

# Vermitteln Suchmaschinen vollständige Bilder aktueller Themen?

## Untersuchung der Gewichtung inhaltlicher Aspekte von Suchmaschinenergebnissen in Deutschland und den USA

Markus Günther\*

*Zielsetzung* — Vor dem Hintergrund von Suchmaschinenverzerrungen sollte herausgefunden werden, ob sich die von Google und Bing vermittelten Bilder aktueller internationaler Themen in Deutschland und den USA hinsichtlich (1) Vollständigkeit, (2) Abdeckung und (3) Gewichtung der jeweiligen inhaltlichen Aspekte unterscheiden.

*Forschungsmethoden* — Für die empirische Untersuchung wurde eine Methode aus Ansätzen der empirischen Sozialwissenschaften (Inhaltsanalyse) und der Informationswissenschaft (Retrievaltests) entwickelt und angewandt.

*Ergebnisse* — Es zeigte sich, dass Google und Bing in Deutschland und den USA (1) keine vollständigen Bilder aktueller internationaler Themen vermitteln, dass sie (2) auf den ersten Trefferpositionen nicht die drei wichtigsten inhaltlichen Aspekte abdecken und dass es (3) bei der Gewichtung der inhaltlichen Aspekte keine signifikanten Unterschiede gibt. Allerdings erfahren diese Ergebnisse Einschränkungen durch die Methodik und die Auswertung der empirischen Untersuchung.

*Schlussfolgerungen* — Es scheinen tatsächlich inhaltliche Suchmaschinenverzerrungen vorzuliegen – diese könnten Auswirkungen auf die Meinungsbildung der Suchmaschinennutzer haben. Trotz großem Aufwand bei manueller, und qualitativ schlechteren Ergebnissen bei automatischer Untersuchung sollte dieses Thema weiter erforscht werden.

*Schlagwörter* — Suchmaschinen, Suchmaschinenverzerrungen, Google, Bing, Inhaltsanalyse, Retrievaltests

### Do search engines provide complete pictures of current topics? Weighting analysis of content aspects of search engine results in Germany and the USA

*Objective* — The goal was to identify potential search engine content bias by comparing pictures of selected current and international events, from Google and Bing across Germany and the US. Criteria were developed for (1) completeness, (2) coverage, and (3) weighting of the particular content aspects.

*Methods* — Empirical analysis was conducted using a hybrid / cross-functional methodology taken from the social sciences (content analysis) and information science (retrieval tests).

*Results* — Both Google and Bing (1) do not provide complete coverage in neither Germany nor the US, (2) both do not cover the three most important content aspects on the first three result positions, and (3) there do not seem to be significant differences regarding the weighting of the content aspects. However, these results are somewhat limited, due to the methodology and the evaluation.

*Conclusions* — The findings indicate that search engine content bias indeed exists. This could have an effect on how public opinions are formed. The topic should be further analyzed, especially in the light of effort required for obtaining results of high quality.

*Keywords* — search engines, search engine bias, Google, Bing, content analysis, retrieval tests

\* Markus Günther, M.A. | Hamburg | m\_guenther1@gmx.de | ORCID: [orcid.org/0000-0001-8548-1613](https://orcid.org/0000-0001-8548-1613)



Diesem Beitrag liegt folgende Abschlussarbeit zugrunde / This article is based upon the following dissertation/thesis:  
 Günther, Markus: *Welches Weltbild vermitteln Suchmaschinen? Untersuchung der Gewichtung inhaltlicher Aspekte von Google- und Bing-Ergebnissen in Deutschland und den USA zu aktuellen internationalen Themen*. Masterarbeit (M.A.), Hochschule für Angewandte Wissenschaften Hamburg, 2015. Volltext: <http://edoc.sub.uni-hamburg.de/haw/volltexte/2016/3329/>

## 1 Suchmaschinenverzerrungen

Nach Lewandowski (2015, S. 272) besteht das ideale Ergebnis einer Suchmaschine zu einer Anfrage aus sämtlichen potenziell relevanten Dokumenten in einer allein nach objektiven Kriterien erfolgten und damit richtigen Reihung. Die Unterschiede eines tatsächlichen Suchergebnisses zu diesem Ideal werden als *Suchmaschinenverzerrungen* (*search engine bias*) bezeichnet. Allerdings kann es keine »neutrale« Suchmaschine geben, da jede Suchmaschine bereits aufgrund ihrer Funktionsweise bestimmten Verzerrungen ausgesetzt ist (vgl. auch Weber 2011, S. 277)

