forschungsfrage benötigt werden. Die Teilnehmer fühlen sich sicherer, wenn ihre Daten respektiert und nur für die Forschung in einem bestimmten Interessenfeld genutzt werden. Die Forscher können die Daten z.B. auf dem Computer des Teilnehmers sammeln und dort durch ein Script so verarbeiten lassen, dass sie anonymisiert zu ihnen gesendet werden. Wenn z.B. ein Script die URL einer Seite einer bestimmten IP-Adresse zuweist und der Forscher nur diese erhält, ist die Privatsphäre gewahrt (vgl. HAWKEY 2009, S. 90). Um dieser Empfehlung nachzukommen, muss der Wissenschaftler von Anfang an wissen, welche Daten er zur Beantwortung seiner Forschungsfrage benötigt (vgl. HAWKEY 2009, S. 91).

Möglichkeit der Datenprüfung (durch Teilnehmer)

Um die Transparenz der Daten für den Teilnehmer zu erhöhen, kann der Forscher ihm Einblick in die gesammelten Daten gewähren und so garantieren, dass nur die vereinbarten Daten übermittelt werden (vgl. HAWKEY 2009, S. 91).

Die Datenschutzmaßnahmen sollten den Teilnehmer nicht in der Ausübung seiner normalen Tätigkeiten behindern.

Es ist zu beachten, dass nicht alle Empfehlungen für alle Studien geeignet und von der Forschungsfrage abhängig sind. Aber die Kontrolle über die

Daten durch den Teilnehmer soll helfen, die Datenschutzbedenken zu mildern. Wichtig ist dabei vor allem, keinen Einfluss auf die Datenschutzeinstellungen des Teilnehmers in seinem Browser zu nehmen und ihm so sein gewohntes Browsingverhalten zu ermöglichen (vgl. HAWKEY 2009, S. 92).

Für die Forscher ist es wichtig, die Möglichkeiten, die das Verhalten der Teilnehmer ändern können, anzunehmen und umzusetzen. Sie sollten bedenken, dass sich das Verhalten der Teilnehmer je nach Ort (zu Hause oder Arbeit) und nach Gerät (Laptop, PC usw.) ändern kann. Zusätzlich könnten unterschiedliche Browser oder Sicherheitseinstellungen innerhalb dieser Umgebungen genutzt werden. Daher muss sichergestellt werden, dass der Datenschutz in all diesen Bereichen ausreichend eingehalten wird, gerade dann, wenn Datenschutzrichtlinien von Firmen oder Organisationen tangiert werden können.

Die Forscher müssen die Datenschutzbelange der Teilnehmer mit ihren Forschungsbedürfnissen und der Beantwortung ihrer Forschungsfrage ausbalancieren (vgl. HAWKEY 2009, S. 94).

8.3 Probanden befragen

In den ergänzenden Befragungen geht es darum herauszufinden, wie sich die Probanden selbst wahrnehmen und darum, ihr Umfeld, ihre Situation oder ihr Verhalten zu analysieren.

Dafür muss der Forscher abhängig von der Forschungsfrage die zu befragende Zielgruppe definieren, ein Befragungsinstrument konstruieren und administrieren, die Ergebnisse analysieren, interpretieren und kommunizieren. Abhängig von der Forschungsfrage bzw. dem Ziel der Befragung können unterschiedliche Methoden der Befragung gewählt werden. Hierzu eignen sich Interviews, Telefonbefragungen, elektronische Befragungen oder Kombinationen aus diesen Methoden. Dabei sollte die Befragung aber immer von der Hypothese und/oder der Forschungsfrage geleitet werden (vgl. RAINIE UND JANSEN 2009, S. 41).

Bei Befragungen im Internet ist es schwierig, ein repräsentatives Beispiel zusammenzustellen, das die gesamte Population widerspiegelt (vgl. RAINIE UND JANSEN 2009, S. 41).

Elektronische Befragungen garantieren dabei den einfachsten Zugang und bieten die beste Vorbereitung der Dateneingabe/-aufbereitung. Ob die Daten qualitativ, quantitativ oder in einer kombinierten Methode ausgewertet werden sollen, hängt wieder von der Forschungsfrage ab (vgl. RAINIE UND JANSEN 2009, S. 42).

Der Forscher muss vor der Befragung ein klares Bild davon haben, was er herausfinden möchte, um sein Befragungsinstrument darauf zu konzentrieren. Dafür sollte er eine Liste mit den benötigten Daten erstellen, um diese dann durch gezielte Fragestellungen von den Probanden zu erheben (vgl. RAINIE UND JANSEN 2009, S. 43).

Auch hier sollten, um die gewünschten Antworten zu generieren, die Teilnehmer informiert werden. Deshalb sollte der Forscher am Anfang des Fragebogens Richtlinien aufführen. Es können aber auch zu einigen spezifischen Fragen ergänzende Erläuterungen gegeben werden, die direkt bei den Fragen stehen. Fragen zu demografischen Daten sollten an das Ende der Befragung gesetzt werden, weil sie die Geduld der meisten Teilnehmer strapazieren und zu falschen oder unvollständigen Antworten führen können. Außerdem sollten Fragen zu einem Gebiet beieinander stehen. Sollte eine Befragung mehrere Gebiete betreffen, sollten diese durch Überschriften kenntlich gemacht werden. Jeder Fragentyp sollte gleich aufgebaut werden und am Anfang der Frage erklärt sein. Dies erhöht die Wahrscheinlichkeit valider Daten (vgl. RAINIE UND JANSEN 2009, S. 43). Zu den Fragentypen, die in einer Befragung verwendet werden können, gehören: Multiple Choice-Fragen (stellen eine begrenzte Anzahl von Auswahlmöglichkeiten) (vgl. RAINIE UND JANSEN 2009, S. 43 f.), skalierte Fragen (Bewertung z.B. nach Schulnotensystem oder „Stimme voll zu“ bis „Stimme gar nicht zu“), sowie offene Fragen (geben keine Antwortmöglichkeiten vor, sind schwer zu analysieren und eher für qualitative Forschung geeignet) (vgl. RAINIE UND JANSEN 2009, S. 44).

8.4 Logfiledaten sammeln und analysieren

Zur Sammlung von Daten aus Logfiles gibt es wie erwähnt verschiedene Methoden (client- oder serverseitig). Welche Vor- und Nachteile diese Methoden der Datensammlung im Allgemeinen haben und was ein Forscher mit den gesammelten Daten herausfinden kann, wird im folgenden Kapitel dargestellt.

Die Sammlung von Daten aus Logfiles ist eine gute Methode, um eine große Menge Daten von vielen Nutzern zu sammeln (vgl. JANSEN 2009, S. 100).

Die Logfiles zeichnen die Interaktionen der Nutzer mit dem jeweiligen System auf. Es handelt sich um Aufnahmen der Suchepisode zwischen der Websuchmaschine und dem Nutzer, der in dieser nach Informationen gesucht hat (vgl. JANSEN 2009, S. 101).

Die Aufzeichnungen der Interaktionen der Nutzer nehmen die Webserver vor. Diese nehmen, mit Hilfe einer Software, die Interaktion eines bestimmten Webbrowsers auf dem Computer des Nutzers mit der Suchmaschine in einem Logfile auf. Die meisten Logfiles sind also serverseitig (vgl. JANSEN 2009, S. 101).

Die Daten aus Logfiles können genutzt werden, um Forschungsfragen zu klären, die sich mit der Interaktion zwischen Webnutzer und Websuchmaschine oder dem Webinhalt während einer Suchsession beschäftigen. Ziel dieser Studien ist es, ein besseres Verständnis der Interaktionen zwischen dem Suchenden, dem Inhalt und dem System zu erhalten (vgl. JANSEN 2009, S. 102). Dazu werden die Charakteristika einer Suchsession betrachtet, um Trends zu isolieren und typische Interaktionen zwischen dem Nutzer und der Suchmaschine zu identifizieren.

Die Interaktionen zwischen dem Nutzer und der Suchmaschine beinhalten mehrere Aktionen wie die Suchanfrageneingabe, die Modifikation einer Suchanfrage, die Ansicht der Ergebnisliste und die Nutzung verschiedener Informationsquellen (z.B. Webseiten, pdf-Dateien, Videos) (vgl. JANSEN 2009, S. 102).

Ein Kritikpunkt der Logfiles ist, dass sie die Empfindungen und Vorstellungen der Nutzer bei der Suche nicht erfassen können. Sie können die eigentlichen Informationsbedürfnisse der Nutzer nicht erfassen und die

Zufriedenheit eines Nutzers mit den Suchergebnissen nicht abschätzen. Logfiles zeigen nur die Aktionen, die ein Nutzer getätigt hat, sagen aber nichts über die Emotionen oder die kognitiven und situationsbedingten Elemente des Suchvorgangs eines Nutzers aus (vgl. JANSEN 2009, S. 103 f.).

Die Stärke dieser Methode liegt in der Unaufdringlichkeit, das heißt, der Nutzer merkt nicht, dass seine Interaktion aufgezeichnet wird. Durch das Hinzufügen weiterer Methoden wird diese Stärke minimiert (vgl. JANSEN 2009, S. 104).

Um zu ermitteln, welche Daten gesammelt werden müssen, muss eine Forschungsfrage artikuliert werden. Diese legt fest, welche Daten gesammelt werden müssen.

Werden die Daten in einem realen System (keine Laborsituation) gesammelt, sollte die Erfassung der Daten den Suchprozess nicht behindern. Die Störung könnte zum Verlust von potenziellen Kunden führen (vgl. JANSEN 2009, S. 105).

Bei der Ausführung einer Interaktion in einer Suchmaschine sollten folgende Faktoren gespeichert werden:

- Zur Identifikation eines individuellen Computers dient ein Cookie. Er wird auf den Computern gespeichert und sendet Informationen vom Webserver zum Webbrowser (so können bspw. die letzten Einstellungen bei einem erneuten Besuch der Webseite abgerufen werden (vgl. HÖCHSTÖTTER 2009, S. 180)).
- Die IP-Adresse des Computers, um den Computer des Nutzers zu identifizieren. Diese wird manchmal als anonyme Nutzercodeadresse zugeteilt.
- Das Datum der Interaktion, wie es von der Suchmaschine aufgezeichnet wurde.
- Die Zeit der Interaktion, z.B. der Zeitpunkt, an dem eine Suchanfrage abgeschickt wurde (vgl. HÖCHSTÖTTER 2009, S. 180 und JANSEN 2009, S. 105).
- Die Suchanfrage, wie sie vom Nutzer eingegeben wurde. Websuchmaschinen nehmen dieses Feld normalerweise immer auf.

- Andere enthaltene Felder sind die Ergebnisseite zu einer Suchanfrage, die die Abstracts der Suchmaschine und die zugehörigen URLs präsentieren, sowie die vom Nutzer genutzte Sprache und die angeklickten Ergebnisse (Click-through) (vgl. JANSEN 2009, S. 105).

Wenn ein Suchender eine Suchanfrage eingibt, sieht er ein Dokument an und kehrt zur Suchmaschine zurück, zeichnet der Webserver diesen zweiten Besuch normalerweise mit der identischen Nutzeridentifikation und der gleichen Suchanfrage auf, aber mit einer neuen Zeit (z.B. Zeit des zweiten Besuchs). Dies ist eine nützliche Information um herauszufinden, wie viele Seiten der Ergebnisseite der Suchende besucht hat, aber leider verdreht es auch die Analyse der Anzahl der Suchanfragen (gleiche Anfragen tauchen doppelt auf). Vor der Analyse muss man also die eingegebenen Suchanfragen von den „Rückkehrern“ trennen. (vgl. JANSEN 2009, S. 108). Diese Trennung erfolgt, indem man die Identifikationsnummern vergleicht und so die unigen Suchanfragen erhält. Bei der Logfileanalyse ist man oft interessiert an den Begriffen und der Begriffsnutzung. Hier ist es oft gut, eigene Tabellen zu erstellen, die die Begriffe und die Häufigkeit des Auftretens enthalten (vgl. JANSEN 2009, S. 109).

In der Datenanalyse sollten zuerst die Charakteristika der Stichprobe analysiert werden.

