Hochschule für Angewandte Wissenschaften Hamburg
Hamburg University of Applied Sciences

DEPARTMENT INFORMATION

Masterarbeit

Kein Klick, kein Treffer? Untersuchung und Analyse von Suchanfragen und Suchergebnisseiten, die keinen Klick auslösten

vorgelegt von

Jessica Drechsler geb. Otte

Studiengang Informationswissenschaft und-management

erster Prüfer: Prof. Dr. Dirk Lewandowski
zweite Prüferin: Prof. Dr. Ulrike Spree

Hamburg, November 2010

Abstract

Die Erforschung von Suchanfragen, die keinerlei Klick auf die Suchergebnisse erhalten haben, ist das Ziel der vorliegenden Arbeit. Zur Untersuchung der Anfragen gibt eine deskriptive Analyse sowie eine Online-Befragung zur Relevanzmessung Aufschluss über diese Suchanfragen und zeigt Gründe für das Unterlassen eines Klicks auf. Dabei macht die Relevanzmessung deutlich, dass das Unterlassen eines Klicks für einen signifikanten Anteil der Suchanfragen bedeutet, dass die Suche bereits auf der Trefferseite beantwortet werden kann. Mit Hilfe weiterer Untersuchungen konnten die Themen und Anfragetypen dieser Suchanfragen erhoben und charakterisiert werden. Der Einsatz der Methodenkombination macht aber auch deutlich, dass Suchanfragen ohne Klick häufig zu irrelevanten Treffern führen und daher unbeantwortet bleiben. Auch hier gibt eine detaillierte Analyse dieser Suchanfragen-Gruppe Aufschluss über Merkmale und Charakteristika dieser Anfragen.

Die Ergebnisse dieser Untersuchung müssen hinsichtlich ihrer Repräsentativität jedoch eingeschränkt werden, da sie nicht auf Grundlage einer Zufallsstichprobe erhoben werden konnten. Des Weiteren führen zeitliche Verzögerungen und die Beschränkung der Relevanzmessung auf die ersten zehn organischen Suchergebnisse dazu, dass die Resultate der Untersuchung nicht auf die Gesamtheit aller Suchanfragen umgelegt werden kann. Stattdessen geben sie Hinweise und Tendenzen für weitere Untersuchungen und machen deutlich, dass auf diesem Gebiet weiterer Forschungsbedarf besteht.

Schlagwörter

Analyse von Suchanfragen, Information Retrieval, Online-Befragung, Relevanzmessung, Suchanfragen ohne Klick, Web-Suche, Web-Suchmaschine

Inhalt

Abbildungsverzeichnis ..... **V**

Tabellenverzeichnis ..... **VII**

Abkürzungsverzeichnis ..... **VIII**

1 Einleitung **1**

1.1 Fragestellung und Hypothesen 3

1.2 Aufbau der Arbeit 7

2 Stand der Forschung **7**

2.1 Suchanfragen ohne Klick 8

2.2 Analyse von Klickdaten 12

2.3 Trefferbeschreibung und -evaluation 13

2.4 Relevanz & Evaluation von Relevanz 17

2.5 Erkenntnisse für die vorliegende Arbeit 19

3 Methodik **24**

3.1 Definition des Untersuchungsgegenstandes 26

3.2 Datenbasis für die deskriptive Analyse 26

3.3 Online-Befragung zur Ableitung des Informationsbedürfnisses 29

3.3.1 Begründung der Methodenwahl *30*

3.3.2 Konzeption *32*

3.3.3 Pretest *34*

3.3.4 Durchführung *35*

3.4 Online-Befragung zur Untersuchung der Relevanz der Suchergebnisse nicht geklickter Suchanfragen 36

3.4.1 Begründung der Methodenwahl *36*

3.4.2 Konzeption *38*

3.4.3 Pretest *42*

3.4.4 Durchführung *44*

4 Ergebnisse 44

4.1 Deskriptive Analyse der Datenbasis 44

4.1.1 Länge der Suchanfragen 45

4.1.2 Nutzung von Suchoperatoren 46

4.1.3 Bestandteile einer URL in der Suchanfrage 48

4.1.4 Vergleich von geklickten und ungeklickten Suchanfragen 49

4.1.4.1 Länge der Suchanfragen 50

4.1.4.2 Bestandteile einer URL in der Suchanfrage 52

4.1.4.3 Themen 53

4.1.4.4 Weitere Eigenschaften von Suchanfragen ohne Klick 58

4.1.5 Zusammenfassung der Ergebnisse der deskriptiven Analyse 61

4.2 Ergebnisse der Online-Befragung zur Ableitung des Informationsbedürfnisses 62

4.2.1 Vorgehen zur Ableitung des Informationsbedürfnisses 63

4.2.2 Vorgehen zur Bestimmung der Aktualität einer Suchanfrage 66

4.3 Ergebnisse der Online-Befragung zur Untersuchung der Relevanz der Suchergebnisse nicht geklickter Suchanfragen 68

4.3.1 Irrelevante Ergebnisse 73

4.3.2 Relevante Ergebnisse 77

4.3.3 Suchergebnis beantwortet Informationsbedürfnis 82

5 Diskussion und Schlussfolgerungen 90

5.1 Probleme und Einschränkungen 90

5.2 Suchanfragen ohne Klick 94

5.3 Empfehlungen & Ausblick 103

6 Fazit 108

Quellen 109

Anhang I 123

Anhang II 128

Anhang III 135

Abbildungsverzeichnis

Abb. 1: Anzahl der Wörter pro Suchanfrage (Quelle: eigene Darst.) 45
Abb. 2: Anzahl der Wörter pro Suchanfragen für Suchanfragen mit und ohne Klick (Quelle: eigene Darst.) 51
Abb. 3: Themen der Suchanfragen mit und ohne Klick (Quelle: eigene Darst.) ... 56
Abb. 4: Überblick über die im Relevanztest vergebenen Relevanzwerte (Quelle: eigene Darst.) 69
Abb. 5: Themen der Suchanfragen, deren ersten zehn Suchergebnisse im Relevanztest als irrelevant gewertet wurden (Quelle: eigene Darst.)..... 74
Abb. 6: Treffer der Suche „[redacted]“ [1](Quelle: SUCHE T-ONLINE.DE 2010). 75
Abb. 7: Suchanfragetypen der Suchanfragen, deren ersten zehn Suchergebnisse im Relevanztest als irrelevant gewertet wurden (Quelle: eigene Darst.) 76
Abb. 8: Anteil der Relevanzwerte der Treffer, deren Suchanfrage mindestens ein relevantes Ergebnis hat (Quelle: eigene Darst.) 78
Abb. 9: Kumulierter Anteil der relevanten Suchergebnisse je Rang (Quelle: eigene Darst.) 80
Abb. 10: Themen der Suchanfragen, die auf der SERP beantwortet werden konnten (Quelle: eigene Darst.) 84
Abb. 11: Anfragetypen der Suchen, die auf der Trefferseite beantwortet werden können (Quelle: eigene Darst.) 85
Abb. 12: Treffer der Suche „[redacted] “ (Quelle: SUCHE T-ONLINE.DE 2010) 87
Abb. 13: Treffer der Suche „66130“ (Quelle: SUCHE T-ONLINE.DE 2010) 88
Abb. 14: Treffer der Suche „wo lebt der jaguar?“ (Quelle: SUCHE T-ONLINE.DE 2010) 88
Abb. 15: Treffer der Suche „barmherzigkeit ist leichter zu üben als“ (Quelle: SUCHE T-ONLINE.DE 2010) 88
Abb. 16: Treffer der Suche „crashkurse“ (Quelle: SUCHE T-ONLINE.DE 2010) 89

[1] Suchanfragen, die Personennamen beinhalten, wurden in der gesamten Arbeit geschwärzt.

Abb. 17: Treffer der Suche „audio 625 usb-pc-headset“ (Quelle: SUCHE T-ONLINE.DE 2010) 89
Abb. 18: Suchergebnisseite der T-Online Suche mit Markierung der unterschiedlichen Ergebnisteile (Quelle: SUCHE T-ONLINE.DE 2010 und eigene Darst.) 91
Abb. 19: Treffer vier der Suche „wo lebt der jaguar?“ (Quelle: SUCHE T-ONLINE.DE 2010) 104
Abb. 20: Suchergebnis von T-Online bei der Suche nach „[redacted]“ (Quelle: SUCHE T-ONLINE.DE 2010) 106
Abb. 21: Einleitungstext der Befragung zur Ableitung des Informationsbedürfnisses (Quelle: eigene Darst.) 126
Abb. 22: Zweite Seite der Befragung zur Ableitung des Informationsbedürfnisses (Quelle: eigene Darst.) 126
Abb. 23: Erste Seite der Befragung zur Ableitung des Informationsbedürfnisses mit Validierungshinweisen (Quelle: eigene Darst.) 127
Abb. 24: Letzte Seite der Befragung zur Ableitung des Informationsbedürfnisses (Quelle: eigene Darst.) 127
Abb. 25: Erste Seite des Relevanztests mit Willkommens- und Aufgabentext (Quelle: eigene Darst.) 132
Abb. 26: Zweite Seite des Relevanztests: Ziel der Suche wird dem Probanden genannt (Quelle: eigene Darst.) 133
Abb. 27: Dritte Seite des Relevanztests: Treffer und Skalen zur Beurteilung des Treffers werden dem Probanden präsentiert (Quelle: eigene Darst.) 133
Abb. 28: Seite nach Abschluss der Beurteilung einer Suchanfrage: Zwischenstand wird dem Probanden kommuniziert (Quelle: eigene Darst.) 134
Abb. 29: Letzte Seite der Befragung: Dank an den Probanden (Quelle: eigene Darst.) 134

Tabellenverzeichnis

Tab. 1: Anzahl der Wörter und Anteil der Ein-Wort-Suchen für Suchanfragen mit und ohne Klick (Quelle: eigene Darst.) 50

Tab. 2: Merkmale geklickte und ungeklickter Suchanfragen im Vergleich (Quelle: eigene Darst.) 59

Tab. 3: Anzahl der Ableitungen pro Suchanfragen (Quelle: eigene Darst.)..... 62

Tab. 4: Beispiele zur Ableitung des Informationsbedürfnisses (Quelle: eigene Darst.) 65

Tab. 5: Aktualitätseinordnung der Suchanfragen (Quelle: eigene Darst.)..... 67

Tab. 6: Erste Übersicht zu den Ergebnissen der Relevanzmessung (Quelle: eigene Darst.) 68

Tab. 7: Auswertung des Relevanztest: Kombination von Relevanzwert und Klickverhalten (Quelle: eigene Darst.) 70

Tab. 8: Gründe für Suchanfragen ohne Klick – Überblick über die wichtigsten Ergebnisse (Quelle: eigene Darst.)..... 102

Abkürzungsverzeichnis

Abb.	Abbildung
CI	Computers or Internet (Computer und Internet)
CTEE	Commerce, travel, employment, or economy (Handel, Reise, Arbeit und Wirtschaft)
EH	Education or humanities (Bildung und Geisteswissenschaften)
ER	Entertainment or recreation (Unterhaltung und Freizeit)
GOV	Government (Regierung und Verwaltung)
HS	Health or sciences (Gesundheit und (Natur-)Wissenschaft)
PC	Personal Computer
PFA	Performing or fine arts (Kunst)
PPT	People, places and things (Personen, Orte und Dinge)
SCER	Society, culture, ethnicity, or religion (Gesellschaft, Kultur, Ethnizität und Religion)
SEO	Search engine optimization (Suchmaschinenoptimierung)
SERP	Search engine result page
SP	Sex or pornography (Sex und Pornographie)
Tab.	Tabelle
UNK	Unknown or other (Nicht bestimmbar und sonstige)
URL	Uniform Resource Locator

1 Einleitung

Die Präsentation von relevanten Treffern zu einer Suchanfrage ist das Ziel einer Web-Suchmaschine. Dabei entscheidet das Urteil des Suchenden, ob das Suchergebnis zu einer Suchanfrage als relevant oder irrelevant gewertet wird. Die Relevanz der präsentierten Treffer für das aktuelle Bedürfnis des Nutzers wiederum beeinflusst seine Zufriedenheit mit der Suchsitzung (vgl. HUFFMANN und HOCHSTER 2007, S. 567; REID 1999, S. 115). Die Evaluation der Suchergebnisse in Beziehung zur Suchanfrage ist also entscheidend, um die Qualität der Ergebnisse bewerten und daraufhin eine Optimierung der Suchfunktion vornehmen zu können. Allerdings ist eine direkte Abfrage der Relevanzeinschätzung des Suchenden häufig nur unter kontrollierten und damit alltagsfernen Bedingungen eines Laborexperiments oder einer Befragung möglich (vgl. z.B. AL-MASKARI et al. 2007, S. 773; GRIESBAUM et al. 2002, S. 210; HAWKING et al. 2001, S. 40; HUFFMANN und HOCHSTER 2007, S. 567). Bei einer solchen Untersuchung besteht daher immer die Gefahr, dass die Urteile der Probanden durch die Beobachtung verzerrt werden (vgl. ALBERS 2009, S. 40 f.; KROMREY 2009, S. 327 f. u. S. 378 f.). Um diese Problematik zu umgehen, wird, neben der Erforschung von Möglichkeiten zur Abgabe eines expliziten Feedbacks, implizites Feedback von Suchenden erhoben und auf seine Tauglichkeit zur Bewertung der Relevanz der Suchergebnisse untersucht (für eine Übersicht verschiedener Studien zum Thema vgl. KELLY und TEEVAN 2003). Als implizites Feedback werden z.B. die Verweildauer auf einer Suchergebnisseite (SERP = Search Engine Result Page), die Zeit, die der Suchende auf der Suchergebnisseite verbringt und die Art des Verlassens der Trefferliste eingesetzt (vgl. FOX et al. 2005. S. 154). Eine weitere Form des impliziten Feedbacks, welche in den letzten Jahren intensiv erforscht wurde, ist die Auswertung von Klickdaten als implizite Relevanzurteile (vgl. z.B. CRASWELL und SZUMMER 2007; GRANKA et al. 2005; JOACHIMS 2002; JUNG et al. 2007). Dabei wird davon ausgegangen, dass die Klicks der Suchenden als implizite Relevanzurteile genutzt werden können, um Auskunft über die Präferenzen der Nutzer über bestimmte Treffer zu geben. Im

Rückschluss würde diese Annahme bedeuten, dass Suchergebnisse, die keinen Klick erhalten, weniger relevant sind als die angeklickten Treffer. Besonders negativ würde dann die Bewertung von Suchanfragen ausfallen, deren Ergebnisse keinerlei Klick erhalten. In diesem Fall müsste davon ausgegangen werden, dass der Nutzer keinen Treffer als relevant bewertet und die Suche daher unbefriedigend verlaufen ist. Diese Überlegungen werden von neueren Suchmaschinenentwicklungen durchkreuzt, denn der Trend geht hier in die Richtung, das Informationsbedürfnis des Nutzers bereits auf der Suchergebnisseite zu befriedigen. Dazu werden bspw. Wetter- und Karteninformationen, Wortdefinitionen, Börsenkurse und Bilder (vgl. z.B. GOOGLE – SUCHFUNKTIONEN 2010) bereits auf der Trefferseite eingebunden. Auf diese Weise muss der Nutzer in einigen Fällen nicht klicken, um die gewünschten Informationen zu erhalten, die Suchergebnisseite reicht aus, um sein Informationsbedürfnis zu befriedigen. Neben den bereits genannten Suchfunktionen, die dem Nutzer die Möglichkeit geben, ohne die Sichtung der dahinter liegenden Webseite auszukommen, bestehen vermutlich noch weitere Suchanfrage(typen), die ohne einen Klick auf den Treffer Informationsbedürfnisse beantworten. Es kann allerdings natürlich nicht davon ausgegangen werden, dass alle Suchanfragen, die keinen Klick erhalten, durch die bloße Sichtung der Ergebnisse beantwortet werden. Thema und Ziel dieser Arbeit ist es daher, die Suchanfragen und die dazugehörigen Suchergebnisseiten, die keinen Klick erhalten haben, näher zu untersuchen und sich dabei nicht nur auf die Suchanfragen zu konzentrieren, deren Antwort bereits auf der Suchergebnisseite gefunden werden konnte, sondern auch auf die Suchanfragen, die den Suchenden aufgrund von irrelevanten Ergebnissen unbefriedigt zurücklassen. Im folgenden Kapitel werden die konkreten Fragestellungen und die dazugehörigen Hypothesen zu dem beschriebenen Arbeitsthema formuliert.

1.1 Fragestellung und Hypothesen

Wie bereits in der Einleitung beschrieben, wird dieser Arbeit das folgende Thema zu Grunde gelegt:

Untersuchung von Suchanfragen und den dazugehörigen Suchergebnisseiten, die keinen Klick auf die ausgelieferten Suchergebnisse erhielten.

Neben deskriptiven Informationen über diese Art der Suchanfrage, soll insbesondere untersucht werden, welche Gründe festgehalten werden können, wenn kein Suchergebnis angeklickt wird. Dabei werden die folgenden Hypothesen als Gründe für eine nicht geklickte Suchanfrage[2] angenommen:

1. Irrelevante Ergebnisse: Es werden nur Treffer(-beschreibungen) präsentiert, die für die Suchanfrage und das dahinter stehende Informationsbedürfnis irrelevant sind.

   Dieser Grund trifft auch dann zu, wenn davon ausgegangen werden kann, dass die Suchanfrage zu keinem passenden Ergebnis führt, da die Anfrage vom Suchenden mit zu wenigen oder nicht zielführenden Suchbegriffen ausgestattet wurde.

   Ebenfalls diesem Grund zuzuordnen sind Suchanfragen, die der Suchende wiederholt absendet, um zu prüfen, ob neue Suchergebnisse zu seinen Suchbegriffen vorliegen. Die wiederholten Suchen, die nicht zu neuen Suchergebnissen führen, können daher für den Suchenden ebenfalls als irrelevant eingestuft werden.

2. Fehler in der Suchanfrage: Die Suchanfrage enthält einen Fehler, welcher von der Suchmaschine erkannt wird und in einem Korrekturvorschlag resultiert. Entscheidet sich der Suchende für den Korrekturvorschlag seiner Suche wird kein Klick auf die Suchergebnisse verzeichnet.

[2] Eine „nicht geklickte Suchanfrage“ bedeutet in dieser Arbeit, dass die Suchergebnisse dieser Suchanfrage keinen Klick erhalten haben. Streng genommen sind die Suchanfragen also nicht ungeklickt, sondern deren Suchergebnisseiten. Um den Lesefluss der Arbeit zu verbessern, wird dieser Begriff eingesetzt.

3. Automatische Suchanfrage: Die Suchanfrage wird automatisch an die Suchmaschine versendet, ein Klick ist dabei nicht beabsichtigt. Ein solches Vorgehen kann z.B. im Bereich der Suchmaschinenoptimierung (SEO) beobachtet werden, hier werden Suchanfrage zur reinen Kontrolle der Suchergebnisposition eines SEO Kunden abgeschickt (vgl. HÜBENER 2009, S. 33f. u. S. 68).

4. Übersehen des/der relevanten Ergebnisse(s): Der Suchende nimmt nur die ersten Ergebnisse der Ergebnisseite wahr, werden diese von ihm als nicht relevant eingeordnet, bricht er seine Suche ab oder startet eine neue Suche. Weitere, evtl. relevante Ergebnisse, welche auf der Suchergebnisseite auf den unteren Postionen zu finden sind, werden übersehen. Ein solches Verhalten konnte in mehreren Studie (vgl. z.B. GUAN und CUTRELL 2007, S. 420; HOTCHKISS et al. 2005, S. 7f.; LORIGO et al. 2006, S. 1128) aufgedeckt werden.

5. Suchanfrage wird durch Trefferbeschreibung beantwortet: Die Suchergebnisseite mit ihren Trefferbeschreibungen enthält bereits die Antwort auf das Informationsbedürfnis des Suchenden, er muss die Antwort nicht durch die Exploration der einzelnen Dokumente heraussuchen. Studien von Stamou und Efthimiadis (vgl. STAMOU und EFTHIMIADIS 2009) und Li et al. (vgl. LI et al. 2009) kommen zu dem Schluss, dass ein signifikanter Anteil der Suchanfragen auf diese Weise beantwortbar ist bzw. beantwortet wird.

6. Abbruch der Suche durch äußere Einflüsse: Der Suchende wird aufgrund einer äußeren Beeinflussung, welche nicht in Zusammenhang mit der Suche steht, davon abgehalten seine Suche zu beenden.

Mit der vorliegenden Arbeit wird nun eine Stichprobe von ungeklickten Suchanfragen hinsichtlich der oben aufgeführten Motive untersucht, um anschließend analysieren zu können, zu welchen Anteilen die genannten Gründe vorliegen, wenn eine Suchanfrage keinen Klick erhält. Da die Situation, die zur letztgenannten Hypothese (Abbruch der Suche durch äußere Einflüsse) führt, nicht nachvollzogen werden kann, wird diese Behauptung vernachlässigt und nicht mit in die Forschung einbezogen.

Die geplante Exploration der Suchanfragen, deren Informationsbedürfnis bereits auf der Suchergebnisseite beantwortet werden kann und daher keinen Klick auslöst, wird den aktuellen Entwicklungen der Web-Suchmaschinen gerecht und kann hinsichtlich dessen Hinweise geben, um zu erforschen, welche Suchanfragen(typen) bereits auf der Suchergebnisseite beantwortet werden können. Wäre es möglich diese Typen von Suchanfragen und ihre Erkennungsmerkmale herauszuarbeiten, könnten daraus Vorschläge für eine verbesserte und nutzerfreundlichere Darstellung dieser Suchergebnisse erarbeitet werden. Demgegenüber steht eine Untersuchung der Suchanfragen, die nicht geklickt wurden, da sie nur irrelevante Ergebnisse lieferten. Daher soll untersucht werden, ob auch hier bestimmte Typen von Anfragen zu erkennen sind. Daraus sollen anschließend Empfehlungen abgeleitet werden, die den Betreibern der Suchmaschine oder dem Suchenden selbst verbesserte Hilfestellungen in einem solchen Fall bieten können, damit die Suche erfolgreich abgeschlossen werden kann. Es ergeben sich also die folgenden Forschungsfragen, welche sich in zwei Bereiche gliedern:

1 Suchanfragen, die auf der SERP beantwortet werden können

- Wie groß ist der Anteil der Suchanfragen, die bereits auf der SERP beantwortet werden?
- Welche Merkmale und Eigenschaften können festgehalten werden, die diese Art der Suchanfrage kennzeichnen (z.B. Thema, Suchanfragetyp)?

- Welche Vorschläge zur Verbesserung der Ergebnisdarstellung von Suchanfragen, die bereits auf der SERP beantwortet werden, können festgehalten werden?

2 Suchanfragen, die nicht auf der SERP beantwortet werden können und irrelevante Ergebnisse liefern

- Wie groß ist der Anteil der Suchanfragen, die keinen Klick erhalten, weil sie irrelevante Ergebnisse liefern?
- Welche Merkmale und Eigenschaften können festgehalten werden, die diese Art der Suchanfrage kennzeichnen (z.B. Thema, Suchanfragetyp, Einsatz von Suchoperatoren, Rechtschreibfehler, Mehrdeutigkeit)?
- Welche Vorschläge zur Verbesserung der Suche können festgehalten werden? Wie könnte dem Suchenden weitergeholfen werden?

Insgesamt wird als Ziel der Forschung festgehalten, dass mit Hilfe diverser Untersuchungen in diesem Bereich angestrebt wird, die Suche zu verbessern, indem Vorschläge zur Anpassung der Suche oder zum Aufbau von Hilfestellungen für den Suchenden entstehen.

1.2 Aufbau der Arbeit

Zur Lösung der Forschungsfrage wird in dem folgenden Kapitel der Stand der Forschung sowie verwandte Studien zum Themengebiet vorgestellt und deren Implikationen für die vorliegende Arbeit dokumentiert. Die Methoden, die zur Klärung des Forschungsvorhabens eingesetzt werden, werden ebenfalls in einem Kapitel detailliert beschrieben und ihr Aufbau sowie ihr Ablauf begründet. Die Dokumentationen der Datenerhebung und Datenauswertung werden ebenfalls im Kapitel „Methodik“ beschrieben. Anschließend werden die Ergebnisse, die aus den erhobenen Daten erfasst werden konnten, aufgegliedert nach der eingesetzten Methode, im Kapitel „Ergebnisse“ detailliert beschrieben und analysiert. Es folgt ein Kapitel, welches Schlussfolgerungen und konkrete Ableitungen aus den Ergebnissen aufbereitet und diskutiert. Am Ende der Arbeit steht ein Fazit, welches die wesentlichen Aspekte der Arbeit noch einmal zusammenfasst.

2 Stand der Forschung

Der folgende Teil der Arbeit gibt Aufschluss über den Stand der Forschung in den diese Arbeit betreffenden Themenbereichen. Der Fokus liegt hier darauf, den aktuellen Forschungsstand für Suchanfragen, deren Ergebnisse nicht geklickt wurden, darzustellen. Verwandte Gebiete aus dem Bereich der zu untersuchenden Forschungsfragen werden ebenfalls thematisiert und erörtert, um die Zusammenhänge zwischen den zugehörigen Forschungsbereichen deutlicher zu machen und die Methodenwahl zu begründen. Daher wird in diesem Kapitel der Forschungsstand zu den Bereichen Analyse von Klickdaten, Untersuchung von Trefferbeschreibung und deren Evaluierung durch den Suchenden sowie Forschungen zum Thema Relevanz und die Evaluation von Relevanz beschrieben.

2.1 Suchanfragen ohne Klick

Es konnten zwei Untersuchungen gefunden werden, welche sich explizit mit Suchanfragen, die keinen Klick auf Treffer erhalten, beschäftigen. Da für die Aufstellung eines Methodenplans und zur Beschreibung der Forschungslücke, die durch die vorliegende Arbeit geschlossen werden soll, eine genauere Kenntnis der bereits durchgeführten Studien in diesem Bereich notwendig ist, werden diese im Folgenden vorgestellt:

Studie: „Good Abandonment in Mobile and PC Internet Search"

Li et al. untersuchen 2009 „abandonment queries", also abgebrochene/aufgegebene Suchanfragen, hinsichtlich ihres Potentials bereits auf der Suchergebnisseite beantwortet zu werden (vgl. LI et al. 2009). Unter „abandonment query" verstehen sie eine Suchanfrage, die in einer Zeitspanne von 24 Stunden keinen Klick erhalten hat und auch nicht durch eine weitere Suchanfrage ergänzt wurde. Ihre Untersuchungen erstrecken sich dabei auf Suchanfragen ohne Klick aus den USA, Japan und China. Sie unterscheiden außerdem nach Suchen per PC oder mobilem Endgerät. Zur Einordnung der Suchanfrage gehen sie im ersten Schritt von der idealen Suchergebnisseite aus und klassifizieren die Suchanfragen als „potential good abandonment" oder „maybe potential good abandonment". Dabei ist eine „potential good abandonment query" definiert als eine Suchanfrage, deren dominierendes Informationsbedürfnis theoretisch auf einer Trefferseite beantwortet werden könnte. Eine „maybe potential good abandonment query" könnte nur möglicherweise auf der SERP beantwortet werden.

Ihre Ergebnisse zeigen für die untersuchte Stichprobe für ungeklickte Suchanfragen aus den USA, die mit Hilfe eines mobilen Endgerätes gestellt wurden, einen Anteil von 38,3 Prozent der Suchanfragen, welche potentiell bereits auf der Trefferseite beantwortet werden könnten. Hingegen stufen sie nur 17 Prozent der ungeklickten Suchanfragen, die per PC Suche abgeschickt werden, als potentiell auf der Suchergebnisseite beantwortbar ein.

In einem nächsten Schritt prüfen sie die Einordnung der Suchanfragen anhand der Suchergebnisseiten der Web-Suchmaschine Google. Von den Anfragen, welche als potentiell auf der Trefferseite beantwortbar klassifiziert wurden und von einem mobilen Endgerät aus abgeschickt wurden, bestätigt sich bei einem Anteil von 36 Prozent der Anfragen diese Einordnung. Bei den PC Suchen ergibt sich hier ein Anteil von 14,3 Prozent.

Sie kategorisieren weiterhin die potentiell auf der Trefferseite beantwortbaren Suchanfragen nach dem zugrundliegenden Informationsbedürfnis, wobei dieses ihnen nicht bekannt ist und daher von den Forschern selbst bestimmt wird. Hierzu entwickeln sie die Kategorien „answer“, „currency“, „click-to-call“, „definition“, „images“ , „local“, „lyrics“, „map“, „celebrities“, „news“, „products“, „person“, „quotation“, „stock“, „sports“, „showtimes“, „spelling”, “SMS”, „translation” und „weather”. Ihr Ziel ist, die häufigsten Informationsbedürfnisse von Suchanfragen, die potentiell zu einem positivem Abbruch führen, herauszustellen und gleichzeitig aufzudecken, welche Kategorien von Informationsbedürfnissen Ausbaupotential zeigen, um zusätzliche „positive Abbrüche“ zu schaffen. Für die mobile Suche in den USA konnten für diese Art der Anfragen die Gruppen „local“ mit einem Anteil von 42,9 Prozent, „answer“ mit 22,3 Prozent und „stock“ mit einem Anteil von 11,9 Prozent herausgestellt werden. Die PC Suche in den USA für Suchanfragen mit potentiell positivem Abbruch führt zu einem Anteil von 28,4 Prozent Suchanfragen in der Kategorie „local“, die Kategorie „answer“ wird mit einem Anteil von 20,5 Prozent der Suchanfragen belegt.

Insgesamt kommen die Forscher zu dem Schluss, dass ein erheblicher Anteil der Suchanfragen ohne Klick auf die Treffer bereits auf der Suchergebnisseite beantwortet werden könnte. Ihre Untersuchung der tatsächlichen Beantwortung dieser Anfragen zeigt, dass hier weiteres Verbesserungspotential der Suchmaschinen besteht. Bezüglich der Unterscheidung zwischen mobilem und fest installiertem Endgerät auf dem die Suche ausgeführt wird, kommen sie zu dem Ergebnis, dass der, im Vergleich zur PC Suche, erhöhte Anteil an „potential good abandonment“ bei mobilen Endgeräten aufgrund der eingeschränkten Möglich-

keiten bei einer Suche entsteht. Sie gehen davon aus, dass der mobil Suchende aufgrund der eingeschränkten Möglichkeiten wahrscheinlich fokussierte und weniger komplexe Anfragen stellt, die dann häufiger direkt auf der Suchergebnisseite beantwortet werden können.

Studie: „Interpreting User Inactivity on Search Results“

Eine weitere Studie, welche sich mit Suchanfragen beschäftigt, die kein Feedback in Form eines Klicks erhalten haben, wurde im Jahr 2009 von Stamou und Efthimiadis (vgl. STAMOU und EFTHIMIADIS 2009) durchgeführt. Im Unterschied zu der bereits vorgestellten Studie von Li et al. werden hier alle Suchanfragen betrachtet, die keinen Klick erhalten haben, es wird nicht darauf geachtet, ob einer Anfrage eine weitere Suchanfrage folgt und somit eine Reformulierung stattfindet. Ziel der Studie ist u.a. die Untersuchung der Informationsbedürfnisse hinter den Anfragen, die keinen Klick erhalten. Insgesamt zielen sie darauf ab, Inaktivität mit Suchergebnissen präziser interpretieren zu können, um auf Grundlage dessen verstärkt User-zentrierte Evaluierungsmethoden vorschlagen zu können.
Als Methode setzen sie eine Befragung gekoppelt mit der Aufzeichnung der Browser-Aktivitäten der Probanden ein. Insgesamt rekrutieren sie sechs Probanden, welche Studenten des Studiengangs Informatik sind. Die Probanden werden in einer Training Session auf die Untersuchung vorbereitet und anschließend werden eine Woche lang ihre Suchen aufgezeichnet. Den Probanden werden dabei drei bzw. vier Fragen per Online-Befragung gestellt. Die ersten Beiden sollen vor dem Abschicken der Suchanfrage, die Letzten nach Sichtung der Suchergebnisse beantwortet werden. In der ersten Frage wird nach dem Informationsbedürfnis des Suchenden gefragt, er hat hier die Wahl zwischen zehn verschiedenen Bedürfnissen, welche von Stamou und Efthimiadis auf Grundlage verschiedener Studien anderer Forscher aufgestellt wurden. Es bestehen die Kategorien „find detailed information about a subject of interest“, „obtain/ download/interact with a resource”, „find a specific URL/ web site”, „find local information”, „look for a person”, „find out about a product”, „get a quick answer to

my query terms", „see if there is any new result retrieved since the last time I issued the query" und „re-find the information I got in one of my previous submissions of the query". Der Proband kann sich nur für ein Informationsbedürfnis entscheiden. In der zweiten Frage wird der User hinsichtlich seiner Absicht ein oder mehrere Ergebnisse anzuklicken befragt (Antwortmöglichkeiten: „Ja", „Nein", „Vielleicht"). Anschließend schickt der User seine Suchanfrage ab und wird daraufhin befragt, ob er ein Ergebnis anklickt. Falls er keines anklickt, erscheint eine Filterfrage, welche den Grund für das nicht anklicken abfragt. Der Proband kann hier zwischen den folgenden Antwortmöglichkeiten wählen: „die Antwort war bereits auf der Suchergebnisseite zu finden", „die Ergebnisse sind irrelevant", „die gefundenen Ergebnisse waren schon bekannt", „die Suche wurde unterbrochen". Auf diese Weise konnten 966 Suchanfragen erhoben werden, von denen 27,02 Prozent keinen Klick erhielten.

Anschließend werden die Informationsbedürfnisse der Suchen mit und ohne Klick aufgeschlüsselt. Suchen, die keinen Klick erhalten haben, haben am häufigsten die Ziele „quick answer" (27,59 Prozent), „local" (21,45 Prozent) und „language-related" (13,03 Prozent). Die geklickten Suchanfragen stammen hingegen zu den größten Teilen aus den Bereichen „informational" (33,47 Prozent), „navigational" (16,74 Prozent) und „transactional" (13,46 Prozent) (vgl. STAMOU und EFTHIMIADIS 2009, S. 6).

Des Weiteren untersuchen Stamou und Efthimiadis die Zufriedenheit der Probanden mit den erhaltenen Suchergebnissen. Sie kommen zu dem Ergebnis, dass Probanden, die informationsorientiert, navigationsorientierte, transaktionsorientierte sowie wiederholte Suchen abschicken, nicht durch die bloße Anzeige der Trefferliste befriedigt werden (vgl. STAMOU und EFTHIMIADIS 2009, S. 8-9). Hingegen sind Probanden, die eine Inaktivität mit den Suchergebnissen beabsichtigten, besonders zufrieden, wenn ihre Suchintention dem Bereich „language-related" oder „quick answer" entspringt (vgl. STAMOU und EFTHIMIADIS 2009, S. 9).

Abschließend kommen Stamou und Efthimiadis zu dem Ergebnis, dass Anfragen, die keine Klicks erhalten, nicht generell als gescheiterte Suchen zu behandeln sind, vielmehr sollte ihre Effektivität hinsichtlich der zugrunde liegenden Suchintention bewertet werden.

2.2 Analyse von Klickdaten

Ein weiterer Bereich, der das Thema der vorliegenden Arbeit berührt, ist die Analyse von Klickdaten zur Optimierung von Web-Suchmaschinen. Wie bereits in der Einleitung angedeutet, wird auf diesem Gebiet viel geforscht (vgl. z.B. CHAO et al. 2009; CRASWELL und SZUMMER 2007; DOU et al. 2008; GOLLAPUDI et al. 2008; GRANKA et al. 2005; JOACHIMS et al. 2007; MACDONALD und OUNIS 2009). Besonders nennenswert sind die Untersuchungen von Granka et al. und Joachims et al. (vgl. GRANKA et al. 2005 und JOACHIMS et al. 2007), da die Forscher Klicks als implizites Userfeedback werten und dahingehend Vorschläge zur Optimierung des Suchalgorithmus ausarbeiten. Dass es sich nur um ein implizites Relevanzurteil handelt, folgern sie aus der Beobachtung, dass das Klickverhalten des Suchenden auf Grund zweier Einflussgrößen verzerrt wird. Einerseits werden Suchende in ihrem Klickverhalten durch das Vertrauen, welches sie der Suchmaschine hinsichtlich deren Fähigkeit relevante Suchergebnisse zu finden, entgegen bringen, beeinflusst (vgl. GRANKA et al. 2005, S. 158). Hier findet eine Beeinflussung des Klickverhaltens durch den Rang, den ein Treffer auf der Suchergebnisseite einnimmt, statt (vgl. CRASWELL et al. 2008, S. 87; JOACHIMS und RADLINSKI 2007, S. 35). Laut Eye-Tracking Untersuchungen der Forscher, sehen sich Nutzer bevor sie klicken die Treffer ab dem dritten Rang nur noch für weniger als die Hälfte aller untersuchten Anfragen an (vgl. JOACHIMS und RADLINSKI 2007, S. 35). Das bedeutet, dass den ersten Positionen einer Suchergebnisseite besonders hohe Aufmerksamkeit entgegen gebracht wird, wohingegen potentiell relevante Treffer, die weiter unten in der Rangfolge auftauchen, unbeachtet bleiben.