Weber (ebd., S. 278–281) teilt die Ursachen von Suchmaschinenverzerrungen in die drei Bereiche Implementierung, Anbieterverhalten und Nutzerverhalten ein. Dabei bezieht sich Implementierung auf die Indexierungs- und Suchalgorithmen von Suchmaschinen, auf rechtliche Vorgaben und auf mögliche Einflussnahmen seitens Suchmaschinenbetreibern. Unter dem Verhalten von Inhalte-Anbietern versteht Weber die *Suchmaschinenoptimierung*. Mit Nutzerverhalten ist gemeint, dass Nutzer die Möglichkeiten von Suchmaschinen oft nicht ausreizen.

Nach Lewandowski (2014, S. 233) werden Suchmaschinenverzerrungen v.a. durch die Verbindung folgender Faktoren zum Problem:

1. die Dominanz des Modells algorithmische Web-Suchmaschine gegenüber anderen Methoden zum Auffinden von Informationen im Web,
2. die Dominanz von Google in diesem Bereich,
3. das Verhalten der Suchmaschinennutzer (kurze Suchanfragen, kaum systematische Ergebnissichtung, geringe Kenntnisse über Suchmaschinen).

Würden Nutzer viele Suchmaschinen verwenden, so würden sich ihre spezifischen Verzerrungen gewissermaßen gegenseitig ausgleichen (außerdem wären systematische Verzerrungen leichter zu entdecken) (vgl. Lewandowski 2015, S. 273).

Nach Weber (2011, S. 282 f.) gibt es einige Gegenstrategien zu Suchmaschinenverzerrungen.

So sollten Alternativen genutzt werden. Dies betrifft nicht nur die Konkurrenz zwischen Google und Bing, sondern insbesondere die Nutzung bereichsspezifischer Suchmaschinen bspw. von Bibliotheken, Verlagen und anderen Informationsanbietern.

Es kann prinzipiell keine unverzerrten Suchergebnisse geben – daher ist es wichtig, sich darüber bewusst zu sein, dass Ergebnisse vielfältigen Einflüssen ausgesetzt sind. In Hinblick auf unterschiedliche Positionierungen lohnt es sich z.B., nicht nur die erste Seite der Ergebnisliste anzusehen.

Viele Nutzer neigen dazu, nur der eigenen Meinung entsprechende Inhalte wahrzunehmen und damit *kognitive Dissonanz* zu vermeiden (vgl. z. B. Weber und Drüeke 2010). Suchmaschinen können diesen Effekt verstärken – doch lässt er sich leicht umgehen, wenn man sich seiner bewusst ist.

Weiter sollte ggf. zusätzlich in anderen Sprachen, etwa Englisch, gesucht werden.

Bereits im Jahr 2000 wiesen Introna und Nissenbaum (2000, 169 f., 175–178, 181 f.) auf Suchmaschinenverzerrungen hin und kritisierten, dass diese das Funktionieren des Webs innerhalb der Gesellschaft einschränken und sowohl seiner grundlegenden Architektur als auch den Werten und Idealen, die es erst zu einem solchen Phänomen hatten werden lassen, zuwiderlaufen würden. Somit würden durch Verzerrungen neben technischen auch politische Themen aufgeworfen werden. Sie warnten davor, die Funktionsweise von Suchmaschinen von den klassischen Marktmechanismen determinieren zu lassen und forderten:

1. die Aufdeckung der Suchmaschinenalgorithmen zum *Indexing*, *Searching* und *Ranking* und
2. eine öffentliche Aufsicht des Suchmaschinenbetriebs.

Zehn Jahre später blickt Granka (2010, S. 364 f.) auf Intronas und Nissenbaums Arbeit zurück und zieht Bilanz. Tatsächlich hatten sich in der Zwischenzeit viele Forscher mit den darin aufgeworfenen Themen

beschäftigt. Dennoch hatten sich Suchmaschinen größtenteils ohne Intronas und Nissenbaums Forderungen entwickelt; sie richteten sich nach einem Online-Informationsmarkt, der von Wettbewerb und Konsumentenverhalten dominiert wird, die Algorithmen wurden nicht im Detail offengelegt und eine öffentliche Aufsicht existiert nur bedingt. Daher will Granka überprüfen, wie diese Themen in der neueren Literatur behandelt werden.