Dazu sollten Faktoren wie die Häufigkeit bestimmter Ausdrücke analysiert werden. Dabei sollten auch die paarweise auftretenden Ausdrücke beachtet werden (vgl. JANSEN 2009, S. 110).

Als Anfrage wird in einer solchen Studie die Zeichenfolge, die in eine Suchmaschine eingegeben wird, definiert. Sie kann aus einem oder mehreren Ausdrücken bestehen. Dabei ist die erste Suchanfrage eines Nutzers die Ausgangsanfrage, die folgenden Anfragen desselben Nutzers sind modifizierte Suchanfragen (vgl. JANSEN 2009, S. 111 f.). Sendet ein Nutzer die gleiche Suchanfrage mehrfach ab, wird diese als identische Anfrage bezeichnet. Diese werden häufig generiert, wenn der Suchende von einer Webseite zurück auf die Suchergebnisseite kommt, um sich evtl. ein anderes Ergebnis anzusehen (vgl. JANSEN 2009, S. 112).

Eine Suchsession, die aufgezeichnet wird, wird in der Forschung als eine Abfolge von Interaktionen zwischen System und Nutzer in einer bestimmten Zeit definiert. Der Nutzer kann dabei eine oder mehrere Suchanfragen an die Suchmaschine senden. Die gewählten Zeitintervalle sind meist kurz (zwischen fünf und 120 Minuten). Dabei ist zu beachten, dass sich jede zeitliche Begrenzung auf die Resultate auswirkt (vgl. JANSEN 2009, S. 112).

Die Sessiondauer ist die totale Zeit, die der Nutzer mit den Interaktionen mit der Suchmaschine verbringt, inklusive der Zeit, die er braucht, um die ersten und darauffolgenden Webdokumente anzusehen, allerdings ohne das letzte Dokument. Die Sessionlänge wird gemessen, indem man die Differenz zwischen der Zeit der Eingabe der (ersten) Suchanfrage und der Zeit, zu der der Nutzer die Suchmaschine zum letzten Mal verlässt, berechnet. Die benötigte Zeit, um das letzte Dokument anzusehen, kann nicht gemessen werden (vgl. JANSEN 2009, S. 112).

Logfiledaten haben den Vorteil, dass man sie preiswert sammeln kann. Die Gesamtkosten setzen sich lediglich aus den Kosten für die Software zum Sammeln der Daten und den Kosten der Speicherung zusammen.

Außerdem zeigt die Methode die unverfälschten Verhaltensweisen der Nutzer, da diese nicht merken, dass ihre Interaktionen gespeichert werden (vgl. JANSEN 2009, S. 114 und HÖCHSTÖTTER 2009, S. 181).

Durch Logfiles lässt sich aber der einzelne Nutzer nicht zweifelsfrei identifizieren, sondern immer nur der individuelle Computer bzw. der genutzte Browser. Sobald dieser von mehreren Personen genutzt wird, wird die Darstellung des Nutzers unpräzise.

Möchte man die Gründe für eine Suche evaluieren, sollte man Logfiles in Verbindung mit qualitativen Methoden nutzen, um die Motivation des Nutzers zu erfassen (vgl. JANSEN 2009, S. 115).

9. Entwicklung neuer Studien

Durch die Analyse der Literatur hat sich gezeigt, dass der Forschungsbedarf im Bereich der Suchmaschinen und der Intention hinter einer Suchanfrage weiterhin groß ist.

Verschiedene Ideen, die Suchanfragen und die zufriedenstellende Präsentation der Suchergebnisse zu erforschen, bringen Ansätze für die Weiterentwicklung von Suchmaschinen hervor.

Ausgehend von den Forschungsergebnissen und den Mängeln in der genutzten Methodik, die die Forscher in ihren publizierten Arbeiten kommunizieren, werden hier Ideen für den weiteren Ausbau dieser Forschungsrichtung gegeben.

Die einzelnen vorgeschlagenen Studien sind dabei immer so aufgebaut, dass erst die Idee hinter dem Vorschlag erklärt und anschließend die Methodik vorgestellt und die benötigte Datengrundlage erläutert wird, um schlussendlich die Durchführbarkeit und die Priorität des Vorschlags zu beleuchten.

Die hier konstruierten Studien werden am Ende priorisiert und beispielhaft miteinander in Verbindung gebracht, sodass aus mehreren kleinen Studien ein größeres Projekt entstehen kann (siehe Anhang 1).

Dabei werden Methoden aus verschiedenen Forschungsbereichen kombiniert, um ein Gesamtbild der Suchintention entwickeln zu können und die verschiedenen Aspekte der Websuche zu beleuchten (vgl. LEWANDOWSKI 2010, S. 3 und JANSEN ET AL 2009, S. 1).

9.1 Entwickelte Studien – Evaluation des Informationsbedürfnisses eines Nutzers

Der größte Unsicherheitsfaktor bei einer Studie zum Thema Intention hinter der Suchanfrage sind die kurzen Suchanfragen, aus denen das Informationsbedürfnis der Nutzer nicht immer zu erkennen ist (vgl. LEWANDOWSKI 2006, S.38). Hinter einer Suchanfrage mit den gleichen Suchworten können verschiedene Intentionen stehen (vgl. KANG UND KIM 2003, S. 66 und DAI ET AL. 2006, S. 8). In den in dieser Arbeit vorgestellten Studien wurden Suchanfragen bewertet und zugeordnet, und zwar nicht von dem „Originalanfragesteller“. Die Aussage, hinter einem bestimmten Suchwort

stehe stets genau ein bestimmtes Informationsbedürfnis, ist schwierig und nicht valide.

I. Laborexperiment zur Überprüfung der Einordnung von Suchanfragen durch Juroren

Um die Intention einer Suchanfrage zu ermitteln, wurden die Suchanfragen in den bereits durchgeführten Studien durch Juroren in die Anfragetypen navigationsorientiert, transaktionsorientiert und informationsorientiert eingeordnet. Meistens sollten sich zwei Juroren über das Informationsbedürfnis einig werden. Ob die Juroren bei ihrer Einordnung das tatsächliche Informationsbedürfnis des Nutzers, der die Anfrage abgeschickt hat, ermittelt haben, konnte nicht überprüft werden.

Eine weitere Studie zu der Intention hinter einer Suchanfrage sollte mit den Originalanfragestellern durchgeführt werden, denn sie wissen, was sie suchen und wie ihr Suchbedürfnis am besten befriedigt werden kann. Das Informationsbedürfnis muss nicht mit der eingegebenen Suchanfrage übereinstimmen, denn es kann Mängel bei der Übertragung des Informationsbedürfnisses geben, das der Nutzer im Kopf hat, zu dem Stichwort, das er eingibt (evtl. ist der Nutzer im Kopf schon weiter, weil er weiß, zu welcher Ressource er möchte) (vgl. BAEZA-YATES ET AL. 2006, S. 106).

Eine solche Studie könnte dazu dienen, die Ergebnisse der bereits abgeschlossenen Studien zu überprüfen und die originären Suchbedürfnisse zu den gängigsten Anfragen zu erheben.

In einem zweiten Schritt sollten Faktoren ermittelt werden, die die Intention des Suchanfragestellers beschreiben.

Um die Faktoren zu überprüfen, sollten sie auf eine größere Stichprobe angewandt und manuell überprüft werden.

1.1 Methodik und Datengrundlage

Das originäre Informationsbedürfnis sollte in einem Mix aus qualitativen Methoden evaluiert werden. Quantitative Methoden können in einem zweiten Schritt angewendet werden, um die Ergebnisse anderer Studien zu durchleuchten.

Die Probanden sollten in einer Laborstudie befragt werden, da hier die Möglichkeit besteht, direkt nach Hintergründen der Suchanfrage bzw. dem Informationsbedürfnis zu fragen (vgl. HÖCHSTÖTTER 2009, S. 178) und z.B. die These von Baeza-Yates et al. zu überprüfen, ob die Nutzer von Suchmaschinen tatsächlich ihr Informationsbedürfnis schlecht übertragen können.

Empfohlen wird eine Teilnehmerzahl von circa 60 Probanden, die möglichst einen Querschnitt der Gesellschaft bilden sollten. Da die Computerarbeitsplätze in einem Labor oft beschränkt sind, sollten hier je nach Verfügbarkeit der Mittel zwei bis drei Durchgänge des Experiments durchgeführt werden, um eine gute Stichprobe zu generieren.

Der Aufbau der Laborstudie sollte eine Einführung beinhalten, damit die Probanden wissen, worum es in der Studie geht. Die Teilnehmer der Studie sollten frei und ohne Einschränkungen suchen können. Dabei kann es sich um Suchanfragen aus den letzten Tagen oder um spontane Einfälle handeln und die Probanden sollten je circa zehn bis 20 Suchanfragen eingeben.

Während ihrer Suche sollten sie befragt werden. Die Probanden sollten nach der Eingabe der Suchanfragen, aber vor dem Abschicken, beschreiben was sie eigentlich finden wollten, um eine Beeinflussung durch die gezeigten Suchergebnisse zu vermeiden.

Nach der Sichtung der Suchergebnisse zu den Suchanfragen sollten sie zu der Zufriedenheit mit der Suchergebnisseite befragt werden (Ist eine Modifikation der Suchanfrage notwendig? Ist das Informationsbedürfnis befriedigt?), und evtl. auf das perfekte Suchergebnis auf der Suchergebnisseite klicken zu können. Dabei würde eine Fehleinschätzung der Webseite, die durch die Beschreibung in der Suchmaschine hervorgerufen, aufgedeckt werden (vgl. HARGITTAI 2002, S.1242).

Außerdem sollten die Interviewer beachten, dass das, von den Probanden Gesagte, oft unterschiedlich sein kann zu dem, was sie wirklich tun (vgl.

ROHER 2008). Z.B. wenn der Nutzer sagt, er möchte eine bestimmte Webseite finden, die er kennt und benennen kann und in den Suchergebnissen auf eine andere Webseite klickt, sollte nachgefragt werden, da die ausgeführten Aktionen des Nutzers nicht mit seiner Beschreibung des perfekten Suchergebnisses übereinstimmen.

Die Fragen an die Probanden sollten nach den Vorgaben von Broder entworfen werden (vgl. BRODER 2002, S. 6). Die Probanden sollten gefragt werden, ob sie eine bestimmte Webseite im Kopf hatten oder mehrere Webseiten benötigten, um ihr Informationsbedürfnis zu befriedigen oder welches Informationsbedürfnis hinter ihrer Suchanfrage stand. Dabei erscheint es nicht nötig, die Anfragetypen nach den Vorgaben Broders zu benennen, da diese hier eher zu Verwirrung der Probanden beitragen könnten.

Wichtig ist, dass die Probanden einander nicht hören können, um eine Beeinflussung und disproportional viele gleiche Suchanfragen zu vermeiden.

Während des Laborexperiments sollten die Suchergebnisseiten sowie die einzelnen Sessions der Probanden aufgezeichnet und den parallel geführten Interviews zugeordnet werden können, um später die nicht relevanten Suchergebnisse der Suchergebnisseiten isolieren zu können. Außerdem sollten Modifikationen der ursprünglichen Suchanfrage sichtbar sein, um nachvollziehen zu können, wann ein Proband unzufrieden mit der Suche war und wie er die Suche verändert hat, um zu einem befriedigenden Ergebnis zu kommen und um diese Aktionen mit den Aussagen gegenüber dem Interviewer vergleichen zu können (vgl. JANSEN ET AL. 2007, S. 1256).

Bei der Verbindung zwischen den Probanden und ihren Suchsessions sollte eine Codierung angewendet werden, um die Privatsphäre der Probanden zu schützen. Dies sollte ihnen zu Beginn des Laborexperiments mitgeteilt werden (vgl. HAWKEY 2009, S. 86).

Für den Fall, dass mehrere Probanden nacheinander den gleichen Computer nutzen, sollten die Einstellungen des Browsers nach jedem Probanden zurückgesetzt werden, um eine Beeinflussung zu vermeiden (vgl. HARGITAI 2002, S. 1241).