Andererseits nennt Granka et al. den „Quality Bias“, der besagt, dass das Klickverhalten von der allgemeinen Qualität der präsentierten Suchergebnisse ab-

hängt (vgl. GRANKA et al. 2005, S. 158). Daher schlussfolgern die Forscher, dass Klicks nur als „pairwise preference statements“ (GRANKA et al. 2005, S. 159) interpretiert werden können. In diesem Kontext wird auch darauf hingewiesen, dass nicht nur geklickte Treffer als Feedbacksignal genutzt werden sollen, sondern auch die Treffer, die nicht geklickt wurden (vgl. GRANKA et al. 2005, S. 159). Die Forscher gehen allerdings nicht weiter darauf ein, auf welche Art und Weise das komplette Unterlassen eines Klicks als Userfeedback gedeutet werden kann. Auch die anderen aufgezählten Studien konzentrieren sich auf den Klick als Feedback und nicht auf sein „Gegenstück“, eine Suchanfrage ohne Klick.

Dennoch geben die Untersuchungen wichtige Hinweise über den Umgang des Suchenden mit der Suchergebnisseite und ihren Treffern sowie natürlich den Ausschlag für die erste Hypothese, dass bei einer Suchanfrage ohne Klick womöglich kein relevantes Ergebnis vorlag.

2.3 Trefferbeschreibung und -evaluation

Die Entscheidung, einen Treffer anzuklicken, kann der Suchende im ersten Schritt nur anhand der Trefferbeschreibung zu einer Suchanfrage fällen. Des Weiteren wird er auch nur die Treffer für einen Klick in Betracht ziehen (können), welche von ihm wahrgenommen werden. Der Umgang mit und die Wahrnehmung der Trefferseite sowie die Darstellung und Bewertung der Trefferbeschreibung nehmen folglich Einfluss auf das (nicht vorhandene) Klickverhalten des Suchenden. So hält Rose 2004 fest, dass das Klickverhalten nicht genutzt werden kann, um die Qualität der Trefferbeschreibungen einzuschätzen, da das Fehlen eines Klicks darauf hinweisen kann, dass es sich entweder um eine perfekte Trefferbeschreibung handelt, die das Informationsbedürfnis des Suchenden befriedigt oder um eine schlechte Trefferbeschreibung, die nicht in der Lage ist, die Relevanz der dahinterliegenden Seite auszudrücken (vgl. ROSE 2004). Diese Einschätzung zeigt bereits, dass die Wahrnehmung und Bewertung der Trefferbeschreibung vielschichtig ist, daher sollen auch für diesen Aspekt der Stand der Forschung und die bereits erhobenen Ergebnisse dargestellt werden.

In mehreren Studien (vgl. z.B. GRANKA et al. 2008, S. 1044; JANSEN und SPINK 2006, S. 257 f.; LORIGO et al. 2006, S. 1127) konnte nachgewiesen werden, dass Suchende nicht bereit sind alle gefundenen Treffer anzusehen, vielmehr evaluieren sie nur die erste Ergebnisseite. Wie bereits im Stand der Forschung zum Bereich der Klickanalyse erwähnt, werden Suchende außerdem in ihrem Suchverhalten von dem dargebotenen Ranking der Suchmaschine beeinflusst, so dass in vielen Fällen nicht einmal alle zehn Ergebnisse der ersten Suchergebnisseite angesehen werden (vgl. JOACHIMS und RADLINSKI 2007, S. 35). Auch Lorigo et al. kommen zu dem Ergebnis, dass in den meisten Fällen nur die am höchsten gerankten Ergebnisse vom User inspiziert werden (vgl. LORIGO et al. 2006, S. 1127). Sie kommen außerdem zu dem Schluss, dass Suchende sich stark auf die Trefferbeschreibung allein verlassen, um die Relevanz einer dahinter liegenden Webseite einzuschätzen (vgl. LORIGO et al. 2006, S. 1127). Sie leiten diese Schlussfolgerung von dem Ergebnis ab, dass in mehr als der Hälfte der untersuchten Suchsessions die Probanden Suchbegriffe abändern, ohne auch nur einen Treffer zu klicken (vgl. LORIGO et al. 2006, S. 1127). Es zeigt sich, dass in nur vier Prozent dieser Fälle alle zehn Treffer angesehen werden, laut Lorigo et al. weist dies daraufhin, dass der Suchende meist sehr schnell, und zwar ohne alle Suchergebnisse zu evaluieren, entscheidet, ob er die aktuelle Suche mit einer neuen oder abgeänderten Suchanfrage abbricht (vgl. LORIGO et al. 2006, S. 1128).

Die Bedeutung der Trefferbeschreibung zur Evaluation der Suchergebnisse kann also als nachgewiesen angenommen werden und ist im Zusammenhang der vorliegenden Untersuchung von großer Bedeutung. Daher sollen im weiteren Verlauf der Stand der Forschung hinsichtlich der Beurteilung von Trefferbeschreibungen sowie die Kriterien, die zu einer solchen Beurteilung führen, beleuchtet werden.

Griesbaum führt 2004 eine Studie zur Evaluation von deutschen Suchmaschinen durch und untersucht u.a. die Übereinstimmung zwischen der Relevanzbewertung der Trefferbeschreibung sowie der dahinter liegenden Webseite. Dabei kommt er zu dem Resultat, dass bei allen untersuchten Suchmaschinen (Google, Lycos, Altavista) in ca. einem Drittel der Fälle die Beurteilungen unterschiedlich ausfielen (vgl. GRIESBAUM 2004). Er vergleicht in einem nächsten Schritt, ob das Relevanzurteil für die Trefferbeschreibung oder für die Webseite besser ausfiel. Für Google wird ein ausgeglichenes Ergebnis festgehalten: bei einem Anteil von 49,86 Prozent der Fälle wurde die Trefferbeschreibung als relevanter bewertet und in 50,14 Prozent der Fälle die tatsächlichen Suchergebnisse. Griesbaum schränkt seine Ergebnisse jedoch hinsichtlich ihrer Validität ein, denn die Beurteilung von Suchergebnis und Trefferbeschreibung wurde von zwei unterschiedlichen Probanden durchgeführt. Des Weiteren finden sich Unterschiede in der Darstellung der Trefferbeschreibungen, da unterschiedliche Suchmaschinen evaluiert wurden. Außerdem merkt Griesbaum an, dass es unterschiedliche Präferenzen bei den Probanden hinsichtlich der verschiedenen Suchmaschinen gibt und dass dahingehend eine Beeinflussung der Beurteilung möglich ist.

Lewandowski bezieht in seine Untersuchung der Retrievaleffektivität von Suchmaschinen ebenfalls die Trefferbeschreibungen mit ein (vgl. LEWANDOWSKI 2008a). Eine Bewertung der Relevanz der Trefferbeschreibungen wurde mit der folgenden Einschränkung vorgenommen: Eine Trefferbeschreibung sollte als relevant bewertet werden, wenn die Beschreibung auf ein relevantes Ergebnis hindeutet, nicht ob die Trefferbeschreibung an sich Informationen enthält, die das Informationsbedürfnis des Suchenden befriedigen (vgl. LEWANDOWSKI 2008a, S. 925).
Der Vergleich von Trefferbeschreibung und tatsächlichem Suchergebnis fällt hier allerdings im Gegensatz zur Studie von Griesbaum detaillierter aus. So werden u.a. die folgenden Messgrößen herangezogen: der Anteil der Ergebnisse, bei denen sowohl die Trefferbeschreibung als auch der Treffer als relevant bewertet wurden; der Anteil der Ergebnisse, dessen Trefferbeschreibung als irrelevant

gewertet wurde, obwohl der dahinter liegende Treffer Relevanz aufwies. Des Weiteren wird der Anteil der Ergebnisse berechnet, die eine relevante Trefferbeschreibung aufweisen, jedoch zu einem nicht relevanten Ergebnis führen (vgl. LEWANDOWSKI 2008a, S. 931ff.).

Insgesamt kommt die Studie zu dem Resultat, dass zwischen den Web-Suchmaschinen Google und Yahoo keine großen Unterschiede bei der Bereitstellung relevanter Suchergebnisse bestehen, jedoch liefert Google den größeren Anteil an relevanten Trefferbeschreibungen. Wobei hier allerdings ein deutlicher Teil (ca. 20 Prozent) zu irrelevanten Ergebnissen führt (vgl. LEWANDOWSKI 2008a, S. 933 und 934). Das Google trotzdem bei Nutzerinterviews gut abschneidet, führt Lewandowski teilweise auf den großen Anteil von durch die Trefferbeschreibung relevant scheinenden Treffern zurück (vgl. LEWANDOWSKI 2008a, S. 933 f.).

Da nun herausgestellt werden konnte, dass Trefferbeschreibungen für die Suche von besonderer Bedeutung sind, interessiert in einem nächsten Schritt, welche Kriterien Suchende heranziehen, um eine Trefferbeschreibung als relevant oder irrelevant zu bewerten.

Crystal und Greenberg führen eine Studie durch, welche versucht die Kriterien herauszuarbeiten, welche Suchende auf dem Gebiet der Gesundheitsthemen nutzen, um sowohl Trefferbeschreibungen als auch die Treffer an sich zu bewerten (vgl. CRYSTAL und GREENBERG 2006). Sie kommen zu dem Schluss, dass Relevanzurteile von Suchenden im Web komplex und vielfältig sind, Suchende beziehen dabei eine große Anzahl verschiedener Kriterien ein (vgl. CRYSTAL und GREENBERG 2006, S. 21). Wobei die Verteilung der verwendeten Kriterien laut Crystal und Greenberg der Long-Tail-Verteilung nach Zipf und Bradford gleicht, denn einige Kriterien werden häufig genutzt, viele Kriterien werden allerdings nur sehr selten herangezogen (vgl. CRYSTAL und GREENBERG 2006, S. 22). Sie schreiben den Trefferbeschreibungen eine beschränkte Effektivität bzgl. der Tätigung von Relevanzurteilen zu, dies führen sie u.a. auf die wenigen Informationen zu-

rück, die in einer Trefferbeschreibung enthalten sind (vgl. CRYSTAL und GREENBERG 2006, S. 25).

Weitere Untersuchungen (vgl. DRORI 2000; DUMAIS et al. 2001) zeigen ebenfalls, dass die Ergänzung der Trefferbeschreibungen durch weitere Metadaten, wie Kategorien oder Keywords, die Usability der Trefferrepräsentation und damit die Relevanzbewertung des dahinter liegenden Ergebnisses weiter verbessern kann.

2008 führte Tittmann im Rahmen seiner Diplomarbeit eine Studie durch, welche sich mit der Struktur der Trefferbeschreibung und deren Beurteilung durch den Nutzer auseinander setzt (vgl. TITTMANN 2008). Im Gegensatz zu anderen Studien in diesem Bereich beschäftigt er sich deskriptiv mit der Zusammensetzung der Trefferbeschreibungen in den gängigen Suchmaschinen Google, Yahoo, MSN, ASK und Seekport. So untersucht er u.a. wie die einzelnen Suchmaschinen die Trefferbeschreibungen erstellen, wie lang die Beschreibung ist und ob sie aus einem vollständigen Satz, einem Ausschnitt aus einem Satz, einzelnen Wörtern oder einer Fremdsprache besteht. Die Beurteilung des letztgenannten Untersuchungsgegenstands zeigt, dass der Einsatz eines vollständigen Satzes in der Trefferbeschreibung von den Probanden in mehr als der Hälfte der Fälle als sehr gut bis gut bewertet wurde (vgl. TITTMANN 2008, S. 40). Die Nutzung eines Ausschnittes aus einem Satz wurde in der Hälfte der Fälle als gut bis zufriedenstellend betrachtet. Schlechter schneidet die Nutzung einzelner Wörter in einer Trefferbeschreibung ab, über die Hälfte der Fälle werden nur als ausreichend bis ungenügend beurteilt (vgl. TITTMANN 2008, S. 40).

2.4 Relevanz & Evaluation von Relevanz

Eine Hypothese zur Begründung der nicht vorhandenen Klicks auf die Trefferliste ist die fehlende Relevanz der Treffer. Daher scheint es auch für diesen Themenbereich sinnvoll, einen Blick auf die Theorie und den Stand der Forschung zu werfen. Zum Einstieg in diesen Bereich wird die Definition des Begriffs der Relevanz in der Informationswissenschaft erläutert, anschließend wird auf das Vorgehen zur Bewertung von Relevanz eingegangen.

In der Informationswissenschaft werden die Begriffe „Relevanz“ und „Pertinenz“ in Hinblick auf die Differenzierung zwischen Informationsbedarf und Informationsbedürfnis unterschieden (vgl. STOCK 2007, S. 68). Relevanz trägt dabei dem objektiven Informationsbedarf Rechnung. In diesem Sinne ist ein Dokument relevant, wenn es mit der Suchanfrage übereinstimmt. Bei einem Vergleich von Übereinstimmung des Informationsbedürfnisses eines Nutzers und dem Dokument wird dagegen von Pertinenz gesprochen.

Die Messung von Relevanz findet, abgesehen von der Ausnahme der algorithmischen Relevanz, immer unter Einbeziehung eines menschlichen Urteils statt (vgl. STOCK 2007, S. 71). Handelt es sich um die Überprüfung einer thematischen Relevanz, werden zur Bewertung unabhängige Experten herangezogen, welche die subjektiven Aspekte eines Nutzers vernachlässigen (vgl. STOCK 2007, S. 71). Abgesehen davon werden Nutzer selbst gebeten eine Relevanzeinschätzung abzugeben.

Wird die Relevanzmessung also durch Nutzer durchgeführt, greift die folgende Aussage:

> Für die Bewertung von Suchmaschinen durch den Nutzer dürfte letztlich die Pertinenz von entscheidender Bedeutung sein. Eine objektive Relevanz können diese nicht messen bzw. beurteilen; Vertrauen in eine Suchmaschine bildet sich nur über die Erfahrung mit der Pertinenz der Ergebnisse. (LEWANDOWSKI 2005, S. 97).

Bei der Bewertung durch den Nutzer ist außerdem zu beachten, dass immer auch Charakterzüge und Erwartungen des Nutzers in die Bewertung mit einfließen (vgl. DUDEK et al. 2006, S. 227).

Wird die Relevanz eines Information Retrieval Systems erhoben, spricht man von einem Retrievaltest. In der Literatur wird darauf hingewiesen, dass dieser sorgfältig geplant werden sollte, um am Ende valide Ergebnisse vorweisen zu können. Tague-Sutcliffe bietet eine Zusammenstellung von Entscheidungen, welche getroffen werden müssen, um eine optimale Testdurchführung zu gewährleisten (vgl. TAGUE-SUTCLIFFE 1992, S. 468-487). Neben der essentiellen Frage, ob überhaupt ein Test nötig ist, weist er daraufhin, dass vor der Untersuchung geklärt werden sollte, welche Art von Test durchgeführt werden soll, wie die Variablen

operationalisiert und welche Datenbanken getestet werden. Des Weiteren macht er deutlich, dass geklärt werden muss, woher die Suchanfragen kommen und wie die Erhebungsdaten gesammelt und analysiert werden sollen (vgl. TAGUE-SUTCLIFFE 1992, S. 468-487). Insgesamt gibt er dem Forschenden zehn Fragen zur Entscheidungsfindung für einen Relevanztest mit auf den Weg. Hawking et al. weisen zusätzlich daraufhin, dass, wenn der Relevanztest an einer Web-Suchmaschine durchgeführt werden soll, es von besonderer Wichtigkeit ist, dass in dem Test reale Informationsbedürfnisse von Nutzern abgebildet werden (vgl. HAWKING et al. 2001, S. 35). Die Erfüllung dieser Voraussetzung für einen Retrievaltest ist eine besondere Herausforderung, wenn lediglich Suchanfragen zur Verfügung stehen, nicht jedoch die zugrunde liegenden Informationsbedürfnisse. Huffmann und Hochster lösen dieses Problem, indem sie zu jeder Suchanfrage mehrere Probanden bitten das Informationsbedürfnis abzuleiten, aufgrund der Mehrdeutigkeit von Suchanfragen erlauben sie auch die Abgabe eines zweiten Informationsbedürfnisses (vgl. HUFFMANN und HOCHSTER 2007, S. 568). Anschließend wird pro Suchanfrage das häufigste erste und zweite Informationsbedürfnis bestimmt und die Sprache der Bedürfnisse normalisiert (vgl. HUFFMANN und HOCHSTER 2007, S. 568).

2.5 Erkenntnisse für die vorliegende Arbeit

In diesem Kapitel werden die für die vorliegende Arbeit wesentlichen Erkenntnisse aus dem oben beschriebenen Stand der Forschung extrahiert und zusammengefasst. Die Reihenfolge der oben aufgeführten Forschungsstände wird dabei aufgegriffen.

Suchanfragen ohne Klick

Die vorgestellten Studien zum Bereich „Suchanfragen ohne Klick“ unterstützen die in der Einleitung getätigte Annahme, dass Suchanfragen, die keine Klicks auslösen, nicht zwangsläufig den Suchenden unbefriedigt zurücklassen. Vielmehr zeigen sie, dass ein signifikanter Anteil der Suchanfragen ohne Klick bereits durch

die Trefferliste beantwortet werden kann. Dementsprechend kann die Vermutung, die der vorliegenden Arbeit zu Grunde liegt, bestätigt werden.

Des Weiteren weisen beide Untersuchungen bei der Kategorisierung der Informationsbedürfnisse hinter den nicht geklickten Suchanfragen ähnliche Suchintentionen auf. So sind sowohl in der Studie von Li et al. als auch in der Untersuchung von Stamou und Efthimiadis die Kategorien „answer“ (bei Stamou und Efthimiadis „quick answer“) und „local“ die am stärksten besetzten Kategorien (vgl. LI et al. 2009, S. 47 und STAMOU und EFTHIMIADIS 2009, S. 6). Diese Ergebnisse liefern bereits Hinweise für die Beantwortung der im Kapitel 1.1 aufgestellten Forschungsfragen nach den Merkmalen und Eigenschaften von auf der SERP beantwortbaren Suchanfragen. Ob die in den beschriebenen Studien erhobenen Themen auch für deutsche Suchanfragen gelten, muss in den Untersuchungen noch herausgefiltert werden.

Allerdings zeigen die vorgestellten Studien hinsichtlich ihrer Repräsentativität Schwachstellen, die in der vorliegenden Arbeit, soweit möglich, umgangen werden sollen. So geben Li et al. an, dass in ihrer Untersuchung viele Beurteilungen enthalten sind und dass sie die Suchanfragen absichtlich so klassifizierten, dass sie eine obere Grenze bei den potentiell auf der Suchergebnisseite beantwortbaren Anfragen einschätzen konnten (vgl. LI et al. 2009, S. 45). Auch fließt bei der Einschätzung des Informationsbedürfnisses lediglich die Ansicht der Forscher mit ein (vgl. LI et al. 2009, S.47).

Stamou und Efthimiadis haben zwar Zugriff auf die realen Informationsbedürfnisse, indem sie diese von den Probanden angeben lassen, allerdings ist ihre Studie hinsichtlich ihrer Generalisierbarkeit eingeschränkt. Diesen „Mangel“ beschreiben sie selbst und führen ihn auf die kleine, homogene Gruppe von sechs Probanden zurück, die aufgrund des belegten Studienfaches überdurchschnittlich kompetent in der Websuche ist (vgl. STAMOU und EFTHIMIADIS 2009, S. 11).

Insgesamt scheint eine weitere Untersuchung in diesem Bereich hinsichtlich des wenig erforschten Untersuchungsgegenstands zweckmäßig. Denn es konnte keine Untersuchung gefunden werden, die sich mit deutschen Suchanfragen in dem ausgewählten Forschungsgebiet beschäftigt, so dass mit der vorliegenden Untersuchung diesbezüglich eine Forschungslücke geschlossen werden kann, indem eventuelle länderspezifische Unterschiede aufgedeckt werden.

Des Weiteren erhält die vorliegende Studie ein Alleinstellungsmerkmal, dadurch dass die Suchanfragen und die Suchergebnisse, die zu keinem Klick führten, hinsichtlich ihrer Merkmale und Eigenschaften untersucht werden. Auf diese Weise soll die Identifikation dieser Anfragen und damit eine an die Nutzerbedürfnisse angepasste Suche/ Suchergebnisseite ermöglicht werden. Die vorgestellten Studien grenzen hier zwar nach Informationsbedürfnis ein, anderweitige Merkmale der Suchanfragen bleiben allerdings unberücksichtigt. Außerdem betrifft diese nähere Untersuchung nur die Suchanfragen. Durch die zusätzliche Berücksichtigung der Suchergebnisse können evtl. weitere Erkennungsmerkmale aufgedeckt werden, die helfen die Suche zu verbessern.

Analyse von Klickdaten

Die Forschungen, die sich mit dem Thema Klickverhalten auseinandersetzen, zeigen, dass ein nicht durchgeführter Klick ein Hinweis auf fehlende relevante Suchergebnisse sein kann. Denn es wird herausgearbeitet, dass Klicks als implizite Relevanzurteile gewertet werden können. Aus diesem Forschungsergebnis leitet sich die Hypothese 1 (Es werden nur Treffer(-beschreibungen) präsentiert, die für die Suchanfrage und das dahinter stehende Informationsbedürfnis irrelevant sind) ab.

Damit wird einer weiteren Hypothese dieser Arbeit, dass ein fehlender Klick auf die Suchergebnisse nicht zwingend bedeutet, dass die Suchergebnisse für den Suchenden irrelevant sind, allerdings nicht widersprochen. Vielmehr zeigen die Studien, dass der Suchende bei der Sichtung der Suchergebnisse, und damit auch bei den möglichen Klicks, vom Rang und von der Qualität der Treffer beeinflusst

wird. Auf diese Weise werden z.B. relevante Treffer, die sich nicht auf den vorderen Rängen der Ergebnisliste befinden, übersehen. Die Studien, die zum Thema Klickverhalten durchgeführt werden, geben folglich wichtige Hinweise auf das Verhalten des Suchenden mit der Liste der Suchergebnisse und münden in der Hypothese 4 (Der Suchende nimmt nur die ersten Ergebnisse der Ergebnisseite wahr, werden diese von ihm als nicht relevant eingeordnet, bricht er seine Suche ab oder startet eine neue Suche. Weitere, evtl. relevante Ergebnisse, welche auf der Suchergebnisseite auf den unteren Postionen zu finden sind, werden übersehen).

Trefferbeschreibung und -evaluation

Die Resultate der vorgestellten Untersuchungen aus dem Bereich der Trefferbeschreibung lassen einige Schlüsse für die vorliegende Arbeit zu:

Sie unterstreichen die bereits auf der Grundlage des aktuellen Forschungsstands zum Thema Analyse von Klickdaten aufgestellte Hypothese, dass die Suchenden, die kein Ergebnis klicken, nur die erste Suchergebnisseite und dort wahrscheinlich nur die am höchsten gerankten Ergebnisse evaluiert haben, bevor sie sich entschieden, eine neue Suche zu starten oder den Suchvorgang komplett abbrachen. D.h., dass evtl. relevante Ergebnisse von der Suchmaschine gefunden wurden, diese jedoch aufgrund einer nicht optimalen Platzierung vom Suchenden übersehen wurden (Hypothese 4).

Des Weiteren wird durch die Darstellung des aktuellen Forschungsstands die Bedeutung der Trefferbeschreibung für die Evaluation des Suchergebnisses noch einmal untermauert. Der Trefferbeschreibung selbst wird allerdings eine mangelhafte Effektivität bei der Beurteilung der Relevanz des Treffers zugesprochen. Legt man dieses Ergebnis auf den vorliegenden Untersuchungsgegenstand um, zeigt sich der Grund für die Verfeinerung der ersten Hypothese. Denn man kann gemäß den beschriebenen Studien in diesem Fall nicht nur davon ausgehen, dass die Suchmaschine auf eine ungeklickte Suchanfrage nur irrelevante Ergebnisse lieferte, stattdessen ist es auch möglich, dass zwar relevante Treffer gefunden

wurden, jedoch durch die Repräsentation vom Suchenden nicht in Betracht gezogen wurden.

Anknüpfend an den gerade beschriebenen Abschnitt, kann außerdem festgehalten werden, dass Trefferbeschreibungen, die einen vollständigen Satz beinhalten, vom Suchenden positiv bewertet werden. Diese Ausführung resultiert ebenfalls in der Hypothese 1, denn ein Treffer könnte dementsprechend negativ bewertet (und damit nicht angeklickt) werden, wenn er dieses Kriterium nicht erfüllt und durch seine Repräsentation vom Suchenden gar nicht in Betracht gezogen wird.

Relevanz & Evaluation von Relevanz

Die beschriebene Theorie zu dem Thema Relevanz und deren Evaluation zeigen, dass es von großer Bedeutung ist, einen Retrievaltest zur Messung von Relevanz sorgfältig zu planen und durchzuführen, um valide Ergebnisse zu erhalten. Insbesondere die Anforderung originäre Informationsbedürfnisse von Suchenden zu nutzen, wenn Web-Suchmaschinen evaluiert werden, ist für die vorliegende Arbeit von besonderer Bedeutung und wird daher im Kapitel 3.4 erneut aufgenommen und diskutiert.

Ebenfalls beachtet werden sollte, dass die Relevanzbewertung von Suchergebnissen durch die Nutzer immer subjektiv ist, eine objektive Bewertung der Ergebnisse kann nicht vorgenommen werden. Dieser Aspekt unterstreicht noch einmal die verschiedenartigen Einflüsse, die bei einer Bewertung und damit auch bei der Entscheidung ein Ergebnis anzuklicken, eine Rolle spielen. Es kann z.B. nicht automatisch gefolgert werden, dass ein Ergebnis nicht angeklickt wurde, da es für den Suchenden irrelevant war. Möglich ist hier auch eine Beeinflussung des Suchverhaltens durch Vorkenntnisse oder Vorurteile des Suchenden sowie äußerer Einflüsse. Dies sollte bei der Ableitung der Ergebnisse aus einem Retrievaltest nicht unbeachtet bleiben.

3 Methodik

Dem Ziel der Arbeit, die Untersuchung von Suchanfragen, die keinen Klick auf die ausgelieferten Suchergebnisse erhielten, soll mit Hilfe einer Kombination verschiedener Untersuchungen näher gekommen werden. Dieses Kapitel beschreibt und begründet die eingesetzten Methoden und schildert detailliert Ablauf und Durchführung der einzelnen Untersuchungen. Die folgende Übersicht gibt Aufschluss über die genutzten Methoden:

Deskriptive Analyse

Zur Charakterisierung von Suchanfragen, die keinen Klick erhalten haben, wird in einem ersten Schritt eine deskriptive Analyse Aufschluss über Eigenschaften und Merkmale von nicht geklickten Suchanfragen geben. Dazu wird eine Stichprobe ungeklickter Suchanfragen auf Merkmale wie Länge der Anfragen, Fehler u.ä. untersucht. Um hier Besonderheiten herausarbeiten zu können, soll ein Vergleich zu Suchanfragen, die Klicks auf die Treffer erhalten haben, erfolgen.

Online-Befragung zur Untersuchung der Relevanz der Suchergebnisse von nicht geklickten Suchanfragen

Zur Herausarbeitung, welche der aufgestellten Hypothesen zutrifft, um das Unterlassen der Klicks zu begründen, wird eine Befragung zur Relevanzmessung konzipiert. Der Fokus liegt darauf die Relevanz der Trefferbeschreibungen von Suchanfragen, die keinen Klick erhielten, zu erheben und damit die bestehenden Möglichkeiten des Suchenden, eine Antwort auf sein Informationsbedürfnis zu finden, offenzulegen.

Die Probanden sollen dabei mit Hilfe eines Online-Fragenbogens mit den nicht geklickten Suchanfragen konfrontiert werden und die Relevanz der dazugehörigen Suchergebnisse beurteilen. Mit Hilfe der Ergebnisse soll abgeleitet werden, zu welchen Anteilen Suchanfragen nicht geklickt werden, da das Ergebnis bereits auf der Suchergebnisseite zu finden ist. Zusätzlich kann mit dieser Untersuchung auch eingegrenzt werden, mit welchem Anteil Anfragen vorliegen, die den Su-

chenden unbefriedigt zurückließen, da kein relevantes Ergebnis gefunden wurde. Aus den beiden beschriebenen Gruppen von Suchanfragen, die nach der Erhebung zur Verfügung stehen würden, sollen schlussendlich weitere Analysen Aufschluss über Merkmale, Eigenschaften und Typen der einzelnen Suchanfragegruppen geben. Der genaue Ablauf der Methode wird in Kapitel 3.4 beschrieben.

Um die Relevanzmessung valide durchführen zu können, ist es allerdings bei einem solchen Test, wie auch in Kapitel 2.4 aufgezeigt, notwendig dem Proband die Informationsbedürfnisse zu den einzelnen Suchanfragen vorlegen zu können. Da sich die vorliegende Datenbasis jedoch, neben den Suchanfragen, nur auf die Anzahl der Suchen und der getätigten Klicks auf die organischen Suchergebnisse sowie auf die Textanzeigen beschränkt, wird vorab eine Online-Befragung zur Ableitung des Informationsbedürfnisses konzipiert und durchgeführt (siehe Kapitel 3.3).

Die Methodenkombination dieser Arbeit gliedert sich dementsprechend in die folgenden Bereiche:

- Eine deskriptive Analyse der zur Verfügung stehenden Datenbasis sowie ein Vergleich deskriptiver Merkmale von Suchanfragen und Suchergebnisseiten, die keinen Klick erhalten haben.
- Eine Online-Befragung zur Ableitung der Informationsbedürfnisse der Suchanfragen zur Vorbereitung der Relevanzmessung der ungeklickten Suchanfragen.
- Eine Relevanzmessung zur Bewertung der Relevanz der Suchergebnisse der nicht geklickten Suchanfragen.

Nachdem zu Beginn dieses Kapitels der Untersuchungsgegenstand detailliert definiert wird, werden anschließend die gewählten Methoden in der oben aufgeführten Reihenfolge aufgegriffen, um ihre Konzeption und ihren Ablauf zu beschreiben und zu begründen.

3.1 Definition des Untersuchungsgegenstandes

Forschungsgegenstand dieser Arbeit sind Suchanfragen, die keinen Klick auf ihre Ergebnisse erhalten haben. Die Definition des Untersuchungsgegenstands gleicht also der Studie von Stamou und Efthimiadis, denn in ihrer Studie wird ebenfalls darauf verzichtet, die nachfolgenden Suchanfragen mit einzubeziehen (vgl. STAMOU und EFTHIMIADIS 2009, S. 2) und sich so nur auf Suchanfragen zu beschränken, die komplett „abgebrochen“ wurden. Grund für diese Entscheidung ist in der vorliegenden Untersuchung, dass alle Suchanfragen interessieren, die eine Inaktivität des Suchenden mit der SERP nach sich ziehen. Auf diese Weise fallen die Anfragen, die anschließend eine Reformulierung erfahren, nicht aus dem Untersuchungsschema heraus und können ebenfalls hinsichtlich ihrer Merkmale untersucht werden, um später auch für diese Anfragen Hilfestellungen entwickeln zu können.

3.2 Datenbasis für die deskriptive Analyse

Für die vorliegende Untersuchung steht eine Stichprobe von Suchanfragen, die im April 2010 an die Suche des Portals T-Online abgeschickt wurden (*http://www.t-online.de*), zur Verfügung. Die organischen Suchergebnisse werden in Form des Indexes bei diesem Portal von Google zur Verfügung gestellt, Textanzeigen und zusätzliche Dienste, wie Ergebnisse auf dem Portal selbst, werden von T-Online bereitgestellt. Die Suche des Portals T-Online hält in Deutschland einen Suchmaschinenmarktanteil von 2 Prozent, das Portal befindet sich damit auf dem vierten Platz nach den Suchmaschinen Google (89,6 Prozent), Bing (2,5 Prozent) und Yahoo (2,4 Prozent) (vgl. WEBHITS 2010). Das T-Online Portal, von dem aus die Websuche aufgerufen werden kann, ist laut Angaben der Arbeitsgemeinschaft Online Forschung in Deutschland das reichweitenstärkste Angebot im Internet, T-Online erreicht eine Netto-Reichweite von 27,09 Millionen Unique Usern (vgl. AGOF - ANGEBOTSRANKING 2010).

Die vorliegende Datengrundlage wurde mit Hilfe eines Logfiles erfasst. Diese Datenerfassungsmethode bietet im Vergleich zu Feldstudien, Laborexperimenten oder Panelstudien die beste Skalierbarkeit, das Speichern der Daten ist dabei unkompliziert und stört den Suchenden nicht, was in einer nicht verzerrten Beobachtung des Users resultiert (vgl. TANG et al. 2007, S. 3). Weiterer Vorteil ist, dass das Logfile eine große Vielfalt an Aufgaben, Suchanfragen und User Experience beinhaltet, was nur schwer anhand anderer Datenquelle erhoben werden kann (vgl. TANG et al. 2007, S. 3). Diese Vielfalt birgt allerdings wiederum einen Nachteil dieser Erhebungsmethode, denn die Daten können verrauscht sein, da sie, neben den Anfragen der Nutzer, auch Einträge von Robots, Spam und Erhebungsausfälle enthalten (vgl. TANG et al. 2007, S. 3). Tang et al. bringen den Nachteil der Datenerfassung per Logfile auf den Punkt: „First and foremost, logs can only measure the how and the what, rather than the why“ (TANG et al. 2007, S. 3). Aus dieser Methode der Datenerfassung resultiert auch das Problem, dass zwar authentische Suchanfragen vorliegen, warum der Suchende diese abgeschickt hat, also sein Informationsbedürfnis, bleibt allerdings unklar. Diesem Problem wird mit der in Kapitel 3.3 erläuterten Methode begegnet.

Die vorhandene Datenbasis besteht aus einer Millionen Suchanfragen, welche die eine Millionen meistgesuchten Anfragen an die T-Online Suche im April 2010 darstellen. Dabei erstreckt sich die Suchhäufigkeit über Suchanfragen, die im April über 300.000 Suchen verzeichnen konnten bis zu Anfragen, die im April sechsmal gesucht wurden.

Da aufgrund des Bradford Gesetzes davon ausgegangen werden muss, dass wenige Suchanfragen häufig und viele Suchanfragen selten gestellt werden (vgl. BATES 2002), zeigt sich hier ein Mangel von wenig oder nur einmal gestellten Suchanfragen. Dies unterstreicht, dass es sich nicht um eine Zufallsstichprobe handelt. Als Konsequenz muss mit einer Verzerrung der Ergebnisse gerechnet werden, weswegen von den Resultaten der Untersuchung nicht direkt auf die Grundgesamtheit aller abgeschickten Suchen geschlossen werden kann.

Die fehlende Zufallsauswahl der Stichprobe macht die folgenden Konsequenzen bei der Nutzung der Top-Millionen Suchanfragen für die Untersuchung denkbar:

a. Einerseits könnte, im Vergleich zu den mehrfach gesuchten Anfragen, eine größere Anzahl von Suchanfragen ohne Klick bei den nur einmal gesuchten Anfragen vorhanden sein, die aufgrund der Ziehung der Top Millionen Anfragen wegfallen. Gerade weil diese Anfragen einzigartig, da nur einmal gesucht, sind, könnten sich hier evtl. Besonderheiten aufdecken lassen, die durch die bloße Betrachtung der mehrfach abgeschickten Suchanfragen untergehen.
b. Andererseits kann bei den mehrfach gesuchten und nicht geklickten Anfragen evtl. davon ausgegangen werden, dass die Urteile der Suchenden „sicherer“ sind, da sich jeweils mehrere Suchende dazu entschieden haben für eine Suchanfrage keinen Treffer anzuklicken.

Die beschriebenen Annahmen zeigen, dass Gefahren, aber auch Chancen bei der Nutzung der vorliegenden Stichprobe auftreten können. Da nicht die Möglichkeit besteht auf eine Zufallsstichprobe auszuweichen, muss mit den möglichen Konsequenzen gearbeitet werden. Darauf muss insbesondere bei der späteren Ableitung von Ergebnissen und Empfehlungen geachtet werden.

Um die Verzerrung der Untersuchung besser eingrenzen zu können,

sollen auch die durch die deskriptive Analyse (siehe Kapitel 4.1) herausgearbeiteten Charakteristika der zur Verfügung stehenden Stichprobe herangezogen werden. Hier soll ein Vergleich mit bestehenden Untersuchungen von Suchanfragen Auffälligkeiten oder Gemeinsamkeiten aufdecken. Auf diese Weise soll herausgearbeitet werden, ob die Stichprobe mit Zufallsstichproben vergleichbar ist oder ob ihre Merkmale deutlich von den Charakteristika bereits untersuchter Zufallsstichproben abweichen.

Abschließend kann an dieser Stelle bereits festgehalten werden, dass die Problematik der fehlenden Zufallsstichprobe nicht umgangen werden kann und die Ergebnisse anschließend dementsprechend interpretiert werden müssen.

3.3 Online-Befragung zur Ableitung des Informationsbedürfnisses

Ein wesentlicher Schritt bei der Aufbereitung der Stichprobe für ihren Einsatz in der geplanten Relevanzuntersuchung ist die Ableitung des Informationsbedürfnisses, welches hinter den Suchanfragen liegt. Denn ein Information Retrieval System wird erst vor dem Hintergrund genutzt, ein wahrgenommenes Bedürfnis nach Information zu stillen (vgl. SHNEIDERMAN et al. 1998, S. 95f.). Daher sprechen Hawking et al. auch die Empfehlung aus, bei einem Retrievaltest reale Informationsbedürfnisse von Nutzern abzubilden (vgl. HAWKING et al. 2001, S. 25), um eine valide Testanordnung zu erreichen. Die vorliegende Datenbasis, die anhand eines Logfiles gewonnen wurde, enthält allerdings nicht die Intentionen zu den einzelnen Suchanfragen. Daher wird eine Befragung zur Ableitung des Informationsbedürfnisses konzipiert, deren Methode im folgenden Text erläutert und anschließend begründet wird. In den nächsten Kapiteln werden dann Aufbau und Ablauf dieser Befragung erläutert.