Bzgl. ihrer Algorithmen haben Suchmaschinenbetreiber nach Granka (ebd., 365 f., 368 f.) grundlegende Funktionsweisen bekanntgegeben, halten die Gewichtungen jedoch aus zwei Gründen geheim:

1. Vollständige öffentliche Algorithmenkenntnis würde zu einem erheblichen Spam-Problem führen, das den Vorteil für die Nutzer überwiegen würde.
2. Es gäbe keinen Wettbewerb mehr: Alle Suchmaschinenbetreiber würden dieselben Algorithmen verwenden, d.h. das Ranking wäre überall gleich, was eine erheblich geringere Informationsauswahl für die Nutzer zur Folge hätte.

Introna und Nissenbaum gingen davon aus, dass durch ihre Forderungen Websitebetreiber mehr Macht erhalten und die Nutzer die Suche besser verstehen würden. Es ist jedoch fraglich, ob den Nutzern mit vollständiger Kenntnis der Algorithmen viel mehr als mit dem Wissen über ihre Prinzipien geholfen wäre. Tatsächlich hilfreich wäre ein teilweises Aufdecken, sodass die Nutzer sehen könnten, falls eine Suchmaschine von ihren organischen Algorithmen abweicht.

Nach Granka (ebd., S. 369 ff.) gingen Introna und Nissenbaum bei ihrer Kritik an einem Suchmarkt

bzgl. der Auswahl von Suchmaschinen von einem zu simplen Nutzerverhalten aus – z.B. zeigt Eyetracking, dass Nutzer sich diesbezüglich unterschiedlich verhalten. Das fördert die Entwicklung eines stabilen Online-Informationsmarktes, v.a. eines solchen, auf dem die besten Suchmaschinen am meisten besucht werden. Wenn sich Suchmaschinenbetreiber also ethisch am Nutzerverhalten orientieren, bedarf es keiner Regulierung.

Dass sich Muster von Mediendominanz und -eigentum aus dem Offline-Bereich (»rich-get-richer-Effekte«) auch im Online-Bereich zeigen, hat die Forschung bestätigt: Websitebetreiber mit mehr Budget können bessere Sites realisieren und ziehen damit mehr Nutzer an. Allerdings wird in diesem Forschungsbereich i.d.R. mit aggregierten Skalen gearbeitet, die auf das Massenpublikum eingehen und nicht auf *Long-Tail*-Suchanfragen (also spezifischere).

Weitere Argumente gegen einen Suchmarkt sind die niedrigen Ausprägungen von Wettbewerb und Auswahl. Dies trifft allerdings eher hypothetisch zu, denn wenn Nutzer unzufrieden sind, wechseln sie die Suchmaschine, und mehr als die Hälfte nutzt mehr als eine.

So vielfältig wie Suchmaschinenverzerrungen sind auch die Untersuchungen zu ihnen, siehe z.B. Joachims et al. 2007, Brown et al. 2008, Yue et al. 2010, Jeong et al. 2012 und White 2013 (White und Horvitz 2015).

Inhaltliche Suchmaschinenverzerrungen sind allerdings noch vergleichsweise wenig erforscht – diese Untersuchung bietet einen Ansatz zum Nachweis überregionaler Verzerrungen.

## 2 Ziele der empirischen Untersuchung

Suchmaschinen haben eine enorme gesellschaftliche Bedeutung (vgl. Lewandowski 2015, S. 2) und sind der beliebteste Dienst des Internets; von 82 Prozent aller deutschen Internetnutzer werden sie mindestens wöchentlich genutzt (vgl. van Eimeren und Frees 2014, S. 386 f.; auch AGOF 2015, S. 11). In anderen Ländern, etwa den USA, sieht es ähnlich aus (vgl. Purcell et al. 2012, S. 3, 5–8). Weltweit werden jeden Monat mehr als 175 Mrd. Anfragen an Suchmaschinen gestellt; mehr als 5,8 Mrd. pro Tag bzw. mehr als 2 Mio. pro Sekunde (vgl. Lewandowski 2015, S. 3).

49 Prozent der Deutsch sprechenden Onliner suchen ebenfalls mindestens wöchentlich nach aktuellen Nachrichten (vgl. van Eimeren und Frees 2014, S. 386 f.); dabei sind Suchmaschinen für 42 Prozent der wichtigste Zugangsweg (vgl. Hölzig und Hasebrink 2014, S. 535). Bereits 2003 wurde geschätzt, dass das Internet neben den klassischen Massenmedien Print, Hörfunk und Fernsehen »seine Funktion als viertes tagesaktuelles Medium gefunden hat« (van Eimeren, Gerhard et al. 2003, S. 358). Im Zeitraum 2009 bis 2013 hat das Internet insgesamt insbesondere zulas-

ten von Fernsehen und Tageszeitung an Bedeutung gewonnen (vgl. TNS Infratest 2013, 8 f., 30 f.).