Abschließend sollten die Probanden nach ihrer Interneterfahrung, ihrem Alter, Geschlecht und weiteren demografischen Daten befragt werden (vgl. JANSEN UND MCNEESE 2005, S. 1487 f.). So kann überprüft werden, ob sich die Intention hinter einer Suchanfrage tendenziell unterscheidet, bspw. je nach Geschlecht.

Den Probanden sollte ein finanzieller Anreiz für die Teilnahme an der Laborstudie geboten werden, denn so ist die Bereitschaft zur Teilnahme ebenso höher wie die Wahrscheinlichkeit, ergiebige Daten zu bekommen (vgl. HARGITTAI 2002, S. 1241).

Um Faktoren zur Identifikation der Anfragetypen ermitteln zu können, müssen die Suchanfragen manuell überprüft werden. Nicht alle Probanden werden ihr Informationsbedürfnis gleich benennen, selbst wenn sie dasselbe meinen.

Sind Faktoren ermittelt, sollten diese an einer großen Stichprobe überprüft werden. Dazu sollte eine valide Stichprobe (ca. 25.000 Suchanfragen), die nach der Häufigkeit³ ausgewählt wurde, anhand der ermittelten Faktoren in die drei Anfragetypen eingeteilt werden. Hier wird bei einer (manuellen) Plausibilitätsprüfung die Verlässlichkeit der Faktoren ermittelt.

Die Logfiles sollten für die Durchführung der beschriebenen Studie folgende Daten enthalten:

Die Suchanfrage der einzelnen Probanden inklusive Zeitstempel, um eine chronologische Reihenfolge der (eventuell) modifizierten Suchanfragen herstellen zu können.

Die Stichprobe zur Überprüfung der entwickelten Faktoren benötigt keine weiteren Angaben außer den Suchworten der Probanden.

³ Hier sollten keine selten gestellten Suchanfragen aus dem sogenannten Long Tail enthalten sein, sondern populäre Suchanfragen, die bspw. mindestens einmal am Tag gestellt werden.

1.2 Priorität und Durchführbarkeit

Nachdem in verschiedenen Studien bemängelt wurde, dass die Suchanfragen nicht von den Suchenden selbst klassifiziert wurden und da die Forscher sich darüber im Klaren sind, dass hinter einem Suchwort mehr als ein Informationsbedürfnis stehen kann, scheint eine Studie, in der die Suchenden direkt nach ihrem Bedürfnis befragt werden überfällig. Dieser Studie wird daher eine hohe Priorität eingeräumt.

Eine solche Studie ist gut durchführbar. Die Probanden werden in der Laborsituation zwar nicht nach sozial unerwünschten Themen z.B. aus dem Bereich der Pornografie suchen. Für eine Beispielsammlung von Daten sind diese auch nicht von Belang und es geht hier eher um die subjektive Bewertung des einzelnen Probanden.

Nach der Auswertung der Daten ist (bei der Einordnung in die von Broder erstellten Anfragetypen) ein Vergleich mit den Daten bereits durchgeführter Studien gut möglich. Dieser würde die Bedenken, die die Forscher in ihren Studien selbst äußern, bestätigen oder zerstreuen (vgl. DAI ET AL. 2006, S. 8 und JANSEN ET AL. 2007, S. 1258).

Durch die relativ geringe Probandenanzahl und die Veränderungen im Verhalten durch die Anwesenheit des Interviewers wird diese Erhebung nicht repräsentativ sein (vgl. HÖCHSTÖTTER 2009, S. 178), aber eine solche Datensammlung kann die Grundlage für die weitere Forschung auf diesem Gebiet bereitstellen (kann weiterhin Fremdbewertung genutzt werden oder können Juroren keine fremden Suchanfragen bewerten?).

Die Überprüfung der ermittelten Faktoren anhand einer großen Stichprobe kann außerdem zeigen, ob die Übertragung des Informationsbedürfnisses auf allgemeine Suchanfragen möglich ist.

In dieser Arbeit bildet diese Studie die Grundlage für einige weitere Studien mit einem von den Suchenden selbst klassifizierten Datenbeispiel. Sollte dies der Fall sein, wird auf Studie I „Laborexperiment zur Überprüfung der Einordnung von Suchanfragen durch Juroren“ verwiesen.

II. Klickdaten zu Identifikation des Informationsbedürfnisses

Lee et al. nutzten 2005 die Klicks der Suchenden aus der Vergangenheit. Sie untersuchten, auf welche Ergebnisse die vorherigen Nutzer klickten und versuchten so, das Informationsbedürfnis abzuleiten. Hier wurde das Informationsbedürfnis, das hinter einer Suchanfrage stehen könnte, zum Vergleich zur automatischen Klassifikation ebenfalls von Juroren zugeordnet.

Der Ansatz, die Klickdaten zur Identifikation der Anfragetypen hinzuzuziehen, ist vielversprechend (vgl. DAI ET AL. 2006, S. 8 und OTTE UND VON MACH 2009, S. 32 f.). Allerdings sollte hier ein Datenbeispiel, wie in Studie I „Laborexperiment zur Überprüfung der Einordnung von Suchanfragen durch Juroren“ beschrieben, vorliegen, um Vergleiche mit Suchanfragen, die von den tatsächlich Suchenden eingeordnet wurden, ziehen zu können. In der Studie von Lee et al. wurde ein sehr eingeschränktes Datenset untersucht, das nur Anfragen eines Universitätsdepartments enthielt (vgl. LEE ET AL. 2005, S. 391 f.). Dieses sollte durch generelle Anfragen ersetzt werden, da diese eher die Informationsbedürfnisse der Internetpopulation bzw. der Suchmaschinennutzer widerspiegeln.

II.1 Methodik und Datengrundlage

Der Klickdatenanalyse sollte die Erstellung eines Datensets, wie in Studie I beschrieben, vorangehen.

Zur Untersuchung der Klicks kann ein unternehmensinternes Tool der Deutschen Telekom genutzt werden, das bspw. auch schon in vorangegangenen Studien verwendet wurde und zu validen Ergebnissen führte. Dieses Tool erlaubt Einsicht in die Klickdaten zu einzelnen Suchanfragen je Trefferposition. Es können sowohl die Treffer der Suchergebnisseite als auch die Klicks auf die bezahlten Anzeigen angesehen werden (vgl. OTTE UND VON MACH 2009, S. 33).

Der Untersuchung liegen die Klickdaten in Verbindung mit den Anfragetypen, die Broder 2002 entwickelte zu Grunde. Danach sollten die Klicks einer Suchanfrage mit einem navigationsorientierten Informationsbedürfnis auf einen Treffer entfallen, während die informationsorientierten Suchen implizieren, dass Informationen gesucht werden, die mit hoher Wahrscheinlichkeit nicht auf einer Seite zu finden sind. Die Klickverteilung sollte

sich in den Ergebnissen der Klickdatenanalyse ablesen lassen, denn bei Lee et al. entfallen bei navigationsorientierten Anfragen circa dreiviertel der Klicks auf das erste Ergebnis, während bei den informationsorientierten Anfragen die Verteilung eher gleichmäßig auf die ersten zehn Treffer verteilt ist (vgl. LEE ET AL. 2005, S. 395).

Transaktionsorientierte Anfragen könnten sich an vermehrten Klicks auf die Anzeigen festmachen lassen, tendenziell stellen dies auch Otte und von Mach fest (vgl. OTTE UND VON MACH 2009, S. 44 f.).

Aus den Logfiles werden für diese Studie die folgenden Daten benötigt. Die aus den Suchlogs gefilterten Daten, die durch die Fragesteller selbst bewertet wurden.

Die gesamte Session eines Nutzers, um die Klickhistorie nachvollziehen zu können und zu sehen, ob eine Anfrage modifiziert wurde oder der Nutzer sich, nachdem er bspw. auf das erste Ergebnis der Suchergebnisseite geklickt hat, noch weitere Ergebnisse angesehen hat, so lassen sich Faktoren für die Einteilung in die verschiedenen Anfragetypen finden und nicht zufriedenstellende Suchergebnisse identifizieren.

Desweiteren wird der Zugang zu dem Tool der Deutschen Telekom benötigt, um die Klickdaten der Suchanfragen zu extrahieren bzw. diese untersuchen zu können.

II.2 Priorität und Durchführbarkeit

Es bietet sich an, die Klickdaten in einer Untersuchung, in der es um die Intention hinter einer Suchanfrage geht einzubeziehen, um eventuelle Klickmuster für die unterschiedlichen Anfragetypen zu ermitteln. Auf diese Weise könnte eine automatische Zuordnung einer Suchanfrage zu einem Anfragetypen generiert werden.

Besonders im Vordergrund würde hier stehen, dass bei transaktionsorientierten Suchanfragen vermutlich eher auf die Anzeigen geklickt wird als bei den navigationsorientierten und informationsorientierten Suchanfragen. In der Umsetzung würde dies bedeuten, dass Suchwörter für transaktionsorientierte Suchanfragen gefunden werden würden und die Schaltung von Werbung für bestimmte Wörter optimiert werden könnte (vgl. SCHMIDT-MÄNZ 2007, S. 186). Suchende mit einem navigationsorientierten oder in-

formationsorientierten Informationsbedürfnis fühlen sich durch Werbung eher gestört.

Die Durchführung der Klickdatenanalyse mit einem entsprechenden Vergleichsset der Informationsbedürfnisse sollte unbedingt durchgeführt werden, um eine automatische Zuordnung der Suchanfragen zu einem Anfragetyp durch gefundene Klickmuster zu ermöglichen.

III. Überprüfung der Suchanfragen-Klassifikation nach Broder 2002

Betrachtet man die rasante Entwicklung des Internets und der dahinterstehenden Technologien, scheint eine Überprüfung der 2002 entwickelten Anfragetypen navigationsorientiert, transaktionsorientiert und informationsorientiert sinnvoll. Die Klassifikation wurde auch in anderen Studien genutzt, wenn auch teilweise die Anfragetypen anders benannt wurden. So bleibt es doch bei einer Einteilung in Kategorien, die das Informationsbedürfnis des Suchmaschinennutzers evtl. nicht (mehr) widerspiegelt (vgl. z.B. KANG UND KIM 2003, S. 64 f. und ROSE UND LEVINSON 2004 S. 13). Es könnte durchaus möglich sein, dass die Suchmaschinen die Nutzer werden zufriedenstellen können, indem die Kategorisierung der Suchanfragen nicht mehr nach Anfragetypen, sondern bspw. nach Themen vorgenommen wird. Dies sind jedoch nur Vermutungen, die durch eine Überprüfung der Kategorisierung bzw. eine Neuordnung der Suchanfragen gefestigt werden sollten.

III.1 Methodik und Datengrundlage

Um die Kategorien zu überprüfen oder neue Ansätze für das Clustern von Suchanfragen zu finden ist eine Gruppendiskussion sinnvoll. Da diese zur Ideengenerierung dienen und es bei dieser Studie eben um die Generierung von neuen Hypothesen und der Überprüfung von bestehenden Kriterien geht, ist diese qualitative Methode am geeignetsten (vgl. WINTER 2000).

Eine Gruppe sollte aus höchstens 20 Probanden bestehen, ideal zur Anregung einer ergiebigen Diskussion sind zehn Teilnehmer (vgl. WINTER 2000).

Die Probanden für eine solche Aufgabe sollten nicht aus dem Bereich des Information Retrieval kommen, da sich herausgestellt hat, dass diese zu „vorbelastet“ im Kategorien-Denken sind (vgl. OTTE UND VON MACH 2009, S. 65). Ideal wäre eine Gruppe, die selbst Suchmaschinen nutzt, zwar kaum Vorbildung über dahinter liegende Techniken (bspw. Suchmaschinen-optimierung) hat, sich aber trotzdem mit der Suchergebnisseite auseinandersetzt und z.B. die Ads (er)kennt.

In Bezug auf die sonstigen demografischen Faktoren wie Alter, Geschlecht, Bildung, Einkommen oder auch Lebenswandel sollte die Gruppe in sich homogen sein. Die Meinungen zu dem untersuchten Thema können aber durchaus heterogen zusammengesetzt sein (vgl. WINTER 2000).