Die Online-Umfrage soll eine Einschätzung des Informationsbedürfnisses von unterschiedlichen Probanden pro Suchanfrage einholen, um anhand von Mehrheitsurteilen das jeweilige Informationsbedürfnis ableiten zu können. Die Befragung ist so konzipiert, dass die Probanden jeweils einen Fragebogen mit einer bestimmten Anzahl von Suchanfragen mit der Aufforderung erhalten, das typische Ziel hinter dieser Suchanfrage aufzuschreiben. Um dem Anteil von mehrdeutigen Suchanfragen (vgl. SONG et al. 2009, S. 228) gerecht zu werden, soll den Probanden auch die Möglichkeit eingeräumt werden ein zweites Informationsbedürfnis festzuhalten. Dieses ist dem ersten Informationsbedürfnis gleichrangig. Zusätzlich wird die geschätzte Aktualität der Suchanfrage erhoben. Eine genaue Einsicht in die Konzeption der Befragung kann in Kapitel 3.3.2 genommen werden.

3.3.1 Begründung der Methodenwahl

Die Methode der Online-Befragung wurde gewählt, da sie den großen Vorteil bietet, dass sie besonders schnell durchgeführt und ausgewertet werden kann (vgl. DIEKMANN 2010, S. 522). Außerdem ist diese Form der Befragung meist kostengünstiger als ihre schriftliche Version. Mit der Abbildbarkeit der gewünschten Fragen, erfüllen die aufgezählten Eigenschaften alle Kriterien, die die Art der Erhebungsmethode für die vorliegende Aufgabenstellung erfüllen sollte. Nachteile, wie eine geringe Rate von Beantwortungen (Responserate) (vgl. COUPER und COUTTS 2006, S.223f.), werden daher in Kauf genommen. Durch eine persönliche Ansprache der Probanden wird allerdings der Versuch unternommen, eine niedrige Responserate zu umgehen.

Die Ableitung des Informationsbedürfnisses der Suchanfrage soll frei erfolgen. Zwar beschäftigen sich Forscher schon seit einiger Zeit mit der Identifikation und dem Verstehen der Ziele, die User in der Websuche verfolgen und schlagen Kategorien und Taxonomien von Nutzerzielen (vgl. z.B. BRODER 2002; CALDERON-BENAVIDES et al. 2010; ROSE und LEVINSON 2004;) sowie deren automatische Klassifizierung (vgl. z.B. JANSEN und BOOTH 2010; LEE et al. 2004; LU et al. 2006) vor. Die Konsultation dieser Studien liefert allerdings lediglich die Erkenntnis, dass die Kategorien, die zur Einteilung der Suchintentionen dienen, je nach Forscher variieren. Die Einteilung in navigations-, informations- und transaktionsorientierte Suchanfragen nach Broder (vgl. BRODER 2002) ist in ihren Grundzügen zwar zu erkennen, jedoch gibt es immer wieder Abweichungen (vgl. CALDERON-BENAVIDES et al. 2010; KANG und KIM 2003; ROSE und LEVINSON 2004). Da folglich keine Einigung über die bestehenden Suchanfragetypen erfolgt, wird geschlussfolgert, dass möglicherweise weitere oder andere Kategorien bestehen. Aus diesem Grund wird dem Probanden, im Gegensatz zu dem Vorgehen von Stamou und Efthimiadis (vgl. STAMOU und EFTHIMIADIS 2009, S. 3f.), zur Ableitung der Suchintention nicht eine Auswahl an Kategorien vorgelegt, um das Informationsbedürfnis zu bestimmen. Stattdessen sind die Probanden angehalten frei das Informationsbedürfnis der Suchanfragen abzuleiten. Mit dieser Methode gehen keine In-

formationsbedürfnisse verloren und es besteht nachfolgend immer noch die Möglichkeit, die abgeleiteten Informationsbedürfnisse zu Kategorien zu verengen.

Andere Methoden, die bereits eingesetzt wurden, um Informationsbedürfnisse abzuleiten, wie die zuletzt angeklickte URL (vgl. DOWNEY et al. 2008, S. 451) oder das Klickverhalten auf der Suchergebnisseite (vgl. z.B. LEE et al. 2004, S. 394 f.; VON MACH und OTTE 2009, S. 50f.) können ebenfalls bei der Ableitung nicht eingesetzt werden, da kein Klick ausgelöst wurde.

Um nachträglich eine möglichst realistische Ableitung der Informationsbedürfnisse zu erreichen, werden mehrere Probanden eingesetzt, die das Informationsbedürfnis ohne Vorgaben in Form von Kategorien o.ä. ableiten. Dieses Verfahren wurde von Huffman und Hochster ebenfalls angewandt (vgl. HUFFMANN und HOCHSTER 2007). Sie bitten zu jeder Suchanfrage fünf Probanden das Informationsbedürfnis frei abzuleiten, aufgrund der Mehrdeutigkeit von Suchanfragen erlauben sie die Abgabe eines zweiten Informationsbedürfnisses. Anschließend wird pro Suchanfrage das häufigste erste und zweite Informationsbedürfnis bestimmt und die Sprache der abgegebenen Bedürfnisse normalisiert (vgl. HUFFMANN und HOCHSTER 2007, S. 568). In Anbetracht der Wichtigkeit der Ableitung des Informationsbedürfnisses wird auch in der vorliegenden Studie eine Konsultation mehrerer Probanden pro Suchanfrage vorgenommen, um aufgrund einer Mehrheitsentscheidung das Bedürfnis hinter der Suchanfrage ableiten zu können. Dieses Vorgehen soll einer Beeinflussung der Ergebnisse der Arbeit durch die Meinung der Forscher vorbeugen, wie sie bspw. in der Studie von Li et al. (siehe Kapitel 2.1 und vgl. LI et al. 2009) durch die von den Forschern vorgenommene Klassifizierung der Suchanfragen angenommen werden kann.

Die Aktualität der Suchanfrage wird ebenfalls abgefragt, um eine mögliche Änderung der Suchergebnisse einschätzen zu können. Da die zur Verfügung stehende Stichprobe im Monat April 2010 gezogen wurde und die Abfrage der Relevanz erst im August 2010 durchgeführt werden konnte, ist eine Änderung der Sucher-

gebnisse bei aktuellen Suchthemen wahrscheinlich. Daher soll auf diese Weise mit Hilfe der Mehrheitsentscheidungen erhoben werden, welche Suchanfragen einen besonderen Aktualitätsbezug haben, um dies bei der Analyse der Ergebnisse beachten zu können. Weitere Hinweise zu dieser Problematik können in Kapitel 4.2.2 und 5.1 nachgelesen werden.

3.3.2 Konzeption

Als Datenbasis der mit einem Informationsbedürfnis auszustattenden Suchanfragen wird eine Zufallsstichprobe von 600 ungeklickten Suchanfragen aus der in Kapitel 3.2 erläuterten Datenbasis gezogen. Es wird die Fallzahl von 600 Suchanfragen gewählt, da dies eine ausreichend große Anzahl für repräsentative Aussagen darstellt. Gleichzeitig ist aber auch gewährleistet, dass ausreichend Probanden angesprochen werden können, damit jede Suchanfrage bearbeitet werden kann.

Diese Stichprobe wird für die Untersuchung allerdings sowohl von allen Anfragen, die als sinnlos eingestuft werden, als auch von fremdsprachigen Anfragen und von Anfragen, die in fremdsprachige Suchergebnisseiten resultieren, befreit. Auf diese Weise soll sichergestellt werden, dass die Anfragen und die Suchergebnisse, mit denen die Probanden konfrontiert werden, von allen Probanden verstanden werden. Des Weiteren werden anstößige Anfragen aus der Stichprobe eliminiert. Um zu sinnvollen Untersuchungsergebnissen zu gelangen, werden außerdem alle Suchanfragen, die nicht in organischen Ergebnissen resultieren, aus der Stichprobe für die Untersuchung entfernt. Denn in diesem Fall ist deutlich, dass der Suchende kein Ergebnis angeklickt hat, da er kein Suchergebnis angezeigt bekommt. Da Tang et al. darauf hinweisen, dass Daten, die per Logfile erhoben werden, verrauscht sein können, da sie, neben den Anfragen der Nutzer, auch Einträge von Robots u.ä. enthalten (vgl. TANG et al. 2007, S. 3), wurden auch Anfragen, die sehr wahrscheinlich automatisiert abgeschickt wurden, entfernt. Die jeweiligen Anteile der hier aufgezählten Merkmale der Suchanfragen können in Kapitel 4.1.4.4 genau nachgelesen werden.

Nach Entfernung der beschriebenen Suchanfragen wurde die Stichprobe wieder aufgefüllt, so dass für die Befragung eine Datenbasis von 600 Suchanfragen vorlag.

Die Umsetzung der Befragung zur Ableitung des Informationsbedürfnisses erfolgt unter den in Kapitel 3.3 beschriebenen Gesichtspunkten, demnach werden die Probanden in einem online präsentierten Fragebogen gebeten frei das Informationsbedürfnis für die angezeigte Suchanfrage abzuleiten. Der genaue Aufbau der Befragung wird im folgenden Text beschrieben:

Zur Einstimmung auf die Untersuchung wird ein „Willkommenstext“ verfasst, der das Ziel hat, den Teilnehmer über den Inhalt der Befragung und ihren Zweck zu informieren. Außerdem wird in diesem Text erläutert, welche Aufgabe der Proband in der Befragung zu lösen hat. Die Aufgabenstellung lautet:

Bitte schreiben Sie zu jeder Suchanfrage auf, welches typische Ziel Ihrer Meinung nach hinter dieser Suchanfrage steckt.

Zum besseren Verständnis wird die Aufgabe mit zwei Beispielen versehen, die einerseits eine Suchanfrage mit einem und andererseits eine Suchanfrage mit zwei möglichen Informationsbedürfnissen zeigen (vgl. Anhang II).

Die folgenden Seiten des Fragebogens fragen die möglichen Informationsbedürfnisse von acht Suchanfragen ab, dazu wird zu Beginn der Seite die Suchanfrage präsentiert, anschließend folgt die konkrete Fragestellung:

Bitte schreiben Sie zu jeder Suchanfrage das Ihrer Meinung nach typische Ziel des Suchenden auf. Sollte es ein zweites mögliches Ziel geben, schreiben Sie dies bitte im zweiten Feld auf.

Nachfolgend werden dem Probanden zwei nummerierte Textfelder zur Niederschrift der Informationsbedürfnisse geboten.

Die zweite Frage befasst sich mit der Aktualität der Suchanfrage und bietet dem Probanden fünf verschiedene Antwortmöglichkeiten zur Auswahl:

Wie hoch schätzen Sie den Aktualitätsbezug der Suchanfrage ein? Das Thema der Suchanfrage …
O … ist sehr aktuell.
O … ist aktuell.
O … ist weniger aktuell.
O … hat keinen aktuellen Bezug.
O Kann ich nicht einschätzen.

Es kann nur eine der fünf Antworten ausgewählt werden.

Am Ende der Befragung steht eine Danksagung für die Teilnahme und die E-Mailadresse der Forscherin für Rückfragen und Anmerkungen zur Befragung.

Die Befragung wurde mit Hilfe des Online-Umfrage Tools von SurveyMonkey (vgl. SURVEYMONKEY 2010) umgesetzt.

3.3.3 Pretest

Der Pretest mit vier Probanden zeigt, dass ein Beispiel zur Veranschaulichung der Aktualität einer Suchanfrage von Nöten ist (vgl. Anhang II). Des Weiteren werden der Einleitungstext sowie die einzelnen Frageseiten übersichtlicher gegliedert. Außerdem konnte die Anzahl der abgefragten Suchanfragen pro Proband auf zehn Anfragen erhöht werden, da die Probanden die Aufgabe acht Suchanfragen abzuleiten schnell lösen konnten und auf Rückfrage äußerten, dass eine Integration zwei weiterer Suchanfragen in die Befragung möglich ist. Die Screenshots im Anhang II illustrieren die einzelnen Seiten der Befragung nach dem Pretest.

3.3.4 Durchführung

Die Befragung wird mit Hilfe von 60 Fragebögen umgesetzt, in jedem der 60 Fragebögen werden dieselben zehn Fragen, allerdings ausgestattet mit unterschiedlichen Suchanfragen gestellt. Die Distribution der Befragungen erfolgt als Link per E-Mail an die Probanden. In der E-Mail wurde das Thema der Befragung kurz angekündigt und auf eine anonyme Durchführung hingewiesen (siehe Anhang II). Die ausgewählten Probanden haben alle Erfahrungen mit einer Web-Suchmaschine. Dieses Kriterium wurde der Teilnahme vorausgesetzt, denn es wird angenommen, dass sie dadurch besser in der Lage sind sich in den ursprünglich Suchenden hineinzuversetzen.

Die Befragung wurde am 05.07.2010 gestartet und konnte bis zum 04.08.2010 von den angeschriebenen Probanden ausgefüllt werden.
Um die bei Online-Befragungen oftmals niedrige Responserate (vgl. COUPER und COUTTS 2006, S.223f.) zu umgehen, wurde eine Woche nach der Auslieferung der Befragung eine Erinnerungsnachricht versandt. Alle Probanden, die noch nicht an der Befragung teilgenommen hatten wurden damit an die Umfrage erinnert und nochmals gebeten teilzunehmen. Nach 14 Tagen wurden die Fragebögen, die keinerlei Rücklauf erfahren hatten, an Ersatzadressen weiterer Probanden weitergeleitet. Diese wurden, falls sie ebenfalls nicht an der Befragung teilnahmen, nach fünf Tagen ebenfalls an die Befragung erinnert.

3.4 Online-Befragung zur Untersuchung der Relevanz der Suchergebnisse nicht geklickter Suchanfragen

Die gewählte Methode sieht vor, dass der Proband, ebenfalls in Form einer Online-Befragung, die Relevanz der ersten zehn organischen Suchergebnisse der nicht geklickten Anfragen hinsichtlich des abgeleiteten Ziel der Suche bewertet. Außerdem soll die Frage beantwortet werden, ob er das betreffende Suchergebnis anklicken würde, ein Klick nicht notwendig ist, da die Trefferbeschreibung selbst bereits Antwort auf das Informationsbedürfnis gibt oder ob er den Treffer nicht anklickt, da er nicht relevant ist. Die sich anschließenden Kapitel begründen diese Methodenwahl und zeigen detailliert den Ablauf der Relevanzmessung auf.

3.4.1 Begründung der Methodenwahl

Die Wahl fiel auf die oben beschriebene Methode des online durchgeführten Relevanztests, da die aufgeworfenen Forschungsfragen auf diese Weise am effektivsten untersucht werden können. Denn durch die Wiedervorlage der Suchergebnisseite, die ein Suchender ohne einen Klick bedacht hat an die Probanden, kann die Suchergebnisseite im Nachhinein näher untersucht und damit auch beurteilt werden. Es geht dabei nicht darum, dass Verhalten des ursprünglich Suchenden zu simulieren, sondern vielmehr darum, die SERP bei nicht geklickten Treffern zu erforschen und damit aufzudecken, welche Möglichkeiten der originär Suchende gehabt hätte, um sein Informationsbedürfnis zu stillen.

Eine Alternative zur beschriebenen Methode wäre, eine eigene Erhebung von Daten durchzuführen, wie dies bei Stamou und Efthimiadis getan wurde (vgl. STAMOU und EFTHIMIADIS 2009 und Kapitel 2.1). Diese Methode hätte den Vorteil, dass der Suchende wirklich ein Informationsbedürfnis verspürt und sein Verhalten mit der SERP direkt beobachtet werden könnte. Aus mehreren Gründen wurde aber auf eine eigene Erhebung von Suchanfragen, Informationsbedürfnis und Klickverhalten verzichtet. Zum einen würde es den zeitlichen und organisatorischen Rahmen dieser Arbeit sprengen, denn um repräsentative Ergebnisse erheben zu können, müsste eine hohe Anzahl von Probanden eingebunden und

ihr Verhalten aufgezeichnet werden. Des Weiteren könnte die Beobachtungssituation, die automatisch entstehen würde, wenn das Suchverhalten der Probanden aufgezeichnet wird, zu einer Verzerrung der Ergebnisse führen. Diese Vermutung stützt sich auf Ausführungen zur Problematik einer offenen Beobachtung. Diese kommen zu dem Schluss, dass mit einem nicht natürlichen Verhalten gerechnet werden muss, wenn Personen offen beobachtet werden (vgl. ALBERS 2009, S. 40 f.; HAWKEY 2009, S.82; KROMREY 2009, S. 327 f., S.378 f.).
Dem Schwachpunkt der gewählten Methode, dass die Probanden nicht die originär Suchenden sind, wird durch den Einsatz mehrerer Probanden pro Suchanfrage zur Ableitung des Informationsbedürfnisses begegnet. Auf diese Weise werden Mehrheitsentscheidungen möglich, die das Aufstellen des Informationsbedürfnisses valider gestalten sollen (siehe auch Kapitel 3.3).

Die Entscheidung, die Studie in Form einer Online-Befragung durchzuführen, hat dieselben Gründe, wie sie bereits im Kapitel zur Ableitung des Informationsbedürfnisses (siehe Kapitel 3.3.1) genannt werden: Zum einen können auf diese Weise kostengünstig eine große Anzahl von Fragebögen generiert werden, auch müssen keine Interviewer bereitgestellt werden, die die Befragung mündlich durchführen oder bei einer schriftlichen Befragung, die Fragebögen verteilen und zum Mitwirken motivieren. Insgesamt ist diese Art der Befragung daher kostengünstiger als ihre schriftliche Form (vgl. DIEKMANN 2010, S. 522). Ein weiterer Vorteil der Online-Befragung ist die Schnelligkeit, mit der sie durchgeführt und ausgewertet werden kann (vgl. DIEKMANN 2010, S. 522). Schlussendlich ist die Zielgruppe, die befragt werden soll, und zwar Suchende im Internet, online zu erreichen.
Die Methode der Online-Befragung birgt allerdings auch Nachteile, die an dieser Stelle nicht unberücksichtigt gelassen werden sollen. Ein Nachteil ist, dass gängige Motivationstechniken nur schwer auf Online–Befragungen angewandt werden können (vgl. DIEKMANN 2010, S. 528). Es kann also kein weiterer Einfluss auf den Befragten ausgeübt werden, nachdem er den Online-Fragebogen präsentiert

bekommen hat. Aus diesen Gründen sind, wie bereits beschrieben, die Responseraten von Online-Befragungen oftmals sehr gering (vgl. COUPER und COUTTS 2006, S.223 f.). Ein weiterer Nachteil ist, dass bei einer selektiven Form der Online-Befragung keine Anteile und Mittelwerte für die gesamte Population geschätzt werden können (vgl. DIEKMANN 2010, S. 525). In dem vorliegenden Fall ist dies allerdings nicht problematisch, da die Aussagen nicht generalisiert auf eine bestimmte Gruppe übertragen werden sollen. Trotz der genannten Nachteile der Online-Befragung überwiegen für das vorliegende Forschungsvorhaben die Stärken dieser Befragungsform, da diese den Anforderungen an die geplante Untersuchung in Form von Kosten, Zeit und vor allem Umsetzbarkeit der Fragestellung gerecht werden.

3.4.2 Konzeption

Das Konzept des Relevanztests sieht vor, dass der Proband in einem ersten Schritt begrüßt, über den Zweck der Forschung aufgeklärt und auf die Einhaltung von Datenschutzbestimmungen hingewiesen wird. Außerdem werden ihm der Ablauf der Befragung und seine Aufgabe in zwei Punkten erläutert:

1. Beurteilen Sie die Relevanz des Treffers, indem Sie bewerten, ob das Suchergebnis das Ziel der Suchanfrage erfüllt. Berücksichtigen Sie bitte jeden Treffer unabhängig von den folgenden bzw. vorherigen Treffern.
2. Nachdem Sie die Relevanz des Treffers beurteilt haben, teilen Sie bitte mit, ob Sie den Treffer anklicken würden oder ob die Beschreibung des Treffers ausreicht, um das Ziel der Suche zu erreichen.

Damit der Proband die Aufgaben der Befragung korrekt lösen kann, veranschaulichen ihm zwei Beispiele den zweiten Teil der Aufgabe (kompletter Einleitungstext ist im Anhang III zu finden).

Anschließend wird dem Proband eine Suchanfrage mitsamt dem zugehörigen Informationsbedürfnis präsentiert. Einzeln werden dazu die ersten zehn organischen Suchergebnisse der ersten Suchergebnisseite in zufälliger Reihenfolge eingeblendet. Grund für diese Beschränkung auf die ersten zehn Treffer ist, dass in mehreren Studien (vgl. z.B. GRANKA et al. 2008, S. 1044; JANSEN und SPINK 2006, S.

257f.; LORIGO et al. 2006, S. 1127) nachgewiesen werden konnte, dass Suchende nicht bereit sind alle gefundenen Treffer anzusehen, vielmehr evaluieren sie größtenteils nur die erste Ergebnisseite (in den meisten Suchmaschinen also die ersten zehn Treffer). Aus den aufgeführten Untersuchungen kann also geschlossen werden, dass die ersten zehn Ergebnisse ausreichen, um eine Suche für den geplanten Relevanztest nachzustellen. Ein weiterer Effekt dieser geringeren Anzahl von zu bewertenden Suchergebnissen ist, dass der Proband evtl. auch motivierter ist und nicht so schnell ermüdet, wie es z.B. bei der Bewertung von 20 oder mehr Suchergebnissen der Fall sein könnte.

Aufgrund weiterer Untersuchungen, die zeigen, dass User vom dargebotenen Ranking der Suchmaschine beeinflusst werden und aus diesem Grund in vielen Fällen nur die am höchsten gerankten Ergebnisse vom User inspiziert werden (vgl. JOACHIMS und RADLINSKI 2007, S. 35; LORIGO et al. 2006, S. 1127), wird davon Abstand genommen, die SERP in ihrer ursprünglichen Form abzubilden, vielmehr sollen die Treffer einzeln zur Relevanzbewertung vorgelegt werden. Auf diese Weise kann sichergestellt werden, dass alle zehn Treffer vom Probanden (theoretisch) die gleiche Aufmerksamkeit erhalten und dementsprechend bewertet werden. Für dieses Vorgehen spricht auch, dass in einer Untersuchung von Lorigo et al. festgestellt wurde, dass bei nur vier Prozent aller Suchanfragen der Stichprobe, die keinen Klick erhielten, alle zehn Treffer angesehen werden. Laut Lorigo et al. weist dies daraufhin, dass der Suchende meist sehr schnell, und zwar ohne alle Suchergebnisse zu evaluieren, entscheidet, ob er die aktuelle Suche mit einer neuen oder abgeänderten Suchanfrage abbricht (vgl. LORIGO et al. 2006, S. 1128). Da es in der vorliegenden Untersuchung nicht darum geht, dass Suchverhalten zu simulieren, sondern die Suchergebnisse zu den ungeklickten Suchanfragen zu untersuchen, wäre es hinderlich, wenn nicht jedes Suchergebnis einzeln bewertet würde.

Dass die Trefferbeschreibungen dem Probanden in einer zufälligen Reihenfolge vorzulegen sind, unterstreichen die oben aufgeführten Argumente ebenfalls. Denn durch die Mischung des Rankings können Lerneffekte, die die Reihenfolge

der Ergebnispräsentation beim Betrachter auslöst (vgl. GORDON und PATHAK 1999, S. 147; GRIESBAUM et al. 2002, S. 210 f.), umgangen werden.

Zur Einschätzung der Relevanz eines Treffers wird der Proband bei der Anzeige von jedem Suchergebnis dazu aufgefordert die Relevanz des Treffers anhand einer fünfstufigen Skala einzuordnen. Die konkrete Fragestellung lautet hier:

Wie relevant ist dieser Treffer?				
Je höher der Wert, desto relevanter der Treffer.				
O	O	O	O	O
0	1	2	3	4

Auf diese Weise kann der Proband den Treffer nicht nur als irrelevant oder als relevant einordnen, vielmehr kann eine weitere Relevanzabstufung durch die mittleren Skalenwerte vorgenommen werden. Die Entscheidung fiel für diese Art der Bewertung, da nicht angenommen wird, dass alle relevanten Dokumente für den Nutzer den gleichen Wert haben. Vielmehr sind wahrscheinlich einige Dokumente als teilweise relevant, andere wiederum als sehr relevant einzuordnen. Eine weitere Möglichkeit der Relevanzeinordnung wäre eine binäre Zuordnung. Um die beschriebenen Abstufungen abbilden zu können, wird jedoch auf diese Darstellung verzichtet. Des Weiteren konnte in verschiedenen Studien herausgearbeitet werden, dass Probanden bei der Beurteilung von Relevanz anhand einer Skala von 0 bis 100 einem Treffer bereits ab einem Wert von 41 bzw. 46 (absolute) Relevanz zuschreiben (vgl. STOCK 2007, S. 74, zit. nach EISENBERG und HU 1987 und Janes 1991). Hier wird also nicht wie vermutet der Mittelwert von 50 genutzt, um zwischen einem (absolut) relevanten und weniger relevanten Treffer zu unterscheiden. Der niedrige(re) Schwellenwert zur Beurteilung der Relevanz weist daraufhin, dass ein Treffer mit einer höheren Wahrscheinlichkeit als relevant eingestuft wird, als das er als irrelevant eingestuft wird. Diesem Effekt soll ebenfalls mit der Nutzung der Skala entgegen getreten werden, denn auf diese Weise kann der Proband einen Treffer auch als nur teilweise relevant einordnen.
Zudem kann eine Skala zur Relevanzeinordnung im vorliegenden Kontext detailliertere Hinweise auf die Relevanz der Treffer bei ungeklickten Suchergebnissei-

ten geben. So kann z.B. vermutet werden, dass, falls Ergebnisse als relevant einsortiert werden, diesen evtl. nur ein sehr geringer Relevanzwert auf der Skala zugeschrieben wird. Durch die Nutzung der beschriebenen Skala können folglich differenziertere Ergebnisse über die Qualität der Treffer einer ungeklickten Suchanfrage eingeholt werden.

Mit Darstellung einer weiteren Frage wird der Proband außerdem vor die Entscheidung gestellt, ob er den Treffer anklicken würde, ihn aufgrund fehlender Relevanz nicht anklicken würde oder ob die bloße Ansicht des Treffers ausreicht, um das vorliegende Informationsbedürfnis zu befriedigen. Dazu wird dem Proband die folgende Frage gestellt:

Würden Sie den Treffer anklicken oder reicht die Trefferbeschreibung aus, um das Ziel der Suche zu erreichen?
O Ja, ich würde den Treffer anklicken.
O Nein, ich würde nicht klicken. Die Ansicht des Treffers reicht, denn er erfüllt das Ziel der Suche.
O Nein, ich würde nicht klicken, da der Treffer nicht relevant ist.

Ist diese Frage beantwortet, wird dem Probanden das nächste Suchergebnis präsentiert. Auf diese Weise werden alle organischen Treffer der ersten SERP abgefragt. Wenn alle Suchergebnisse für eine Anfrage abgefragt wurden, wird mit der Abfrage der nächsten Suchanfrage begonnen. Jeder Proband soll auf diese Weise mit zehn Suchanfragen konfrontiert werden. D.h., dass jeder Teilnehmer 100 Treffer zu bewerten hat.

Wie aus der Beschreibung des Untersuchungsablaufs ersichtlich, werden dem Probanden zur Relevanzbewertung (nur) die Trefferbeschreibungen vorgelegt. Eine zusätzliche Abfrage der Relevanz des dahinter liegenden Dokuments ist nicht vorgesehen, da auch der originär Suchende nicht das Suchergebnis selbst, sondern nur die Trefferbeschreibung evaluierte und auf Grundlage dessen die Entscheidung getroffen hat, keinen Treffer anzuklicken. Verschiedene Studien (vgl. z.B. GRIESBAUM 2004; LEWANDOWSKI 2008a) zeigen jedoch, dass die Relevanzbewertung von Trefferbeschreibung und Suchergebnis oft nicht über-

einstimmen. D.h. es kann nicht zwingenderweise auf eine schlechte Suchleistung der Suchmaschine geschlossen werden, wenn für einige Suchanfragen das Ergebnis erzielt wird, dass kein Klick erfolgte, da alle Ergebnisse als nicht relevant eingestuft wurden. In diesem Fall könnte z.B. auch angenommen werden, dass die Fähigkeit der Suchmaschine, adäquate Trefferbeschreibungen zu generieren, eingeschränkt ist und hinter den nicht relevanten Trefferbeschreibungen relevante Dokumente liegen.

Der Relevanztest wurde mit Hilfe eines eigens für Relevanzmessungen von Web-Suchmaschinen programmierten Tools umgesetzt. Die Besonderheit dieses Tool ist die Möglichkeit eine Suchanfrage anzulegen und ihre Ergebnisse bei der gewünschten Suchmaschine mit Hilfe eines sogenannten Scrapers einzuholen. Die auf diese Weise eingeholten Ergebnisse in Zusammenhang mit Suchanfragen und Ziel der Suche können anschließend dem Probanden komfortabel in einer Online-Befragung präsentiert werden. Eine Ansicht der Befragung kann in Anhang III eingesehen werden.

3.4.3 Pretest

Vor der Durchführung des Relevanztests wurde ein Pretest durchgeführt, um eventuell auftretende Probleme aufzudecken und zu beheben. Die teilnehmenden Probanden wurden lediglich nach dem Kriterium „Erfahrungen mit einer Web-Suchmaschine“ ausgewählt. Dieses Kriterium wird für die Teilnahme vorausgesetzt, da davon ausgegangen wird, dass Personen, die selbst die Web-Suche nutzen, routinierter und damit einfacher und schneller die Relevanz der Treffer einschätzen können.

Der Pretest deckte die im Folgenden aufgelisteten Probleme auf, die anschließend durch den Programmierer des Tools mit den beschriebenen Lösungen behoben wurden:

- Die Probanden konnten die Treffer im Pretest nicht auf den ersten Blick sehen. Da die Treffer im unteren Bereich des Bildschirms angezeigt werden (siehe auch Abb. 27 im Anhang III), verschiebt sich die Darstellung der Treffer bei einem kleineren Bildschirm, einer Öffnung des Relevanztests in einem kleinen Fenster und bei Nutzung des Internet Explorers. Hier konnten die zu bewertenden Treffer dann nur noch durch scrollen sichtbar gemacht werden. Die auf den ersten Blick fehlenden Treffer verunsicherten die Testpersonen im Pretest und ließen sie die Befragung frustriert abbrechen.

  Die Darstellung des Relevanztests wurde daher so eingestellt, dass eine dynamische Anpassung der Darstellung erfolgen konnte. Auf diese Weise wurde die Wahrscheinlichkeit erhöht, dass die Treffer gleich auf den ersten Blick ersichtlich wurden. Zusätzlich wurde in dem Anschreiben an die Probanden darauf hingewiesen, dass sich der Test in einem maximierten Fenster am besten ausfüllen lässt (siehe auch Anhang III).
- In den Tests wurde ersichtlich, dass in einigen Browserversionen die Umlaute, die in Suchanfrage oder der Beschreibung des Informationsbedürfnisses vorkommen, nicht korrekt umgewandelt wurden. Daher wurden alle Umlaute durch die HTML Version dieser Buchstaben ersetzt.
- Des Weiteren wurde ersichtlich, dass Suchen, die zu weniger als zehn Ergebnissen führten, im Tool noch nicht korrekt dargestellt werden konnten. Aus diesem Grund wurden diese Anfragen aus der Stichprobe entfernt.

Außerdem wurde im Pretest deutlich, dass es sinnvoll ist, die Probanden darauf hinzuweisen, dass das Tool erst vor Kurzem fertig gestellt wurde und daher noch Fehler auftreten könnten. Aus diesem Grund wurde dies im Anschreiben an die Probanden erwähnt und gebeten, Fehler an die Forscherin zu melden.

3.4.4 Durchführung

Die Probanden wurden mit Hilfe eines Links zur LogIn Seite des Relevanztools geleitet. Hier konnten sie, nach Eingabe eines Zugangscodes an der Relevanzmessung teilnehmen. Gemeinsam mit dem Link wurde den Testteilnehmern außerdem eine kurze Einführung mit dem Hinweis auf eine anonyme Durchführung und der Zugangscode per E-Mail zu geschickt (siehe Anhang III).

Die Befragung startete am 27.08.2010 und wurde bis zum 19.09.2010 durchgeführt. Um einen höheren Rücklauf zu erhalten, wurden am 08.09.2010 erneut E-Mails an die Befragten versandt, die daran erinnerten an der Befragung teilzunehmen.

4 Ergebnisse

In diesem Kapitel werden die Ergebnisse, die mit Hilfe der unterschiedlichen Methoden zur Untersuchung der nicht geklickten Treffer erhoben werden konnten, beschrieben. Die Präsentation der Ergebnisse folgt dabei der in Kapitel 3 gewählten Reihenfolge. Eine zusammenfassende Interpretation und Auswertung aller Resultate dieser Arbeit kann in dem sich anschließenden Kapitel 5 nachgelesen werden.

4.1 Deskriptive Analyse der Datenbasis

Die Stichprobe der ca. eine Millionen Suchanfragen enthält einen Anteil von 65.535 Suchanfragen, die weder einen Klick auf die organischen noch einen Klick auf die Textanzeigen erhalten haben. Die Datenbasis enthält also einen Anteil von 6,56 Prozent nicht geklickter Suchanfragen.

Ein Vergleich, mit den Anteilen nicht geklickter Suchanfragen in früheren Studien zeigt wesentlich höhere Werte. Callan et al. halten bspw. fest, dass bis zu 50 Prozent der Suchen nicht einen Klick erhalten (vgl. CALLAN et al. 2007, S. 27). Bei der eigenen Erhebung von Suchen und Klickverhalten im Jahr 2009 erhalten Stamou

und Efthimiadis einen Anteil von 27 Prozent ungeklickter Suchanfragen. Auch wenn die beiden Vergleichswerte sich unterscheiden, unterstreichen sie doch die in Kapitel 3.2 unter Punkt a) aufgeführte Vermutung, dass unter den Suchanfragen, die nicht in den Top Millionen Anfragen enthalten sind, eine größere Anzahl von nicht geklickten Suchanfragen zu finden ist, die durch die fehlende Zufallsstichprobe wegfällt. Ob die Stichprobe der Top Millionen sich auch hinsichtlich weiterer Merkmale von einer Zufallsstichprobe von Suchanfragen unterscheidet, wird in dem folgenden Abschnitt behandelt.

4.1.1 Länge der Suchanfragen

Im Durchschnitt besteht eine Suchanfrage der Stichprobe aus 2,1 Wörtern. Dieser Wert liegt im Durchschnitt, denn er entspricht der Suchanfragenlänge anderer Studien. Schmidt-Mänz stellt eine Übersicht verschiedener Studien, die die Analyse von Suchanfragen zum Ziel haben, zusammen. Hier zeigt sich, dass die Länge der Suchanfragen zwischen 1,6 und 2,9 Wörtern schwankt, ein klarer Trend ist nicht erkennbar (vgl. SCHMIDT-MÄNZ 2007, S. 27f.). Die Abb. 1 gibt detaillierter Aufschluss über die Anzahl der Wörter pro Suchanfragen für die vorliegende Stichprobe.

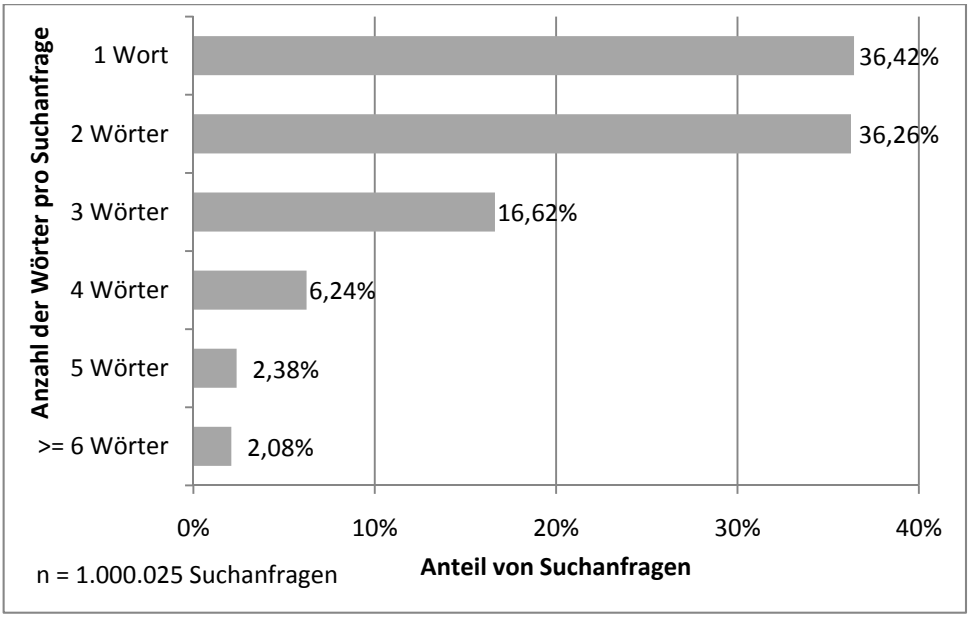

Abb. 1: Anzahl der Wörter pro Suchanfrage (Quelle: eigene Darst.)