Allerdings sind die Ergebnisse von Suchmaschinen nie »neutral«, sondern immer in irgendeiner Weise verzerrt (vgl. Abschnitt 1).

Angesichts der massenhaften Nutzung von Suchmaschinen, u.a. eben, um sich über das Tagesgeschehen zu informieren und von Suchmaschinenverzerrungen, stellt sich die Frage, was für Bilder Suchmaschinen eigentlich von aktuellen (also »neuen«) Themen vermitteln – genauer, ob und ggf. inwiefern inhaltliche Verzerrungen bestehen. Diese Fragestellung wurde auf unterschiedliche, aber kulturell ähnliche Regionen ausgeweitet: Wie unterscheiden sich diese Bilder etwa in Deutschland und den USA?

Die Forschungsfragen und -hypothesen der empirischen Untersuchung beziehen sich auf drei Themenbereiche: die Vollständigkeit, die Abdeckung und die Gewichtung inhaltlicher Aspekte von Suchergebnissen.

#### 1. Vollständigkeit:

F1.1: Vermitteln Google und Bing in Deutschland und den USA in Hinsicht auf die wichtigen inhaltlichen Aspekte vollständige Bilder aktueller internationaler Themen?

H1.1: Google und Bing vermitteln in Deutschland und den USA in Hinsicht auf die wichtigen inhaltlichen Aspekte in mindestens 75 Prozent der Fälle vollständige Bilder aktueller internationaler Themen.

F1.2: Gibt es bei den durch Google und Bing vermittelten Bildern aktueller internationaler Themen in Deutschland und den USA in Hinsicht auf die wichtigen inhaltlichen Aspekte Unterschiede in puncto Vollständigkeit?

H1.2: Durch die Ähnlichkeit der Suchmaschinen und der Kultur- und Sprachräume gibt es bei den durch Google und Bing vermittelten Bildern aktueller internationaler Themen in Deutschland und den USA in Hinsicht auf die wichtigen inhaltlichen Aspekte in puncto Vollständigkeit in mindestens 75 Prozent der Fälle keine Unterschiede.

#### 2. Abdeckung:

F2: Werden von Google und Bing in Deutschland und den USA bereits auf den ersten Trefferpositionen die wichtigsten inhaltlichen Aspekte aktueller internationaler Themen abgedeckt?

H2.1: Auf Trefferposition 1 werden von Google und Bing in Deutschland und den USA die drei wichtigsten inhaltlichen Aspekte aktueller internationaler Themen abgedeckt.

H2.2: Bei Berücksichtigung der ersten drei Trefferpositionen werden von Google und Bing in Deutschland und den USA mindestens 75 Prozent der wichtigsten inhaltlichen Aspekte aktueller internationaler Themen abgedeckt.

#### 3. Gewichtung:

F3: Gibt es bei den durch Google und Bing vermittelten Bildern aktueller internationaler Themen in Deutschland und den USA signifikante Unterschiede bei der Gewichtung der wichtigen inhaltlichen Aspekte?

H3: Wegen der Ähnlichkeit der Suchmaschinen und der Kultur- und Sprachräume gibt es bei der Gewichtung der wichtigen inhaltlichen Aspekte durch Google und Bing vermittelter Bilder aktueller internationaler Themen in Deutschland und den USA keine signifikanten Unterschiede.

## 3 Methodik der empirischen Untersuchung

Die Methodik stellt eine Kombination von Ansätzen der empirischen Sozialwissenschaften und der Informationswissenschaft, nämlich von Inhaltsanalyse und Retrievaltests, dar. Im Folgenden wird daher zuerst knapp dieser Hintergrund behandelt. Danach wird das grundlegende Vorgehen erläutert. Abschließend wird die Methodik anhand einer Einzelauswertung und korrespondierend zu den Forschungsfragen und -hypothesen (vgl. Abschnitt 2) demonstriert.

### 3.1 Hintergrund

Früh definiert die *Inhaltsanalyse* als »eine empirische Methode zur systematischen und intersubjektiv nachvollziehbaren Beschreibung inhaltlicher und formaler Merkmale von Mitteilungen, meist mit dem Ziel einer darauf gestützten interpretativen Inferenz auf mitteilungsexterne Sachverhalte« (2011, S. 133).