Diese Gruppe ist am geeignetsten, um zu sehen, ob sie die Suchanfragen in ähnliche Kategorien einteilt wie sie in der Forschung genutzt werden, oder ob sie einen neuen Ansatz findet, der in der Praxis die Informationsbedürfnisse besser befriedigt. Dabei ist zu beachten, dass die Gruppe in einer Einführung nicht darüber informiert werden sollte, dass es bereits eine Kategorisierung der Suchanfragen gibt, damit sie unbelastet in die Studie geht.

Die Suchanfragen, die hier verwendet werden, sollten den Probanden zur Verfügung gestellt werden, damit sie sich einen Überblick über die existierenden Suchworte verschaffen können und sich erste Gedanken zu möglichen Clustern machen. So vorbereitet können sie in die Gruppendiskussion einsteigen.

Die Ideen der Probanden sollten durch einen Moderatoren festgehalten und sichtbar für alle vermerkt werden (bspw. an einer Tafel). Die Ergebnisse sollten abfotografiert oder abgeschrieben werden auch um zu vermeiden, dass Informationen verloren gehen, wenn gravierende Änderungen vorgenommen werden (wenn z.B. die Probanden feststellen, dass die Einordnung nicht das gesamte Spektrum der Suchanfragen abdecken kann und sie einen neuen Ansatz finden müssen).

Zusätzlich kann die Diskussion auf Tonband oder Video aufgezeichnet werden und später transkribiert werden (vgl. WINTER 2000).

Wichtig ist es, den Probanden eine Beschränkung der zu erstellenden Kategorien aufzuerlegen, um eine Vielzahl von Kategorien mit sehr

wenigen Suchanfragen zu vermeiden (vgl. OTTE UND VON MACH 2009, S. 65).

Der Moderator einer solchen Diskussion sollte geschult und in der Lage sein, die Gruppe immer wieder auf das eigentliche Thema zurückzubringen. Eine Gruppe entwickelt oft eine Eigendynamik und kommt vom Thema ab oder beißt sich an einzelnen Aspekten fest. Außerdem dominieren oft einzelne Mitglieder mit einem hohen Redeanteil die Diskussion, während andere Teilnehmer sich gar nicht äußern. Die Ergebnisse einer Gruppendiskussion hängen maßgeblich von der Moderation ab, die nicht zu rigide sein und die freie Gesprächsentfaltung nicht hemmen darf, aber gleichzeitig immer wieder den Zweck der Diskussion in den Fokus rückt (vgl. WINTER 2000).

Für die Durchführung einer solchen Studie wird eine Stichprobe von circa 50 Suchanfragen benötigt, die keine weiteren Informationen aus den Logfiles enthalten sollten. 50 Suchanfragen sind ausreichend, um eine Einteilung vorzunehmen und die Probanden nicht zu überfordern. Es sollte sich dabei um reale Suchanfragen handeln.

III.1.a Alternative Gruppendiskussion

Eine Alternative zu der oben beschriebenen Gruppendiskussion kann eine Einteilung der Probanden in verschiedene Gruppen sein. Sie sollten nach Vorbildung/Interneterfahrung aufgeteilt werden.

Die Kleingruppen sollten sich Gedanken über eine mögliche Themenbildung machen.

In einem zweiten Schritt sollten die Probanden ihre Themen vorstellen und erklären, warum sie sich für diese Einteilung entschieden haben.

Mit Hilfe eines Moderators kann dann eine Diskussion über die einzelnen Themenfelder der verschiedenen Gruppen abgehalten werden. Die Probanden können ihre Entscheidungen verteidigen und bei Übereinstimmungen kann man die Beweggründe der einzelnen Gruppen für die Bildung einer Kategorie vergleichen.

Während der Kategorienfindung und der Diskussion sollten die Argumente und evtl. verworfene Ideen aufgezeichnet werden.

Aus der Diskussion der einzelnen Ideen können am Ende gemeinsame Kategorien für die Einteilung der Suchanfragen gefunden werden.

III.2 Priorität und Durchführbarkeit

Das schwierigste an dieser Studie scheint das Finden geeigneter Probanden zu sein. Die Gefahr dabei ist, dass sie überfordert sein und außer einer thematischen Kategorisierung keine Ergebnisse liefern könnten. Durch geeignete Moderation und Impulse kann diesem Problem begegnet werden.

Ist die richtige Gruppenkonstellation gefunden und ein geeignetes Suchanfragenset zusammengestellt, sollte die Einführung und die Moderation den Probanden so weit helfen, dass sie brauchbare Ergebnisse hervorbringen. Die Gruppendiskussion kann in mehreren Durchgängen durchgeführt werden, so dass es z.B. fünf Gruppendiskussionen mit jeweils neuen Probanden gibt.

Da die erneute Untersuchung der gefundenen und viel genutzten Anfragetypen das Bestehen derselben bestätigen oder alternativ neue Ansätze für eine geeignete Trefferpräsentation auf den Suchergebnisseiten der Suchmaschinen liefern kann, wird dieser Studie eine relativ hohe Priorität zugeschrieben.

IV. Tagebuchstudie „Onlineaktivität“

Dai et al. stellten nach Beendigung ihrer Studie fest, dass es für die Evaluierung des Informationsbedürfnisses eines Nutzers sinnvoll wäre, die gesamte Session eines Nutzers zu beobachten (vgl. DAI ET AL. 2006, S.3). Dazu sollten alle Aktivitäten eines Nutzers aufgezeichnet werden, jede Handlung vom Beginn bis zum Ende einer Session. Die Vorgehensweisen des jeweiligen Nutzers in Verbindung mit dem Ergebnis, das sein Informationsbedürfnis befriedigt, könnten Rückschlüsse auf die Kategorisierung der Suchanfrage geben.

Beobachtet man die kompletten Onlineaktivitäten des Probanden könnte man z.B. transaktionsorientierte Anfragen an der durchgeführten Transaktion am Ende der Session erkennen. Recherchen zu einer möglichen Transaktion könnte man an Schlüsselverben wie „kaufen“ erkennen (vgl.

KANG UND KIM 2003, S. 70 und ASHKAN UND CLARKE 2009, S. 800). Wenn der Nutzer seine Suche unterbricht und später wieder aufnimmt, kann dies ein Hinweis darauf sein, dass er mit den Ergebnissen der ersten Suche nicht zufrieden war. Ebenfalls sichtbar wird der Wechsel zu einer anderen Suchmaschine.

Ist eine statistisch relevante Anzahl solcher Sessions von verschiedenen Probanden für eine spezielle Suchanfrage gesammelt worden, kann die Suchanfrage dem jeweiligen Anfragetyp zugeordnet werden (vgl. DAI ET AL. 2006, S.3).

IV.1 Methodik und Datengrundlage

Um eine statistisch relevante Anzahl von Sessions für eine Suchanfrage zu erhalten, werden eine entsprechende Probandenanzahl und eine gewisse Zeit der Datenaufnahme benötigt.

Um aussagekräftige Ergebnisse zu erhalten, wird empfohlen, zu 60 Suchanfragen jeweils drei bis vier Informationsbedürfnisse zu sammeln. Dazu wird eine Probandenanzahl von mindestens 70 Personen benötigt, da die Suchanfragen in der Stichprobe von mehreren Personen unabhängig gestellt werden müssen.

Um die Bewegungen und Aktivitäten der Probanden aufzuzeichnen, wird empfohlen, eine Software auf den Computern der Probanden zu installieren. Zusätzlich sollten sie ein Tagebuch über die durchgeführten Suchen führen (vgl. ROSE UND LEVINSON 2004, S. 19), in dem sie festhalten, was sie suchen wollten und ob ihre Suche erfolgreich war (vgl. OTTE UND VON MACH 2009, S. 67). Gerade wenn eine informationsorientierte Anfrage abgeschickt wird, wird sich der Proband mehrere Ergebnisse anschauen. Das Tagebuch wird dabei behilflich sein, die Ergebnisse der Suchergebnisseite als gut oder unbefriedigend einordnen zu können. In den Tagebüchern sollten die Probanden Fragen beantworten, die beschreiben, wonach sie gesucht haben und wie sie die Suche ausgeführt haben, um ihr Informationsbedürfnis zu befriedigen (vgl. TEEVAN ET AL. 2004, S. 416). Die Sessions (die Bewegungen im jeweiligen Browser) der Probanden sollten aufgezeichnet und codiert mit den Tagebuchaufzeichnungen verbunden werden, den Probanden sollte die Anonymisierung ihrer Daten

mittgeteilt werden, um den Bedenken bezüglich des Datenschutzes zu begegnen (vgl. HAWKEY 2009, S. 80).

So können die Aktivitäten mit der Zufriedenheit des Nutzers verglichen werden und es können Bewegungs- und/oder Klickmuster für einzelne Anfragen und später für einzelne Anfragetypen entwickelt werden.

Alternativ zu den Tagebucheinträgen könnte eine solche Studie in einem Labor durchgeführt werden, in dem die Aufzeichnungssoftware bereits installiert ist. Die Tagebücher sollten durch Fragebögen ersetzt werden.

Allerdings wird davon abgeraten, in dieser Situation mit Interviewern zu arbeiten, um die Konzentration der Probanden auf die Suche nicht zu stören.

Hier ist ebenfalls wieder zu beachten, dass die Probanden in diesen beobachteten Situationen eventuell anders agieren als sie es unbeobachtet tun würden (vgl. HÖCHSTÖTTER 2009, S. 178).

Für diese Studie werden codierte Logfiles benötigt, bei denen die IP-Adresse und sonstige Daten, die Rückschlüsse auf den Probanden zulassen könnten, entfernt und bspw. durch Codenamen oder -nummern ersetzt wurden. Außerdem sollten die Suchanfragen eines Probanden in eine chronologische Reihenfolge gebracht werden können, um Modifizierungen der Suchanfrage oder die erneute Eingabe zu einem späteren Zeitpunkt sehen zu können (vgl. ROSE UND LEVINSON 2004, S.16). Dafür sollte ein Zeitstempel in den Daten zu finden sein.

Wichtig ist, dass diese Codierung auf die Tagebücher bzw. Fragebögen übertragen wird, damit diese den Logfiles zugeordnet werden können.

Außerdem sollte zu Beginn der Studie eine Informationsveranstaltung abgehalten werden, in der die unterschiedlichen Anfragetypen erklärt werden.

So haben die Probanden die Möglichkeit, ihre Anfrage einem oder mehreren Anfragetypen zuzuordnen und die Forscher haben, nach der Entwicklung von Faktoren, die eine Suchanfrage einem Anfragetypen zuzuordnen können, einen Vergleich mit der Originalintention des Antragstellers (vgl. OTTE UND VON MACH 2009, S. 67).

IV.2 Priorität und Durchführbarkeit

Um Probanden bei einer solch zeitaufwändigen Studie zur Mitarbeit zu gewinnen, sollte ein finanzieller Anreiz geboten werden.

Technisch und methodisch ist diese Studie durchführbar. In früheren Forschungen wurde bereits erwähnt, dass hinter einer Suchanfrage bzw. einem Suchwort mehrere Informationsbedürfnisse stehen können (vgl. KANG UND KIM 2003, S. 66 und DAI ET AL. 2006, S. 8).

Bei dieser neuen Studie könnte eruiert werden, dass eine Suchanfrage viele verschiedene Informationsbedürfnisse haben kann, die je nach Nutzer unterschiedlich sind und dass die Ableitung der Intention durch Juroren in bereits durchgeführten Studien dadurch als nicht valide bezeichnet werden muss.

Interessant kann die Studie für die Platzierung der Werbung werden, wenn bei bestimmten Suchworten eher auf bezahlte Anzeigen geklickt wird als bei anderen. Dies könnte zu einem besseren Einsatz von Werbung und damit zu mehr Umsatz für die werbetreibenden Unternehmen führen (vgl. DAI ET AL 2006, S. 8). Hier kann herausgefunden werden, ob es Anfragen gibt, bei denen alle Nutzer mit einem Klick auf eine bezahlte Anzeige zufrieden sind und ob es Klickmuster gibt, die eine kommerzielle Anfrage identifizieren. Bei diesen Anfragen sollte die Werbung prominent platziert werden, da sie mit höherer Wahrscheinlichkeit geklickt wird (vgl. ASHKAN UND CLARKE 2009, S. 800).