Fast gleich viele Suchanfragen bestehen aus einem Wort (36,42 Prozent) bzw. zwei Wörtern (36,26 Prozent). Eine Studie von Jansen et al. zeigt eine ähnliche Verteilung, auch hier stellen die Anfragen, die aus einem oder zwei Worten bestehen, mit jeweils 31 Prozent fast identisch große Anteile der Stichprobe dar (vgl. JANSEN et al. 2000, S. 213). Der Anteil der Ein-Wort Anfragen in der vorliegenden Stichprobe ist jedoch im Vergleich zu bereits durchgeführten Studien leicht erhöht, in den von Schmidt-Mänz zusammengetragenen Studien erstrecken sich die Anteile von Ein-Wort-Anfragen von 20,4 Prozent bis 33 Prozent (vgl. SCHMIDT-MÄNZ 2007, S. 27). Dementsprechend kleiner sind in der vorliegenden Stichprobe die Anteile der Anfragen, die aus drei Worten bestehen: 16,62 Prozent der Suchanfragen der Stichprobe erfüllen dieses Kriterium. Anfragen mit vier Worten stellen einen Anteil von 6,24 Prozent und Suchanfragen mit fünf Worten einen Anteil von 2,38 Prozent an der Stichprobe. Mehr als sechs Worte enthält nur ein Anteil von 2,08 Prozent der Suchen.

Insgesamt kann an dieser Stelle mit Hinblick auf vergleichbare Untersuchungen geschlossen werden, dass die Stichprobe hinsichtlich der Länge der Suchanfragen keine nennenswerten Besonderheiten aufweist.

4.1.2 Nutzung von Suchoperatoren

Um die Stichprobe weiter zu charakterisieren und den Vergleich mit bereits durchgeführten Logfile-Analysen weiterführen zu können, wird die Nutzung von Suchoperatoren erhoben. Die T-Online Suche unterstützt die Suchoperatoren AND (auch als Symbol +), NOT (auch als Symbol -) und OR (auch als Symbol |). Mehrere Begriffe können mit der Phrasensuche unter Nutzung der Anführungszeichen verbunden werden (vgl. T-ONLINE.DE SUCHE 2010).

Der Suchoperator AND bzw. das Symbol + wurde in 4.421 Suchanfragen angewandt. Dabei wurde darauf geachtet, dass der Operator als Verknüpfung eingesetzt wurde, d.h. das Anfragen, die zu Beginn oder am Ende der Suchanfrage AND oder + nutzen, nicht mit einbezogen wurden. Dies stellt einen Anteil von 0,44 Prozent der Suchanfrage der Stichprobe dar.

NOT bzw. das Symbol – wurden in 71.405 Suchanfragen eingesetzt (7,14 Prozent), auch hier wurden Anfragen, die zu Beginn oder am Ende der Suchanfrage NOT oder das Symbol – einsetzten, nicht einbezogen.

Der Suchoperator OR bzw. das Symbol |, konnten in 70 Suchanfragen entdeckt werden. Dies stellt einen Anteil von 0,006 Prozent aller Suchanfragen der Stichprobe dar. Auch hier durfte der Operator nicht am Anfang oder Ende des Wortes genutzt werden.

Insgesamt bleibt an dieser Stelle anzumerken, dass bei der Nutzung der Operatoren AND, NOT und OR nicht sicher geschlossen werden kann, dass diese wirklich als Operatoren eingesetzt wurden. Vielmehr lässt eine grobe Durchsicht der Suchanfragen vermuten, dass diese Worte nicht mit der Absicht als Suchoperator zu fungieren eingesetzt wurden. Gleiches gilt für den Einsatz der Symbole als Operatoren, besonders das Minuszeichen wird bei Suchanfragen, die URLs enthalten häufig verwendet und steht daher höchstwahrscheinlich nicht für den Suchoperator NOT. Die aufgeführten Zweifel lassen sich nicht eindeutig nachweisen, trotzdem sollten die aufgeführten Anteile nur unter Vorbehalt betrachtet werden. Aufgrund der beschriebenen Unsicherheit können die Operatoren nicht zum Vergleich mit weiteren Logfile-Analysen herangezogen werden.

Als weiterer Suchoperator wurde das Anführungszeichen zur Verwendung der Phrasensuche untersucht. In vorangegangen Studien zeigt sich hier ein Anteil von ca. fünf Prozent der Anfragen, die als Phrasen gestellt wurden (vgl. SPINK et al. 2001, S. 229). Die vorliegende Datenbasis enthält jedoch keine einzige Phrasensuche. An dieser Stelle kann nur gemutmaßt werden, dass dies an der fehlenden Zufallsauswahl der Stichprobe liegt. Es wäre z.B. möglich, dass die Top Millionen Suchanfragen bei diesem Merkmal eine Verzerrung auslösen, da die „gängigen“, häufig gesuchten Anfragen ohne diesen Operator gestellt werden.

Zusammenfassend muss festgehalten werden, dass die Untersuchung von Suchoperatoren nicht den gewünschten Vergleich mit bereits veröffentlichten Studien ermöglicht und daher keine weiteren Hinweise für eine Verzerrung der Stichprobe gesammelt werden konnten.

4.1.3 Bestandteile einer URL in der Suchanfrage

Ein weiteres Merkmal, welches untersucht wurde, ist der Anteil der Suchen, die eine URL oder Teile einer URL enthalten. Dabei wurde nach 34 verschiedenen Domainendungen in Kombination mit einem Beginn der Anfrage mit „www.“ gesucht. Insgesamt wurden 50.409 Anfragen identifiziert, die mit „www.“ beginnen und eine URL typische Endung wie z.B. „.de“ oder „.com“ enthielten. Dies machte einen Anteil von 5,04 Prozent aller gestellten Suchanfragen aus. Dabei wurden am häufigsten Anfragen, die mit „www.“ begannen und mit „.de“ (77,40 Prozent) und „.com“ (14,56 Prozent) endeten, gestellt. Filtert man nur nach Anfragen, die mit „www.“ beginnen, zeigt sich ein Anteil von 6,18 Prozent der Stichprobe, wird nur auf die Endung „.de“ geachtet, erhält man einen Anteil von 6,64 Prozent.

Die Untersuchung dieses Merkmals ist insbesondere für den im folgenden Kapitel stattfindenden Vergleich von geklickten und nicht geklickten Suchanfragen interessant. Denn Suchanfragen, die Bestandteile einer URL enthalten, weisen stark daraufhin, dass der Suchende zu einer bestimmten Website navigieren möchte (vgl. THUROW und MUSICA 2009, S. 46). Damit können diese Anfragen als navigationsorientierte Suchanfragen klassifiziert werden. Eine navigationsorientierte Anfrage ist nach Broder eine Suchanfrage, die das Ziel hat zu einer bestimmten Webseite zu navigieren, die dem Suchenden bereits bekannt ist oder bei der angenommen wird, dass diese Seite existiert. Das Informationsbedürfnis des Suchenden kann in diesem Fall mit nur einem relevanten Treffer gestillt werden (vgl. BRODER 2002, S. 5). Im bereits angesprochenen Vergleich ist interessant, wie groß der Anteil der Suchanfragen, die eine URL oder Teile einer solchen enthalten, ist, wenn kein Suchergebnis angeklickt wurde. Denn in einer Studie von Lewandowski konnte herausgefunden werden, dass Google in über 80 Prozent der Fälle in der Lage ist bei einer navigationsorientierten Suchanfrage den korrekten Treffer auf die erste Position der SERP zu bringen (vgl. LEWANDOWSKI 2009, S. 6). Da die T-Online Suche von Google bereitgestellt wird, kann dieses Forschungsergebnis auch auf diese Suche übertragen werden. Daher sollte die Gruppe der

Anfragen, die URLs oder Teile einer solchen enthält, bei den nicht geklickten Anfragen relativ klein sein.

Abschließend muss an dieser Stelle gefolgert werden, dass die Herausarbeitung der Charakteristika der zur Verfügung stehenden Stichprobe für die Eingrenzung der Verzerrung keine nennenswerten Einsichten erbracht hat. Zwar zeigt die durchschnittliche Länge der Suchanfragen, dass die Stichprobe hier im Durchschnitt liegt. Andererseits konnten einige Merkmale nicht verglichen werden, da entweder keine Vergleichswerte gefunden werden konnten (z.B. Anteil der Suchanfragen, die keinen Klick auslösten) oder eine Auswertung des Merkmals nicht sauber ausgeführt werden konnte (z.B. Anteil der Suchanfragen, die Suchoperatoren nutzen). Daher gilt weiterhin die Aussage aus Kapitel 3.2, dass die Problematik der fehlenden Zufallsstichprobe nicht umgangen werden kann und die Ergebnisse anschließend dementsprechend interpretiert werden müssen.

4.1.4 Vergleich von geklickten und ungeklickten Suchanfragen

Um evtl. vorhandene Unterschiede zwischen geklickten und nicht geklickten Suchanfragen herausarbeiten zu können, wird die Stichprobe der Top Millionen Suchanfragen aufgeteilt in Suchanfragen mit Klick und Suchanfragen ohne Klick. Dabei erhält man 934.490 Suchanfragen (93,45 Prozent), die mindestens einen Klick auf die organischen Suchergebnisse oder die Textanzeigen erhalten haben und 65.535 Suchanfragen (6,55 Prozent), die keinen Klick erhalten haben. Mit der folgenden deskriptiven Untersuchung dieser beiden Stichproben sollen im Vergleich Besonderheiten herausgearbeitet werden.

4.1.4.1 Länge der Suchanfragen

Die Tab. 1 gibt Aufschluss über die Anzahl der Wörter pro Suchanfrage für die beiden Suchanfrage-Gruppen geklickte und ungeklickte Suchen.

Kategorie	**Suchanfragen mit Klick** (n=934.490 Anfragen)	**Suchanfragen ohne Klick** (n=65.535 Anfragen)
Durchschnittliche Anzahl von Worten	2,07 Wörter	2,63 Wörter
Anteil Ein-Wort Anfragen	37,15 %	26,07 %

Tab. 1: Anzahl der Wörter und Anteil der Ein-Wort-Suchen für Suchanfragen mit und ohne Klick (Quelle: eigene Darst.)

Vergleicht man die beiden Gruppen, erkennt man, dass die nicht geklickten Suchanfragen durchschnittlich aus mehr, nämlich 2,63 Wörtern, bestehen als die geklickten Anfragen. Diese weisen eine durchschnittliche Wortlänge von 2,07 Wörtern auf. Dementsprechend ist der Anteil der Ein-Wort Anfragen bei den nicht geklickten Suchanfragen um mehr als 10 Prozent niedriger als derselbe Anteil der geklickten Anfragen.

Die weiteren Wort-Anteile der nicht geklickten und der geklickten Suchanfragen können der Abb. 2 entnommen werden.

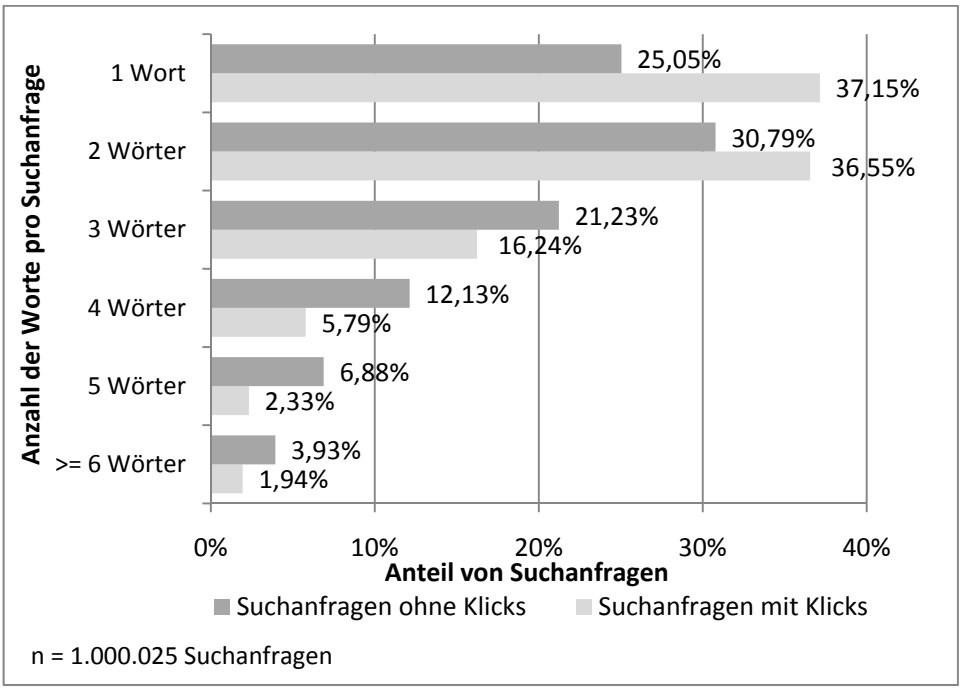

Abb. 2: Anzahl der Wörter pro Suchanfragen für Suchanfragen mit und ohne Klick (Quelle: eigene Darst.)

Über die unterschiedliche durchschnittliche Länge der Suchanfragen mit und ohne Klick können an dieser Stelle nur Vermutungen angestellt werden. Bailey et al. prüfen den Zusammenhang zwischen Art des Informationsbedürfnisses und der Länge der Suchanfrage. Sie kommen zu dem Ergebnis, dass das Informationsbedürfnis immer unspezifischer und inhaltlich weiter gefasst ist, wenn die Länge der Suchanfrage abnimmt (vgl. BAILEY et al. 2007, S. 2). Dementsprechend könnte gefolgert werden, dass hinter längeren Anfragen ein spezifischeres Informationsbedürfnis steckt, welches allerdings in den vorliegenden Fällen nicht durch einen Klick bedacht wurde. Wenn man nun davon ausgeht, dass die Suchenden keinen Treffer anklickten, weil die Ergebnisse irrelevant waren, könnte dies damit begründet werden, dass die gängigen Web-Suchmaschinen darauf ausgelegt sind, weit gefasste, kurze Anfragen zu beantworten. Denn diese kurzen Anfragen, ohne Einsatz von Suchoperatoren, stellen auch den größten Anteil al-

ler Suchen an eine Suchmaschine dar (vgl. SCHMIDT-MÄNZ 2007, S. 28). Auch das Unternehmen Google, was mit einem Marktanteil von 90,2 Prozent die in Deutschland am häufigsten genutzte Web-Suchmaschine betreibt (vgl. WEBHITS 2010), fordert in seiner Hilfe zur Web-Suche unter dem Titel „Tipps für bessere Suchen“ bspw. dazu auf, die Suchanfragen so einfach wie möglich zu halten und das Gesuchte mit so wenigen Wörtern wie möglich zu beschreiben (vgl. GOOGLE WEB-SUCHE HILFE 2010).

Es wird an dieser Stelle also festgehalten, dass nicht geklickte Suchanfragen durchschnittlich mehr Wörter und damit wahrscheinlich weniger weit gefasste Bedürfnisse beinhalten als die geklickten Suchanfragen. Weiter kann gemutmaßt werden, dass sie aufgrund dessen nicht in das gängige Anfragenschema der Web-Suchmaschinen passen und daher keine relevanten Ergebnisse geliefert werden (können). Dieser Behauptung wird in den folgenden Kapiteln nachgegangen.

4.1.4.2 Bestandteile einer URL in der Suchanfrage

Anfragen, deren Suchergebnisse nicht geklickt wurden, beginnen in 4,04 Prozent der Fälle mit dem URL Bestandteil „www.“ und enden mit einer typischen URL Endung wie „.de“ oder „.com“. Nutzt man nur die Endung „.de“ als Indikator für eine navigationsorientierte Suchanfrage, ergibt sich hier ein Anteil von 6,9 Prozent aller nicht geklickten Anfragen, nutzt man hingegen nur den Bestandteil „www.“ zeigt sich ein 6,07 prozentiger Anteil an der Stichprobe der ungeklickten Suchanfragen.

Einen etwas größeren Anteil von Suchanfragen, die eine URL oder Bestandteile einer solchen enthalten, zeigt sich bei der Untersuchung der geklickten Suchanfragen. Hier beginnen 5,09 Prozent der Suchen mit dem URL typischen Bestandteil „www.“ und enden mit einer der typischen URL Endungen. Filtert man auch hier nur die Suchanfragen heraus, die die URL typische Endung „.de“ enthalten, ergibt sich ein Anteil von 6,62 Prozent aller geklickten Suchanfragen, die Suchen, die mit dem Bestandteil „www.“ beginnen machen einen Anteil von 6,19 Prozent der geklickten Suchen aus.

Die oben aufgeführte Vermutung, dass der Anteil der navigationsorientieren Suchanfragen in der Stichprobe der Suchanfragen, die keinen Klick auslösten, wesentlich geringer sein müsste, als in der Stichprobe der Suchanfragen, die einen Klick auslösten, bestätigt sich nicht. Der Anteil in der Stichprobe der Suchanfragen mit Klick ist mit ca. einen Prozent nur leicht erhöht. Der fehlende, aber erwartete große Unterschied zwischen geklickten und nicht geklickten Anfragen könnte sich damit erklären, dass in dieser Arbeit nur anhand der URL Bestandteile herausgefiltert wurde, ob es sich um eine navigationsorientierte Anfrage handelt oder nicht. Da sich eine navigationsorientierte Anfrage noch anhand weiterer Kriterien, wie z.B. der Verwendung eines Domainnamens (z.B. „Facebook“ oder „Wer kennt wen“) oder der Suche nach dem Namen eines Unternehmens (z.B. „Daimler AG“, „Coca Cola GmbH“), ausmachen lässt, ist mit der gewählten Identifikation nur ein kleiner Bereich der navigationsorientierten Anfragen abgedeckt. Daher kann nur gemutmaßt werden, dass sich der Anteil aller navigationsorientierten Anfragen stärker unterscheiden würde.

Abschließend kann gefolgert werden, dass der Anteil an Anfragen, die keinen Klick auslösten und URL Bestandteile enthalten, überrascht. Es wurde ein wesentlich geringer Anteil erwartet, da davon ausgegangen wurde, dass dem Suchenden die gesuchte Website bekannt ist und die Suchterme daher präziser zu dem gewünschten Ergebnis führen sollten, als bei anderen Anfragen.

4.1.4.3 Themen

Um weitere Unterschiede zwischen Suchanfragen mit und ohne Klick zu erheben, wird aus beiden Suchanfragen-Gruppen jeweils eine Zufallsstichprobe von 600 Anfragen gebildet, welche nach verschiedenen Kriterien untersucht wird. Die kleinere Stichprobe ermöglicht einen detaillierteren Blick auf die Suchanfragen und ihre Ergebnisse, um hier Besonderheiten im Vergleich herausarbeiten zu können. Außerdem dient diese Stichprobe auch als Basis für die weiteren Untersuchungen in den Kapiteln 3.3 und 3.4.

Ein erstes Kriterium, welches für Suchanfragen, die keinen Klick erhielten, anhand der kleineren Stichprobe untersucht werden soll, ist die thematische Einordnung dieser Suchen. Es gilt herauszufinden, ob sich die Themen der ungeklickten Anfragen von denen der geklickten Anfragen unterscheiden. Heben sich die Themen der Anfragen ohne Klick stark von den Themen der Anfragen mit ausgelöstem Klick ab, könnte, aufbauend auf der ersten Hypothese, die besagt, dass Suchanfragen nicht geklickt werden, da ihre Treffer für das vorliegende Informationsbedürfnis irrelevant sind, gemutmaßt werden, dass die für die ungeklickten Suchanfragen häufig vorkommenden Themen von Web-Suchmaschinen nicht zufriedenstellend abgedeckt werden (können).

Die thematische Einteilung der Suchanfragen wird nach denselben Kategorien vorgenommen, wie sie 2001 während einer Logfile Untersuchung von Spink et al. entwickelt wurden (vgl. SPINK et al. 2001, S. 233). Grund für die Wahl dieser Themen ist, dass diese auch in einer Studie zur Untersuchung von Suchanfragen an deutschen Web-Suchmaschinen wieder aufgegriffen wurden (vgl. LEWANDOWSKI 2006a, S. 34) und daher ein Vergleich erleichtert wird.
Die folgenden elf Themenbereiche wurden von Spink et al. (vgl. SPINK et al. 2001, S. 233) erstellt und kommen nun für die vorliegenden Stichproben der Suchanfragen mit sowie ohne Klick zum Einsatz. Die Übersetzung der Themen ist Lewandowski 2006 entnommen (vgl. LEWANDOWSKI 2006a, S. 34):

- People, places and things (PPT; Personen, Orte und Dinge),
- Computers or Internet (CI; Computer und Internet),
- Commerce, travel, employment, or economy (CTEE; Handel, Reise, Arbeit und Wirtschaft),
- Entertainment or recreation (ER; Unterhaltung und Freizeit),
- Health or sciences (HS; Gesundheit und (Natur-)Wissenschaft),
- Sex or pornography (SP; Sex und Pornographie),
- Government (GOV; Regierung und Verwaltung),
- Education or humanities (EH; Bildung und Geisteswissenschaften),

- Society, culture, ethnicity, or religion (SCER; Gesellschaft, Kultur, Ethnizität und Religion),
- Performing or fine arts (PFA; Kunst),
- Unknown or other (UNK; nicht bestimmbar und sonstige).

Bei einer solchen Untersuchung ist zu beachten, dass eine Einordnung nach Themen immer subjektiv geprägt ist (vgl. SPINK et al. 2001, S. 233; LEWANDOWSKI 2006a, S. 34). Denn wie in Kapitel 3.2 beschrieben, liefert ein Logfile lediglich Suchanfragen, ein direkter Rückschluss auf die Ziele und Intentionen der jeweiligen Suche ist nicht möglich. Um der subjektiven Beeinflussung während der thematischen Einsortierung zu begegnen, werden in den bereits angesprochenen Studien jeweils zwei Juroren pro Suchanfrage eingesetzt. Entscheiden diese sich für unterschiedliche Themen, wird das abschließende Urteil in einer Diskussion geklärt (vgl. LEWANDOWSKI 2006a, S. 34; SPINK und JANSEN 2004, S. 68). Die Themensortierung für die vorliegende Stichprobe wird nach demselben Muster angefertigt, auch hier kommen zwei Juroren zum Einsatz, welche die Suchanfragen im ersten Schritt unabhängig voneinander Themen zuordnen. Bei Uneinigkeit folgt eine Diskussion der jeweiligen Anfrage.

Die thematische Sortierung der Themen zeigt für die Suchanfragen ohne Klick, dass das Thema „Personen, Orte und Dinge“ hier die am stärksten besetzte Themenklasse ist (siehe Abb. 3). Hierauf entfallen 29,67 Prozent der Suchanfragen aus der Stichprobe. An dieser Stelle muss angemerkt werden, dass in diese Klasse hauptsächlich Personennamen zugeordnet wurden. Aus diesem Grund wird vermutet, dass hier keine Klicks erfolgen, da durch den fehlenden Einsatz der Phrasensuche häufig nur der gesuchte Vor- oder Nachname in den Suchergebnissen auftritt, die Suche liefert dementsprechend viele irrelevante Treffer. Weiteren Aufschluss über diese stark ausgeprägte Themengruppe soll die Relevanzauswertung in Kapitel 4.3 ergeben.

Im Vergleich zu den Suchanfragen, deren Ergebnisse mindestens einen Klick erhalten haben, fällt auch die Besetzung der Klasse „Nicht bestimmbar und sonstige“ besonders ins Auge. Dieser Themenbereich ist ebenfalls, mit 24,5 Prozent, stark vertreten. Es bleibt anzumerken, dass hier nur Suchanfragen einsortiert wurden, die nicht bestimmbar waren, da die Wörter und Wortkombinationen den Juroren nicht geläufig sind. Die starke Ausprägung dieser Klasse lässt sich evtl. damit begründen, dass die Suchen der ungeklickten Suchanfragen häufig durch Fehler geprägt sind, die den Sinn des Wortes entstellen und so eine Suche nach dem wirklich gemeinten Begriff nicht mehr möglich ist. Allerdings lässt sich auf diese Weise nur ein Anteil der unbestimmbaren Suchanfragen begründen, eine weitere Erklärung für diese starke Besetzung konnte nicht gefunden werden. Weitere Unterschiede zeigen sich außerdem bei den Themen „Unterhaltung und Freizeit“, „Gesundheit und (Natur-)Wissenschaft“ und „Sex und Pornographie“ sowie „Bildung und Geisteswissenschaften“. In diesen Themen ist jeweils die Gruppe der geklickten Suchanfragen stärker vertreten. Die weiteren Werte für die einzelnen Themen können der Abb. 3 entnommen werden.

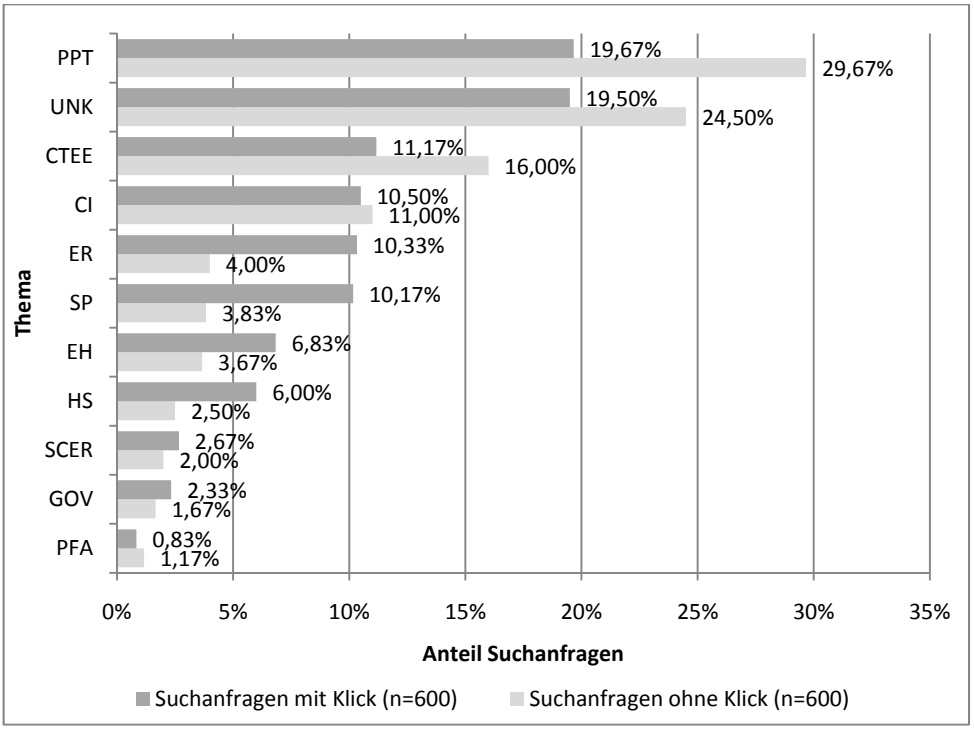

Abb. 3: Themen der Suchanfragen mit und ohne Klick (Quelle: eigene Darst.)

Ein Vergleich der Themenzuordnung für beide Stichproben (n=1.200 Suchanfragen), also sowohl die ungeklickten als auch die geklickten Anfragen, mit der bereits erwähnten Untersuchung von Lewandowski (vgl. LEWANDOWSKI 2006a, S. 39) zeigt, dass sich hier die am stärksten besetzten Themen unterscheiden. In der Untersuchung von Lewandowski nimmt das Thema „Handel, Reise, Arbeit und Wirtschaft“ mit 29 Prozent den ersten Rang ein, für die vorliegende Stichprobe ist es die Klasse „Personen, Orte und Dinge“ mit 24,58 Prozent der Suchen. Als zweitstärkste Kategorie wird in der Untersuchung von Lewandowski das Thema „Nicht bestimmbar und sonstige“ besetzt (20 Prozent), im Falle der Anfragen in der vorliegenden Arbeit teilen sich den zweiten Rang die Klassen „Handel, Reise, Arbeit und Wirtschaft“ (17,83 Prozent) und „Nicht bestimmbar und sonstige“ (17,83 Prozent).
Ein Vergleich mit der Untersuchung von Spink aus dem Jahr 2002 (vgl. SPINK und JANSEN 2004) zeigt mehr Ähnlichkeiten, denn auch in dieser Untersuchung ist das Thema „Personen, Orte und Dinge“ mit 49,27 Prozent der Suchanfragen der Stichprobe auf den ersten Platz, gefolgt von der Klasse „Handel, Reise, Arbeit und Wirtschaft“ (12,52 Prozent). Die Kategorie „Nicht bestimmbar und sonstige“ wird bei Spink nicht ausgewiesen und kann daher nicht verglichen werden.

Insgesamt zeigen sich beim Vergleich der Untersuchungen zur thematischen Klassifikation einige Gemeinsamkeiten aber auch Unterschiede. Eine Begründung für die Unterschiede in der Themenverteilung der verglichenen Untersuchungen lässt sich nicht sicher feststellen. Die Unterschiede könnten sich aus den verschiedenen Zeitpunkten, an denen die Untersuchungen durchgeführt wurden ergeben oder an einem unterschiedlichen Nutzerverhalten liegen. Diese Gründe werden auch in der Untersuchung von Lewandowski angenommen, da auch hier ein Vergleich der Themenzuordnungen Unterschiede aufzeigt (vgl. LEWANDOWSKI 2006a, S. 39).
Für die Suchanfragen ohne Klick lässt sich in der Gegenüberstellung zu den Anfragen mit Klick jedoch festhalten, dass hier die Klassen „Personen, Orte und

Dinge“ und „Nicht bestimmbar und sonstige“ durch ihre starke Besetzung besonders auffallen. Die zu Beginn des Kapitels angebrachte Folgerung, dass die für die ungeklickten Suchanfragen häufig vorkommenden Themen von Web-Suchmaschinen nicht zufriedenstellend abgedeckt werden, kann an dieser Stelle nicht unterstrichen werden, da ein großer Anteil der Suchanfragen als nicht bestimmbar klassifiziert wurde. Dennoch kann vermutet werden, dass die Suche nach Personen nicht zufriedenstellend von Web-Suchmaschinen abgedeckt wird.

4.1.4.4 Weitere Eigenschaften von Suchanfragen ohne Klick

Zur weiteren Untersuchung der Suchanfragen ohne Klick bieten sich die folgenden Merkmale an:

a) Kein Ergebnis: Gibt es organische Suchergebnisse zu der Suchanfrage? Und wenn nein, werden Anzeigen angezeigt?
b) Anstößige Suchanfrage: Handelt es sich um eine anstößige Suchanfrage? Darunter fallen in dieser Studie pornographische, gewalt- und drogenverherrlichende Suchanfragen.
c) Fremdsprachige Suche/ Suchergebnisse: Handelt es sich um eine fremdsprachige Suchanfrage oder eine Suchanfrage, die überwiegend zu fremdsprachigen Ergebnissen führt? Englische Suchanfragen und Suchergebnisse wurden separat erfasst.
d) Suchanfrage mit Korrekturvorschlag: Hat die Suchanfrage einen Korrekturvorschlag („Meinten Sie: …“) zur Folge?
e) Suchergebnisseite mit Anzeige(n): Werden zu der Suchanfrage Anzeigen eingeblendet?
f) Suchergebnisseite mit Ergebnis(sen) von T-Online: Wird auf der Suchergebnisseite auf Ergebnisse des T-Online Portals hingewiesen?
g) Suchergebnisseite mit Maps-Ergebnis: Ist die Suchergebnisseite mit Ergebnissen der Map-Suche (Anzeige einer Karte) ausgestattet?
h) Suchanfrage ergibt keinen Sinn: Ergibt die Suchanfrage einen Sinn? Als sinnlos wurden z.B. Suchanfragen definiert, die nur aus Sonderzeichen

oder Wiederholungen eines einzigen Buchstabens bestanden. Des Weiteren fallen hierunter Suchanfragen mit groben Rechtschreibfehlern, die den Sinn der Suche nicht mehr sicher erschließbar machen.

i) Automatisierte Suchanfrage: Handelt es sich um eine Suchanfrage, die ohnehin nicht mit der Absicht gestellt wurde, ein Informationsbedürfnis durch einen Klick zu befriedigen, da es sich um eine Abfrage zur Suchmaschinenoptimierung handelt?

Diese Kriterien wurden einerseits aufgestellt, um Merkmale und Eigenschaften ungeklickter Suchanfragen im Vergleich zu erheben, andererseits bereits zur Bereinigung der Stichprobe als Vorbereitung auf die folgenden Untersuchungen (siehe Kapitel 3.3 und 3.4).

Die Tab. 2 gibt Aufschluss über die Anteile der oben aufgestellten Merkmale der nicht geklickten und der geklickten Suchanfragen.

	Suchanfragen ohne Klick n= 600		**Suchanfragen mit Klick** n= 600	
Merkmal	**Anzahl Suchen**	**in %**	**Anzahl Suchen**	**in %**
Kein Ergebnis davon mit Werbeanzeige(n)	39 11	6,5 % 1,83 %	1 1	0,17 %
Anstößige Suchanfrage	29	4,83 %	52	8,67 %
Fremdsprachige Suche/ Suchergebnisse	43	7,17 %	17	2,83 %
Englischsprachige Suche/ Suchergebnisse	74	12,33 %	36	6 %
Suchanfrage mit Korrekturvorschlag	114	19 %	79	13,17 %
Suchergebnisseite mit Anzeige(n)	417	69,5 %	467	77,83 %
Suchergebnisseite mit Ergebnis(sen) von T-Online	175	29,17 %	131	21,83 %
Suchergebnisseite mit Maps-Ergebnis	7	1,17 %	21	3,5 %
Suche ergibt keinen Sinn	4	0,66 %	0	0 %
Automatisierte Suchen	63	10,5 %	0	0 %

Tab. 2: Merkmale geklickte und ungeklickter Suchanfragen im Vergleich (Quelle: eigene Darst.)

Im Vergleich der beiden Stichproben zeigen sich einige Unterschiede zwischen geklickten und nicht geklickten Suchanfragen. Diese Aspekte werden im Folgenden herausgegriffen und erläutert. In den Fällen, in denen sie bereits dazu beitragen, die in dieser Arbeit aufgestellten Hypothesen zu unterstützen, wird dies ausdrücklich betont.

Die nicht geklickten Suchanfragen zeigen, im Vergleich zu den geklickten Anfragen, häufiger kein organisches Suchergebnis (6,5 Prozent). Dies unterstützt die Hypothese 1 zur Erklärung einer nicht geklickten Anfrage. Diese besagt, dass Suchanfragen keinen Klick auslösen, da sie nur zu irrelevanten Ergebnissen führen. In diesem Fall liegen keine organischen Ergebnisse vor, weswegen von irrelevanten Ergebnissen gesprochen werden kann.

Interessant ist auch der erhöhte Anteil fremdsprachiger Suchanfragen in der Stichprobe der nicht geklickten Suchanfragen. Zusammengenommen sind 19,5 Prozent der Anfragen, die im Anschluss an die Suche keinen Klick erhielten, nicht in deutscher Sprache verfasst oder resultieren in überwiegend fremdsprachigen Ergebnissen. Im Gegensatz dazu weisen die Anfragen, die einen Klick erhielten, hier nur einen Anteil von 8,83 Prozent der Suchanfragen auf.
Der erhöhte Anteil fremdsprachiger Suchergebnisse in der Stichprobe der nicht geklickten Suchanfragen lässt sich an dieser Stelle nicht begründen. Es kann lediglich gemutmaßt werden, dass fremdsprachige Suchergebnisse vom Suchenden als irrelevant eingestuft werden, da sie aufgrund der unbekannten Sprache nicht verstanden werden können. Eine Studie von Hansen und Karlgren zeigt außerdem, dass die Einschätzung der Relevanz gefundener Dokumente in einer Fremdsprache, auch wenn sie auf nahezu muttersprachlichem Niveau beherrscht wird, länger dauert, als die Relevanzeinschätzung eines Dokuments in der Muttersprache (vgl. HANSEN und KARLGREN 2005, S. 631). Eine Einschätzung der Relevanz ist für ein fremdsprachiges Dokument also aufwendiger und wird demensprechend evtl. eher gescheut, weswegen kein Klick ausgelöst wird. Es wird ver-

mutet, dass dieses Verhalten eintritt, wenn die Suchanfrage mit der Absicht gestellt wurde deutsche Suchergebnisse zu erhalten, die Trefferliste jedoch größtenteils aus fremdsprachigen Ergebnissen besteht.

Des Weiteren wird Hypothese 2 unterstützt, die eine Suchanfrage, die keinen Klick auslöste damit begründet, dass die Suche einen Korrekturvorschlag der Suchmaschine auslöste. Falls der Suchende den Korrekturvorschlag anstatt eines Suchergebnisses anwählt, wird für die Suche kein Klick verzeichnet. Diese Hypothese wird bekräftigt, da ein Anteil von 19 Prozent der Anfragen, die nicht geklickt wurden, in einem Korrekturvorschlag resultiert. Bei diesen Fällen kann folglich vermutet werden, dass die Suchenden sich dafür entschieden einen Korrekturvorschlag anstatt eines Ergebnisses anzuklicken. Für Suchanfragen, die einen Klick auslösen, ist der Anteil mit 13,17 Prozent geringer.