Nach diesem Autor besteht der Ablauf einer Inhaltsanalyse standardisiert aus vier Phasen.

In der Planungsphase wird das Untersuchungsziel bestimmt, es werden theorie- und empiriegeleitet Hypothesen gebildet, auf dieser Basis wird die Grundgesamtheit des Untersuchungsmaterials (z.B. Zeitungen) bestimmt und ggf. eine Stichprobe gezogen (ebd., S. 147–153). In der Entwicklungsphase werden Kategorien gebildet – diese geben an, auf welche Merkmale der untersuchten Mitteilungen sich die Analyse beziehen soll und welche konkreten Texteinheiten als ähnlich betrachtet werden (ebd., S. 153–163). In der Testphase werden Codierungen zur Probe und zu Reliabilitäts- und Validitätstests durchgeführt (ebd., S. 163–198). In der Anwendungsphase wird das Textmaterial systematisch oder ggf. nach Zufallskriterien auf die Codierer verteilt, damit sich eventuelle Codierer-Verzerrungen neutralisieren (ebd., S. 198–207). Bei der Auswertung werden die erhobenen Daten mit Hilfe statistischer Auswertungsmethoden derart weiterverarbeitet, dass Schlussfolgerungen im Sinne der Hypothesen möglich sind.

Nach Lewandowski (2015, S. 215–221; ausführlicher s. Lewandowski 2011) sind *Retrievaltests* eine Methode mit langer Tradition, um die *Retrieval-effektivität* von Information-Retrieval-Systemen zu messen. Mit dem Aufkommen von Suchmaschinen wurde sie für diese angepasst.

Mit Retrievaleffektivität ist die Fähigkeit einer Suchmaschine gemeint, auf Anfragen relevante Dokumente auszugeben. Sie wird meist durch die *Precision* bestimmt, die angibt, wie gut es einer Suchmaschine gelingt, ausschließlich relevante Ergebnisse auszugeben. Sozusagen das Gegenstück ist der *Recall*, der den Anteil der gefundenen relevanten Dokumente an der Gesamtzahl der im Datenbestand vorhandenen relevanten Dokumente misst – seine Bestimmung ist in der Praxis allerdings kaum möglich.

Für die meisten Tests wird ein (z.T. leicht modifizierter) Standardaufbau verwendet, wie er aus der Information-Retrieval-Literatur und den Evaluierungsinitiativen (v.a. TREC (*Text REtrieval Conference*)) (vgl. Harman und Voorhees 2006)) bekannt ist: Eine bestimmte Anzahl von Suchanfragen wird an unterschiedliche Suchmaschinen gesendet, von denen jeweils eine bestimmte Anzahl Ergebnisse anonymisiert, gemischt und dann Juroren zur Relevanzbeurteilung vorgelegt wird. Danach werden die Ergebnisse wieder ihren Systemen und Treffer-

positionen zugeordnet. I.d.R. wird ihre Precision gemessen.

Der Aufbau von Retrievaltests orientiert sich meistens an den Schritten von Tague-Sutcliffe (1992). Die Spezifika von Suchmaschinen haben Gordon und Pathak (1999) und (darauf aufbauend) Hawking et al. (2001) berücksichtigt. Das typische Retrievaltest-Design besteht aus folgenden Schritten:

1. Auswahl von Suchanfragen/-aufgaben,
2. Senden der Suchanfragen an die Suchmaschinen,
3. Erfassung und Speicherung der Ergebnisse,
4. Mischen der Ergebnisse; Unkenntlichmachung ihrer Herkunft,
5. Bewertung der Ergebnisse durch Juroren,
6. Zusammenführung der Bewertungen und Zuordnung zu den Suchmaschinen,
7. Auswertung der Ergebnisse.

### 3.2 Vorgehen

Die Grundidee der Untersuchung war es, aktuelle internationale Themen aus der »analogen Welt« anhand aktueller Zeitungen zu identifizieren und in der »digitalen Welt« in Form von Suchergebnissen auf inhaltliche Verzerrungen hin zu untersuchen.

Bei der Themenfindung wurden jeweils die drei auflagenstärksten Tageszeitungen aus Deutschland und den USA – *Bild*, *Süddeutsche Zeitung* und *Frankfurter Allgemeine Zeitung* (FAZ) bzw. *USA Today*, *Wall Street Journal* (WSJ) und *New York Times* (vgl. Statista 2015d bzw. 2015a) – vom 28. April 2015 in Hinsicht auf gemeinsame Themen gesichtet.