9.2 Entwickelte Studien – Klassifikation von Suchanfragen nach Thema

Die Zuordnung von Suchanfragen zu verschiedenen Themen hat in der Vergangenheit gezeigt, dass die Suchmaschinen jedes Themenfeld bedienen können müssen (vgl. BRODER 2002, S. 9 und LEWANDOWSKI 2006, S. 39).

Jansen und Spink verglichen in ihrer Studie 2006 das Vorkommen verschiedener Themen mit früheren Ergebnissen. Sie stellten fest, dass sich die Themen der Websuche verändern bzw. verschieben (vgl. JANSEN UND SPINK 2006, S. 260 und CALDERON- BENAVIDES ET AL. 2010, S. 21).

Die Analyse der Themenfelder bietet die Möglichkeit, die Suchergebnisse an diese anzupassen. Es besteht die Möglichkeit, dass die Suchergebnisseiten nicht nur je nach Anfragetyp angepasst werden sollten, sondern auch nach Themen unterteilt unterschiedliche Ergebnisse präsentieren sollten.

V. Überprüfung der Themenzugehörigkeit der Suchanfragen

Die Veränderungen in den Themen der Websuche sollten wieder mit den Originalanfragestellern untersucht werden. Bisher wurden die Suchanfragen von Juroren in die verschiedenen Themen eingeordnet (LEWANDOWSKI 2006, S. 5; JANSEN UND SPINK 2006, S. 258 f. und OTTE UND VON MACH 2009, S. 20). Unstimmigkeiten der Juroren wurden dabei durch Diskussionen beseitigt. Otte und von Mach hatten die Möglichkeit, die unterschiedlichen Zuordnungen der Juroren einsehen und vergleichen zu können (vgl. OTTE UND VON MACH 2009, S. 21). Dabei stellten sie fest, dass Suchanfragen zwei verschiedenen Themen zugeordnet werden können, die beide einen Sinn ergeben würden, wenn man die Intention des Fragestellers nicht kennt (vgl. OTTE UND VON MACH 2009, S. 21 und CALDERON-BENAVIDES ET AL. 2010, S. 24).

Die Vielfältigkeit der Themen ist in den Studien ausreichend und sollte, schon der Vergleichbarkeit wegen, übernommen werden. So können die Einordnungen durch Juroren mit den Einordnungen der real Suchenden verglichen und Rückschlüsse auf die Validität der früher getätigten Kategorisierung vorgenommen werden.

V.1 Methodik und Datengrundlage

Diese Studie kann hervorragend mit Studie I „Laborexperiment zur Überprüfung der Einordnung von Suchanfragen durch Juroren“ verbunden werden. Die an Studie I teilnehmenden Probanden sollen befragt werden und können so auch gleichzeitig zu der Themenzugehörigkeit ihrer Suchanfrage befragt werden.

Dazu sollten in der Laborstudie die Themenfelder gut sichtbar angebracht werden und der Proband neben der Frage nach seinem Informations-

bedürfnis auch danach befragt werden, welchem Thema er die Suchanfrage zuordnen würde.

Anhand der Zuordnungen der Probanden können Faktoren ermittelt werden, die ein Thema erkennen lassen. Diese Faktoren sollten anhand einer großen Stichprobe überprüft werden.

Zur Durchführung dieser Studien werden aus den Logfiles die Suchanfragen der einzelnen Probanden mit Zeitangabe, um eine chronologische Reihenfolge herstellen zu können benötigt.

V.2 Priorität und Durchführbarkeit

In Verbindung mit Studie I „Laborexperiment zur Überprüfung der Einordnung von Suchanfragen durch Juroren“ ist diese Studie gut durchführbar.

Die Überprüfung der durch Juroren den Themen zugeordneten Suchanfragen ist sinnvoll, um eine Richtung für die weitere Forschung auf diesem Gebiet vorgeben zu können.

Zur Überprüfung der ermittelten Faktoren der Themenzugehörigkeit einer Suchanfrage sollte ebenfalls die verwendete Stichprobe aus Studie I herangezogen werden.

9.3 Entwickelte Studien – Suchverhalten von Suchmaschinennutzern/Einsatz qualitativer Methoden

Einen anderen Ansatz der Einteilung von Suchanfragen untersuchte Schmidt-Mänz in ihrer Studie. Sie untersuchte verschiedene Ereignisse und deren Faktoren. Sie schlägt vor, Suchterme unter der Beachtung der verschiedenen Eigenschaften von Dauerbrennern, Impulsen und Events zu untersuchen (vgl. SCHMIDT-MÄNZ 2007, S. 183).

Da diese Art der Einteilung noch nicht hinreichend erforscht wurde, aber als vielversprechend anzusehen ist, sollte hier eine große Studie angelegt werden.

VI Entwicklung optimaler Suchergebnisseiten

In dieser Studie sollen die gesuchten Themen nach Dauerbrenner, Impuls und Event eingeteilt werden. Diese Einteilungen können einerseits zur Findung geeigneter Vokabeln für die Webseiten dienen, um diese zu optimieren, andererseits zur Einteilung der Suchanfragen und einer, mit zeitlichen Aspekten verbundenen, unterschiedlichen Präsentation der Suchergebnisse auf der Suchergebnisseite.

Diese Ereignisse mit einem festen Datum (Events) oder solche, die immer aktuell sind (Dauerbrenner) oder die unvorhersehbaren (Impulse) (vgl. SCHMIDT-MÄNZ 2007, S. 160) könnten unterschiedliche Suchergebnisseiten benötigen.

Dies sollte anhand einer Studie eruiert werden, in der qualitative und quantitative Methoden zum Einsatz kommen.

VI.2 Methodik und Datengrundlage

Um eine solche Studie durchführen zu können, sollten bekannte Ereignisse und deren Bezeichnungen genutzt werden. Die Impuls-Schlagwörter sollten sich dabei auf ein Ereignis beziehen, das zeitnah geschehen ist.

Um zu überprüfen, ob die unterschiedlichen Ereignisse verschiedene Suchergebnisseiten brauchen, um das Informationsbedürfnis des Suchenden zu befriedigen, wird der Einsatz einer qualitativen Methode empfohlen. Um Faktoren für die Einteilung der Suchanfragen unter zeitlichen Aspekten zu finden, sollte eine Stichprobe (ca. 30.000 Suchanfragen) aus den Logfiles genommen werden, deren Klickhistorie (Klicks auf welches Ergebnis, Ergebnis organisch oder Anzeige) und Häufigkeiten über einen längeren Zeitraum aufgezeichnet wurden. Außerdem sollte hier die ganze Session des Nutzers zur Verfügung stehen, um die getätigten Klicks und evtl. Modifikationen der Suchanfrage erkennen zu können.

Anhand dieser Daten sollten Faktoren ermittelt werden, die die Events, Dauerbrenner und Impulse unterscheidbar machen. Danach sollte untersucht werden, wie lange sich ein hohes Interesse an bspw. einem Impuls-Schlagwort hält und ob sich das Klickverhalten der Probanden über die Zeit ändert.

Ob bei dieser Art der Einteilung der Suchanfragen unterschiedliche Suchergebnisseiten zu einer Befriedigung des Suchenden führen, sollte durch eine qualitative Methode herausgefunden werden.

In diesem Fall dient die Studie zur Ideenfindung und zur Entwicklung neuer Ansätze. Hier geht es darum, durch die Probanden neue Eindrücke für die Gestaltung der Suchergebnisseite zu bekommen und die durch die Studie gewonnenen Einflüsse weiterzuverwenden und zu optimieren (vgl. GAVER ET AL. 1999, S.29).

Es sollen noch keine Lösungen designt werden, sondern Möglichkeiten neuer Lösungsansätze gefunden werden (vgl. GAVER ET AL. 1999, S.25). Die Grundidee für diese Studie wurde aus dem bibliothekarischen Bereich übernommen. Ritter und Goel führten in der Carnegie Library of Pittsburgh eine Studie durch, um dort mit qualitativen Studien die Services und Angebote zu verbessern (RETTIG UND GOEL 2005, S. 17 f.). Dazu machten die Nutzer Fotos und beschrieben, wie sie sich in den gezeigten Situationen fühlten (Abb. 1 „I am stupid...“). Auf die gemachten Bilder schrieben sie, was sie stört (Abb. 1 rote Schrift) und was sie verbessern würden bzw. was ihnen fehlt (Abb. 1 schwarze Schrift) (vgl. RETTIG UND GOEL 2005, 28 f.)

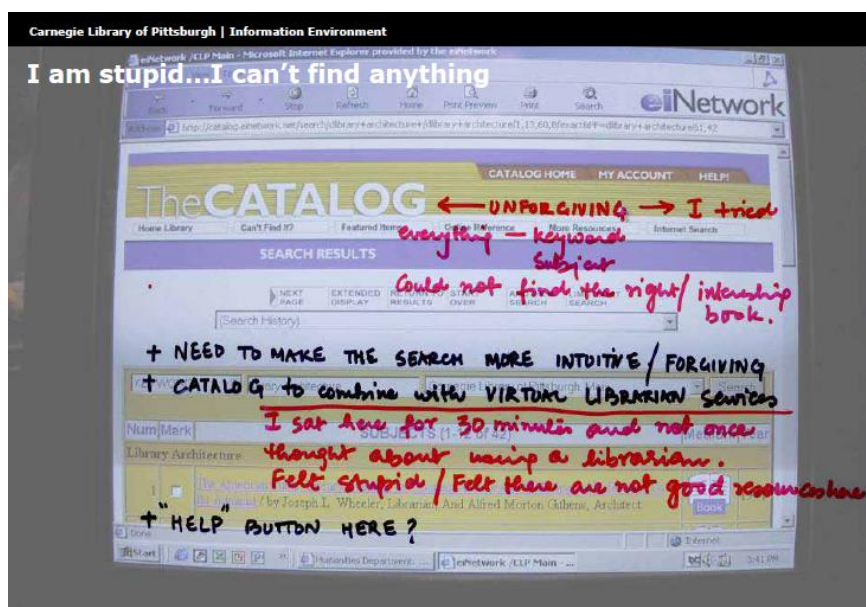


Abbildung 1: Nutzerfoto mit Beschriftung, eigene Darst., 2010 nach RETTIG UND GOEL 2005, S. 31

Statt die Probanden in der hier entwickelten Studie Fotos anfertigen zu lassen, wie die ideale Suchergebnisseite bei z.B. einer Impuls-Suche aussehen sollte, sollen sie hier ihre eigene ideale Trefferseite basteln.

Es wird empfohlen, diese Studie mit zehn Probanden nacheinander durchzuführen (vgl. GAVER ET AL. 1999, S.22).

Den Probanden sollten verschiedene Pappkarten mit Symbolen bekannter sozialer Dienste und angedeuteten Suchergebnisbeschreibungen (Abstracts auf der Suchergebnisseite – Blindtexte) zur Verfügung gestellt werden. Dann sollen ihnen Suchanfragen aus den verschiedenen Bereichen der zeitlichen Einordnung genannt werden. Die Probanden müssten aus den Pappkarten ihre optimale Suchergebnisseite zusammensetzen. Es sollte ihnen möglich sein, auf einer oder mehreren Pappkarten eigene Ideen/Wünsche für die Suchergebnisseite zu notieren, bspw. aus Datenbanken oder anderen Ressourcen.

Bei dieser Studie kann herauskommen, dass z.B. bei Impuls-Suchen Einträge aus Foren zu diesem Thema oder Echtzeitergebnisse aus sozialen Netzwerken wie Twitter, Facebook oder anderen Web 2.0-Diensten am interessantesten für die Suchenden sind und neben aktuellen Nachrichten zur Stillung ihres Informationsbedürfnisses ausreichen.

Nach der ersten Ideensammlung durch diese Probanden können für die unterschiedlichen „zeitlichen Anfragetypen“ verschiedene Suchergebnisseiten designt werden, welche dann anhand weiterer Studien auf ihre Anwendbarkeit untersucht werden sollten. So kann auch herausgefunden werden, wo Werbung platziert werden sollte und bei welchem dieser Anfragetypen die Werbung am ehesten toleriert wird (vgl. HÖCHSTÖTTER 2007, S. 140).