4.1.5 Zusammenfassung der Ergebnisse der deskriptiven Analyse

Der Vergleich von Suchanfragen, die einen bzw. keinen Klick auslösten, zeigt einige Besonderheiten der nicht geklickten Anfragen auf, die an dieser Stelle aufgelistet werden:

Suchanfragen, die keinen Klick auslösen,…

- … sind im Durchschnitt länger als Suchanfragen, die einen Klick auslösen, nämlich 2,63 Wörter lang.
- … führen häufiger zu keinem organischen Suchergebnis (6,5 Prozent).
- … weisen einen erhöhten Anteil von fremdsprachigen Suchanfragen bzw. Suchergebnissen auf (19,5 Prozent).
- … resultieren häufiger in einem Korrekturvorschlag der Suchmaschine (19 Prozent).
- … lassen sich in den meisten Fällen den Themen „Personen, Orte und Dinge“ (29,67 Prozent) und „nicht bestimmbar und sonstige“ (24,50 Prozent) zuordnen.

4.2 Ergebnisse der Online-Befragung zur Ableitung des Informationsbedürfnisses

Die in Kapitel 3.3.4 beschriebenen Maßnahmen zur Aktivierung der Probanden führten dazu, dass insgesamt 1.427 Antworten zu den 600 Suchanfragen der Online-Befragung eingeholt werden konnten. Dabei erhielten die meisten Anfragen, nämlich 71 Prozent der Anfragen der Stichprobe, von zwei unterschiedlichen Probanden eine Einschätzung des Informationsbedürfnisses, während drei Ableitungen von unterschiedlichen Probanden pro Anfrage nur bei einem Anteil von 26,3 Prozent eingeholt werden konnten. Vier oder mehr Ableitungen des Informationsbedürfnisses von unterschiedlichen Probanden werden in nur 2,67 Prozent der Fälle festgehalten. Tab. 3 gibt detailliert Aufschluss über die Anzahl der Antworten pro Suchanfrage.

Anzahl der Antworten pro Anfrage	**Anzahl der Anfragen**	**in %**
2	426	71 %
3	158	26,33 %
>= 4	16	2,66 %
Gesamt	600	100 %

Tab. 3: Anzahl der Ableitungen pro Suchanfragen (Quelle: eigene Darst.)

Nach der Befragung konnten 478 Anfragen (79,67 Prozent) mit einem Informationsbedürfnis oder zwei Informationsbedürfnissen versehen werden, wobei die Zuordnung eines zweiten Informationsbedürfnisses nur selten vorkam: 20 Anfragen wurden von den Probanden mit einem zweiten Informationsbedürfnis bedacht, dies stellt einen Anteil von 3,33 Prozent an der Stichprobe dar. Bei den restlichen 122 Anfragen (20,33 Prozent) fand keine Einigung der Probanden statt oder die Begriffe der Suchanfrage waren dem Probanden nicht geläufig. Diese Anfragen wurden hinsichtlich ihrer Themen und Eigenschaften genauer untersucht, um herauszufinden, ob Besonderheiten im Vergleich zu den Anfragen, welche abgeleitet werden konnten, zu entdecken sind. Es zeigten sich die fol-

genden Auffälligkeiten: Im Vergleich zur übrigen Stichprobe der Suchanfragen ohne Klick entfällt hier ein wesentlich größerer Anteil auf die Themenkategorie „nicht bestimmbar und sonstige“ (28,69 Prozent), diese Anfragen sind also nicht bestimmbar. Betrachtet man die Suchen dieser Themeneinordnung detaillierter, zeigt sich, dass sie dieser Kategorie zugeordnet werden, da die Suchbegriffe entweder nicht bekannt sind, Rechtschreibfehler den Sinn der Suche soweit entstellen, dass keine zuverlässige Annahme über das Thema der Suche ausgesprochen werden kann oder die Kombination der (zwar bekannten) Worte keinerlei ersichtlichen Sinn ergibt. Des Weiteren wurde die Eigenschaft untersucht, ob die Anfragen einen Rechtschreibfehler enthalten oder bei der Suche ein Korrekturvorschlag ausgelöst wird. Es wird vermutet, dass dieser Anteil erhöht ist und damit zum Unverständnis der Anfragen beiträgt. Nach einer Prüfung wird deutlich, dass ein Anteil von 20,49 Prozent der Suchen einen Korrekturvorschlag auslösen, bei den Anfragen, deren Ziel abgeleitet werden konnte, liegt der Anteil bei 16,27 Prozent. Die vorher ausgesprochene Vermutung bestätigt sich folglich.
An dieser Stelle kann geschlussfolgert werden, dass als Gründe für eine fehlende Ableitung in vielen Fällen festgehalten werden kann, dass die Suchanfragen nicht bestimmbar sind, da sie unbekannte Worte oder Wortkombinationen enthalten oder falsch geschrieben wurden.

4.2.1 Vorgehen zur Ableitung des Informationsbedürfnisses

Die Befragung zur Ableitung des Informationsbedürfnisses ist so konzipiert, dass pro Anfrage mehrere Meinungen eingeholt werden sollen, damit eine valide Ableitung der Suchintention möglich ist. Ziel ist es drei Antworten pro Anfrage einzuholen. Die aufgeführten Ergebnisse zeigen jedoch, dass in den meisten Fällen nur zwei Probanden pro Suchanfrage aktiviert werden konnten. In diesen Fällen wurde das Informationsbedürfnis nur abgeleitet, wenn beide Probanden bei der Ableitung übereinstimmen. Das gleiche Verhältnis an Antworten, um ein Informationsbedürfnis abzuleiten, nämlich zwei Probanden mit gleichen Antworten, ist auch für die Ableitung bei drei oder mehr Antworten pro Anfragen vorgese-

hen. Auf diese Weise kann die aufgestellte Bedingung, ein Informationsbedürfnis nur festzuhalten, wenn mehrere Probanden dafür gestimmt haben, eingehalten werden.

Ein zweites Informationsbedürfnis wird festgehalten, wenn zwei weitere Antworten der Anfrage übereinstimmen. Dabei wird nicht darauf geachtet, ob eines dieser beiden Informationsbedürfnisse als erstes oder zweites Ziel der Suche vom Probanden deklariert wurde (siehe Beispiel Nr. 3 & 4, Tab. 4).

Falls mehrere, ähnliche (aber nicht gleiche) Ableitungen von den Probanden abgegeben wurden, wird ein allgemeineres Ziel abgeleitet, welches die Antworten aller Probanden umfasst (siehe Beispiel Nr. 1, Tab. 4).

Wenn aber davon ausgegangen werden kann, dass ein spezielleres Ziel, welches von einem Proband genannt wird, in einem allgemeineren Ziel eines anderen Probanden enthalten ist, wird versucht, das speziellere Ziel mit in das Informationsbedürfnis aufzunehmen (siehe Beispiel Nr. 2 & 5, Tab. 4). Auf diese Weise werden die Antworten zu den Informationsbedürfnissen nicht verfälscht, vielmehr werden die abgeleiteten Ziele der Suche für den Relevanztest anschaulicher und damit besser bewertbar.

Falls keine der Antworten mit einer anderen Antwort übereinstimmt, wird kein Informationsbedürfnis festgehalten (siehe Beispiel Nr. 6, Tab. 4). Dies kam in 18,17 Prozent der Fälle vor (109). Geben zwei oder mehr Probanden an, dass sie mit dem Suchbegriff nichts anfangen können, wird auch hier kein Informationsbedürfnis abgeleitet. 2,17 Prozent der Anfragen wurden auf diese Weise bewertet (13 Anfragen).

	Suchanfrage[3]	Antworten[4]				Informationsbedürfnis
1	Seo tutorial	Der Suchende sucht eine Anleitung zur Suchmaschinenoptimierung.	Der Suchende sucht nach einer Anleitung zum SEO.	Auffinden eines SEO Tutorials, um Hintergrundinfos zu bekommen.	Suche nach Informationen zur Suchmaschinenoptimierung eigener Webseiten.	Eine Anleitung zur Suchmaschinenoptimierung finden.
2	[redacted]	Suche zu der Person [redacted]	Es werden biografische Daten zu der Person [redacted] gesucht.	Es wird nach den Kontaktdaten der Person [redacted] gesucht.		Informationen zu der Person "[redacted] ", z.B. Kontaktdaten oder biografische Daten.
3	sekundarschule am lerchenfeld schönebeck	Die Internetseite der "sekundarschule am lerchenfeld schönebeck".	Es werden die Kontaktdaten (Adresse, Telefonnummer) der Schule gesucht.	Allgemeine Informationen über die "sekundarschule am lerchenfeld schönebeck".	Es wird die Homepage der Schule gesucht, um dort zum Beispiel den Stundenplan runterzuladen.	1.Die Internetseite der "Sekundarschule am Lerchenfeld Schönebeck". 2.Informationen über die "Sekundarschule am Lerchenfeld Schönebeck", z.B. Adresse oder Telefonnummer.
4	scanwizard für xp	Der Download eines Programms (Scanwizard) für Windows XP.	Download Programme	Informationen, Bewertungen, Funktionen über das Programm (Scanwizard) für Windows XP.	Infos zu Scanwizard für XP.	1.Der Download eines Programms (Scanwizard) für Windows XP. 2.Informationen über das Programm (Scanwizard) für Windows XP.
5	flachdach bauordnung aufständerung	Jemand sucht Informationen dazu, ob die Bauordnung die Aufständerung eines Flachdachs erlaubt.	Suche nach einer bestimmten rechtlichen Regelung.			Informationen, ob die Bauordnung rechtlich die Aufständerung eines Flachdachs erlaubt.
6	haustürgeschäft freunde bekannte verwandt	Suchende sucht Informationen über Haustürgeschäfte, die er mit Freunden, Bekannten usw. mache kann/ will.	Der Suchende möchte Informationen darüber wie er Freunde, Bekannte oder Verwandte für ein Haustürgeschäft begeistern kann.	Es handelt sich evtl. um einen Namen eines Geschäftsmodells, einer Website o.ä.		keine Einigung

Tab. 4: Beispiele zur Ableitung des Informationsbedürfnisses (Quelle: eigene Darst.)

[3] Die Suchanfragen wurden nach Angabe der Logfile-Datei übernommen, weswegen Rechtschreibfehler auftreten können. Suchanfragen, die Personennamen enthalten, wurden geschwärzt.

[4] Die Antworten wurden nach Angabe der Probanden in der Online-Befragung übernommen, daher können Rechtschreib- oder Grammatikfehler auftreten.

4.2.2 Vorgehen zur Bestimmung der Aktualität einer Suchanfrage

Die Frage nach der Aktualität einer Suchanfragen wurde mit 747 Antworten bedacht, die sich auf die Anzahl von 478 Anfragen, deren Informationsbedürfnis abgeleitet werden konnte, verteilen. Die Hälfte der Anfragen (239 Anfragen) erhielt unterschiedliche Aktualitätseinschätzungen. Davon wiederum konnten 80 Anfragen (13,33 Prozent) keiner bestimmten Aktualitätsstufe zugeordnet werden, da sie sowohl die Einschätzung erhielten aktuell zu sein (Antworten: „ist aktuell“ und/oder „ist sehr aktuell“) als auch die Einschätzung nicht aktuell zu sein (Antworten „hat keinen aktuellen Bezug“ und/oder „ist weniger aktuell“). Aus diesen unterschiedlichen Einordnungen der Aktualität der Suchanfragen kann gefolgert werden, dass entweder bei den Probanden, die die gleiche Suchanfrage bewerteten, unterschiedliche Meinungen über den jeweiligen Aktualitätsbezug einer Suchanfrage bestehen oder aber ein unterschiedlicher Kenntnisstand über aktuelle Ereignisse herrschte.

Die restlichen Anfragen, deren Aktualitätseinschätzung der Probanden sich unterschied, konnten bzgl. ihrer Tendenz als „aktuell“ oder „nicht aktuell“ einsortiert werden. Dabei wurde eine Suche als aktuell eingestuft, wenn die Antworten „ist aktuell“ und „ist sehr aktuell“ ausgewählt wurden. Als nicht aktuell gilt eine Anfrage, wenn sich die Probanden für die Kategorien „ist weniger aktuell“ und „hat keinen aktuellen Bezug“ entschieden. Bei der Zuordnung beider Aktualitätsstufen wurde das Ankreuzen der Antwort „kann ich nicht einschätzen“ nicht berücksichtigt.

Nach einer Analyse der erhobenen Daten nach den beschriebenen Kriterien konnten 135 der Anfragen mit einer unterschiedlichen Bewertung der Aktualität als nicht aktuell eingeordnet werden. 24 Anfragen wurden als aktuell bezeichnet. Die restlichen 239 Anfragen konnten eindeutig einer Kategorie zugeordnet werden, da sich alle Probanden für eine bestimmte Antwort entschieden. Hier wurden zwölf Anfragen als nicht einschätzbar einsortiert, 154 Anfragen wurden als „hat keinen aktuellen Bezug“ kategorisiert und 38 Anfragen wurden als weniger

aktuell eingeordnet. 33 Anfragen bekamen die Zuordnung „ist aktuell“ zugeschrieben und zwei Anfragen wurden als sehr aktuell bezeichnet.

Die Tab. 5 gibt einen Überblick über die Aktualitätseinordnungen der 478 Anfragen.

Aktualitätsstufe	**Mehr als eine Einordnung** n=239 Anfragen	**Eindeutige Einordnung** n=239 Anfragen	**Gesamt**
Hat keinen aktuellen Bezug.	135 (28,24 %)	154 (32,22 %)	327 (68,41 %)
Ist weniger aktuell.		38 (7,95 %)	
Ist aktuell.	24 (5,02 %)	33 (6,90 %)	59 (12,34 %)
Ist sehr aktuell.		2 (0,42 %)	
Kann ich nicht einschätzen.	-	12 (2,51 %)	92 (19,25 %)
Keine Einigung	80 (16,74 %)	-	
Gesamt	239 (50 %)	239 (50 %)	478 (100 %)

Tab. 5: Aktualitätseinordnung der Suchanfragen (Quelle: eigene Darst.)

Die Tabelle macht ersichtlich, dass über die Hälfte der Anfragen (68,41 Prozent) als nicht aktuell bezeichnet werden, lediglich 12,34 Prozent der Anfragen wird ein aktueller Bezug bescheinigt. Die restlichen Anfragen (19,25 Prozent) konnten keiner Aktualitätsstufe zugeordnet werden, da die Probanden die Aktualität der Anfrage entweder nicht einschätzen konnten oder weil die Meinungen der Probanden zur Aktualität der Suche zu weit auseinander lagen.

4.3 Ergebnisse der Online-Befragung zur Untersuchung der Relevanz der Suchergebnisse nicht geklickter Suchanfragen

Die Online-Befragung zur Ableitung des Informationsbedürfnisses stattete 498 Anfragen mit einer Suchintention aus. Für den Relevanztest standen anschließend 482 Anfragen zur Relevanzmessung bereit, da Anfragen, die nicht zehn Ergebnisse erbrachten oder ein Semikolon enthielten abgezogen werden mussten, weil sie vom Relevanztool nicht verarbeitet werden konnten.

Für 426 der 482 Anfragen konnte eine Einschätzung der Relevanz und des Klickverhaltens für alle zehn Positionen ihrer ersten Suchergebnisseite eingeholt werden. Die Responserate liegt also bei 88,38 Prozent.

D. h., dass 56 Anfragen (11,61 Prozent der Anfragen) keine Relevanzeinschätzung bekamen und in Folge dessen in den weiteren Auswertungen nicht berücksichtigt werden können.

Ursprünglich war vorgesehen, dass jede im Relevanztool vorhandene Suchanfrage, die noch keine Antwort in Form einer Relevanzeinschätzung erhalten hat, bei der Abfrage priorisiert wird. Aufgrund eines Fehlers im Relevanztool wurden den Probanden jedoch auch Anfragen vorgelegt, welche bereits eine Antwort erhalten hatten. Auf diese Weise erhielten 119 Anfragen (27,93 Prozent der Anfragen) von zwei Probanden eine Einschätzung der Relevanz, die übrigen 307 Anfragen wurden von je einem Probanden bewertet.

Tab. 6 listet die beschriebenen Zahlen auf und gibt damit einen ersten Überblick über die Ergebnisse des Relevanztests.

Kriterium	**Anzahl**	**Anzahl in %**
Anzahl der Anfragen im Relevanztool	482	100 %
Anzahl der Anfragen mit Relevanzeinschätzung	426	88,59 %
Von einem Probanden beantwortet	307	72,07 %
Von zwei Probanden beantwortet	119	27,93 %

Tab. 6: Erste Übersicht zu den Ergebnissen der Relevanzmessung (Quelle: eigene Darst.)

Einen detaillierteren Überblick über die Resultate des Relevanztests gibt die Abb. 4. Sie zeigt, wie häufig die Relevanzwerte von 0 bis 4 vergeben wurden.

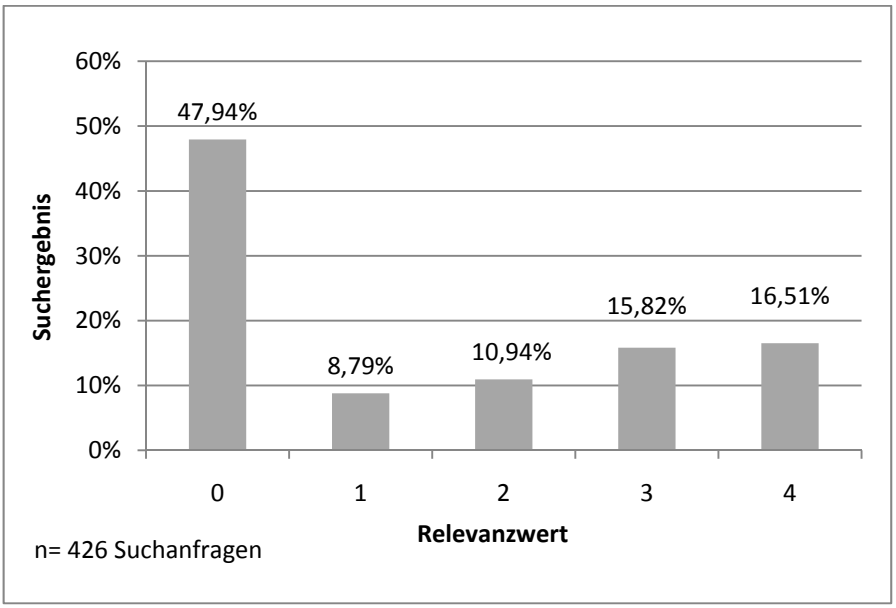

Abb. 4: Überblick über die im Relevanztest vergebenen Relevanzwerte (Quelle: eigene Darst.)

Dabei ist ersichtlich, dass ein hoher Anteil der Suchergebnisse als irrelevant, nämlich mit einem Relevanzwert von Null, bewertet wurde. Dieser Anteil beläuft sich auf 47,94 Prozent aller bewerteten Suchergebnisse. Relevanz wurde hingegen 52,06 Prozent der beurteilten Suchergebnisse zugesprochen. Dabei muss dieses Ergebnis detaillierter betrachtet werden, denn die Relevanz verteilt sich hier auf vier Werte, wobei die Relevanzzuschreibung stärker ist, je höher der gewählte Wert ist. Es kann angemerkt werden, dass die Vergabe eines Relevanzwertes von eins in 8,79 Prozent und die Vergabe eines Wertes von zwei bei 10,94 Prozent der Suchergebnisse ausgesprochen wurde. Eine Relevanzzuschreibung mit dem Wert 3 wurde in 15,82 Prozent der Fälle vorgenommen und eine Zuschreibung des Relevanzwertes von 4 traf auf 16,51 Prozent der Suchergebnisse zu.

Um die Relevanzwerte detaillierter untersuchen zu können, wird in einem nächsten Schritt die obige Auswertung mit den Angaben über das potentielle Klickverhalten des Probanden kombiniert (siehe Tab. 7).

	Ja, ich würde den Treffer anklicken.		**Nein, ich würde nicht klicken, da der Treffer nicht relevant ist.**		**Nein, ich würde nicht klicken. Die Ansicht des Treffers reicht, denn er erfüllt das Ziel der Suche.**	
	Anzahl der Treffer	in %	Anzahl der Treffer	in %	Anzahl der Treffer	in %
Relevanzwert 0						
2.613	2	0,08%	2.595	99,31%	16	0,61%
Relevanzwert 1						
479	62	12,94%	397	82,88%	20	4,18%
Relevanzwert 2						
596	376	63,09%	181	30,37%	39	6,54%
Relevanzwert 3						
862	799	92,69%	19	2,20%	44	5,10%
Relevanzwert 4						
900	795	88,33%	8	0,89%	97	10,78%

Tab. 7: Auswertung des Relevanztest: Kombination von Relevanzwert und Klickverhalten (Quelle: eigene Darst.)

Durch die Kombination dieser Variablen zeigt sich, dass der Relevanzwert 0 in fast allen Fällen bedeutet, dass der Proband den Treffer irrelevant findet und ihn aus diesem Grund nicht anklicken würde (99,31 Prozent). Die restlichen Prozentpunkte, die bei einem Relevanzwert von Null auf die Antworten „Ja, ich würde den Treffer anklicken.“ Und „Nein, ich würde nicht klicken. Die Ansicht des Treffers reicht, denn er erfüllt das Ziel der Suche.“ entfallen, können als nichtig gewertet werden. Es kann davon ausgegangen werden, dass der Proband hier aus Versehen oder mit Absicht eine falsche Antwort wählte, denn für die beschriebenen beiden Antworten wäre es notwendig den Treffer Relevanz zuzuschreiben.

Der Relevanzwert 1 muss detaillierter betrachtet werden, denn hier wurde dem Suchergebnis bereits eine gewisse Relevanz zugeschrieben. Dieser Wert ist allerdings in den meisten Fällen so gering, dass der Treffer nicht angeklickt würde, da

er als nicht relevant (genug) angesehen wird (82,88 Prozent). Ein Anteil von 12,94 Prozent der Suchergebnisse würde allerdings trotz des geringen Relevanzwertes vom Probanden angeklickt werden, 4,18 Prozent der Treffer, die mit 1 als Relevanzwert bedacht wurden, erhielten die Antwort, dass die Ansicht des Treffers bereits ausreicht, um das Informationsbedürfnis zu befriedigen.

Ab einem Relevanzwert von 2 zeigt sich, dass der Anteil der Suchergebnisse, die angeklickt würden, ansteigt und bereits mehr als die Hälfte der Suchergebnisse in dieser Kategorie ausmacht (63,09 Prozent). Das Suchergebnis aufgrund fehlender Relevanz nicht anklicken würden hier aber immer noch 30,37 Prozent der Probanden. Der Anteil der Anfragen, die bereits auf der Suchergebnisseite beantwortet werden können, steigt hier auf 6,54 Prozent an.

Bei einem Relevanzwert von 3 würden fast alle Treffer angeklickt, denn 92,69 Prozent der Suchergebnisse werden mit der Antwort „Ja, ich würde den Treffer anklicken“ bedacht. 5,1 Prozent der Anfragen werden hier als bereits auf der Suchergebnisseite beantwortbar beschrieben. Der restliche Anteil von 2,2 Prozent der Anfragen wird als irrelevant bezeichnet und daher nicht angeklickt.

Die Betitelung eines Suchergebnisses mit dem Relevanzwert 4 zeigt, dass der Anteil der Treffer, die angeklickt würden (88,33 Prozent) zu Gunsten der Treffer, die bereits auf der Suchergebnisseite beantwortet werden können und daher nicht angeklickt werden (10,78 Prozent), sinkt. Der Anteil der nicht angeklickten Treffer fällt auf 0,89 Prozent. Hier kann davon ausgegangen werden, dass der Proband aus Versehen oder mit Absicht eine falsche Antwort auswählte, denn für die Antwort einen Treffer aufgrund der fehlenden Relevanz nicht anzuklicken, wäre es notwendig dem Treffer keine oder nur einen geringen Relevanzwert zuzuschreiben.

Insgesamt zeigt diese Auswertung, dass je höher der Relevanzwert ist, der einem Suchergebnis zugeordnet wird, die Wahrscheinlichkeit das dieser Treffer angeklickt wird steigt. Bereits ab einem Relevanzwert von 2 werden über die Hälfte der so einsortierten Suchergebnisse angeklickt. Dieses Ergebnis war zu erwarten, denn wie bereits in Kapitel 3.4.2 beschrieben, zeigen Studien, dass Probanden bei der Beurteilung von Relevanz anhand einer Skala von 0 bis 100 einem Treffer bereits ab einem Wert von 41 bzw. 46 (absolute) Relevanz zuschreiben (vgl. Stock 2007, S. 74, zit. nach EISENBERG und HU 1987 und JANES 1991). Der niedrige Schwellenwert zur Beurteilung der Relevanz kann auch in der vorliegenden Untersuchung beobachtet werden.
Außerdem zeigt sich, dass der Anteil der Anfragen, die bereits auf der Trefferseite beantwortet werden können, mit steigendem Relevanzwert ebenfalls ansteigen. Diese Entwicklung war ebenfalls zu erwarten, denn wenn das Ziel einer Suche bereits auf der Trefferseite zu finden ist, kann davon ausgegangen werden, dass dieses Ergebnis vom Suchenden als besonders relevant bewertet wird: die gesuchte Antwort ist gefunden und zusätzlich muss keine weitere Aktion erfolgen, um die Antwort zu erhalten.
Eine Ausnahme bildet allerdings der Relevanzwert 3. Hier steigt der Anteil der bereits auf der Trefferseite beantwortbaren Anfragen nicht, sondern fällt leicht ab. Eine Erklärung für diese Entwicklung konnte nicht gefunden werden.

Um auf die in der Einleitung dieser Arbeit aufgestellten Hypothesen zur Begründung, warum eine Suchanfrage keinerlei Klick erhalten hat, eingehen zu können, sollen an dieser Stelle die einzelnen Suchanfrage-Gruppen, die mit Hilfe des Relevanztests erhoben und abgegrenzt werden konnten, untersucht werden. In Anlehnung an die Hypothesen konnten die folgenden Gruppen von Suchanfragen bzw. ihren Ergebnissen ausgemacht werden:

- Irrelevante Ergebnisse: Suchanfragen, deren ersten zehn organischen Ergebnisse für die Suchanfrage und das dahinter stehende Informationsbedürfnis als irrelevant bewertet wurden.

- Relevante Ergebnisse: Eines oder mehrere Suchergebnisse der ursprünglich nicht geklickten Suchanfrage wird durch den Probanden als relevant bewertet und würde angeklickt werden.
- Suchergebnisseite beantwortet Informationsbedürfnis: Die Suchergebnisseite mit ihren Trefferbeschreibungen wird von den Probanden als ausreichend zur Beantwortung des Informationsbedürfnisses bewertet.

Im nachfolgenden Text werden die einzelnen Gruppen näher untersucht, um Anteile, Besonderheiten und weitere Merkmale dieser Gruppen herausarbeiten zu können und aufbauend darauf, in dem sich anschließenden Kapitel, Empfehlungen und Schlussfolgerungen für die einzelnen Suchanfrage-Gruppen erarbeiten zu können.

4.3.1 Irrelevante Ergebnisse

Eine Begründung für das Unterlassen eines Klicks auf die Trefferliste lautet, dass nur Treffer bzw. Trefferbeschreibungen präsentiert wurden, die keine Relevanz für das Informationsbedürfnis des Suchenden aufweisen. Eine solche Begründung würde also auf eine Suchanfrage zutreffen, wenn alle zehn organischen Ergebnisse durch den Probanden mit der Antwort „Nein, ich würde nicht klicken, da der Treffer nicht relevant ist.“ bedacht würden. Dieses Kriterium trifft auf 88 Suchanfragen zu, dies stellt einen Anteil von 20,65 Prozent der 426 Suchanfragen dar. Falls eine Anfrage von zwei Probanden bewertet wurde und nur ein Proband alle Suchergebnisse als irrelevant beurteilte, wurde diese Anfrage ebenfalls mit aufgenommen. Wenn beide Probanden alle Treffer einer Anfrage als irrelevant bewerten, wird diese Anfrage nur einmal gezählt.

Um diese Gruppe von Suchanfragen näher beschreiben zu können, werden in einem nächsten Schritt ihre Merkmale und Eigenschaften untersucht. Allerdings muss an dieser Stelle angemerkt werden, dass die Anzahl der 88 Suchen keine ausreichend große Fallzahl darstellt, weswegen nur Hinweise und Tendenzen,

nicht jedoch repräsentative Aussagen über diese Suchanfrage-Gruppe gesammelt werden können.

Als zu untersuchende Merkmale einer Suchanfrage, welche nur zu irrelevanten Ergebnissen führt, wurden im Kapitel 1.1 das Thema, der Suchanfragetyp, Fehler in der Suchanfrage und die Mehrdeutigkeit der Suchanfrage genannt. Der folgende Text beschreibt auf welche Weise diese Merkmale bei den Suchanfragen ohne einen relevanten Treffer auf der ersten Suchergebnisseite ausgeprägt sind.

Die Untersuchung des Merkmals Thema zeigt, dass hier insbesondere die Kategorie „Personen, Orte und Dinge“ (PPT) auffällt, welche die Hälfte aller Suchanfragen ausmacht (siehe auch Abb. 5).

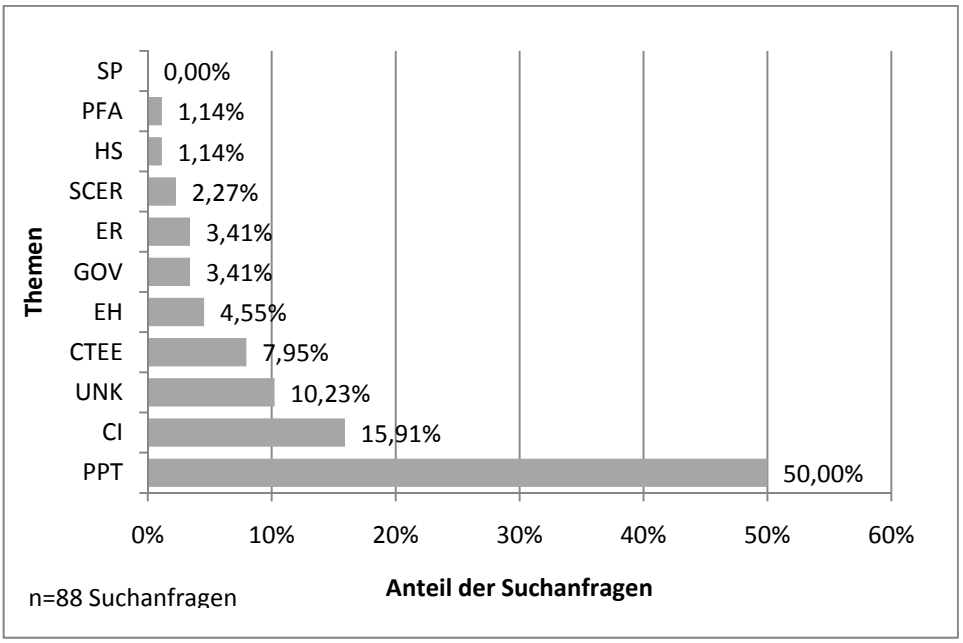

Abb. 5: Themen der Suchanfragen, deren ersten zehn Suchergebnisse im Relevanztest als irrelevant gewertet wurden (Quelle: eigene Darst.)

Eine nähere Betrachtung der Suchen, die dieses Thema zugeordnet bekamen, zeigt, dass der größte Teil dieser Suchen als Personensuchen bezeichnet werden müssen. 38 der 44 Anfragen enthalten ausschließlich den Namen einer Person, in einigen Suchen wird dies auch mit einem Ort oder einer Institution kombiniert.

Diese starke Ausprägung der Suche nach Personen führte dazu, dass auch die restlichen Suchanfragen unabhängig ihres Themas hinsichtlich einer Personensuche untersucht wurden. Es zeigt sich, dass 42 der 88 Suchanfragen, das sind 47,73 Prozent, einen Personennamen enthalten. Aus diesem hohen Anteil an Personensuchen in dieser Suchanfragen-Gruppe kann gefolgert werden, dass diese Suchen nicht zufriedenstellend beantwortet werden konnten. Der Grund, welcher für diese Aussage bereits in Kapitel 4.1.4.3 angeführt wurde, und zwar, dass die Suche unbefriedigend verläuft, da durch den fehlenden Einsatz der Phrasensuche häufig nur der gesuchte Vor- oder Nachname in den Suchergebnissen auftritt und die Suche dementsprechend viele irrelevante Treffer liefert, kann bestätigt werden. Als Beispiel wird die Suche nach der Person „[redacted]“ angeführt. Der zweite Treffer dieser Suche kann in Abb. 6 abgelesen werden.

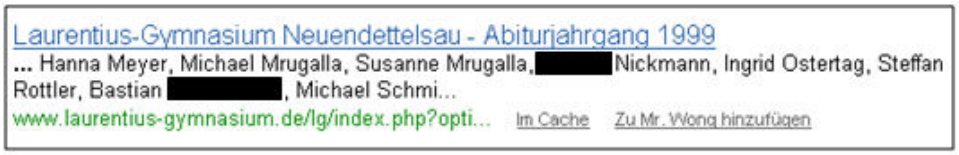

Abb. 6: Treffer der Suche „[redacted]“ (Quelle: SUCHE T-ONLINE.DE 2010)

Es wird ersichtlich, dass hier zwar beide Suchbegriffe vorkommen, jedoch nicht in Kombination miteinander, was das Suchergebnis zu einem Irrelevanten werden lässt.

In einem nächsten Schritt werden die 88 Suchen hinsichtlich ihres Suchanfragetyps von der Forscherin einsortiert. Die Auswertung zeigt sich in Abb. 7.

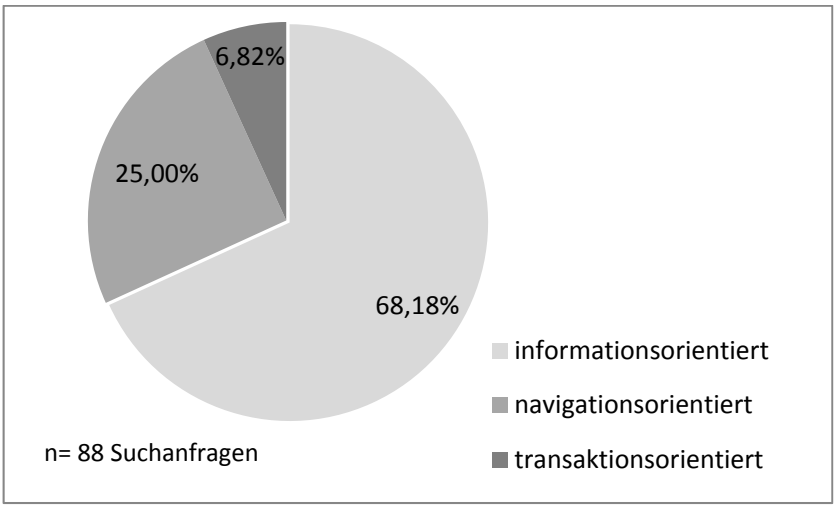

Abb. 7: Suchanfragetypen der Suchanfragen, deren ersten zehn Suchergebnisse im Relevanztest als irrelevant gewertet wurden (Quelle: eigene Darst.)

Ein Vergleich der erhobenen Daten mit früheren Studien zeigt, dass der Anfragetyp „transaktionsorientiert“ unterrepräsentiert ist. In vergangenen Studien stellt er jeweils einen Anteil zwischen 15 und 36 Prozent (vgl. BRODER 2002, S. 6; LEWANDOWSKI 2006a, S. 5; ROSE und LEVINSON 2004, S. 18), in der vorliegenden Studie nur einen Anteil von 6,82 Prozent. Jedoch wird nochmals in Erinnerung gerufen, dass die untersuchte Fallzahl zu klein ist, um repräsentative Äußerungen tätigen zu können, zusätzlich zeigt die weite Spanne für den beschriebenen Anfragetyp, dass keine exakten Ergebnisse vorliegen. Abschließend kann an dieser Stelle festgehalten werden, dass eine Untersuchung nach Anfragetyp zwar Besonderheiten aufzeigt, eine detailliertere Exploration wird aufgrund der geringen Fallzahl jedoch verworfen.

Ein weiteres Untersuchungskriterium der Suchanfragen ohne relevante Treffer ist die Analyse der Anfragen nach Fehlern. Denn es wird gemutmaßt, dass ein Fehler in der Suchanfrage zu unpassenden und damit irrelevanten Suchergebnissen führen könnte. Die Analyse der 88 Anfragen zeigt, dass für einen Anteil von

23,86 Prozent der Suchanfragen, die nur irrelevante Treffer hervorbrachten, ein Korrekturvorschlag von der Suchmaschine ausgelöst wurde. Dieser Anteil ist, im Vergleich zur untersuchten Stichprobe der 600 ungeklickten Treffer (siehe Kapitel 4.1.4.4), erhöht. Dies könnte ein Hinweis darauf sein, dass die ursprünglich Suchenden wirklich den Korrekturvorschlag nutzten, da aufgrund des Fehlers die Suchergebnisse irrelevant waren.

Das Merkmal der Mehrdeutigkeit einer Suchanfrage bzw. des dahinter liegenden Informationsbedürfnisses, welches ebenfalls einen Grund für fehlende relevante Treffer darstellen könnte, trifft auf die vorliegende Suchanfragen-Gruppe nicht zu. Denn dem Problem der Mehrdeutigkeit einer Anfrage wurde bereits mit Hilfe der Umfrage zur Ableitung des Informationsbedürfnisses begegnet. Auf diese Weise wurden potentiell mehrdeutige Suchanfragen bereits bei der Abfrage der Relevanz berücksichtigt.

4.3.2 Relevante Ergebnisse

Als eine weitere Gruppe von Suchanfragen werden die Anfragen untersucht, die laut Probanden zu relevanten Ergebnissen führten. Eine Suchanfrage, die in diese Gruppe fällt, hat für mindestens eines der zehn Ergebnisse einen Relevanzwert von mindestens 1 zugewiesen bekommen. Außerdem sollte dieses Ergebnis vom Probanden mit „Ja, ich würde den Treffer anklicken“ beurteilt werden. Diese Gruppe von Anfragen steht folglich im Gegensatz zu den ursprünglichen Anfragen und ihren (nicht vorhandenen) Klickdaten.