So wurden 14 Themen identifiziert. Tabelle 1 zeigt sie und ihr jeweiliges Vorkommen.

Zu diesen Themen wurden jeweils eine deutsche und eine englische Suchanfrage formuliert, z.B. ‚varoufakis entmachtung / varoufakis disempowerment‘, ‚israel luftangriff syrien grenze / israel airstrike syria border‘ und ‚ostukraine kämpfe mariupol / east ukraine fights mariupol‘.

Die zu untersuchenden Suchmaschinen sollten unabhängig und bedeutend sein, um dem Informations-Massenkonsum möglichst nahe zu kommen.

Nach Statista (2015c bzw. 2015b) wird sowohl in Deutschland als auch in den USA der Suchmaschinenmarkt deutlich von Google dominiert – stärkster Konkurrent ist in beiden Regionen Bing, auch wenn diese Suchmaschine (in Deutschland selbst mit ihrem Partner Yahoo) keine mit Google vergleich-

Nr.	Thema	Zeitungen vom 28.04.2015:					
		BILD	Süddt.	FAZ	USA Today	WSJ	NY Times
1	neuer Sparkurs der Deutschen Bank	x	x	x		x	x
2	Zügelung Varoufakis' nach Vortrag in Riga	x	x	x		x	x
3	Erdbeben Nepal	x	x	x		x	x
4	Geld, das an Hillary Clinton ging		x		x	x	
5	Rolle Großbritanniens hinsichtlich EU und Welt		x	x			x
6	Proteste in Burundi gegen Präsident Nkurunziza		x			x	x
7	Luftangriffe Israels im syrischen Grenzgebiet			x		x	
8	Zusammenschluss der USA und Japans zur besseren Verteidigung			x		x	x
9	Rücktritt des Ministerpräsidenten Südkoreas Lee Wan-Koo			x		x	
10	40. Jährung des Vietnam-Krieg-Endes			x		x	
11	Entwicklungshilfe- und Migrantenproblematik Afrikas			x			x
12	Hitze und Regen durch Klimawandel		x				x
13	Ablehnung des Übernahmeangebots von Teva bzgl. Mylan		x				x
14	Erneut Kämpfe in der Ost-Ukraine; Mariupol			x			x

Tabelle 1: Gemeinsame Themen deutscher und amerikanischer Zeitungen

Nr.	Aspekt	knapp	ausführlich
I	Der Luftangriff	Durch einen isr. Luftangriff wurden mutmaßliche Terroristen getötet, die eine Bombe im isr.-syr. Grenzgebiet platzieren wollten	Am Sonntag(abend) näherten sich in der Nähe von Madschal Schams im von Israel besetzten Teil der Golan-Höhen aus Richtung Syrien 3/4 bewaffnete Männer mit einer Bombe / einem Sprengsatz dem Grenzzaun zu Israel; Sie befanden sich auf isr. Territorium, aber auf syrischer Seite des Grenzzauns; Dort wurden sie von einer Patrouille der Isreal Defense Forces (IDF) gesichtet; Anscheinend wollten die Männer die Bombe / den Sprengsatz am Grenzzaun verstecken/platzieren/deponieren & so auf isr. Truppen einen Anschlag verüben; Durch einen Luftangriff mit mehreren Kampfflugzeugen des isr. Militärs wurden die mutmaßlichen Terroristen/Attentäter/Extremisten getötet & der mögliche Anschlag verhindert; Es ist von 3/4 Toten die Rede; Isr. Medien berichteten, dass 2 der 4 Männer Söhne eines früheren polit. Gefangenen aus dem Drusen-Dorf Magdal Shams auf Israels Seite der Golanhöhen waren, die Familie verließ Syrien in den 1980ern
II	Reaktionen zum Luftangriff von isr. Seite	Netanjahu lobte den Einsatz	Israel machte keine Angaben zu Opfern; Netanjahu lobte den Einsatz / die am Luftangriff beteiligten Truppen für ihre »Wachsamkeit« / ihr »schnelles & präzises« Eingreifen; Er twitterte: »Jeder Versuch, unseren Soldaten & Zivilisten zu schaden, wird mit einer entschiedenen Reaktion beantwortet wie die Militäraktion heute Abend, die eine Terrorattacke vereitelte«; Ein IDF-Sprecher sagte, man werde sich gegen alle Angriffe auf Israel wehren
III	Reaktionen zum Luftangriff von syr. Seite	Es übernahm niemand die Verantwortung für den Angriff	Niemand übernahm die Verantwortung für die Attacke aus Syrien; Es wird die Hisbollah verdächtigt; Syrische Staatsmedien berichteten zunächst nicht über den Vorfall; Das syr. Observatorium für Menschenrechte meldete die Tötung 4 syr. Soldaten durch eine vom isr. besetzten Teil der Golanhöhen abgefeuerten Rakete, ob das Geschoss von einem Flugzeug oder Fahrzeug kam, sei unklar