Diese Studie kann auch anhand der bestehenden Anfragetypen durchgeführt werden und könnte zeigen, dass die Weiterleitung auf eine bestimmte Webseite bei navigationsorientierten Anfragen genau das ist, was der Nutzer wünscht.

VI. 3 Priorität und Durchführbarkeit

Die Suchergebnisseite für unterschiedliche Suchbedürfnisse zu optimieren, ist eines der Hauptziele dieses Forschungsbereichs.

Durch die Aufnahme neuer Ansätze zur Ergebnispräsentation und die Kombination verschiedener Methoden könnte diese Studie zwar eine veränderte Art der Unterteilung der Suchanfragen hervorbringen, die den Suchenden aber dennoch zufriedenstellt, indem bspw. Ergebnisse aus verschiedenen (Daten-)Kollektionen angezeigt werden. Aus diesem Grund wird dieser Studie eine hohe Priorität eingeräumt.

10. Empfehlungen für die Bereitstellung von Logfiledaten

Um die Forschung und das Verständnis für das Informationsbedürfnis hinter einer Suchanfrage, die an eine Suchmaschine gestellt wird, zu erhöhen, sollte die Forschung in diesem Bereich weitergeführt werden.

Dies kann mit der Durchführung der oben beschriebenen Studien erfolgen.

10.1 Datencharakteristika

Aus den Logfiles lassen sich auch Daten gewinnen, anhand derer sich die jeweils genutzte Datenstichprobe überprüfen lässt (vgl. JANSEN ET AL. 2009, S. 6 f.).

Dabei handelt es sich um Charakteristika, die die Beschaffenheit der Stichprobe beschreiben und sie vergleichbar mit anderen Stichproben machen. Dazu zählen die Wörter die pro Anfrage verwendet wurden (vgl. SCHMIDT-MÄNZ 2007, S. 27). In Verbindung mit der Einteilung in Anfragetypen und/oder Themen kann die Aufgliederung Faktoren aufzeigen, die zur Erkennung des jeweiligen Typs/Themas beitragen. So sind navigationsorientierte Anfragen oft die kürzesten in einer Untersuchung, da hier eine bestimmte Webseite im Kopf des Nutzers ist, die er finden möchte. Dafür benötigt er nur (etwas mehr als) ein Wort (OTTE UND VON MACH 2009, S. 15). Um eine Untersuchung zur Einteilung der Suchanfragen durchzuführen, werden die Suchanfragen aus den Logdateien benötigt (so wie sie vom Nutzer eingegeben wurden).

Außerdem sollte die Nutzung von Phrasensuche und Operatoren untersucht werden. Unter einer Phrasensuche versteht man üblicherweise Begriffe, die gemeinsam in Anführungszeichen stehen (vgl. SCHMIDT-MÄNZ 2007, S. 79, SPINK ET AL. 2001, S. 229 und OTTE UND VON MACH 2009, S. 16).

Als Operatoren werden die Booleschen Operatoren genannt. Sie werden als AND, OR oder NOT definiert und geben Aufschluss darüber, wie verortet die Nutzer in der (Online-)Recherche sind (vgl. SPINK ET AL. 2001, S. 229).

Auch die Nutzung von Wortpaaren sollte beachtet werden, um so z.B. Personensuchen zu erkennen (vgl. JANSEN 2009, S. 110).

10.2 Logfiledaten für entwickelte Studien

Um die beschriebenen Forschungsvorhaben realisieren zu können, sollten bestimmte Daten aus den Logfiles für die Forschung zur Verfügung gestellt werden.

Vorausgesetzt wird dabei, dass in den Logfiledaten Faktoren wie die IP-Adresse (ggf. durch eine Codierung ersetzt), das Datum und die Uhrzeit sowie die Suchanfragen des jeweiligen Nutzers enthalten sind (vgl. HÖCHSTÖTTER 2009, S. 180, KAMMERER 2002 und vgl. TANG ET AL. 2007, S. 2), so dass z.B. die Möglichkeit besteht, die Suchanfragen eines Nutzers in eine chronologische Reihenfolge zu bringen.

Die Logfiledaten zeigen so auch die Klickhistorie des jeweiligen Nutzers, von welcher Ergebnisseite er auf welche Webseite geklickt hat und evtl. von dieser wieder auf die Suchergebnisseite zurückkehrte. Die Modifizierung oder komplette Neueingabe einer Suchanfrage wird so auch festgehalten.

Bei den verschiedenen qualitativen Studien ist es wichtig, dass die Daten der jeweiligen Nutzer mit den anderen genutzten Materialien (Tagebücher etc.) in Verbindung gebracht werden können, um die Auswertung der Daten möglich zu machen. Hierzu nutzt man am besten Aufzeichnungssoftware, um die Interaktionen der Nutzer festzuhalten. In einer Laborstudie können die Daten auch je nach genutztem Computer den einzelnen Probanden zugeordnet werden.

Für eine Studie, die sich mit der Einteilung unter zeitlichen Aspekten beschäftigt, werden Logfiles benötigt, die die Suchanfragedaten über einen langen Zeitraum aufgezeichnet haben. Hier sollte die Möglichkeit bestehen, die Suchanfragen nach Schlagworten zu ordnen und nach der Zeit, zu der die Anfrage abgeschickt wurde sowie die zugehörige Klickhistorie in die Forschung mit einbeziehen zu können.

Aus der Klickhistorie sollte hervorgehen, ob es sich um einen organischen Treffer oder um eine bezahlte Anzeige aus der Suchergebnisliste handelt.

Unabdingbar für die weitere Forschung ist die Nutzung der Session der einzelnen Nutzer. Sie zeigen das gesamte Set von Suchanfragen eines Nutzers, die er über die aufgezeichnete Zeit eingeben hat. Dieses Set kann aus einer Suchanfrage, aus mehreren unigen Anfragen oder aus modifizierten Suchanfragen bestehen (vgl. SPINK ET AL. 2001, S. 3).

Die Beachtung des gesamten Suchverlaufs und der getätigten Klicks sowie der Rückkehr auf die Suchergebnisseite, nachdem eine Webseite angeklickt wurde, zeigt, wie der jeweilige Nutzer letztendlich sein Informationsbedürfnis befriedigt hat (vgl. KANG UND KIM 2003, S. 70).

Um ein gutes Bild der Suche eines Nutzers und somit auch seines Informationsbedürfnisses zu erhalten, sollten die Suchanfrage, die Ergebnisse der Suchmaschine, die geklickten Ergebnisse und weitere Suchen oder Aktionen des Nutzers aufgezeichnet werden (vgl. ROSE UND LEVINSON 2004, S.16).

Die einfache Betrachtung des eingebenden Suchwortes reicht nicht aus, um die hinter der Suchanfrage stehende Intention des Suchenden zu vermuten (vgl. CALDERON- BENAVIDES ET AL. 2010, S. 24).

Außerdem lässt sich auch an den Sessions nicht die Zufriedenheit eines Nutzers oder die tatsächliche Intention hinter der Suchanfrage ablesen. Zur Erfassung der Empfindungen oder subjektiven Elemente eines Suchvorgangs sollten daher ergänzende Methoden wie Fragebögen, Online-Befragungen oder Ähnliches genutzt werden (vgl. JANSEN 2009, S. 103 f.).

Einer der wichtigsten Faktoren der entwickelten Studien ist die Einteilung von Suchanfragen durch den Originalanfragesteller, egal ob die Zuordnung nach Anfragetyp oder nach Thema. In der Literatur wurde der Aspekt der

durch Juroren zugeordneten Suchanfragen immer wieder als Mangel der genutzten Methode benannt (vgl. z.B. KANG UND KIM 2003, S. 70, CALDERON- BENAVIDES ET AL. 2010, S. 22 ff., DAI ET AL. 2006, S. 8 und JANSEN ET AL. 2007, S. 1262). Daher wird den Forschungsarbeiten, die mit der Einordnung von Suchanfragen durch den Originalanfragesteller arbeiten, eine besonders hohe Priorität eingeräumt. Die Einteilung in die Anfragetypen navigationsorientiert, transaktionsorientiert und informationsorientiert sollte ebenfalls überprüft und gegebenenfalls durch andere Klassifizierungen der Suchanfragen ersetzt werden. Den Studien, die sich mit der Klassifikation der Suchanfragen und der Entwicklung neuer Einteilungen oder der Verbesserung der Suchergebnisseite beschäftigen, wird daher ebenfalls eine hohe Priorität zugewiesen.

11. Fazit

Die Arbeiten im Bereich der Suchmaschinenforschung, die sich mit der Intention eines Suchenden hinter seiner Suchanfrage beschäftigen, haben Studien und Artikel hervorgebracht, die interessante Ansätze zu weiteren Forschungen aufzeigen. Fast jeder der Forscher zeigt in den Veröffentlichungen auf, welche Schwächen in seiner Methodik, seiner Datenstichprobe oder seiner Ergebnisfindung vorhanden waren.

Durch die hier durchgeführte Analyse der wichtigsten und innovativsten Arbeiten konnten diese Mängel gesammelt dargestellt werden und führten zu neuen Ideen oder Vorschlägen, die Studien in abgewandelter Form zu wiederholen bzw. deren Ergebnisse durch die Abänderung der Methode zu überprüfen.

Allerdings können hier die unterschiedlichen Benennungen der Suchanfragetypen und die Unterschiede in den genutzten Methoden nicht behoben werden, wodurch die Vergleichbarkeit der Studien oft schwierig ist (vgl. JANSEN ET AL. 2007, S. 1254). Hier hätte es von Anfang an die Einigung auf gültige und einheitliche Benennungen geben müssen, um einen Vergleich der Ergebnisse zu garantieren.

Die untersuchten Suchanfragen in den verschiedenen Studien werden oft nur ein oder zwei Kategorien zugeordnet. So wird außer Acht gelassen, dass sie durchaus mehrere Bedeutungen haben können und die Betracht-

tung des gesamten Informationsbedürfnisses ist nicht gegeben (vgl. CALDERON- BENAVIDES ET AL. 2010, S. 21). Baeza-Yates et al. gehen davon aus, dass es bei der Formulierung des Informationsbedürfnisses bei der Eingabe einer Suchanfrage in eine Suchmaschine Mängel seitens des Benutzer geben kann (vgl. BAEZA-YATES ET AL. 2006, S. 106). Dies trägt dazu bei, dass die Nutzer unterschiedliche Schlagwörter für dasselbe Informationsbedürfnis gebrauchen und sich die durch die Suchmaschine generierten Ergebnisse in ihrer Qualität stark unterscheiden (vgl. KANG UND KIM 2003, S. 64). Die tatsächliche Zufriedenheit und die Intention hinter der Suchanfrage kann nur dann herausgefunden werden, wenn die Nutzer direkt während der Suche befragt werden und zu ihren eigenen Suchanfragen, Informationsbedürfnissen und ihrer Zufriedenheit mit der Suche und den Suchergebnissen Auskunft geben. Baeza-Yates et al. stellten den Juroren, die in ihrer Studie Suchanfragen in Themen einordnen sollten, die geklickten Ergebnisse der Originalanfragesteller zur Verfügung, um ihnen die Einordnung in die Themen zu erleichtern. Die Identifikation des Informationsbedürfnisses fiel so leichter (vgl. BAEZA-YATES ET AL. 2006, S. 108). Jedoch fehlt hier die Möglichkeit der Überprüfung des tatsächlichen Informationsbedürfnisses durch Vergleichsdaten mit Suchanfragen, die durch die Originalanfragesteller zugeordnet wurden, um die Einordnung durch Juroren zu überprüfen.