Es konnten 325 Anfragen ausgemacht werden, die mit einem relevanten und angeklickten Treffer durch den Probanden bewertet wurden. Dies stellt mit 76,11 Prozent die Mehrheit der Anfragen im Relevanztest dar. Zu diesen 325 Suchanfragen lagen je zehn Suchergebnisse vor, dementsprechend wurden hier 3.250 Suchergebnisse durch die Probanden hinsichtlich ihrer Relevanz bewertet. Von diesen erhielten 1.567 Suchergebnisse die Beurteilung, dass sie angeklickt würden, d.h. fast die Hälfte der Suchergebnisse wurde von den Probanden mit

„Ja, ich würde den Treffer anklicken“ bewertet (48,22 Prozent). Dies bedeutet, dass im Durchschnitt fast fünf Ergebnisse pro Suchanfrage angeklickt wurden.
Die Abb. 8 zeigt detailliert die einzelnen Relevanzwerte, die diesen Suchergebnissen zugewiesen wurden. Fast die Hälfte der Ergebnisse bekommt einen relativ hohen Relevanzwert von 3 oder 4 zugewiesen. Der Relevanzwert 1 wird mit einem Anteil von 9,75 Prozent am seltensten vergeben.

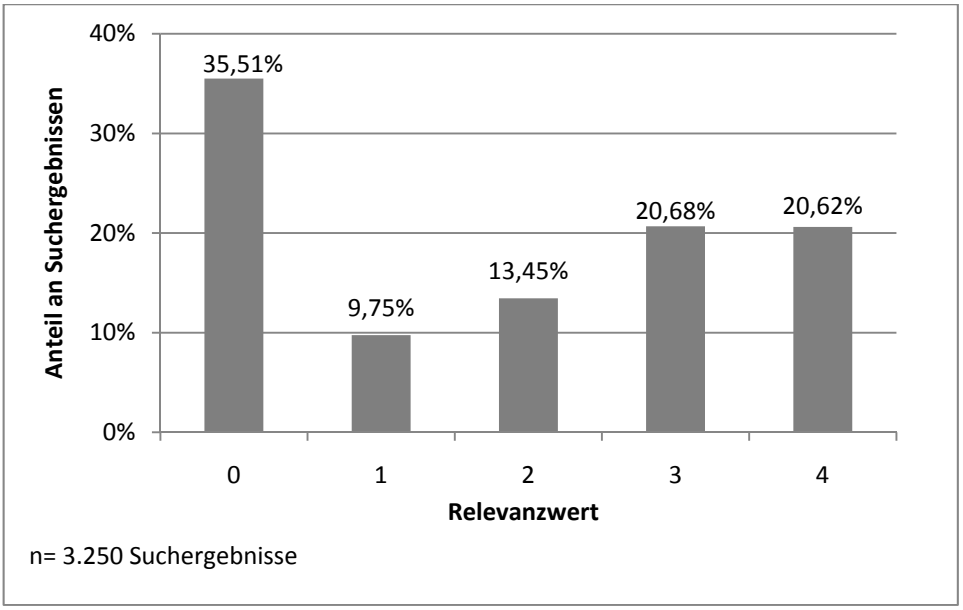

Abb. 8: Anteil der Relevanzwerte der Treffer, deren Suchanfrage mindestens ein relevantes Ergebnis hat (Quelle: eigene Darst.)

Diese Werte überraschen angesichts der Tatsache, dass es sich bei den Suchen um ursprünglich nicht geklickte Anfragen handelt, die nun von den Probanden angeklickt würden. Es werden die folgenden Gründe für dieses Resultat angenommen und in den sich anschließenden Textabschnitten ausgeführt:

a) Der Aktualitätsbezug der betreffenden Suchanfrage führt dazu, dass sich die Suchergebnisse zwischen der Suche des ursprünglich Suchenden und dem Probanden änderten.
b) Das relevante Suchergebnis / Die relevanten Suchergebnisse wurde(n) vom originär Suchenden übersehen, nicht jedoch vom Probanden.

c) Die sich stark unterscheidende Relevanzbewertung der Treffer liegt in der Person des Suchenden / ihren Vorkenntnissen begründet.
d) Die Testsituation beeinflusst die Wahl des Relevanzwertes und das potentielle Klickverhalten der Probanden.

a) Aktualitätsbezug der Suchanfragen

Eine Eigenschaft, welche die große Zahl der Suchen mit als relevant bewerteten Ergebnissen erklären könnte, ist der Aktualitätsbezug der Suchanfragen. Denn handelt es sich um eine aktuelle Anfrage, ist die Wahrscheinlichkeit, dass sich die Suchergebnisse zu dieser Anfrage in dem Zeitraum zwischen dem Abschicken der Anfrage durch den ursprünglich Suchenden und dem Einholen der Ergebnisse für den Relevanztest (drei Monate) geändert haben, hoch (siehe auch Kapitel 5.1). In diesem Fall könnten z.B. zu den vorher irrelevanten Ergebnissen, neue relevante Ergebnisse hinzugekommen sein. Ebenso ist es möglich, dass sich das Ranking der Suchergebnisse änderte und dadurch auf den ersten zehn organischen Treffern nun relevante Treffer sichtbar sind. Eine Prüfung ergab unter den 325 Anfragen einen Anteil von 14,15 Prozent als aktuell eingeschätzte Suchanfragen. 19,69 Prozent der 325 Anfragen konnten keinem Aktualitätsgrad zugeordnet werden. Demnach ist ein Anteil von 66,15 Anfragen als nicht aktuell eingeschätzt worden. Bei 14,15 Prozent der Anfragen (46 Anfragen) könnte also der Aktualitätsbezug der Suche zu einer Änderung der Suchergebnisse geführt haben und so die unterschiedliche Relevanzbewertung durch Suchenden und Proband erklären. Ein Vergleich mit allen im Relevanztest beantworteten Anfragen zeigt, dass sich die Prozentsätze gleichen. Insgesamt bekommen 12,91 Prozent aller Anfragen des Relevanztests Aktualität zugeschrieben, 66,9 Prozent werden als nicht aktuell beschrieben und 20,19 Prozent aller Anfragen des Relevanztest können bezüglich ihrer Aktualität nicht eingeschätzt werden. Hinsichtlich der Aktualität der Anfragen zeigen sich in der in diesem Abschnitt untersuchten Suchanfrage-Gruppe also keine Besonderheiten.

b) Übersehen des relevanten Suchergebnisses

Eine weitere Möglichkeit, den Unterschied zwischen dem Urteil des originär Suchenden und dem Probanden zu erklären ist, dass das relevante Ergebnis vom Suchenden übersehen wurde, da es nicht auf den ersten Positionen der SERP zu finden war. Ein solches Verhalten wurde bereits in verschiedenen Studien nachgewiesen und wird ausführlicher in Kapitel 2.3 diskutiert. Würde ein solches Verhalten als Grund für die nun angeklickten Ergebnisse auftreten, sollte man dies daran erkennen, dass die relevanten Ergebnisse auf den unteren Rängen der SERP anzutreffen sind.

Die Abb. 9 zeigt allerdings, dass für die meisten Suchanfragen die relevanten Ergebnisse auf den ersten Rängen der Suchergebnisseite zu finden sind. Auf der ersten Position finden sich bereits 79,08 Prozent aller relevanten Ergebnisse. Betrachtet man die ersten und zweiten Treffer, finden sich hier bereits knapp 90 Prozent aller relevanten Ergebnisse. Auf den nachfolgenden Rängen kommen jeweils nur wenige relevante Ergebnisse hinzu, wie die Abbildung deutlich macht.

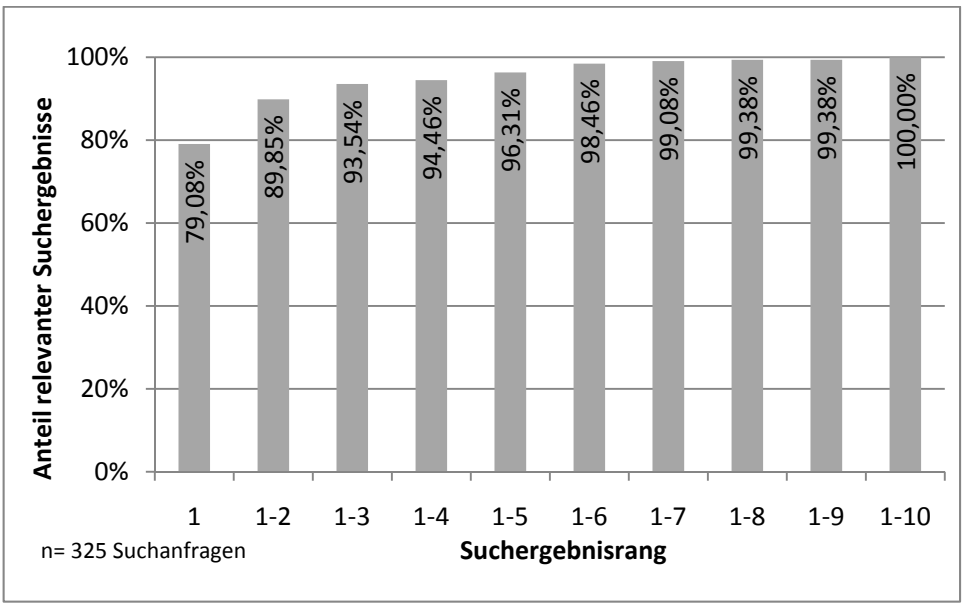

Abb. 9: Kumulierter Anteil der relevanten Suchergebnisse je Rang (Quelle: eigene Darst.)

Es kann also nicht bestätigt werden, dass ein größerer Teil der relevanten Treffer womöglich vom ursprünglich Suchenden übersehen wurde. Da Lorigo et al. in einer Eyetracking-Studie nachweisen konnte, dass in den meisten Fällen nur die am höchsten gerankten Ergebnisse vom User inspiziert werden (vgl. LORIGO et al. 2006, S. 1127), sollte der ursprünglich Suchende durch die Platzierung der relevanten Treffer auf den ersten beiden Positionen der SERP bei fast 90 Prozent der Suchanfragen, die Suchergebnisse nicht übersehen haben. Lediglich bei einem Anteil von 10,15 Prozent der Suchanfragen könnte das relevante Suchergebnis übersehen worden sein, wenn der Suchende die Suchergebnispositionen drei bis zehn nicht beachtete.

c) Persönlichkeit / Vorkenntnisse des Suchenden

Eine weitere, allerdings nicht prüfbare Vermutung für den Grund für die sich stark unterscheidende Relevanzbewertung von ursprünglich Suchenden und Proband könnte in den persönlichen Umständen und Vorkenntnissen der Probanden und der Suchenden liegen. Dudek et al. merken an, dass bei der Bewertung durch den Nutzer zu beachten ist, dass immer auch Charakterzüge und Erwartungen des Nutzers mit einfließen (vgl. DUDEK et al. 2006, S. 227). So wäre bspw. denkbar, dass der ursprünglich Suchende keinen Treffer mit einem Klick bedachte, da die erhaltenen Ergebnisse durch eine vorherige Suche bereits gesichtet wurden und sie daher für ihn keine Relevanz mehr aufwiesen. Des Weiteren wäre es möglich, dass er auf den ersten Blick bemerkt, dass die Formulierung seines Suchbedürfnisses nicht treffend war, da er die Anfrage z.B. zu allgemein gehalten hat. In diesem Fall finden sich zur Anfrage relevante Ergebnisse, jedoch nicht der Art wie sie der Suchende erwartet hat.

Insgesamt muss an dieser Stelle gefolgert werden, dass die beschriebenen Vermutungen nicht nachgeprüft werden können, da eine Ansprache des ursprünglich Suchenden nicht möglich ist. Allerdings wird vermutet, dass die in diesem Abschnitt beschriebenen Gründe zumindest für einige Anfragen zutreffen.

d) Testsituation

Die abweichende Bewertung der Treffer durch den ursprünglich Suchenden und dem Proband könnte teilweise auch durch die Testsituation selbst hervorgerufen werden. Diese zeichnet sich u.a. dadurch aus, dass das Ziel der Suche nicht im Original vorlag, sondern in einer vorangehenden Befragung abgeleitet wurde. Dementsprechend ist es möglich, dass das Informationsbedürfnis trotz des Einsatzes mehrerer Probanden nicht korrekt abgeleitet wurde. Die Bewertung der Relevanz und das potentielle Klickverhalten würden sich dann aufgrund der verschiedenen Informationsbedürfnisse unterscheiden.

Abschließend wird für die Gruppe der Suchanfragen, die im Relevanztest relevante Ergebnisse zugeordnet bekamen, gefolgert, dass der Unterschied im Klickverhalten von ursprünglich Suchenden und Proband teilweise mit dem Aktualitätsbezug der Suche (14,15) und dem Übersehen des relevanten Suchergebnis durch den ursprünglich Suchenden (10,15 Prozent) erklärt werden können. Allerdings muss auch angenommen werden, dass sich ein Anteil der Bewertung der Anfragen aufgrund von Persönlichkeit und Vorkenntnis zwischen Suchenden und Proband unterscheidet. Des Weiteren ist es außerdem möglich, dass durch die Testsituation selbst Verzerrungen auftraten. Die letztgenannten Gründe für die unterschiedliche Bewertung zwischen Proband und originär Suchenden können allerdings nicht nachgeprüft werden.

4.3.3 Suchergebnis beantwortet Informationsbedürfnis

Eine Hypothese zur Begründung der Anfragen ohne Klick ist, dass das Informationsbedürfnis hinter diesen Anfragen bereits auf der Suchergebnisseite durch die bloße Ansicht der Trefferbeschreibung beantwortet werden konnte und daher kein Klick auf das Ergebnis notwendig ist. Diese Bewertung eines Treffers wurde im Relevanztest mit Hilfe der Antwortmöglichkeit „Nein, ich würde nicht klicken. Die Ansicht des Treffers reicht, denn er erfüllt das Ziel der Suche“ angeboten. Insgesamt wurden 83 Anfragen mit dieser Antwort gekennzeichnet (19,48 Pro-

zent). Auf diese 83 Anfragen entfallen 216 Trefferbeschreibungen, die als Antworten auf das Ziel der Suchanfrage gewertet wurden. Allerdings sind diese 83 Anfragen nicht nur mit der Wertung „Nein, ich würde nicht klicken. Die Ansicht des Treffers reicht, denn er erfüllt das Ziel der Suche.“ ausgezeichnet worden, in 68 Fällen erhielten einer oder mehrere der übrigen Treffer dieser Suchen auch die Wertung „Ja, ich würde den Treffer anklicken“. D.h. nur 15 Suchanfragen würden laut dem Relevanztest allein durch die Ansicht des Treffers beantwortet werden. Zwar könnte daraus gefolgert werden, dass die übrigen 68 Suchen potentiell einen Klick bekommen hätten und aus diesem Grund streng genommen keine Erklärung für das Unterlassen des Klicks durch den ursprünglich Suchenden bieten. Dennoch sollen im Folgenden alle 83 Anfragen näher untersucht werden, um, wie in dem Kapitel zur Fragestellung der Arbeit herausgestellt, diese Anfragen und ihre Eigenschaften charakterisieren zu können. Allerdings muss auch an dieser Stelle angemerkt werden, dass die Anzahl der 83 Suchen keine ausreichend große Fallzahl darstellt, weswegen nur Hinweise, nicht jedoch repräsentative Aussagen über diese Suchanfrage-Gruppe gesammelt werden können.

Zur Charakterisierung von Anfragen, die auf der Trefferseite beantwortet werden konnten, werden in der Fragestellung dieser Arbeit die Bereiche Thema der Suche und Suchanfragetyp festgehalten. Eine Analyse der Art des Treffers in Zusammenhang mit der Intention der Suche soll ebenfalls Aufschluss über die Charakteristika der Suchen geben, die bereits auf der Suchergebnisseite beantwortet werden können.

Eine Auswertung der Suchanfragen nach ihrem Thema kann in Abb. 10 abgelesen werden.

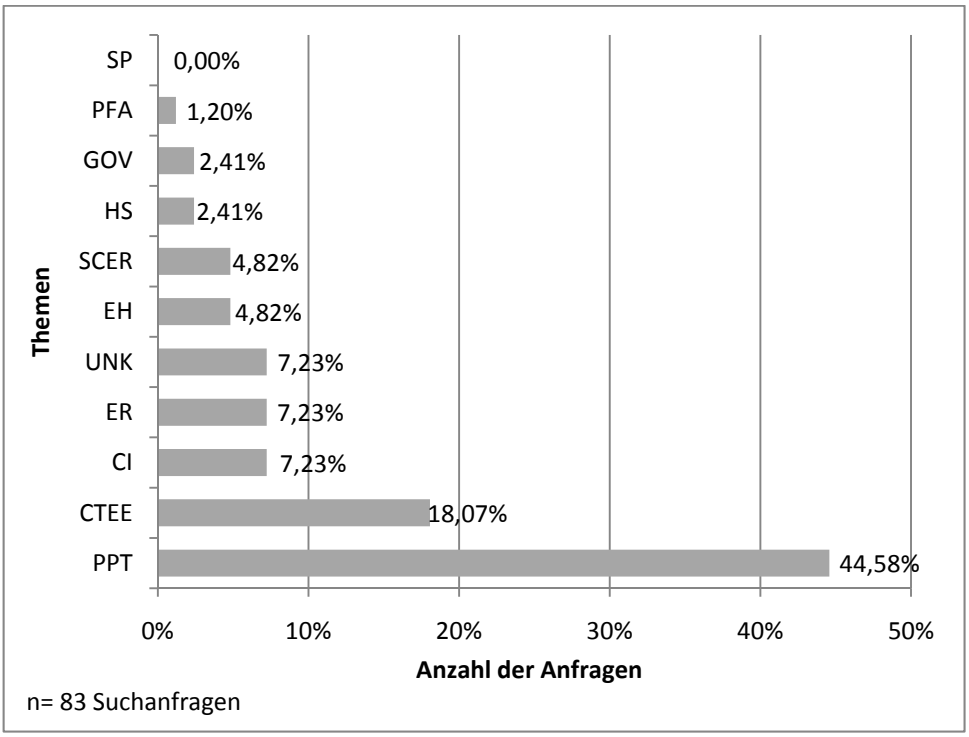

Abb. 10: Themen der Suchanfragen, die auf der SERP beantwortet werden konnten (Quelle: eigene Darst.)

Im Vergleich zur Themenauswertung der Zufallsstichprobe der 600 ungeklickten Suchanfragen zeigt sich ein erhöhter Anteil des Themas „Personen, Orte und Dinge“ (PPT) (44,58 Prozent im Vergleich zu 29,67 Prozent, Näheres in Kapitel 4.1.5.3). Eine detailliertere Analyse zeigt, dass hier 28 der 37 Anfragen einen Personennamen beinhalten. Dieser erhöhte Anteil gleicht dem Anteil für die Gruppe der Suchanfragen ohne ein irrelevantes Ergebnis (siehe Kapitel 4.3.1), auch hier konnte ein hoher Anteil von Personennamen unter den Suchanfragen entdeckt werden. Der Anteil für die auf der Trefferseite beantwortbaren Suchen ist zwar niedriger, im Vergleich zur Zufallsstichprobe aber immer noch stark erhöht. Die in Kapitel 4.1.4.3 aufgestellten Vermutungen, dass eine Suche nach Personen häufig nicht zufriedenstellend verläuft, da ohne Nutzung der Phrasensuche oft nur der gesuchte Vor- oder Nachname in den Suchergebnissen auftritt und die

Suche dementsprechend viele irrelevante Treffer liefert, muss also eingeschränkt werden. Augenscheinlich werden fast ebenso viele Personensuchen bereits durch die Trefferbeschreibung beantwortet.
Die weitere Aufteilung der Suchanfragen auf die unterschiedlichen Themen zeigt keine Besonderheiten.

In einer weiteren Untersuchung teilt die Forscherin die Suchanfragen in die Anfragetypen navigationsorientiert, transaktionsorientiert und informationsorientiert ein. Das Ergebnis kann in Abb. 11 abgelesen werden.

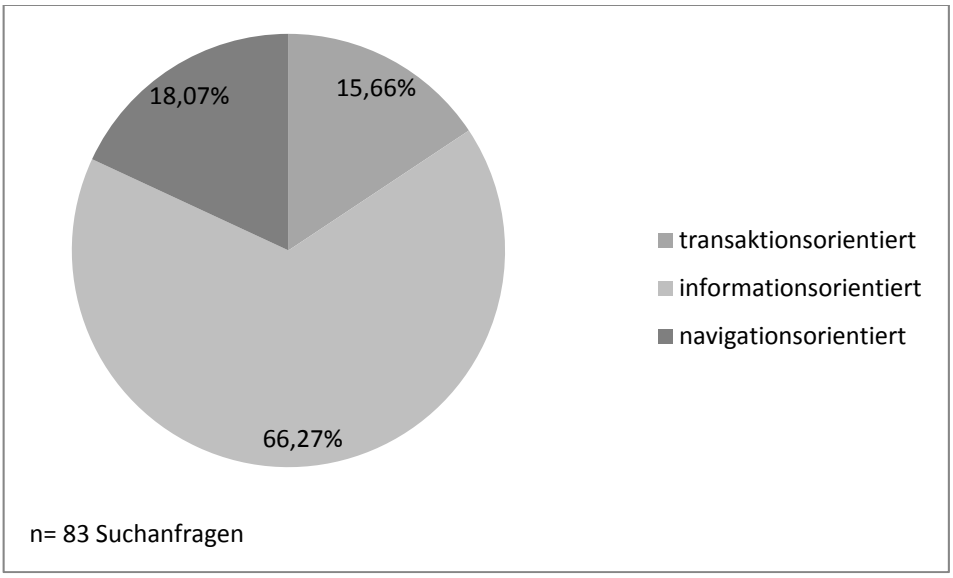

Abb. 11: Anfragetypen der Suchen, die auf der Trefferseite beantwortet werden können (Quelle: eigene Darst.)

Die Mehrheit der Anfragen, nämlich 66,27 Prozent, ist als informationsorientiert eingestuft worden. 18,07 Prozent der Anfragen werden als navigationsorientiert einsortiert und 15,66 Prozent der Anfragen werden als transaktionsorientiert bezeichnet. Überraschend ist der hohe Anteil der navigationsorientierten und transaktionsorientierten Anfragen für diese Suchen. Eine navigationsorientierte Suche zeichnet sich dadurch aus, dass sie durch eine bestimmte, dem Suchenden oft bereits bekannte Website, beantwortet werden kann. Das Ziel der Suche ist es also zu dieser Website zu gelangen, weswegen das Suchergebnis angeklickt

werden müsste. Auch eine transaktionsorientierte Suche bedarf eines Klicks, denn hier geht es darum z.B. einen Newsletter zu abonnieren, einen Kauf abzuschließen oder die Suche auf einer speziellen Seite fortzusetzen. Aus den angeführten Gründen überrascht die Auswertung der Anfragetypen, da vom Proband ausgewählt wurde, dass Ergebnis nicht anklicken zu wollen, da das Ziel der Suche bereits mit Ansicht der Trefferbeschreibung beantwortet werden konnte. Als Gründe für dieses Verhalten kann nur gemutmaßt werden, dass die Probanden den Fragebogen nicht korrekt ausfüllten oder er nicht verstanden wurde. Diese Problematik betrifft 28 Anfragen aus der in diesem Abschnitt beschriebenen Suchanfragen-Gruppe. Eine nähere Untersuchung dieser Anfragen wird im folgenden Abschnitt vorgenommen.

In einem weiteren Schritt sollen die Treffer, die Antwort auf das Ziel der Suche geben, untersucht werden. Dafür wird im Zusammenhang mit der Suchintention herausgearbeitet, welcher Art die Anfrage ist und wie sie die Suchintention beantwortet. Dabei wurden die Trefferbeschreibungen gemeinsam mit dem Ziel der Suche betrachtet, die Treffer näher analysiert und die Anfragen anschließend in Gruppen gegliedert. Die folgenden Gruppen konnten dabei für Anfragen, die auf der SERP beantwortet werden können, herausgearbeitet werden:

- Personensuchen,
- Beantwortung einer Faktenfragen,
- Suche nach Textstellen,
- Definition oder kurze Erklärung eines Wortes,
- Produktdetails.

Im Folgenden werden die einzelnen Gruppen erläutert und mit Beispielen zur Veranschaulichung hinterlegt.

Personensuchen

Die Untersuchung der Anfragen, die als auf der SERP beantwortbar charakterisiert werden konnten, zeigt, dass 28 Anfragen eine Personensuche beinhalten. Das Ziel der Suche war in den meisten Fällen als „Informationen über die Person herausfinden“ definiert worden, in einigen Fällen sollten auch konkreter Kontaktdaten gesucht werden. Eine Analyse der zugehörigen Treffer zeigt, dass die Suchergebnisse, die die Suche auf der SERP beantworten konnten, bereits im Treffer Kontaktdaten, wie die Telefonnummer oder die Adresse der gesuchten Person enthalten (siehe z.B. Abb. 12).

Abb. 12: Treffer der Suche „[redacted] telefonnummer“ (Quelle: SUCHE T-ONLINE.DE 2010)

Ebenfalls geben viele der untersuchten Treffer Auskunft über weitere Informationen zu der gesuchten Person, z.B. auf welche Schule sie gegangen ist, welchem Beruf sie nachgeht oder welche Sportart sie ausübt. Insgesamt konnten 49 der 216 Treffer eine Personensuche bereits durch die Trefferbeschreibung beantworten.

An dieser Stelle muss allerdings berücksichtigt werden, dass Namen nicht einzigartig sind und der ursprünglich Suchende evtl. eine andere Person meinte, die er anhand von ihm bekannten Merkmalen (bestimmter Beruf, Wohnort) identifizieren konnte. Dies ist dem Probanden nicht möglich gewesen, so dass bei diesen Treffern in Betracht gezogen werden muss, dass sie dem ursprünglich Suchenden evtl. keine Informationen lieferten, obwohl sie scheinbar die gesuchte Person enthielten.

Beantwortung einer Faktenfragen

Des Weiteren zeigt sich, dass Suchen, die einfache Fakten abfragen, z.B. „Suche nach dem Ort zu der Postleitzahl 66130“ (siehe Abb. 13) oder „Suche nach dem Lebensraums des Tiers Jaguar“ (siehe Abb. 14), ebenfalls unter die Anfragen fallen, welche bereits auf der SERP beantwortet werden können.

Abb. 13: Treffer der Suche „66130“ (Quelle: SUCHE T-ONLINE.DE 2010)

Hier geben die Treffer durch ganze Sätze oder Teile eines Satzes Antwort auf das Suchbedürfnis.

Abb. 14: Treffer der Suche „wo lebt der jaguar?“ (Quelle: SUCHE T-ONLINE.DE 2010)

Suche nach Textstellen

Eine weitere Art Suchanfragen, bei denen das Suchbedürfnis bereits auf der Trefferseite beantwortet werden kann, sind solche, die nach bestimmten Textstellen suchen, um diese zu vervollständigen oder weitere Informationen dazu zu finden (siehe z.B. Abb. 15)

Abb. 15: Treffer der Suche „barmherzigkeit ist leichter zu üben als“ (Quelle: SUCHE T-ONLINE.DE 2010)

Erklärungen/ Definitionen eines Wortes

Auch die Definition oder Erklärung eines Wortes wird in einigen Fällen direkt auf der SERP gegeben, ein Beispiel ist die Suche nach der Definition des Wortes „Crashkurs“ (siehe Abb. 16).

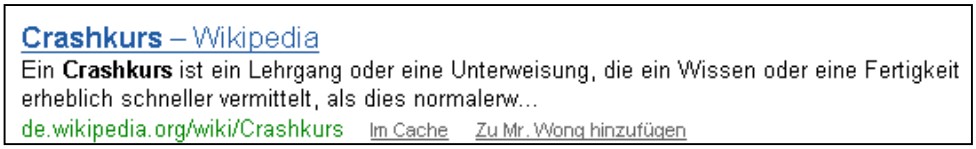

Abb. 16: Treffer der Suche „crashkurse“ (Quelle: SUCHE T-ONLINE.DE 2010)

Produktdetails

Einige Produktsuchen, z.B. nach einem Hotel oder einem technischen Gerät, werden ebenfalls als auf der SERP beantwortbar eingestuft. Der erste Eindruck ist hier, dass der Proband eine falsche Einordnung vorgenommen hat, da es sich dabei um transaktionsorientierte Suchen handelt, die folglich einer weiteren Aktion bedürfen. Allerdings zeigt eine nähere Untersuchung der betreffenden Suchergebnisse, dass hier oft bereits Details zum Produkt genannt werden (siehe Abb. 17) und damit dem Suchenden bereits Informationen zu dem Produkt gegeben werden, die seine Suchintention beantworten könnten (Bspw. wenn ein Preisvergleich angestrebt wird).

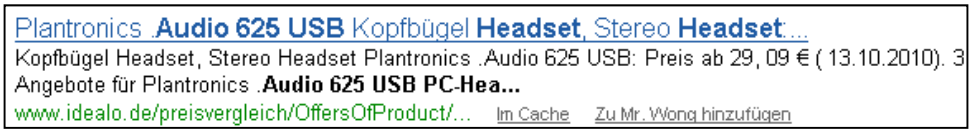

Abb. 17: Treffer der Suche „audio 625 usb-pc-headset“ (Quelle: SUCHE T-ONLINE.DE 2010)

Auch einige navigationsorientierte Anfragen werden als auf der SERP beantwortbar eingestuft, dies konnte bereits bei der Analyse der Suchanfragetypen herausgearbeitet werden. Von der Forscherin wurde angenommen, dass diese Anfragen nicht in diese Suchanfrage-Gruppe eingeordnet würden, da sie angeklickt werden müssen, damit eine Navigation bzw. Weiterleitung auf die gewünschte Seite

stattfinden kann. Eine Untersuchung der Ergebnisse der betreffenden Suchanfragen zeigt, dass in mehr als der Hälfte der Fälle der Treffer der navigationsorientierten Anfrage die gesuchte URL enthält. Vermutlich ordneten die Probanden sie als auf der SERP beantwortbar ein, da die gesuchte URL bereits auf der Trefferseite zu sehen ist. Da eine Einordnung in diese Gruppe trotz gefundener URL nicht korrekt ist, werden diese 15 Anfragen abgezogen. Es kann folglich nur noch ein Anteil von 15,96 Prozent Suchanfragen, die auf der Suchergebnisseite beantwortet werden können, festgehalten werden.

Abschließend kann für die Untersuchung der Trefferbeschreibungen gefolgert werden, dass diese Aufschluss über die Art der Suchen geben konnte, die auf der SERP beantwortet werden können und damit bereits Hinweise für Vorschläge und Empfehlungen für diese Suchanfragen-Gruppe liefert, die in Kapitel 5.3 aufgeführt und diskutiert werden.

5 Diskussion und Schlussfolgerungen

In diesem Kapitel werden die Ergebnisse dieser Arbeit zusammengebracht und diskutiert, um anschließend Ableitungen vorstellen und konkrete Empfehlungen aussprechen zu können. Vorab werden im Kapitel 5.1 die auftretenden Probleme und die sich daraus ergebenden Einschränkungen der vorliegenden Studie aufgelistet und analysiert. Diese fließen dann in die sich anschließende Ergebnisdiskussion und Empfehlungserstellung mit ein.

5.1 Probleme und Einschränkungen

An dieser Stelle soll auf die Probleme und die daraus resultierenden Einschränkungen der vorliegenden Arbeit eingegangen werden. Ebenfalls erläutert werden sollen der Umgang und die Lösung dieser Problematiken.

Beschränkung des Relevanztests auf die organischen Suchergebnisse der ersten SERP

Zur Relevanzbewertung konnten lediglich die organischen Suchergebnisse einbezogen werden. Zusätzlich werden auf der Suchergebnisseite der T-Online Suche allerdings noch Anzeigen (vgl. Abb. 18, Nr. 2), Ergebnisse aus dem T-Online Portal (vgl. Abb. 18, Nr.1) und Boxen mit Suchergebnissen aus Bereichen wie z.B. „Video“ oder „E-Book“ (vgl. Abb. 18, Nr.3) angezeigt.

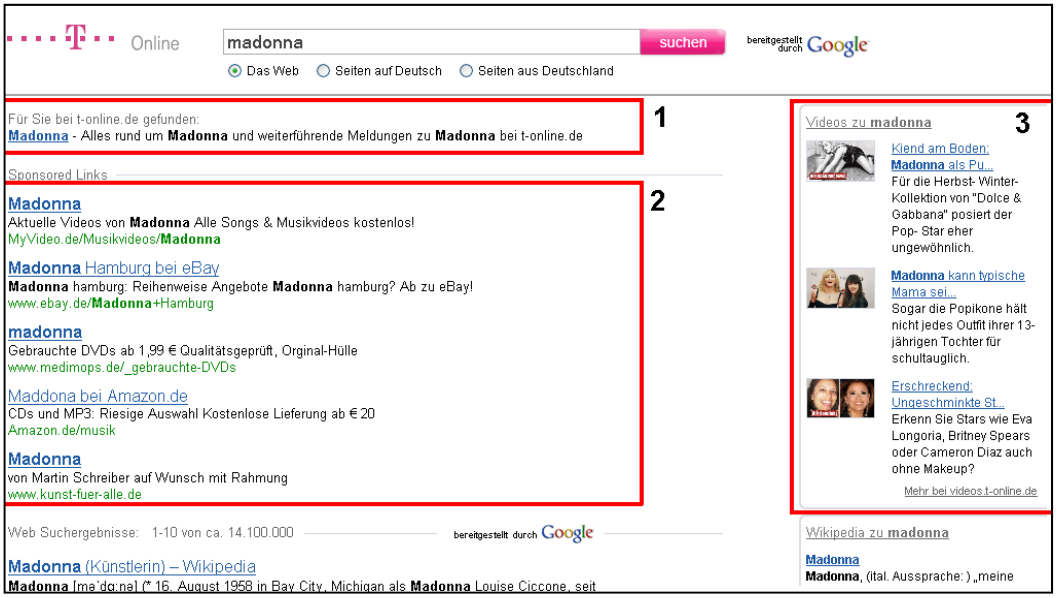

Abb. 18: Suchergebnisseite der T-Online Suche mit Markierung der unterschiedlichen Ergebnisteile (Quelle: SUCHE T-ONLINE.DE 2010 und eigene Darst.)

Die deskriptive Analyse der ungeklickten Suchanfragen zeigt, dass 69,5 Prozent der Suchanfragen der Stichprobe mit mindestens einer Anzeige ausgestattet sind, 29,17 Prozent mit einem Ergebnis des T-Online Portals und bei 1,17 Prozent der Suchanfragen werden zusätzlich zu den organischen Treffern rechts Boxen mit weiteren Treffern eingeblendet. Die fehlende Einbeziehung dieser übrigen Suchresultate lassen vermuten, dass der Anteil der Anfragen, die bereits auf der SERP beantwortet werden können, höher ist, da auch Textanzeigen oder die Suchergebnisboxen der T-Online Suche die Anfragen bereits durch ihre Trefferbeschreibung beantworten können.

Auch der Ausschluss der weiteren Suchergebnisseiten kann das Ergebnis, insbesondere für die Angaben zur Höhe des Anteils der Suchanfragen, die bereits auf der Trefferseite beantwortet werden können, verfälschen. Wie bereits in Kapitel 2.2 ausführlich dargelegt, zeigen diverse Studien, dass sich die meisten Suchenden nur die erste Suchergebnisseite ansehen, weswegen auch in dieser Studie auf eine Bewertung der nachfolgenden Suchergebnisse verzichtet wurde. Dennoch wäre es möglich, dass auch weitere Trefferseiten mit in die ursprüngliche Suche eingeflossen sind und auf diesen die Suchanfrage bereits durch die Trefferbeschreibung beantwortet wurde.

Der Anteil der Anfragen, die auf der Trefferseite zu beantworten wären, kann folglich als eine untere Grenze gesehen werden. Denn durch die Einbeziehung aller weiteren Suchergebnisse sowie weiterer Suchergebnisseiten ist es wahrscheinlich, dass sich der Anteil der auf der Trefferseite beantwortbaren Suchanfragen erhöht.