Tabelle 2: Thema 7 – inhaltliche Aspekte (Ausschnitt)

baren Werte erreicht. Andere Suchmaschinen sind unbedeutend und i.d.R. abhängig. So fiel die Wahl auf google.de und bing.de, sowie google.com und bing.com.

Für die Datenerfassung wurde das an der *Hochschule für Angewandte Wissenschaften Hamburg* entwickelte *Relevance Assessment Tool* (RAT) verwendet. Nach Lewandowski und Sünkler (2012, S. 237 f.) ist das RAT ein Werkzeug zur Durchführung von Retrievaltests und zur weitgehend automatisierten Erfassung, Aufbereitung und Auswertung der Daten. Da die Untersuchung nicht mit Probanden und manuell durchgeführt wurde, ist hier lediglich die *Suchmaschinenscraper*-Komponente des Tools von Interesse. Der Scraper des RATs simuliert den Aufruf einer Suchmaschine und das Abschicken einer Suche durch einen Webbrowser und kann dann Ergebnisseiten (*Search Engine Result Pages*, SERPs) und Ergebnisse selbst erfassen und speichern (für technische Details s. ebd., S. 240 f.).

Bei der empirischen Untersuchung wurden unter Umgehung von Personalisierungen (etwa durch Suchhistorie oder Standort) jeweils die SERPs als lokale Kopien und die ersten zehn »textlichen Ergebnisse« für die Anfragen als Links in Excel-Listen sowie ebenfalls als lokale Kopien erfasst. Textliche Ergebnisse bedeutet organische Ergebnisse und ggf. vorhandene *Universal-Search-News-Treffer* (Nachrichtenergebnisse).

### 3.3 Auswertung

Dem bei Retrievaltests klassischen Problem der Bestimmung des Recalls (s. Abschnitt 3.1) wurde mit einer *Pooling*-Methode begegnet: Zu jedem der Themen wurde aus den inhaltlichen Aussagen der erfassten relevanten Suchmaschinen-Ergebnisse ein Pool aus mehreren »Aspekten« gebildet. Dieser bildete jeweils die Grundgesamtheit, gegen den später die Retrievaleffektivität der einzelnen Suchmaschinen gespiegelt wurde.

Die inhaltlichen Pools wurden gebildet, indem zu jedem Thema alle zugehörigen erfassten Suchergebnisse gesichtet und die darin enthaltenen inhaltlichen Aussagen gesammelt wurden. Da die Ergebnisse und damit die Suchmaschinen letztlich bezüglich des Vorhandenseins und der Ausführlichkeit inhaltlicher Aspekte verglichen werden sollten, wurden diese inhaltlichen Aussagen anschließend

thematisch zu solchen Aspekten gruppiert (die Anzahl dieser Aspekte variierte von vier bis neun). Die Gruppierung der Aspekte bildete gleichzeitig ihre ausführlichste Version. Zusätzlich wurde jeweils eine knappe Version definiert, die lediglich aus den grundlegenden inhaltlichen Aussagen bestand. So wurden Ausführlichkeitsdimensionen aufgespannt, anhand derer dann die einzelnen Ergebnisse hinsichtlich des Vorhandenseins und der Ausführlichkeit der inhaltlichen Aspekte beurteilt werden konnten.

Ab hier soll die Auswertung zu Thema 7 – dem Luftangriff Israels an der syrischen Grenze – als Leitfaden dienen. Tabelle 2 zeigt einen Ausschnitt der fünf inhaltlichen Aspekte dieses Themas.

Anhand der inhaltlichen Aspekte in ausführlicher und knapper Definition wurden die einzelnen Ergebnisse der Suchmaschinen dann mithilfe einer fünfteiligen *Likert-Skala* beurteilt (wobei natürlich die Länge der Texte berücksichtigt wurde):

knapp	eher knapp	neutral	eher ausführlich	ausführlich
-2	-1	0	1	2

Tabelle 3 zeigt die Auswertung der jeweils zehn ersten textlichen Ergebnisse der Suchmaschinen zu Thema 7 hinsichtlich des Vorkommens und der Ausführlichkeit der inhaltlichen Aspekte (s. Tabelle 2). Die Farben markieren gleiche Inhalte – graue Ergebnisse sind irrelevant (z.B. weil sie veraltet waren, andere Themen behandelten, oder sich ohne Registrierung/Anmeldung nicht (ganz) lesen ließen).