Die Entwicklung neuer Methoden zeigt, dass sich einige Schwächen der bereits durchgeführten Studien und Experimente bspw. durch die Auswahl der richtigen Probanden vermeiden lassen können. Als enorm wichtig wird hier angesehen, dass die Fragesteller selbst zu ihrem Informationsbedürfnis hinter der Suchanfrage befragt werden, um die Frage zu klären, ob die Beurteilung einer Suchanfrage und die anschließende Zuordnung zu einem Anfragetyp oder Thema durch Juroren überhaupt möglich ist. Sollte eine solche Überprüfung ein positives Ergebnis hervorbringen, wären die Bedenken bezüglich dieses Vorgehens, die viele Forscher selbst äußern, eliminiert (vgl. z.B. KANG UND KIM 2003, S. 70, CALDERON- BENAVIDES ET AL. 2010, S. 22 ff., DAI ET AL. 2006, S. 8 und JANSEN ET AL. 2007, S. 1262). Ein negatives Ergebnis würde die Forschungsergebnisse der Studien, die diese Methode nutzen, in Frage stellen.

Die Sammlung der Daten aus Logfiles hat den Vorteil, dass die Daten in der natürlichen Umgebung des Nutzers gesammelt werden und die Nutzer, durch die Unaufdringlichkeit der Methode nicht merken, dass ihre Daten gesammelt werden und dementsprechend natürlich handeln (vgl. JANSEN ET AL. 2009, S. 6).

Für die Erforschung, welche Suchterme genutzt werden, wie viele Wörter pro Suchanfrage verwendet wurden, für die Evaluierung sonstiger Datencharakteristika und die Entwicklung von Klickmustern eignen sich die Daten aus Logfiles. Jedoch ist die alleinige Nutzung von Logfiles für die Erforschung der Intention einer Suchanfrage ungeeignet (vgl. TANG ET AL. 2007, S. 4).

Ein Nachteil der Sammlung von Daten aus Logfiles ist, dass die Forscher nicht nachfragen können, ob der Nutzer die gewünschten Suchergebnisse auf der ersten Suchergebnisseite gefunden hat, ob er seine Suche abgebrochen hat oder welche Gründe es gegeben hat, warum er sich nur die erste Ergebnisseite angesehen hat (vgl. JANSEN ET AL. 2009, S. 6).

Wenige Forscher haben in ihren Studien qualitative Methoden angewandt. Da es bei dem Informationsbedürfnis der Suchenden um die subjektive Wahrnehmung jedes einzelnen Nutzers geht, sollte sich die Forschung in Zukunft auch den qualitativen Methoden widmen, um zumindest eine Grundlage der echten Meinung der Menschen hinter der Suchanfrage zu haben. Die aus qualitativen Methoden entwickelten Faktoren oder Ansätze können an großen Datenstichproben verglichen werden. So werden Ergebnisse generiert, die den Bedürfnissen der wirklichen Nutzer einer Suchmaschine entsprechen und nicht denen der Juroren, die fremde Suchanfragen zuordnen.

Hargittai bemängelte diesen Umstand in ihrer Studie von 2002 und suchte Probanden, die keine explizite Vorbildung im Umgang mit Suchmaschinen hatten (vgl. HARGITTAI 2002, S. 1239 f.). Ihre Ergebnisse zeigen, dass die normale Internetpopulation anders sucht als Studenten, die sonst oft als Probanden für Studien eingesetzt wurden (vgl. LEE ET AL. 2005, S. 391 f, SPINK ET AL. 2001, S. 13, JANSEN UND MCNEESE 2005, S. 1486 und OTTE UND VON MACH 2009, S. 61). Forscher sollten bei der Rekrutierung von Probanden darauf achten, einen Querschnitt durch die (Internet)Population

zu erreichen, um die Bedürfnisse aller Nutzergruppen berücksichtigen zu können. Zwar ist dieses Vorgehen aufwendiger als die Anwerbung von Studenten, jedoch ist die Nutzergruppe „Student“ durch das Studium eher befähigt, Informationen zu finden und Recherchen durchzuführen als andere Nutzergruppen. Ihr Wissenstand ist höher und die Versiertheit im Umgang mit dem Internet oft besser als bei anderen Nutzergruppen. Die Ergebnisse solcher Studien spiegeln nicht die Belange anderer Nutzergruppen wider.

Probanden, die freiwillig an Studien teilnehmen, also außerhalb von Seminaren, sind eher bereit, ihre Daten preiszugeben und nicht so gehemmt, ihre Onlineaktivitäten beobachten zu lassen (vgl. HAWKEY 2009, S. 82). Dies bedeutet jedoch nicht, dass sie ihr Verhalten nicht ändern könnten, wenn sie an einer Studie teilnehmen.

Schmidt-Mänz untersuchte 2007 den Umgang und die Suchstrategien verschiedener Nutzergruppen mit dem Internet. Sie definierte die Internetgeneration als alle Personen bis 29 Jahre (vgl. SCHMIDT-MÄNZ 2007, S. 108). Hier wäre es interessant, in einer Studie (bspw. Studie III Überprüfung der Suchanfragen-Klassifikation nach Broder 2002), die Gruppe der „alten“ Internetgeneration der nachwachsende Jugendgeneration gegenüberzustellen. Die Jugend ist noch früher an die Nutzung des Internets herangeführt worden und hat möglicherweise eine andere Art zu recherchieren oder nutzt andere Wege, um Informationen zu finden, als die ältere Internetgeneration. Da schon die ältere Internetgeneration weniger Interesse hat, Zeit in die Recherche zu investieren als andere Nutzergruppen, sollte geprüft werden, ob sich dieser Trend verstärkt hat und die Bereitstellung von Suchergebnisseiten, die die perfekten Suchergebnisse für eine Suchanfrage generiert, noch wesentlicher geworden ist (vgl. SCHMIDT-MÄNZ 2007, S. 110 f.). Da die Gruppe der Jugendlichen die Nutzer von morgen sind, wäre eine Evaluation ihrer Vorlieben ein interessantes Nebenprodukt der Studie.

Die Empfehlung dieser Arbeit lautet, die gesamte Session eines Nutzers aus den Logfiledaten zu nutzen. Die Zuordnung der einzelnen Anfragen zum jeweiligen Nutzer erfolgt durch die IP-Adresse bzw. durch willkürlich

vergebene Identifikationsnummern. Die Bewegungen des Nutzers, wenn er bspw. die Suchergebnisseite verlässt, ein Ergebnis ansieht und danach wieder auf die Suchergebnisseite zurückkehrt, werden von den Logfiles der Suchmaschine (normalerweise) mit der identischen Nutzeridentifikation aufgezeichnet (vgl. JANSEN 2009, S. 102). So lässt sich die Anzahl der angesehenen Suchergebnisse eines Nutzers sowie die Anfragen und die getätigten Aktionen, wie die Modifikation oder Verfeinerung einer Anfrage durch das Hinzufügen oder Entfernen einzelner Wörter, nachvollziehen. Dies vermittelt ein besseres Verständnis der Interaktionen zwischen dem Suchenden und der Suchmaschine. Dadurch können Trends entdeckt werden und die typischen Interaktionen zwischen dem Nutzer und der Suchmaschine identifiziert werden. Man kann z.B. herausfinden, ob bestimmte Dateitypen (pdfs oder Videos) verstärkt bei bestimmten Suchwörtern geklickt werden. Zur Zufriedenheit der Suchmaschinennutzer könnte dann beitragen in diesen bestimmten Fällen diese Dateitypen auf den ersten Rängen der Suchergebnisseite zu finden (vgl. JANSEN 2009, S. 102).

Um der Kritik bei der Nutzung der Logfiles zu begegnen, nämlich, dass diese die Emotionen und Vorstellungen der Nutzer nicht erfassen können, sollte die Erforschung der Logfiles, wenn es um die subjektiven Belange der Nutzer geht, mit einer empirischen Forschungsmethode kombiniert werden (vgl. JANSEN 2009, S. 103 f.). Auch lassen sich die Nutzer in langen Aufzeichnungsphasen nicht zweifelsfrei identifizieren. Die IP-Adresse des Computers bleibt zwar dieselbe, jedoch können in der aufgezeichneten Zeit der Logfiles verschiedene Personen den Computer benutzen (vgl. JANSEN 2009, S. 115). Auch diesem Problem kann mit der Verwendung einer weiteren Methode, die die Aktivitäten eines Nutzers identifiziert (z. B. Tagebücher, aus denen die Suchanfragen oder Aktionen einer Person hervorgehen) begegnet werden.

Zu diesem Schluss kommen auch verschiedene Autoren, die empfehlen, die Klassifikationen von Suchanfragen anhand des gesamten Suchverlaufs und der Handlungen der Nutzer vorzunehmen (vgl. KANG UND KIM 2003, S. 70) oder die Komplexität der Suche in einer Suchmaschine zu sehen und nicht die einfache Handlung (die Eingabe einer Suchanfrage) zu betrachten

(vgl. CALDERON- BENAVIDES ET AL. 2010, S. 24). Rose und Levinson empfehlen, Faktoren, wie die Suchanfrage, die Ergebnisse der Suchmaschine, geklicktes Ergebnis, weitere Suchen oder andere Aktionen des Nutzers immer in eine Studie mit einzubeziehen, um so ein gutes Bild der Suche und des Informationsbedürfnisses eines Nutzers zu bekommen (vgl. ROSE UND LEVINSON 2004, S.16).

Durch die Symbiose aus Wirtschaft und Wissenschaft mit der Hochschule für Angewandte Wissenschaften Hamburg, der Deutsche Telekom und der n-size GmbH entstehen die idealen Voraussetzungen, um die entwickelten Studien durchführen zu können. Die Partner aus der Wirtschaft sind in der Lage, die benötigten Datengrundlagen zur Verfügung zu stellen und die Hochschule für Angewandte Wissenschaften Hamburg kann in verschiedenen Projekten (bspw. Abschlussarbeiten) die benötigten Studien durch Studierende vorbereiten oder durchführen lassen.

Die gewonnen Erkenntnisse helfen den Partnern aus der Wirtschaft, ihre Angebote und Services zu verbessern. Die Deutsche Telekom kann z.B. die Präsentation der Suchergebnisse an neue Erkenntnisse aus den durchgeführten Studien anpassen, dadurch die Nutzerzufriedenheit erhöhen und ist in der Lage, die Nutzer zu halten bzw. neue zu gewinnen.

Literatur

ASHKAN UND CLARKE 2009

Ashkan, Azin; Clarke, Charles L.A : *Term-Based Commercial Intent Analysis*. - In: SIGIR'09, July 19–23, 2009, Boston, Massachusetts, USA. - S. 800 – 801

BAEZA-YATES ET AL. 2006

Baeza-Yates, Ricardo; Calderon-Benavides, Liliana; Gonzalez-Caro, Cristina : *The Intention Behind Web Queries*. In: String Processing and Information Retrieval, 2006, Volume 4209/2006. S. 98 – 109.