Zeitspanne zwischen Erhebung der Stichprobe und der Suchergebnisse

Eine weitere Schwierigkeit der Studie ergibt sich aus der Zeitspanne, die zwischen der Ziehung der Stichprobe und der Einholung der Suchergebnisse für den Relevanztest liegt. Die Stichprobe stammt aus April 2010, die Extraktion der Suchergebnisse konnte aufgrund technischer Gegebenheiten erst im August dieses Jahres stattfinden. Eine Änderung der Suchergebnisse ist folglich möglich. Denn die Aktualität des Indexes einer Web-Suchmaschine ist von großer Bedeutung für die Qualität ihrer Ergebnisse. Die Datenbestände müssen möglichst aktuell gehalten werden, um die Dynamik des Webs abbilden zu können. Dabei ist es aufgrund der Größe des Webs für die Suchmaschinen unbedingt notwendig zu erkennen, welche Seiten oft und welche eher selten aktualisiert werden, um ressourcenorientiert arbeiten zu können (vgl. LEWANDOWSKI 2006b, S. 147 f.). In einer dreijährig angelegten Studie konnte Lewandowski herausarbeiten, dass täglich aktualisierte Webseiten im Jahr 2007 bei Google eine durchschnittliche Aktualisierungsfrequenz von 2,2 Tagen, bei MSN von 5,7 Tagen und bei Yahoo von 2,7

Tagen hatten (vgl. LEWANDOWSKI 2008b, S. 823). Die Webseiten, die nur unregelmäßig aktualisiert wurden, werden auch nicht so häufig im Index der Suchmaschine aktualisiert. Laut der Studie von Lewandowski weist Google für Webseiten, die unregelmäßig aktualisiert werden, eine durchschnittliche Aktualisierungsfrequenz von 14,8 Tagen im Jahr 2007 auf. Der Median von 6 macht allerdings deutlich, dass Google zwischen Seiten, die häufiger aktualisiert werden und Seiten, die weniger häufig aktualisiert werden, unterscheidet (vgl. LEWANDOWSKI 2008b, S. 826). Eine Untersuchung der Stichprobe der täglich aktualisierten Webseiten nach der ältesten Webseite zeigt für 2007 ein Alter von zehn Tagen, im Vergleich dazu steht die älteste gefundene Webseite der Stichprobe der Webseiten, die unregelmäßig aktualisiert werden mit einem Alter von 175 Tagen (vgl. LEWANDOWSKI 2008b, S.824).

Für die vorliegende Untersuchung machen die Zahlen deutlich, dass eine Änderung der Suchergebnisse in dem angesprochenen Zeitrahmen von drei bis vier Monaten bei aktuellen Themen sehr wahrscheinlich ist. Aus diesem Grund wurde der Aktualitätsbezug für jede Anfrage abgefragt, eine Änderung der Suchergebnisse ist also für die als aktuell einzuschätzenden Suchanfragen weitaus höher als für Anfragen die keinen aktuellen Bezug haben. Daher darf dies bei einer Interpretation der Ergebnisse nicht unberücksichtigt bleiben.

Keine Zufallsstichprobe

Die Problematik der fehlenden Zufallsstichprobe konnte nicht umgangen werden, vielmehr musste festgestellt werden, dass sich die Stichprobe der Top Millionen Suchanfragen hinsichtlich des Anteils der ungeklickten Suchanfragen stark von einer Zufallsstichprobe unterscheidet. Denn der in der vorliegenden Arbeit erhobene Anteil von 6,56 Prozent Suchanfragen, die keinen Klick erhalten, ist wesentlich geringer als bei den verglichenen Zufallsstichproben. Dementsprechend muss gefolgert werden, dass ein großer Anteil von Suchanfrage ohne Klick durch die Verwendung der Top Millionen Suchanfragen wegfallen. Zwar zeigt der weitere Vergleich verschiedener Merkmale der vorliegenden Datenbasis mit Zu-

fallsstichproben von Suchanfragen keine Auffälligkeiten, jedoch lassen sich einige Merkmale, wie die Nutzung von Suchoperatoren oder die Anteile der Anfragetypen nicht sicher bestimmen.
Abschließend muss an dieser Stelle festgehalten werden, dass die Problematik der fehlenden Zufallsstichprobe nicht umgangen werden kann und die in dieser Arbeit herausgestellten Ergebnisse nur auf diese Datenbasis umgelegt werden können.

5.2 Suchanfragen ohne Klick

Durch eine deskriptive Analyse ungeklickter Suchanfragen und zwei Online-Befragungen, die die Relevanzmessung ungeklickter Suchergebnisse zum Ziel hatten, konnte der Zielsetzung dieser Arbeit, Erkenntnisse über Suchanfragen zu gewinnen, die keinen Klick auf der Suchergebnisseite auslösten, nachgekommen werden. Im folgenden Text werden die bisherigen Erkenntnisse zusammengefügt und diskutiert.

Suchanfragen, die auf der SERP beantwortet werden können

Die in der Einleitung aufgegriffene Aussage, wonach die Klicks der Suchenden als implizite Relevanzurteile gewertet werden können, war ein Auslöser dieser Arbeit. Denn betrachtet man diese Aussage im Rückschluss, müsste die Relevanzbewertung einer Suchanfrage ohne Klick besonders negativ ausfallen. Diese Schlussfolgerung konnte mit der vorliegenden Arbeit negiert werden. Erhält eine Suchanfrage keinen Klick deutet dies nicht zwangsläufig auf eine Suche ohne relevantes Ergebnis und damit auf einen unbefriedigte Suche hin. Vielmehr konnte mit Hilfe des Relevanztests festgestellt werden, dass ein wesentlicher Anteil der Anfragen der Stichprobe durch die bloße Ansicht des Treffers beantwortet werden konnte. Ein Anteil von 15,96 Prozent der Suchanfragen der Stichprobe wurde als auf der Suchergebnisseite beantwortbar eingestuft und bestätigt damit die Hypothese 5 (Suchanfrage wird durch Trefferbeschreibung beantwortet). Ein direkter Vergleich dieses Anteils mit bisher durchgeführten Studien

ist nicht möglich, da jeweils unterschiedliche Stichproben gebildet wurden. Insgesamt kann aber geschlossen werden, dass die Forscher (vgl. LI ET AL. 2009; STAMOU und EFTHIMIADIS 2009) ebenfalls zu dem Schluss kommen, dass eine Suchanfrage ohne Klick nicht zwingend eine gescheiterte Suche ist. Die Höhe des Anteils der bereits auf der SERP zu beantwortenden Suchanfragen ist dabei in der vorliegenden Arbeit eher als Tendenz denn als repräsentative Aussage zu verstehen, wurde doch in Kapitel 3.2 festgehalten, dass die genutzte Datenbasis keine Zufallsstichprobe ist und daher die Ergebnisse nicht auf alle Suchanfragen umgelegt werden können. Des Weiteren konnten alle Ergebnisse, die zusätzlich zu den organischen Treffern eingeblendet wurden, nicht mit in die Evaluation der Relevanz einbezogen werden. Es kann also nur vermutet werden, dass der Anteil der Suchanfragen, die bereits auf der Trefferseite zu beantworten sind, noch höher liegt.

In der vorliegenden Arbeit konnte außerdem herausgestellt werden, dass einige Fragestellungen besonders häufig ohne Klick auf das Suchergebnis beantwortet werden können. Dazu gehören: Informationen über Personen, (einfache) Faktenfragen, Suche nach Textstellen, Definitionen und Worterklärungen sowie die Suche nach Produkten. Ein Vergleich dieser Kategorien mit den Untersuchungen von Li et al. und Stamou und Efthimiadis zeigt Gemeinsamkeiten, aber auch Unterschiede. So ist den Studien gemein, dass für die Suchen, die bereits auf der SERP beantwortet werden können, eine Kategorie angeführt wird, die auf die Beantwortung einer Frage abzielt. Li et al. nennen diese Kategorie „answer“, Stamou und Efthimiadis „quick answer“. In der vorliegenden Studie wird sie als Antwort auf eine Faktenfrage bezeichnet. Dass diese Kategorie in allen Studien als ein Suchanfragetyp identifiziert wird, der bereits auf der SERP beantwortbar ist, unterstreicht seine Bedeutung.
Es bestehen allerdings auch Unterschiede bei der Kategorisierung der Suchanfragen, die auf der SERP beantwortbar sind. Bei Li et al. wird bspw. die Kategorie „local“ als stärkste Kategorie identifiziert, während Stamou und Efthimiadis hier

u.a. die Kategorie „language-related“ hervor stellen. In der vorliegenden Untersuchung kommen diese Kategorien nicht vor, stattdessen werden viele Suchen nach Personen verzeichnet, die bereits durch die Trefferbeschreibung beantwortet werden können. Hier ist allerdings anzumerken, dass Personennamen nicht einzigartig sind und das daher in einigen Fällen davon ausgegangen werden muss, dass nicht die korrekte Person im Treffer genannt wurde. Diese könnte nur der originär Suchende identifizieren, welcher allerdings nicht zur Verfügung stand. Abschließend kann zur Kategorisierung der Suchanfragen, die bereits durch die Trefferbeschreibung beantwortet werden, geschlossen werden, dass hier Unterschiede zwischen den Studien vorliegen, welche sich teilweise durch länderspezifische Differenzen als auch durch die bereits erläuterte fehlende Zufallsstichprobe erklären lassen könnten. Eine sichere Schlussfolgerung ist nicht möglich.

Suchanfragen ohne Klick mit irrelevanten Treffern

Die Untersuchung der ungeklickten Suchanfragen zeigte aber auch, dass bei einem wesentlichen Anteil der Anfragen nicht davon ausgegangen werden kann, dass die Suche mit einer geeigneten Trefferbeschreibung beantwortet werden kann. So zeigt die deskriptive Analyse der Suchanfragen ohne Klick, dass ein Anteil von 4,67 Prozent dieser Suchen überhaupt kein Suchergebnis aufweist. Dies macht deutlich, dass hier selbstverständlich auch kein relevanter Treffer gefunden werden kann. Suchanfragen, die mindestens einen Klick auslösten, hatten im Vergleich nur in 0,17 Prozent der Fälle kein Suchergebnis zu verzeichnen.
Auch der Relevanztest macht deutlich, dass ein ungeklickter Treffer oft bedeutet, dass die Treffer(-beschreibungen) irrelevant sind. So werden alle zehn Suchergebnisse von 20,65 Prozent der untersuchten Suchanfragen im Relevanztest als irrelevant bewertet. Dies bestätigt Hypothese 1 (Es werden nur Treffer(-beschreibungen) präsentiert, die für die Suchanfrage und das dahinter stehende Informationsbedürfnis irrelevant sind). Auch bei diesem Wert muss beachtet werden, dass es sich bei der Datenbasis nicht um eine Zufallsstichprobe handelte

und keine weiteren Ergebnisse außer den organischen Treffern hinzugezogen wurden. Es kann nur vermutet werden, dass sich der Anteil der Suchanfragen mit irrelevanten Treffern verkleinern würde, wenn man alle Treffer einer Suchergebnisseite hinzuziehen würde.
Suchanfragen ohne Klick enthalten außerdem häufig Fehler, denn in 19 Prozent der Fälle lösen sie einen Korrekturvorschlag der Suchmaschine aus. Die Untersuchung der Suchanfragen, die nur in irrelevante Treffer münden zeigt, dass dieser Anteil hier sogar noch einmal ansteigt, und zwar auf 23,86 Prozent der Suchen. Es ist also sehr wahrscheinlich, dass ein Anteil der Suchanfragen ohne Klick folglich mit fehlerhaften Suchbegriffen gestellt wurde und aufgrund dessen zu keinem relevanten Ergebnis führte. Dies bestätigt die Hypothese 2, die besagt, dass ein Grund für das nicht Klicken ein Fehler in der Suche ist, welcher von der Suchmaschine erkannt wird und in einem Korrekturvorschlag resultiert. Entscheidet sich der Suchende für den Korrekturvorschlag seiner Suche wird folglich kein Klick auf die Suchergebnisse verzeichnet.

Durch die deskriptive Analyse der Suchanfragen der ungeklickten Treffer im Vergleich zu den Suchanfragen der geklickten Treffer konnte ein möglicher weiterer Grund für die fehlenden relevanten Treffer herausgearbeitet werden. Es wurde festgestellt, dass Suchanfragen ohne Klick durchschnittlich aus mehr Wörtern bestehen als die ungeklickten Suchanfragen, nämlich aus 2,6 Wörtern, während die geklickten Treffer durchschnittlich nur 2 Wörter aufweisen. Die Diskussion dieses Ergebnisses in Kapitel 4.1.4.1 kommt zu dem Schluss, dass die längeren ungeklickten Anfragen höchstwahrscheinlich ein spezifischeres Informationsbedürfnis beinhalten, welches nicht in das gängige Anfrageschema der Suchmaschinen passt, welche dazu auffordern Anfragen so einfach und kurz wie möglich zu halten (vgl. GOOGLE WEB-SUCHE HILFE 2010). Die ungeklickten Anfragen, deren Ergebnisse im Relevanztest keine Relevanz zugeschrieben wurde, verzeichnen eine durchschnittliche Suchanfragenlänge von 2,6 Wörtern, während die Anfragen, denen Relevanz zugeschrieben wurde eine etwas kürzere durchschnittliche

Suchanfragenlänge von 2,5 Wörtern aufweisen. Dies kann als weiterer Beleg für die aufgeführte Hypothese gewertet werden.

Die deskriptive Analyse weist durch die Untersuchung der fremdsprachigen Suchanfragen und Suchergebnisse außerdem daraufhin, dass diese in der Stichprobe der ungeklickten Suchanfragen mit einem Anteil von 19,5 Prozent der Suchanfragen stark vertreten sind. Im Vergleich zu den Suchanfragen, die mindestens einen Klick erhalten, zeigen hier nur 8,83 Prozent der Suchen eine fremdsprachige Suchanfrage oder überwiegend fremdsprachige Suchergebnisse. Da diese Anfragen für die beiden Online-Befragungen herausgenommen werden mussten, können hier keine weiteren Schlussfolgerungen gezogen werden. Es könnte zwar die Möglichkeit bestehen, dass für diese Anfragen das Ziel der Suche bereits durch die Trefferbeschreibung beantwortet werden konnte und daher kein Klick erfolgte, aufgrund der in Kapitel 4.1.4.4 angeführten Untersuchungen zu fremdsprachigen Suchergebnissen wird jedoch eher die Vermutung gestützt, dass der hohe Anteil fremdsprachiger Suchanfragen / Suchergebnisse darauf hinweist, dass zumindest für die Anfragen, dessen Suchergebnisse in einer dem Suchenden fremden Sprache ausgegeben werden, die Ergebnisse nicht zufriedenstellend sind und daher keinen Klick erhielten.

Die weiteren Analysen ergeben, dass insbesondere die Suche nach einer Person bei den Anfragen ohne relevanten Treffer thematisch stark vertreten ist, was zeigt, dass dieses Thema zwiespältig auftritt. Auf der einen Seite ist es bei den Anfragen, die bereits auf der SERP beantwortet werden können, eines der Themen mit den meisten Suchanfragen, andererseits ist es jedoch auch stark bei den Suchanfragen vertreten, deren Treffer als irrelevant bewertet werden. Die Diskussion der Personensuchen, die auf der Trefferseite beantwortet werden können, kommt jedoch zu dem Schluss, dass nicht sicher gefolgert werden kann, ob wirklich die korrekte Person gefunden wurde, da Namen häufig in gleicher Zusammensetzung auftreten. Aus diesem Grund wird die Schlussfolgerung gezo-

gen, dass Personensuchen häufig nicht zufriedenstellend beantwortet werden können.

Suchanfragen ohne Klick mit relevanten Treffern

Die im Relevanztest untersuchten Suchanfragen ohne Klick erhielten mehrheitlich, nämlich in 76,11 Prozent der Fälle, die Einschätzung mindestens einen relevanten Treffer, welcher auch angeklickt würde, aufzuweisen. Wie bereits in vorangegangenen Kapiteln betont, geht es bei dem Relevanztest nicht darum, dass Verhalten des ursprünglich Suchenden zu simulieren, sondern vielmehr darum, die SERP bei nicht geklickten Treffern zu erforschen und damit aufzudecken, welche Möglichkeiten der originär Suchende gehabt hätte, um sein Informationsbedürfnis zu stillen. Demnach müsste gefolgert werden, dass der ursprünglich Suchende bei 76,11 Prozent der Suchanfragen, deren Ergebnisse keinen Klick vom Suchenden erhalten haben, doch die Möglichkeit gehabt hätte ein relevantes Ergebnis unter den ersten zehn organischen Treffern zu entdecken. Das trotzdem kein Ergebnis angeklickt wurde, kann teils mit den sich ändernden Suchergebnissen für aktuelle Suchanfragen (14,15 Prozent) und teils mit relevanten Suchergebnissen, welche nicht auf den Top Positionen der SERP zu finden sind (10,15 Prozent), erklärt werden. Letztgenannte Treffer könnten übersehen worden sein, wie auch in Hypothese 4 (Übersehen des/der relevanten Ergebnisse(s)) dieser Arbeit vermutet. Ein weiterer Grund für den Unterschied im (potentiellen) Klickverhalten des Suchenden und des Probanden können aus der ersten Hypothese dieser Arbeit abgeleitet werden. Hier wird als Grund für das nicht anklicken eines Treffers angenommen, dass der Suchende eine wiederholte Suche abschickte, aber kein Ergebnis anwählte, da keine neuen Treffer eingeholt werden konnten und demzufolge kein relevanter Treffer vorhanden war. Diese Annahme lässt sich nicht belegen, die Hinweise deuten aber daraufhin, dass ein bestimmter Anteil der Anfragen aus diesem Grund vom ursprünglich Suchenden keinen Klick erhielt. Weitere Gründe für das sich stark unterscheidende Klickverhalten von Suchendem und Proband kann mit dem Umstand erläutert werden, dass die äu-

ßeren Umstände und die Vorkenntnisse des jeweils Urteilenden sich unterschieden.

Bevor auf Grundlage der hier aufgelisteten Schlussfolgerungen in Kapitel 5.3 Empfehlungen für die Anpassung der Suche abgeleitet werden, gibt die Tab. 8 noch einmal einen Überblick über die wichtigsten Ergebnisse, erfasst die erhobenen Gründe für das Unterlassen eines Klicks und führt zusätzlich die eingesetzte Methode, den Beleg und die jeweiligen Einschränkungen auf.

Nr.	Gründe für Suchanfragen ohne Klick	Methode	Belege & Anmerkungen	Einschränkungen
1	Ungeklickte Suchanfragen können durch eine Trefferbeschreibung auf der SERP beantwortet werden.	Relevanz-messung	Der Relevanztest zeigt, dass ein Anteil von 15,96 Prozent der Suchanfragen ohne Klick bereits durch die Trefferbeschreibung beantwortet werden kann und nicht angeklickt werden muss.	Keine Zufallsstichprobe. Es wurden keine weiteren Ergebnisse außer den organischen Treffern hinzugezogen.
2	Ungeklickte Suchanfragen führen nur zu irrelevanten Treffern.	Relevanz-messung	Der Relevanztest zeigt, dass ein Anteil von 20,65 Prozent der Suchanfragen besteht, die keinen Klick erhalten, da sie nur zu irrelevanten Ergebnissen führen.	Keine Zufallsstichprobe. Es wurden keine weiteren Ergebnisse außer den ersten zehn organischen Treffern hinzugezogen.
2.1	Ungeklickte Suchanfragen sind häufig fremdsprachig oder weisen auf fremdsprachige Ergebnisse hin. Die Ergebnisse werden nicht verstanden oder eine Relevanzeinschätzung in einer Fremdsprache wird gescheut.	Deskriptive Analyse	Suchanfragen ohne Klick enthalten 19,5 Prozent fremdsprachige Suchanfragen oder Suchanfragen, die zu fremdsprachigen Suchergebnisse führen.	Keine Zufallsstichprobe. Keine Prüfung durch Befragungen möglich.
2.2	Ungeklickte Suchanfragen enthalten weniger weit gefasste Informationsbedürfnisse, welche nicht in das gängige Anfrageschema der Web-Suchmaschinen passen.	Deskriptive Analyse, Relevanz-messung	Ungeklickte Suchanfragen sind durchschnittl. 2,6 Wörter lang, geklickte Anfragen 2 Wörter. Der Relevanztest zeigt, dass Suchen, die nur zu irrelevanten Treffern führen, durchschnittl. längere Suchanfragen beinhalten.	Keine Zufallsstichprobe. Fallzahlen sind zu klein für repräsentative Aussagen.

2.3	Ungeklickte Suchanfragen enthalten häufig Fehler und führen aufgrund dessen zu keinem relevanten Ergebnis.	Deskriptive Analyse, Relevanzmessung	Suchanfragen ohne Klick lösen wesentlich mehr Korrekturvorschläge aus (19 Prozent), als es Suchanfragen mit Klick tun (13,17 Prozent). Der Relevanztest zeigt, dass für die Suchen, die nur in irrelevante Treffer münden, der Anteil der Korrekturvorschläge auf 23,86 Prozent der Anfragen ansteigt.	Keine Zufallsstichprobe. Fallzahlen sind zu klein für repräsentative Aussagen.
2.4	Ungeklickte Suchanfragen sind häufig Suchen nach Personen, die ohne Einsatz von Phrasen und Alleinstellungsmerkmalen der Person zu irrelevanten Ergebnissen führen.	Deskriptive Analyse, Relevanzmessung	Die deskriptive Analyse zeigt, dass ungeklickte Suchanfragen in 29,67 Prozent eine Suche nach einer Person sind. Der Relevanztest ergibt, dass für Suchanfragen, die in irrelevanten Treffer münden, 47,73 Prozent der Suchanfragen aus Personensuchen bestehen.	Keine Zufallsstichprobe. Fallzahlen der Auswertung des Relevanztests sind zu klein für repräsentative Aussagen.

Tab. 8: Gründe für Suchanfragen ohne Klick – Überblick über die wichtigsten Ergebnisse (Quelle: eigene Darst.)

5.3 Empfehlungen & Ausblick

Die vorliegende Arbeit findet für Suchanfragen, die keinerlei Klick auslösen, verschiedene Gründe, wobei zwei Gründen besondere Aufmerksamkeit zugesprochen wird: Den Suchanfragen, die keine Klicks auslösen, da sie nur zu irrelevanten Treffern führen und den Anfragen, die keine Klicks bekommen, da sie bereits durch eine Trefferbeschreibung auf der SERP beantwortet werden können. Alle weiteren erhobenen Gründe (siehe auch Tab. 8, Nr. 2.1 bis 2.4) können einem dieser Gründe untergeordnet werden. Die Empfehlungen und der Ausblick auf zukünftige Forschung, die in diesem Kapitel ausgesprochen werden, sind aus diesem Grund auch in diese beiden Gruppen aufgeteilt.

Suchanfragen, die auf der SERP beantwortet werden können

Die vorliegende Arbeit zeigt, dass die Treffer, die bereits die Antwort auf das Informationsbedürfnis beinhalten, als besonders relevant für den Suchenden bewertet werden. Dementsprechend gilt es, Anfragen, die potentiell auf der Ergebnisseite beantwortet werden können, zu erkennen, um die Trefferdarstellung anzupassen. In dieser Arbeit wurden verschiedene Kategorien von Suchanfragen herausgearbeitet, die auf der Suchergebnisseite beantwortet werden können. Eine Kategorie, welche in allen vorgestellten sowie in der vorliegenden Studie zum Thema Suchanfragen ohne Klick aufgezeigt wurde, ist eine Suche, die schnell bzw. kurz Antwort auf eine Frage gibt. Erkennungsmerkmale für diese Kategorie zu erarbeiten fällt schwer, da die Kategorie sehr weit gefasst ist. Allerdings fallen sowohl in der vorliegenden Arbeit als auch in der Arbeit von Li et al. natürlich gestellte Fragen, wie z.B. „Wo lebt der Jaguar“ oder „who is the lead singer for tonic“ (LI et al. 2009, S. 46) als Beispiele für diese Kategorie auf. Als Erkennungsmerkmal könnten in diesem Fall folglich Frageworte dienen. Jedoch gilt es auf diesem Gebiet weiter zu forschen, um weitere Anfragen dieser Kategorie erkennen und daraufhin die Suche anpassen zu können.

Eine Empfehlung, um die Suche für die beschriebene Kategorie anzupassen, ist die Trefferbeschreibungen zu erweitern, um mehr Informationen, die die Ant-

wort auf die Anfrage enthalten (könnten), darzustellen. Auch scheint es bei diesen Anfragen sinnvoll die Trefferbeschreibung mit ganzen Sätzen auszustatten, auf diese Weise ist die Antwort auf die Anfrage im richtigen Kontext gesetzt und kann vom Suchenden verwertet werden. Die Vorschläge für die Verbesserung der Darstellung sollen am Beispiel der Suchanfrage „wo lebt der Jaguar“ verdeutlicht werden, dessen vierter Treffer in Abb. 19 zu sehen ist.

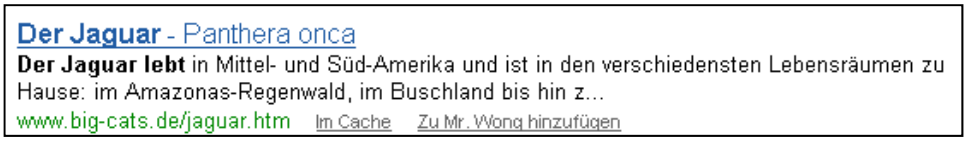

Abb. 19: Treffer vier der Suche „wo lebt der jaguar?“ (Quelle: SUCHE T-ONLINE.DE 2010)

Der Treffer gibt die Antwort auf die Suchanfrage in einem ganzen Satz, wobei das Ende des Satzes abgeschnitten ist. Die Antwort auf den Lebensraum des Jaguars ist zwar mit dem Satzteil „Der Jaguar lebt in Mittel-und Südamerika“ bereits gegeben, jedoch könnte der Suchende eine detailliertere und damit hochwertigere Antwort erhalten, wenn der Satz zu Ende geführt und er damit über die verschiedenen Lebensräume des Jaguars aufgeklärt würde.

Die weiteren Anfragekategorien, die in dieser Arbeit für Suchanfragen, die bereits auf der SERP beantwortet werden können, aufgestellt wurden, werden in diesem Kapitel nicht berücksichtigt, da sie in den anderen Studien nicht aufgedeckt werden konnten oder nur in geringen Fallzahlen auftraten.

Daraus leitet sich wiederum die Empfehlung ab, vermehrt Forschung auf dem Gebiet der Suchanfragen, die bereits auf der Trefferseite beantwortet werden können, zu leisten, um dem Suchenden mit Hilfe gesicherter Aussagen schnell und ohne weiteren Klick Antworten auf seine Anfragen zu ermöglichen. Die Studien zeigen, dass bereits ein signifikanter Anteil der Anfragen auf der SERP beantwortet werden kann, jedoch noch Ausbaupotential besteht. Um Anregungen für verbesserte und damit nutzerfreundlichere Darstellung für diese Anfragen zu

erlangen, könnte eine weitere Befragung von Suchenden Aufschluss über Suchanfragen geben, deren Antworten bereits durch die Trefferbeschreibung gegeben werden sollten. Der Einsatz einer weiteren Befragung oder auch einer Gruppendiskussion könnte außerdem Ideen und Wünsche zu (neuen) Darstellungsformen der Treffer, die die Antwort auf die Anfrage enthalten, geben. Außerdem könnte das Wissen, welches bereits in dem Forschungsbereich der Informationsextraktion gesammelt wurde und z.B. in Question-Answering-Systemen Anwendung findet, mit den Erkenntnissen über Suchanfragen, die keinen Klick erhalten, da sie bereits auf der Suchergebnisseite beantwortet werden können, kombiniert werden. Auf diese Weise könnte die Erkennung der Anfragen, die auf der SERP beantwortet werden können, vorangetrieben werden und gleichzeitig Suchergebnisse geliefert werden, die die Anfrage des Suchenden beantworten ohne das eine weitere Aktion notwendig ist.

Suchanfragen, die nur zu irrelevanten Treffern führen

Die weitere Suchanfragen-Gruppe, die in dieser Arbeit erforscht wurde, sind die Suchanfragen, die keinen Klick erhielten, da sie nur zu irrelevanten Treffern führten. Als Erkennungsmerkmale konnte hier das Thema der Personensuche sowie Fehler in der Suchanfrage identifiziert werden. Aus diesen Erkenntnissen wurden die folgenden Empfehlungen abgeleitet:

Dass viele Anfragen zu Personen nicht zufriedenstellend beantwortet werden, liegt oft daran, dass zwar Vor- und Nachname der gesuchten Person gefunden werden, diese jedoch nicht in Kombination auftreten (siehe z.B. Abb. 6).

Dem Suchenden könnte daher bei Erkennung einer Personensuche der Einsatz des Suchoperators der Phrasensuche vorgeschlagen werden. Das Erkennen einer Person wird auf dem T-Online Suchportal bereits jetzt eingesetzt, um den Dienst „Ruflotse“ anzubieten. Die Abb. 20 zeigt dies am Beispiel der Suche nach der Person „[redacted]“.

Abb. 20: Suchergebnis von T-Online bei der Suche nach „[redacted]“ (Quelle: SUCHE T-ONLINE.DE 2010)

Es konnte außerdem erhoben werden, dass Suchanfragen, die zu keinen relevanten Ergebnissen führen, häufig Korrekturvorschläge auslösen. Daher kann gefolgert werden, dass sie Fehler enthalten. Da diese bereits mit einem Vorschlag zur Korrektur bedacht werden, ist dieses Problem also bereits erkannt und mit einer Lösungsmöglichkeit ausgestattet. Eine weitere Möglichkeit Fehler in den Suchanfragen zu vermeiden, wäre die Nutzung einer automatischen Vervollständigung der Suchanfragen.

Insgesamt kann also gefolgert werden, dass der Suchende, im Falle einer Suchanfrage, die nur zu irrelevanten Treffern führt, Hilfestellungen angeboten bekommen sollte, damit die Suche trotzdem erfolgreich abgeschlossen werden kann. Um weitere (verbesserte) Hilfestellungen zu erarbeiten, wäre es für zukünftige Forschung daher sinnvoll Such-Sessions zu untersuchen. Auf diese Weise könnte erforscht werden, wie Anfragen, die vorab zu keinem relevanten Treffer führten, verändert werden, damit sie zum Erfolg führen. Die Suchhilfe könnte mit einer solchen Untersuchung an den Ursachen für die irrelevanten Treffer ansetzen und ihre Funktion dahingehend anpassen.

Die vorliegende Untersuchung zeigt außerdem, dass die Beurteilung vieler Suchergebnisse sehr unterschiedlich ausfällt. Gründe hierfür wurden ausführlich diskutiert. Eine Empfehlung, um die Einschätzung der Suchanfragen zu erleichtern, wäre es die Trefferbeschreibungen zu ergänzen. Untersuchungen von Drori und Dumais et al. (vgl. DRORI 2000; DUMAIS et al. 2001) zeigen, dass eine Ergänzung der Trefferbeschreibung z.B. durch Metadaten wie Kategorien oder Keywords, die Usability der Trefferrepräsentation und damit die Relevanzbewertung der Treffer erhöht. Auf diese Weise könnten die relevanten/ irrelevanten Treffer vom Suchenden einfacher identifiziert werden.

Zum Abschluss des Kapitels soll noch einmal betont werden, dass die Untersuchung aufgrund der verschiedenen Problematiken, nur Anstöße und Hinweise für weitere Forschung geben, jedoch keine repräsentativen Aussagen für die Gesamtheit der Suchanfragen aufstellen kann. Aus diesem Grund wurden auch nur Empfehlungen ausgesprochen, wenn die dazugehörigen Ergebnisse zusätzlich von anderen Studien bestätigt werden konnten oder nicht von einem der aufgezeigten Problematiken dieser Studie betroffen sind.
Für weitere Forschungen auf diesem Gebiet wird daher empfohlen mit größeren Fallzahlen und einer Zufallsstichprobe zu arbeiten. Des Weiteren wird vorgeschlagen das Suchbedürfnis direkt vom Suchenden zu erheben, denn als ein Grund für den großen Unterschied zwischen Beurteilung der Treffer durch Proband und Suchenden wird vermutet, dass das Informationsbedürfnis des ursprünglich Suchenden nicht korrekt abgeleitet werden konnte. Allerdings müsste der Problematik eines verfälschten Informationsbedürfnisses durch das Bewusstsein des Probanden, dass seine Suchanfrage untersucht wird, noch mit einer geeigneten Methode begegnet werden. Die weiteren eingesetzten Methoden in dieser Arbeit haben sich als praktikabel und zielführend erwiesen und können daher auch für künftige Forschungen in diesem Gebiet empfohlen werden.

6 Fazit

Das übergeordnete Ziel der Arbeit, Suchanfragen, die keinen Klick erhalten, zu erforschen, konnte mit der vorliegenden Untersuchung erreicht werden. Durch die Kombination von deskriptiver Analyse und Relevanzmessung konnten die in der Einleitung der Arbeit aufgestellten Gruppen von Suchanfragen näher untersucht und charakterisiert werden. In Folge dessen bestätigte sich, dass das Unterlassen eines Klicks nicht bedeuten muss, dass kein relevanter Treffer gefunden wurde, vielmehr konnte herausgearbeitet werden, dass ein signifikanter Anteil der Suchanfragen der Stichprobe durch die bloße Ansicht der Trefferbeschreibung beantwortet wurde. Demgegenüber steht ein Anteil von Suchanfragen, die keinen Klick erhalten, da sie nur zu irrelevanten Treffern führen. Auch dieser Grund konnte mit Hilfe der in dieser Studie durchgeführten Untersuchungen für das nicht vorhandene Klickverhalten bestätigt werden. Weitere Explorationen stellen Besonderheiten und Merkmale dieser beiden Suchanfrage-Gruppen heraus und führen anschließend zu konkreten Empfehlungen für die Web-Suche und die zukünftige Forschung auf diesem Gebiet.

Zwar konnten die zuvor aufgestellten Aufgaben in der vorliegenden Arbeit gelöst werden, jedoch bleibt anzumerken, dass aufgrund diverser Einschränkungen der Untersuchungen, wie in Kapitel 5.1 aufgezeigt, den Ergebnissen nicht vorbehaltlos Repräsentativität zugesprochen werden kann. Ziel der Forschung war es jedoch auch die Repräsentativität der Ergebnisse im Forschungsbereich der Suchanfragen ohne Klick zu steigern. Dies konnte allerdings aus offensichtlichen Gründen, wie z.B. der fehlenden Zufallsstichprobe oder den teilweise zu geringen Fallzahlen, nicht verhindert werden. Die Ergebnisse der vorliegenden Arbeit können daher nicht als gesicherte Aussagen gewertet werden. Stattdessen müssen sie als Tendenzen und Hinweise genutzt werden, die für zukünftige Forschung auf diese Gebiet wichtige Anhaltspunkte liefern und von denen bereits erste Empfehlungen für Suchende und Betreiber der Web-Suchmaschinen abgeleitet werden können.