BRODER 2002

Broder, Andrei: *A taxonomy of web search*. In: SIGIR Forum 36 (2002), Nr. 2., S. 3-10 – Online- Ressource:
<http://www.acm.org/sigir/forum/F2002/broder.pdf>. - letzter Abruf: 2010-10-18

CALDERON- BENAVIDES ET AL. 2010

Calderon- Benavides, Liliana; Gonzalez-Caro, Cristina und Baeza-Yates, Ricardo : *Towards a Deeper Understanding of the User's Query Intent*. – In: SIGIR 2010 Workshop on Query Representation and Understanding. Geneva, Switzerland 2010. – S. 21 – 24. Online Ressource:
http://ciir.cs.umass.edu/sigir2010/qru/QRU_proceedings.pdf. letzter Abruf: 2010- 11- 30

DAI ET AL. 2006

Dai, Honghua, Zhao, Lingzhi, Nie, Zaiqing, Wen, Ji-Rong, Wang, Lee, Li, Ying: *Detecting Online Commercial Intention (OCI)*. World Wide Web Conference, May 23-26. 2006, Edinburgh, Scotland. – S. 1-9. – Online Ressource: <http://research.microsoft.com/en-us/um/people/znie/WWW2006-oci.pdf>. letzter Abruf: 2011-01-11

GAVER ET AL. 1999

Gaver, Bill; Dunne, Tony; Pacenti, Elena : *Design : Cultural probes*. -In: interactions 6 (1999), Nr. 1. S. 21 – 29. Online Ressource:
<http://portal.acm.org/citation.cfm?doid=291224.291235>. letzter Abruf: 2011-02-04

HARGITTAI 2002

Hargittai, Eszter : *Beyond Logs and Surveys: In-Depth Measures of People's Web Use Skills*. In: Journal of the American Society for Information Science and Technology, 53(2002), Nr. 14. S.1239–1244

HAWKEY 2009

Hawkey, Kirstie : *Privacy Concerns for Web logging Data*. In: Jansen, Bernard J.; Spink, Amanda, Taksa, Isak : Handbook of research on Web log analysis. Hershey, PA : Information Science Reference, 2009. S. 80 – 98

HÖCHSTÖTTER 2007

Höchstötter, Nadine : *Suchverhalten im Web* : Erhebung, Analyse und Möglichkeiten. In: Information Wissenschaft und Praxis 58(2007), Nr. 3, S. 135-140. Online Ressource: http://www.bui.haw-hamburg.de/pers/ulrike.spree/ws2007_2008/suchverhalten.pdf. letzter Abruf: 2011-02-04

HÖCHSTÖTTER 2009

Höchstötter, Nadine : *Methoden der Erhebung von Nutzerdaten und ihre Anwendung in der Suchmaschinenforschung*. – In: Lewandowski, Dirk (Hrsg.) : Handbuch Internet Suchmaschinen. Heidelberg : AKA Verlag, 2009. S. 175 – 202. Auch als Online Ressource: <http://eprints.rclis.org/bitstream/10760/12739/1/Datenerhebung.pdf>. Letzter Arbuf: 2011-02-01

JANSEN 2009

Jansen, Bernard J. : *The Methodology of Search log Analysis*. In: Jansen, Bernard J.; Spink, Amanda, Taksa, Isak : Handbook of research on Web log analysis. Hershey, PA : Information Science Reference, 2009. S. 100 – 123

JANSEN ET AL. 2007

Jansen, Bernard J.; Booth, Danielle L.; Spink, Amanda: *Determining the informational, navigational, and transactional intent of Web queries*. In: Information Processing and Management 44 (2008), S. 1251–1266 Online-Resource: http://faculty.ist.psu.edu/jjansen/academic/pubs/jansen_user_intent.pdf. letzter Abruf: 2010-12-04

JANSEN ET AL. 2009

Jansen, Bernard J.; Spink, Amanda, Taksa, Isak : *Research and Methodological Foundations of Transaction Log Analysis*. In: Jansen, Bernard J.; Spink, Amanda, Taksa, Isak : Handbook of research on Web log analysis. Hershey, PA : Information Science Reference, 2009. S. 1 – 16

JANSEN UND MCNEESE 2005

Jansen, Bernard J. ; McNeese, Michael D. : *Evaluating the Effectiveness of and Patterns of Interactions With Automated Searching Assistance*. In: Journal of the American Society for Information Science and Technology, 56 (2005), Nr. 14, S. 1480–1503

JANSEN und SPINK 2006

Jansen, Bernard J.; Spink, Amanda: *How are we searching the World Wide Web? A comparison of nine search engine transaction logs*. In: Information Processing and Management 42 (2006), S. 248–263

KAMMERER 2002

Kammerer, Peter : *Das Logfile* : Grundlage statistischer Auswertung. - Petershagen / Eggersdorf, 2002–04-19. Online Ressource: http://www.kammerer-electronic.de/download/statistik_logfile.html. - letzter Abruf: 2011- 02-02

KANG UND KIM 2003

Kang, In-Ho ; Kim, GilChang: *Query Type Classification for Web Document Retrieval*. In: *SIGIR'03*, July 28–August 1, 2003. Toronto, Canada, S. 64-71
Online Ressource:
<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.58.5400&rep=rep1&type=pdf>. letzter Abruf: 2010-12-11

LEE ET AL. 2005

Lee, Uichin; Liu, Zhenyu; Cho, Junghoo : *Automatic Identification of User Goals in Web Search*. In: *WWW '05 Proceedings of the 14th international conference on World Wide Web (2005)*, S. 391 – 400
Online Ressource:
<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.59.3810&rep=rep1&type=pdf>. letzter Abruf: 2010-12-11

LEWANDOWSKI 2006

Lewandowski, Dirk: *Themen und Typen der Suchanfragen an deutsche Web-Suchmaschinen*. In: Lehner, Franz; Nösekabel, H.; Keinschmidt, P. [Hrsg.]: *Multikonferenz Wirtschaftsinformatik 2006*. S. 33-43 - Online Ressource: http://www.bui.haw-hamburg.de/fileadmin/user_upload/lewandowski/doc/Themen_Typen_Suchanfragen.pdf. letzter Abruf: 2010-10-18

LEWANDOWSKI 2010

Lewandowski, Dirk : *Query Understandin in Web Search* : Vorhabenbeschreibung : Projektantrag. – Hochschule für Angewandte Wissenschaften Hamburg : Hamburg, 2010

OTTE UND VON MACH 2009

Otte, Jessica; von Mach, Sonja: *Analyse von klassifizierten Suchanfragen mittels qualitativer und quantitativer Methoden* : Projektarbeit im Rahmen des Research & Venture Projekts an der Hochschule für Angewandte Wissenschaften Hamburg : Projektdokumentation. - Hamburg, Hochschule für Angewandte Wissenschaften Hamburg : 2009. - 94 S.

RAINIE UND JANSEN 2009

Rainie, Lee ; Jansen, Bernhard J.: *Surveys as a complementary method for web log analysis*. In: Jansen, Bernard J.; Spink, Amanda, Taksa, Isak : *Handbook of research on Web log analysis*. Hershey, PA : Information Science Reference, 2009. S. 39 - 64

RETTIG UND GOEL 2005

Rettig, Marc; Goel, Aradhana : *Designing for Experience* : Frameworks and Project Stories Designing Stories. - Washington, D.C., 2005. Online Ressource:
http://darmano.typepad.com/for_blog/rettiggoel.uxWeek.8.25.05.pdf. letzter Abruf: 2011-02-07

ROHER 2008

Rohrer, Christian : *When to Use Which User Experience Research Methods*. Jakob Nielsen, 2008. – Online Resource:
<http://www.useit.com/alertbox/user-research-methods.html>. - letzter Abruf: 2011-02-06

ROSE UND LEVINSON 2004

Rose, Daniel E. ; Levinson, Danny: *Understanding User Goals in Web Search*. In: WWW 2004, May 17–22, 2004, New York, New York, USA. S.13 - 19

SCHMIDT-MÄNZ 2007

Schmidt-Mänz, Nadine: *Untersuchungen des Suchverhaltens im Web: Interaktion von Internetnutzern mit Suchmaschinen*. Hamburg : Dr. Kovac Verl., 2007
ISBN 978-3-8300-2725-6

SPINK ET AL. 2001

Spink, Amanda ; Wolfram, Dietmar ; Jansen, Bernhard J. ; Saracevic, Tefko: *Searching the web : Thepublic and their queries*. In: Journal of the American Society for Information Science and Technology 52 (2001), Nr. 3, S. 226 – 234 Auch als Online Ressource:
<http://comminfo.rutgers.edu/~tefko/JASIST2001.pdf>. letzter Abruf: 2010-12-22

TANG ET AL. 2007

Tang, Diane ; Grimes, Carrie ; Russell , Daniel M.: *Query logs are not enough*. In: Proc. WWW 2007 Workshop on Query Log Analysis (2007), S. 1-7
Auch als Online Ressource: http://www2007.org/workshops/paper_51.pdf. - letzter Abruf: 2011-02-02

TEEVAN ET AL. 2004

Teevan, Jaime, Alvarado, Christine, Ackerman, Mark S., Karger, David R. : *The Perfect Search Engine Is Not Enough : A Study of Orienteering Behavior in Directed Search*. – In: Proceedings of the SIGCHI conference on Human factors in computing systems. Vienna/Austria, 2004. S. 415 422

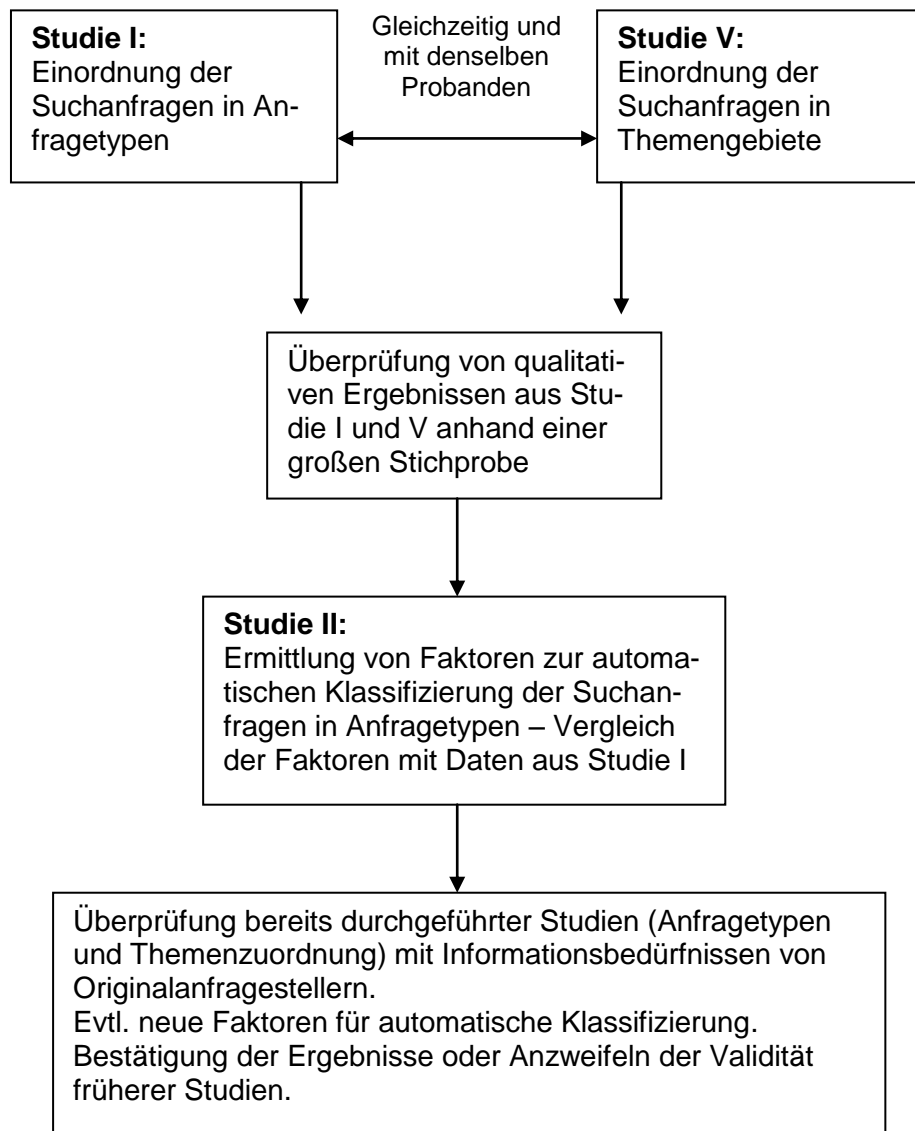
WINTER 2000

Winter, Stefanie : *Gruppendiskussion*. Universität Karlsruhe, 2000-05-19.
Online Ressource: http://imihome.imi.uni-karlsruhe.de/ngruppendiskussion_b.html. letzter Abruf: 2011-02-01

Anhang

1. Beispielhafte Kombination der entwickelten Studien

In der folgenden Illustration ist die Verbindung mehrerer Studien zu einem größeren Projekt dargestellt. Sie beruht auf den schriftlichen Ausführungen in Kapitel 9, Studie I, II und V.



Eidesstattliche Versicherung

Ich versichere, die vorliegende Arbeit selbständig ohne fremde Hilfe verfasst und keine anderen Quellen und Hilfsmittel als die angegebenen benutzt zu haben. Die aus anderen Werken wörtlich entnommenen Stellen oder dem Sinn nach entlehnten Passagen sind durch Quellenangabe kenntlich gemacht.

Hamburg, den 18. Februar 2011