Quellen

AGOF - ANGEBOTSRANKING 2010

Arbeitsgemeinschaft für Online Forschung e.V.: *Angebotsranking*. - URL: http://www.agof.de/angebotsranking.619.html. - Abruf: 2010-08-21

ALBERS 2009

Albers, S. [Hrsg.]: *Methodik der empirischen Forschung*. 3., überarb. und erw. Aufl.. Wiesbaden : Gabler, 2009

AL-MASKARI et al. 2007

Al-Maskari, Azzah ; Sanderson, Mark ; Clough, Paul: The relationship between IR effectiveness measures and user satisfaction. In: *Proceedings of the ACM SIGIR International Conference on Research and Development in Information Retrieval* (2007), S. 773–774

BAILEY et al. 2007

Bailey, Peter ; Phan, Nina ; Wilkinson, Ross: Understanding the Relationship of Information Need Specificity to Search Query Length. In: *SIGIR'07,* July 23–27, 2007, Amsterdam, The Netherlands

BATES 2002

Bates, Marcia J.: After the Dot-Bomb: Getting Web Information Retrieval Right This Time. In: *First Monday* 7 (2002), Nr.7. - URL: http://131.193.153.231/www/issues/issue7_7/bates/. – Abruf: 2010-09-12

BRODER 2002

Broder, Andrei: A taxonomy of web search. In: *SIGIR Forum* 36 (2002), Nr. 2, S. 3-10

CALDERON-BENAVIDES et al. 2010

Calderon-Benavides, L. ; Gonzalez-Caro, C. ; Baeza-Yates, R. : In Towards a Deeper Understanding of the User's Query Intent. In: *SIGIR 2010 Workshop on Query Representation and Understanding,* Geneva : ACM, S. 21-24

CALLAN et al. 2007

Callan, J. ; Allan, J. ; Clarke, Ch.L.A. ; Dumais, S. ; Evans, D.A. ; Sanderson, M. ; Zhai, Ch.: Meeting of the MINDS: an information retrieval research agenda. In: *ACM SIGIR Forum*, 41 (2007), Nr. 2, S. 25-34

CHAO et al. 2009

Chao, Liu ; Guo, Fan ; Wang, Yi-Min: Efficient Multiple-Click Models in Web Search. In: *WSDM 2009 : Proceedings of the Second International Conference on Web Search and Web Data Mining*. New York : ACM, 2009, S. 124-131

COUPER und COUTTS 2006

Couper, Mick P. ; Coutts, Elisabeth: Online-Befragung: Probleme und Chancen verschiedener Arten von Online-Erhebungen. In: Diekmann, Andreas [Hrsg.]: *Methoden der Sozialforschung*. Kölner Zeitschrift für Soziologie und Sozialpsychologie 44 (2004) (Sonderheft), S. 217-243

CRASWELL et al. 2008

Craswell, Nick ; Zoeter, Onno ; Taylor, Michael ; Ramsey, Bill: An Experimental Comparison of Click Position-Bias Models. In: *WSDM '08 : Proceedings of the first ACM international conference on Web search and data mining*. New York : ACM, 2008, S. 87-94

CRASWELL und SZUMMER 2007

Craswell, Nick ; Szummer, Martin: Random Walks on the Click Graph. In: *SIGIR'07*, July 23–27, 2007, Amsterdam, The Netherlands, S. 1-8

CRYSTAL und GREENBERG 2006

Crystal, Abe ; Greenberg, Jane: Relevance Criteria Identified by Health Information Users During Web Searches. In: *Journal of the American Society for Information Science and Technology (JASIST*) 57 (2006), Nr. 10, S. 1-52

DIEKMANN 2010

Diekmann, Andreas: *Empirische Sozialforschung : Grundlagen, Methoden ,Anwendungen*. Orig.-Ausg., vollst. überarb. und erw. Neuausg. 2007, 21. Aufl.. Reinbek bei Hamburg : Rowohlt-Taschenbuch-Verl., 2010 (rowohlts enzyklopädie)

DOU et al. 2008

Dou, Zhicheng ; Song, Ruihua; Yuan, Xiaojie ; Wen, Ji-Rong: Are Click-through Data Adequate for Learning Web Search Rankings? In: *CIKM'08 : Proceeding of the 17th ACM conference on Information and knowledge management*. New York : ACM, 2008, S. 73-82

DOWNEY et al. 2008

Downey, Doug ; Dumais, Susan ; Liebling, Dan ; Horvitz, Eric: Understanding the Relationship between Searchers' Queries and Information Goals. In: *CIKM'08*, October 26–30, 2008, Napa Valley, California, USA, S. 449-458

DRORI 2000

Drori, Offer: The Benefits of Displaying Additional Internal Document Information on Textual Database Search Result Lists. In: *Proceedings of the fourth European Conference on Research and Advanced Technology for Digital Libraries - ECDL2000*. Springer (2000), S: 69-82. - URL: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.33.2883&rep=rep1&type=pdf. – Abruf: 2010-09-12

DUDEK et al. 2006

Dudek, Debra ; Mastora, Anna ; Landoni, Monica: Is Google the answer? A study into usability of search engines. In: *Library Review* 56 (2007), Nr. 3, S. 224-233

DUMAIS et al. 2001

Dumais, Susan ; Cutrell, Edward ; Chen, Hao: Optimizing Search by Showing Results In Context. In: *SIGCHI'01,* March 31-April 4, 2001, Seattle, WA, USA, S. 1-10

EISENBERG und HU 1987

Eisenberg, M.B. ; Hu, X.: Dichotomous relevance judgments and the evaluation of information systems. In: *ASIS '87 Proceedings of the 50th ASIS Annual Meeting, Medford, NJ: Learned Information*, S. 66-70

FOX et al. 2005

Fox, Steve ; Karnawat, Kuldeep ; Mydland,Mark ; Dumais, Susan ; White, Thomas: Evaluating Implicit Measures to Improve Web Search. In: *ACM Transactions on Information Systems* 23 (2005) Nr. 2, S. 147–168

GOLLAPUDI et al. 2008

Gollapudi, Sreenivas ; Das Sarma, Atish ; Ieong, Samuel: Bypass Rates: Reducing Query Abandonment using Negative Inferences. In: *KDD'08,* August 24–27, 2008, Las Vegas, Nevada, USA. S. 177-185

GOOGLE – SUCHFUNKTIONEN 2010

Google Inc.: *Suchfunktionen*. - URL: http://www.google.de/intl/de/help/features.html. – Abruf: 2010-07-23

GOOGLE INSIGHTS FOR SEARCH 2010

Google Inc.: *Google Insights for Search : Web-Suche Interesse Deutschland, Apr. 2010*. – URL: http://www.google.com/insights/search/#geo=DE&date=4%2F2010%201m&cmpt=q. – Abruf: 2010-07-23

GOOGLE WEB-SUCHE HILFE 2010

Google Inc.: *Hilfe für die einfache Suche : Grundlagen der Google Suche : Web-Suche Hilfe*. - URL: http://www.google.de/support/websearch/bin/answer.py?hl=de&answer=134479. - Abruf: 2010-07-24

GORDON und PATHAK 1999

Gordon, Michael ; Pathak, Praveen: Finding information on the World Wide Web: the retrieval effectiveness of search engines. In: *Information Processing & Management* 35 (1999), Nr.2, S. 141-180

GRANKA et al. 2005

Granka, Laura ; Joachims, Thorsten ; Pan, Bing ; Hembrooke, Helene ; Gay, Geri: Accurately Interpreting Clickthrough Data as Implicit Feedback. In: *SIGIR'05*, August 15–19, Salvador, Brazil, 2005, S. 154-161

GRANKA et al. 2008

Granka, Laura ; Lorigo, Lori ; Haridasan, Maya ; Brynjarsdottir, Hrönn ; Xia, Ling ; Joachims, Thorsten ; Gay, Geri ; Pellacini, Fabio ; Pan, Bing: Eye tracking and online search: Lessons learned and challenges ahead. In: *Journal of the AmericanSociety for Information Science and Technology* 59 (2008), S. 1041-1052

GRIESBAUM 2004

Griesbaum, Joachim: Evaluation of three German search engines: Altavista.de, Google.de and Lycos.de. In: *Information Research*, Vol. 9 (2004) Nr. 4. - URL: http://informationr.net/ir/9-4/paper189.html – Abruf: 2010-11-03

GRIESBAUM et al. 2002

Griesbaum, Joachim ; Rittberger, Marc ; Bekavac, Bernard: Deutsche Suchmaschinen im Vergleich: AltaVista.de, Fireball.de, Google.de und Lycos.de. In: Hammwöhner, Rainer ; Wolff, Christian ; Womser-Hacker, Christa (Hrsg.*): Information und Mobilität : Optimierung und Vermeidung von Mobilität durch Information : Proceedings des 8. Internationalen Symposiums für Informationswissenschaft*. Konstanz : UVK Verlagsgesellschaft mbH, 2002. S. 201 – 223

GUAN und CUTRELL 2007

Guan, Zhiwei ; Curtrell, Edward: An Eye Tracking Study of the Effect of Target Rank on Web Search. In: *CHI 2007 Proceedings*, April 28-May 3, 2007, San Jose, CA, USA. S. 417-420

HANSEN und KARLGREN 2005

Hansen, Preben ; Karlgren, Jussi: Effects of foreign language and task scenario on relevance assessment. In: *Journal of Documentation* 61 (2005), Nr. 5, S. 623 - 639

HAWKEY 2009

Hawkey, Kirstie: Privacy Concerns for Web Logging Data. In: Jansen, Bernard J. ; Spink, Amanda ; Taksa, Isak [Hrsg.]: *Handbook of Research on Web Log Analysis*. Hershey, Pa. [u.a.] : Information Science Reference, 2009, S. 80-98

HAWKING et al. 2001

Hawking, David ; Craswell, Nick; Bailey, Peter ; Griffiths, Kathleen: Measuring Search Engine Quality. In: *Information Retrieval* 4 (2001), Nr. 1, S. 33-59

HOTCHKISS et al. 2005

Hotchkiss, Gord ; Alston, Steve ; Edwards, Greg: *Eye Tracking Study* : *How Searchers See and Click on Google Search Results*. Enquiro Search Solutions Inc., 2005

HÜBENER 2009

Hübener, Markus: *Suchmaschinenoptimierung kompakt: Anwendungsorientierte Techniken für die Praxis*. Heidelberg [u.a.] : Springer, 2009

HUFFMAN und HOCHSTER 2007

Huffman, Scott B. ; Hochster, Michael: How Well does Result Relevance Predict Session Satisfaction? In: *Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval*. ACM: New York, 2007, S. 567-573

JANES 1991

Janes, J.W.: The binary nature of continuous relevance judgments: a study of users' perceptions. – In: *Journal of the American Society for Information Science* 42 (1991), S. 754-756

JANSEN et al. 2000

Jansen, Bernard J. ; Spink, Amanda ; Saracevic, Tefko: Real life, real users, and real needs: A study and analysis of user queries on the web. In: *Information Processing and Management* 36 (2000), Nr. 2, S. 207-227

JANSEN und BOOTH 2010

Jansen, Bernard J. ; Booth, Danielle: Classifying Web Queries by Topic and User Intent. In: *CHI 2010*, April 10–15, 2010, Atlanta, Georgia, USA

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 & Management* 42 (2006), S. 248-263

JOACHIMS 2002

Joachims, Thorsten: Optimizing search engines using clickthrough data. In: *Proceedings of the ACM Conference on Knowledge Discovery and Data Mining* (KDD), 2002, S. 1-10

JOACHIMS et al. 2007

Joachims, Thorsten ; Granka, Laura ; Pan, Bing ; Hembrooke, Helene ; Radlinski, Filip ; Gay, Geri: Evaluating the Accuracy of Implicit Feedback from Clicks and Query Reformulations in Web Search. In: *ACM Transactions on Information Systems*, 25 (2007), Nr. 2, S. 1-26

JOACHIMS und RADLINSKI 2007

Joachims, Thorsten, Radlinski, Filip: Search engines that learn from implicit feedback. In: *Computer* 40 (2007), S. 34-40

JUNG et al. 2007

Jung, Seikyung ; Herlocker, Jonathan L. ; Webster, Janet: Click data as implicit relevance feedback in web search. In*: Information Processing and Management* 43 (2007), S.791–807

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

KELLY und TEEVAN 2003

Kelly, Diane ; Teevan, Jaime: Implicit feedback for inferring user preference: a bibliography. In: *ACM SIGIR Forum* 37 (2003), Nr. 2, S. 18-28

KROMREY 2009

Kromrey, Helmut [Hrsg.]: *Empirische Sozialforschung : Modelle und Methoden der standardisierten Datenerhebung und Datenauswertung*. 12., überarb. und erg. Aufl.. Stuttgart : Lucius & Lucius, 2009

LEE et al. 2004

Lee, Uichin ; Liu, Zhenyu ; Cho, Junghoo: Automatic identification of user goals in web search. In: *UCLA Computer Science* (2004), S. 391- 400

LEWANDOWSKI 2005

Lewandowski, Dirk: *Web Information Retrieval : Technologien zur Informationssuche im Internet*. Frankfurt am Main : DGI, 2005 (Deutsche Gesellschaft für Informationswissenschaft und Informationspraxis: Reihe Informationswissenschaft der DGI, Bd. 7)

LEWANDOWSKI 2006a

Lewandowski, Dirk: Themen und Typen der Suchanfragen an deutsche Web-Suchmaschinen. In: Lehner, Franz; Nösekabel, H.; Keinschmidt, P. [Hrsg.]: *Multikonferenz Wirtschaftsinformatik 2006 (MKWI '06),* Proceedings 20. - 22. Februar 2006, Universität Passau. Berlin: Gito [Lecture Notes in Informatics], S. 33-43

LEWANDOWSKI 2006b

Lewandowski, Dirk: Aktualität als erfolgskritischer Faktor bei Suchmaschinen. In: *IWP - Information: Wissenschaft und Praxis 57* (2006), Nr. 3, S. 141-148. - URL: http://www.bui.haw-hamburg.de/fileadmin/user_upload/lewandowski/doc/Aktualitaet_erfolgskritischer_ Faktor_IWP2006.pdf. – Abruf: 2010-09-08

LEWANDOWSKI 2008a

Lewandowski, Dirk: The retrieval effectiveness of web search engines: considering results descriptions. In: *Journal of Documentation* 64 (2008), Nr. 6, S. 915-937

LEWANDOWSKI 2008b

Lewandowski, Dirk: A three-year study on the freshness of web search engine databases. In: *Journal of Information Science* 34 (2008), Nr.6, S. 817-831

LEWANDOWSKI 2009

Lewandowski, Dirk: T*he retrieval effectiveness of search engines on navigational queries.* Hamburg, Hochschule für Angewandte Wissenschaften Hamburg, Fakultät Design, Medien und Information, Department Information, 2008 [preprint of an article accepted for publication in ASLIB Proceedings 61(2009)]

LI et al. 2009

Li, Jane ; Huffman, Scott B. ; Tokuda, Akihito: Good Abandonment in Mobile and PC Internet Search. In: *SIGIR'09,* July 19–23, Boston, Massachusetts (2009), S. 43-50

LORIGO et al. 2006

Lorigo, Lori ; Pan, Bing ; Hembrooke, Helene ; Joachims, Thorsten ; Granka, Laura ; Gay, Geri: The influence of task and gender on search and evaluation behavior using Google. In: *Information Processing and Management* 42 (2006), S. 1123–1131

LU et al. 2006

Lu, Yumao ; Peng, Fuchun ; Li, Xin ; Ahmed, Nawaaz: Coupling feature selection and machine learning methods for navigational query identification. In: *CIKM'06 : Proceedings of the 15th ACM international conference on Information and knowledge management*. New York : ACM, 2006, S. 682-689

MACDONALD und OUNIS 2009

Macdonald, Craig ; Ounis, Iadh: Usefulness of Quality Click-through Data for Training. In: *WSCD '09: Proceedings of the 2009 workshop on Web Search Click Data*. New York: ACM, 2009, S. 1-5

REID 1999

Reid, Jane: A Task-Oriented Non-Interactive Evaluation Methodology for Information Retrieval Systems. In: *Information Retrieval* 2 (2000), Nr. 1, S. 115-129

ROSE 2004

Rose, Daniel E.: Why Is Web Search So Hard… to Evaluate? In: *Journal of Web Engineering* 3 (2004), Nr. 3 & 4, S. 171-181

ROSE und LEVINSON 2004

Rose, D.E. and Levinson, D.: Understanding user goals in web search. In: *WWW '04: Proceedings of the 13th international conference on World Wide Web* (2004), S. 13-19.

SCHMIDT-MÄNZ 2007

Schmidt-Mänz, Nadine: *Untersuchungen des Suchverhaltens im Web: Interaktion von Internetnutzern mit Suchmaschinen*. Hamburg : Dr. Kovac Verl., 2007

SHNEIDERMAN et al. 1998

Shneiderman, Ben, Byrd, Donald ; Croft, Bruce: Sorting out searching: A user-interface framework for text searches. In: *Communications of the ACM* 41 (1998), Nr. 4, S. 95-98

SONG et al. 2009

Song, Ruihua ; Luo, Zhenxiao ; Nie, Jian-Yun ; Yu, Yong ; Hon, Hsiao-Wuen: Identification of ambiguous queries in web search. In: *Information Processing and Management* 45 (2009), S. 216-229

SPINK et al. 2001

Spink, Amanda ; Wolfram, Dietmar ; Jansen, Bernhard J. ; Saracevic, Tefko: Searching the web : The public and their queries. In: *Journal of the American Society for Information Science and Technology* 52 (2001), Nr. 3, S. 226 – 234. - URL: http://comminfo.rutgers.edu/~tefko/JASIST2001.pdf. - Abruf: 2010-07-23

SPINK und JANSEN 2004

Spink, Amanda ; Jansen, Bernard J.: *Web Search: Public Searching of the Web*. Dordrecht [u.a.]: Kluwer Academic Publ., 2004 (Information science and knowledge management ; 6)

STAMOU und EFTHIMIADIS 2009

Stamou, Sofia ; Efthimiadis, Efthimis N.: Interpreting User Inactivity on Search Results**.** In: *Proceedings of the 32nd European Conference on IR Research (ECIR) on Advances in Information Retrieval*, Milton Keynes, UK (2010), S. 1-12

STOCK 2007

Stock, Wolfgang G.: Information Retrieval : *Informationen suchen und finden*. München : Oldenbourg, 2007 (Einführung in die Informationswissenschaft, Bd. 1)

STROHMAIER et al. 2007

Strohmaier, M.; Lux, M.; Granitzer, M.; Scheir, P.; Liaskos, S. ; Yu, E.: How do users express goals on the web? - An exploration of intentional structures in web search. In: *'We Know'07 International Workshop on Collaborative Knowledge Management for Web Information Systems, in conjunction with WISE'07*, Nancy, France (2007), S. 67-78

SUCHE T-ONLINE.DE 2010

Deutsche Telekom AG: *Suche t-online.de.* - URL: http://suche.t-online.de/ – Abruf: 2010-09-10

SURVEYMONKEY 2010

SurveyMonkey: Softwaretool für Onlineumfragen & -fragebögen. - URL: http://de.surveymonkey.com/ - Abruf: 2010-11-03

TAGUE-SUTCLIFFE 1992

Tague-Sutcliffe, Jean: The pragmatics of information retrieval experimentation, revisited. In: *Information Processing & Management* 28 (1992), Nr. 4, S. 467-490

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

THUROW und MUSICA 2009

Thurow, Shari ; Musica, Nick: *When search meets web usability*. Indianapolis : New Riders, 2009

TITTMANN 2008

Tittmann, Lukas: *Zur Struktur der Trefferbeschreibungen in Web-Suchmaschinen und deren Beurteilung durch den Nutzer*. Hamburg, Hochschule für Angewandte Wissenschaften, Studiendepartment Information, Dipl.-Arb., 2008

T-ONLINE.DE SUCHE 2010

Deutsche Telekom AG: *t-online.de Suche : Service- und Hilfeseiten*. – URL: http://themen.t-online.de/toi/html/de/Hilfe_Suche_FAQ5.htm. – Abruf 2010-11-24

VON MACH und OTTE 2009

Von Mach, Sonja ; Otte, Jessica: Identifikation von navigationsorientierten und kommerziellen Suchanfragen anhand einer Klickdatenanalyse. In: *HAW Abstracts in Information Science and Services* (haiss) 1 (2009), Nr. 1, S. 39-52. - URL: http://www.bui.haw-hamburg.de/fileadmin/haiss/haiss2009.pdf. – Abruf: 2010-11-03

WEBHITS 2010

WebHits: Web-Barometer. - URL: http://www.webhits.de/deutsch/index.shtml?webstats.html. – Abruf: 2010-08-21

Anhang I

Online-Befragung zur Ableitung des Informationsbedürfnisses

Willkommenstext

Herzlich Willkommen!

Es freut mich sehr, dass Sie sich an meiner Befragung beteiligen möchten!

Thema der Befragung ist es, die Ziele von Suchanfragen, die an Web - Suchmaschinen gestellt werden, näher zu erforschen.

Dazu wird Ihnen in dem folgenden Fragebogen jeweils eine Suchanfrage präsentiert. Dabei handelt es sich um Suchanfragen, die tatsächlich an eine Web-Suchmaschine gestellt wurden. Aus diesem Grund können Rechtschreibfehler enthalten sein.
Bitte schreiben Sie zu jeder Suchanfrage auf, welches typische Ziel Ihrer Meinung nach hinter dieser Suchanfrage steckt. Nutzen Sie bitte keine Hilfsmittel, wie z.B. eine Suchmaschine, um das Ziel der Suche abzuleiten. Sollte Ihnen eine Suchanfrage mehrdeutig erscheinen, können Sie auch ein mögliches zweites Ziel aufschreiben.

Dazu zwei Beispiele:
Suchanfrage: *download skype.*
Ein Ziel könnte lauten: Der Suchende hat das Ziel, das Programm „skype“ herunterzuladen.

Suchanfrage: *Historienfilme.*
Ein Ziel könnte lauten: Es wird eine Definition gesucht, die erklärt was ein Historienfilm ist.

Ein zweites Ziel könnte sein: Es wird ein Angebot von Historienfilmen gesucht, aus dem der Suchende einen Film auswählen und kaufen will.

Anschließend bitte ich Sie, die Aktualität des Themas der Suchanfragen auf einer Skala einzuschätzen.
Dabei wäre eine Anfrage wie z.B. *WM 2010* als sehr aktuell einzuschätzen, da es sich um ein gerade laufendes Ereignis handelt, zu dem momentan immer neue Informationen hinzukommen.
Die Suchanfrage *Definition Megabyte* wird hingegen als nicht aktuell eingeschätzt. Hier wird davon ausgegangen, dass es sich um ein Thema handelt, welches von aktuellen Ereignissen nicht beeinflusst wird und daher momentan nicht mit neuen Informationen ergänzt wird.

Zum Start der Befragung klicken Sie bitte auf „Weiter“.
Start der Umfrage [LINK]

Text des E-Mail Anschreibens

Betreff: Bitte um Teilnahme an einer Online-Befragung (Zeitaufwand: ca. 10 min.)

Guten Tag [*Name*],

für meine Masterarbeit in dem Studiengang Informationswissenschaft und -management an der HAW Hamburg führe ich eine Online-Befragung zum Thema Suchanfragen an Web-Suchmaschinen durch. Ich würde mich sehr freuen, wenn Sie daran teilnehmen. Die Beantwortung der Fragen dauert ca. 10 Minuten und wird selbstverständlich anonym durchgeführt. Um teilzunehmen, klicken Sie bitte den folgenden Link an:

[LINK]

Sollten Sie Rückfragen zur Befragung haben, beantworte ich diese gerne per E-Mail:
Jessica.Otte@haw-hamburg.de.

Mit freundlichen Grüßen,
Jessica Otte

8%

Herzlich Willkommen!

Es freut mich sehr, dass Sie sich an meiner Befragung beteiligen möchten!

Thema der Befragung ist es, die Ziele von Suchanfragen, die an Web-Suchmaschinen gestellt werden, zu erforschen.

Dazu wird Ihnen in dem folgenden Fragebogen jeweils eine Suchanfrage präsentiert. Dabei handelt es sich um Suchanfragen, die tatsächlich an eine Web-Suchmaschine gestellt wurden. Aus diesem Grund können Rechtschreibfehler enthalten sein.

Bitte schreiben Sie zu jeder Suchanfrage auf, welches typische Ziel Ihrer Meinung nach hinter dieser Suchanfrage steckt.
Nutzen Sie bitte keine Hilfsmittel, wie z.B. eine Suchmaschine, um das Ziel der Suche abzuleiten.
Sollte Ihnen eine Suchanfrage mehrdeutig erscheinen, können Sie auch ein mögliches zweites Ziel aufschreiben.

Dazu zwei Beispiele:
Suchanfrage: *download skype*.
Ein Ziel könnte lauten: Der Suchende hat das Ziel, das Programm „skype" herunterzuladen.

Suchanfrage: *Historienfilme*.
Ein Ziel könnte lauten: Es wird eine Definition gesucht, die erklärt was ein Historienfilm ist.
Ein zweites Ziel könnte sein: Es wird ein Angebot von Historienfilmen gesucht, aus dem der Suchende einen Film auswählen und kaufen will.

Bitte schätzen Sie anschließend die Aktualität des Themas der Suchanfrage auf einer Skala ein.
Dabei wäre eine Anfrage wie z.B. *WM 2010* als sehr aktuell einzuschätzen, da es sich um ein gerade laufendes Ereignis handelt, zu dem momentan immer neue Informationen hinzukommen.
Die Suchanfrage *Definition Megabyte* wird hingegen als nicht aktuell eingeschätzt. Hier wird davon ausgegangen, dass es sich um ein Thema handelt, welches von aktuellen Ereignissen nicht beeinflusst und daher momentan nicht mit neuen Informationen ergänzt wird.

Zum Start der Befragung klicken Sie bitte auf „Weiter".

Weiter

Abb. 21: Einleitungstext der Befragung zur Ableitung des Informationsbedürfnisses (Quelle: eigene Darst.)

17%

Suchanfrage: **ronald reagan**

Bitte schreiben Sie zu der Suchanfrage das Ihrer Meinung nach typische Ziel des Suchenden auf.
Sollte es ein zweites mögliches Ziel geben, schreiben Sie dies bitte im zweiten Feld auf.

1.
2.

Wie hoch schätzen Sie den Aktualitätsbezug der Suchanfrage ein?
Das Thema der Suchanfrage…

- … ist sehr aktuell.
- … ist aktuell.
- … ist weniger aktuell.
- … hat keinen aktuellen Bezug.
- Kann ich nicht einschätzen.

Zurück Weiter

Abb. 22: Zweite Seite der Befragung zur Ableitung des Informationsbedürfnisses (Quelle: eigene Darst.)

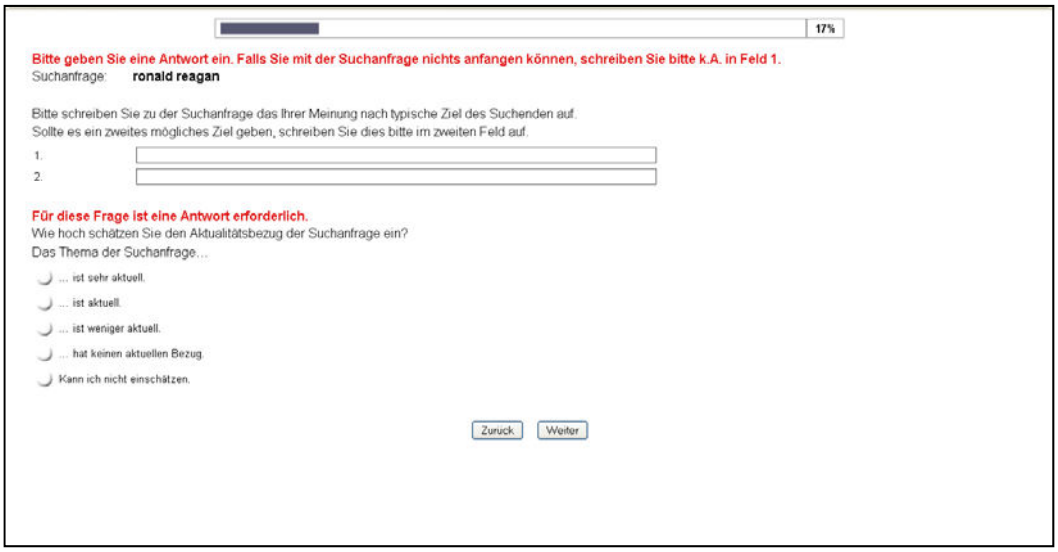

Abb. 23: Erste Seite der Befragung zur Ableitung des Informationsbedürfnisses mit Validierungshinweisen (Quelle: eigene Darst.)

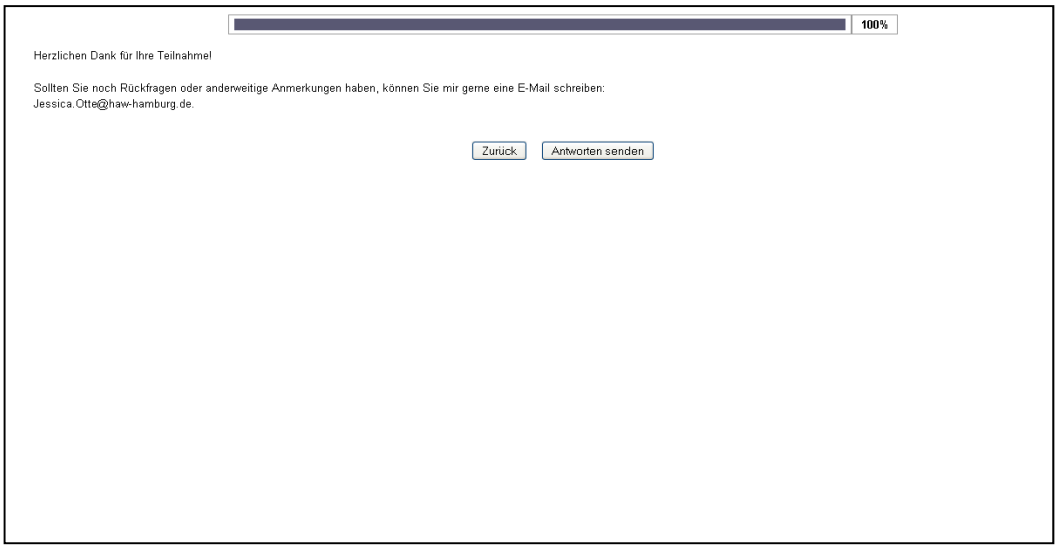

Abb. 24: Letzte Seite der Befragung zur Ableitung des Informationsbedürfnisses (Quelle: eigene Darst.)

Anhang II

Online-Befragung zur Untersuchung der Relevanz der Suchergebnisse nicht geklickter Suchanfragen

Willkommenstext

Herzlich Willkommen!

Es freut mich sehr, dass Sie sich an meiner Befragung, die ich im Rahmen meiner Masterarbeit durchführe, beteiligen!

Ziel der Untersuchung ist es, die Relevanz von Treffern einer Websuchmaschine zu bewerten.
Ihre Aufgabe ist es auf einer Skala anzugeben, für wie relevant sie einen Treffer halten. Eine detaillierte Beschreibung der Aufgabe lesen Sie bitte im folgenden Abschnitt mit dem Titel „Ihre Aufgabe“ nach.

Sollten Sie Rückfragen oder Anmerkungen zur Befragung haben, können Sie mir gerne eine E-Mail schreiben an: Jessica.Otte@haw-hamburg.de.

Vielen herzlichen Dank und beste Grüße,
Jessica Otte

Aufgabentext

Um die Relevanz der Treffer zu bewerten, werden im folgenden Fragebogen jeweils eine Suchanfrage und das dazugehörige Ziel der Suche aufgeführt.
Bitte beurteilen Sie die ersten zehn Treffer der Suche, diese werden im unteren Teil des Bildschirms einzeln in zufälliger Reihenfolge angezeigt. Bei der Beurteilung berücksichtigen Sie bitte:

1. Beurteilen Sie die Relevanz des Treffers, indem Sie bewerten, ob das Suchergebnis das Ziel der Suchanfrage erfüllt. Berücksichtigen Sie bitte jeden Treffer unabhängig von den folgenden bzw. vorherigen Treffern.

2. Nachdem Sie die Relevanz des Treffers beurteilt haben, teilen Sie bitte mit, ob Sie den Treffer anklicken würden oder ob die Beschreibung des Treffers ausreicht, um das Ziel der Suche zu erreichen.

Um Aufgabe 2 zu verdeutlichen, folgen zwei Beispiele:

Suchanfrage: Öffnungszeiten Uni Bibliothek Hamburg
Ziel: Die Öffnungszeiten der Universitätsbibliothek Hamburg erfahren.

Treffer:
Willkommen im Bibliothekssystem Universität Hamburg: Stabi ...
Die Bibliothek ist Mo-Fr von 9-21 Uhr, Sa-So von 10-21 Uhr geöffnet ... Berendsohn-Arbeitsstelle für Exilliteratur / Ossietzky-Lesesaal ...
www.sub.uni-hamburg.de

Die Trefferbeschreibung enthält bereits das Ziel der Suche (die Öffnungszeiten der Bibliothek), der Treffer muss daher nicht angeklickt werden.

Suchanfrage: Hamburg Veranstaltungen 2010

Ziel: Eine Übersicht über die Veranstaltungen, die 2010 in Hamburg stattfinden.

Treffer:

Veranstaltungskalender - Juli 2010 - Veranstaltungskalender ...
Veranstaltungen der Region Hamburg im Juli 2010: Finden Sie alle aktuellen Veranstaltungen und Termine in Hamburg und dem Alten Land, ebenso im Hamburger ...
www.hamburg.city-map.de/03010000

In diesem Fall ist das Ziel der Suche nicht im Treffer abzulesen, hier ist das Anklicken des Treffers notwendig, um an die gewünschten Informationen zu kommen.

Zum Start der Befragung klicken Sie bitte auf „Start“.

Text des E-Mail Anschreibens

Hallo,

für meine Masterarbeit untersuche ich Suchanfragen, die an Web-Suchmaschinen gestellt wurden. An der ersten Untersuchung, die ich zu diesem Thema bereits durchgeführt habe, hast Du dich ja eventuell schon beteiligt. Dafür danke ich Dir ganz herzlich!

Die folgende zweite Untersuchung ist eine Relevanzbewertung von Suchergebnissen einer Websuchmaschine. Ich würde mich sehr freuen, wenn Du auch an dieser Befragung teilnehmen würdest. Danach ist auch Schluss mit den Befragungen:-)

Die Beantwortung der Fragen dauert ca. 10 Minuten und wird selbstverständlich anonym durchgeführt. Um teilzunehmen, klick bitte den folgenden Link an und gib den Zugangscode ein. Damit du die zu bewertenden Treffer gleich im Blick hast, öffnest du die Umfrage am besten in einem maximierten Fenster.

http://www.electronicpassion.de/rat/
Zugangscode: lWL8ln

Vielen lieben Dank für Deine Unterstützung und viele Grüße,
Jessica

PS: Das Umfragetool wurde erst vor kurzem entwickelt, daher können leider noch kleinere Probleme auftreten. Wenn Dir etwas auffällt, kannst du mir gerne Bescheid geben. Danke!

Ansichten des Relevanztests

Relevance Assessment Tool

Hochschule für Angewandte Wissenschaften Hamburg
Hamburg University of Applied Sciences

Herzlich Willkommen:

Es freut mich sehr, dass Sie sich an meiner Befragung, die ich im Rahmen meiner Masterarbeit durchführe, beteiligen!

Ziel der Untersuchung ist es, die Relevanz von Treffern einer Websuchmaschine zu bewerten.
Ihre Aufgabe ist es auf einer Skala anzugeben, für wie relevant sie einen Treffer halten. Eine detaillierte Beschreibung der Aufgabe lesen Sie bitte im folgenden Abschnitt nach.

Sollten Sie Rückfragen oder Anmerkungen zur Befragung haben, können Sie mir gerne eine E-Mail schreiben an:
Jessica.Otte@haw-hamburg.de.

Vielen herzlichen Dank und beste Grüße,
Jessica Otte

Ihre Aufgabe:

Um die Relevanz der Treffer zu bewerten, werden im folgenden Fragebogen jeweils eine Suchanfrage und das dazugehörige Ziel der Suche aufgeführt.
Bitte beurteilen Sie die ersten zehn Treffer der Suche, diese werden im unteren Teil des Bildschirms einzeln in zufälliger Reihenfolge angezeigt. Bei der Beurteilung berücksichtigen Sie bitte:

1. Beurteilen Sie die Relevanz des Treffers, indem Sie bewerten, ob das Suchergebnis das Ziel der Suchanfrage erfüllt. Berücksichtigen Sie bitte jeden Treffer unabhängig von den folgenden bzw. vorherigen Treffern.
2. Nachdem Sie die Relevanz des Treffers beurteilt haben, teilen Sie bitte mit, ob Sie den Treffer anklicken wurden oder ob die Beschreibung des Treffers ausreicht, um das Ziel der Suche zu erreichen.

Um Aufgabe 2 zu verdeutlichen, folgen zwei Beispiele:

Suchanfrage: Öffnungszeiten Uni Bibliothek Hamburg
Ziel: Die Öffnungszeiten der Universitätsbibliothek Hamburg erfahren.

Treffer:
Willkommen im Bibliothekssystem Universität Hamburg: Stabi ...
Die Bibliothek ist Mo-Fr von 9-21 Uhr, Sa-So von 10-21 Uhr geöffnet ... Berendsohn- Arbeitsstelle für Exilliteratur / Ossietzky-Lesesaal ...
www.sub.uni-hamburg.de

Die Trefferbeschreibung enthält bereits das Ziel der Suche (die Öffnungszeiten der Bibliothek), der Treffer muss daher nicht angeklickt werden.

Suchanfrage: Hamburg Veranstaltungen 2010
Ziel: Eine Übersicht über die Veranstaltungen, die 2010 in Hamburg stattfinden.

Treffer:
Veranstaltungskalender - Juli 2010 - Veranstaltungskalender ...
Veranstaltungen der Region Hamburg im Juli 2010: Finden Sie alle aktuellen Veranstaltungen und Termine in Hamburg und dem Alten Land, ebenso im Hamburger ...
www.hamburg.city-map.de/03010000

In diesem Fall ist das Ziel der Suche nicht im Treffer abzulesen, hier ist das Anklicken des Treffers notwendig, um an die gewünschten Informationen zu kommen.

Start

Abb. 25: Erste Seite des Relevanztests mit Willkommens- und Aufgabentext (Quelle: eigene Darst.)

Abb. 26: Zweite Seite des Relevanztests: Ziel der Suche wird dem Probanden genannt (Quelle: eigene Darst.)

Abb. 27: Dritte Seite des Relevanztests: Treffer und Skalen zur Beurteilung des Treffers werden dem Probanden präsentiert (Quelle: eigene Darst.)

Abb. 28: Seite nach Abschluss der Beurteilung einer Suchanfrage: Zwischenstand wird dem Probanden kommuniziert (Quelle: eigene Darst.)

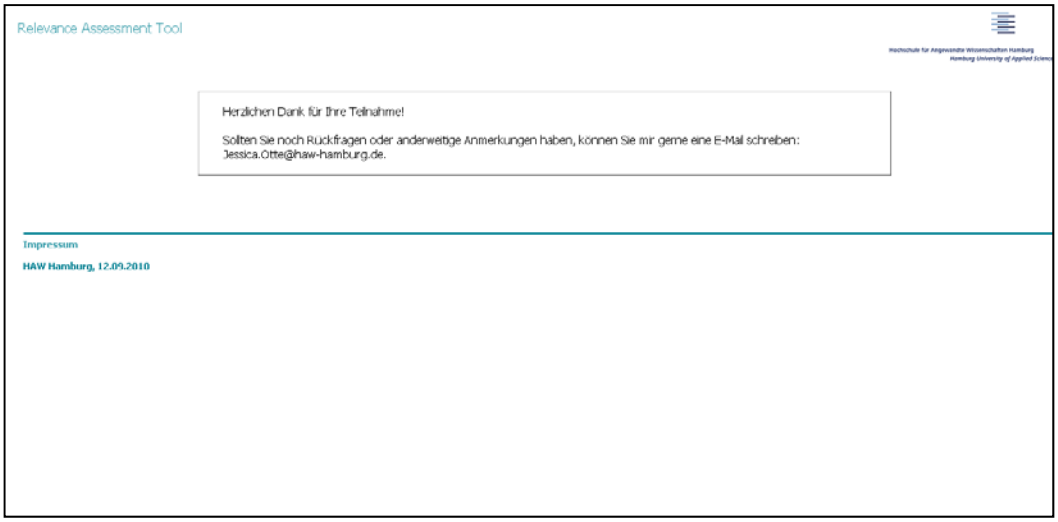

Abb. 29: Letzte Seite der Befragung: Dank an den Probanden (Quelle: eigene Darst.)

Anhang III

Eidesstattliche Erklärung

Ich versichere, die vorliegende Arbeit selbstständig 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,

Jessica Drechsler geb. Otte