Z.B. deckte also das erste google.com-Ergebnis drei der fünf Aspekte ab: bspw. I, den Luftangriff, eher ausführlich. Bspw. wurde Aspekt III, Reaktionen zum Luftangriff von syrischer Seite, in diesem Text nicht behandelt. Dass wie hier Universal-Search-News-Treffer (häufig die ersten drei) oft nochmals als organische Ergebnisse ausgegeben wurden, zog sich durch alle Auswertungen, der vorliegende Fall gleicher Inhalte (unter unterschiedlichen URLs und Titeln) bei bing.com ist ein extremer.

Um den Auswertungsaufwand einzugrenzen, wurden als nächstes unter der Annahme, dass Ausführlichkeit Wichtigkeit bedeutet, jeweils die drei wichtigsten Aspekte identifiziert. Anschließend wurde das Vorkommen dieser wichtigen Aspekte in allen Ergebnissen gezählt. Das Resultat ist in Tabelle 4 dargestellt.

		google.de									
Aspekt	1	2	3	4	5	6	7	8	9	10	
I	1	1	1	1	1			1			
II	1	1	1	1	1			0			
III	-2	1		1							
IV	-1	-1	1	-1	1						
V	0	0	0	0	0			0			

		bing.de									
Aspekt	1	2	3	4	5	6	7	8	9	10	
I	1	1	0	1	1	1	1	1	0	1	
II	0	1	0	1	1	0	0	1	0	0	
III			1	-2			-1		1		
IV	1	0	1	-1	0	1		1	1	1	
V	0		1	0		0		0	1	0	

		google.com									
Aspekt	1	2	3	4	5	6	7	8	9	10	
I	1		2	1	2		1	2	1		
II	0		1	1	1		1	1	1		
III			1	2	1			0	2		
IV	-1		-2	2	-2			-2	2		
V			0	2	0			2	2		

		bing.com									
Aspekt	1	2	3	4	5	6	7	8	9	10	
I			1		2	1	1	1	1	1	
II			1		1	1	1	1	1	1	
III			2			2		2	2	2	
IV			2			2		2	2	2	
V			2		0	2		2	2	2	

Tabelle 3: Thema 7 – Vorkommen und Ausführlichkeit der inhaltlichen Aspekte

I	21
II	19
V	17
III	16
IV	15

Tabelle 4: Thema 7 – Wichtige Aspekte mit der Häufigkeit ihres Auftretens

In diesem Fall gehörten alle Aspekte mindestens einmal zu den wichtigsten drei Aspekten der Ergebnisse, daher zählten alle zu den wichtigen Aspekten.

Als nächstes wurde untersucht, ob und wie die wichtigen Aspekte von den Suchmaschinen innerhalb der ersten zehn textlichen Ergebnisse abgedeckt wurden. In diesem Fall wurden sie von allen Suchmaschinen abgedeckt. (Das war die Ausnahme – i.d.R. deckten die Suchmaschinen, manchmal sogar beide einer Region, einen wichtigen Aspekt oder mehrere nicht ab.)

An dieser Stelle können für Thema 7 die beiden Forschungsfragen zur Vollständigkeit (s. Abschnitt 2) beantwortet werden.

Hypothese 1.1 kann nicht widerlegt werden: Google und Bing vermitteln in Deutschland und den USA in Hinsicht auf die wichtigen inhaltlichen Aspekte in mindestens 75 Prozent der Fälle vollständige Bilder aktueller internationaler Themen. Somit kann Forschungsfrage 1.1 bestätigt werden: Google und Bing in Deutschland und den USA vermitteln in Hinsicht auf die wichtigen inhaltlichen Aspekte vollständige Bilder aktueller internationaler Themen.

Hypothese 1.2 kann ebenfalls nicht widerlegt werden: Bei den durch Google und Bing vermittelten Bildern aktueller internationaler Themen in Deutschland und den USA gibt es in Hinsicht auf die wichtigen inhaltlichen Aspekte in puncto Vollständigkeit in mindestens 75 Prozent der Fälle keine Unterschiede. Somit kann Forschungsfrage 1.2 verneint werden: Bei den durch Google und Bing vermittelten Bildern aktueller internationaler Themen in Deutschland



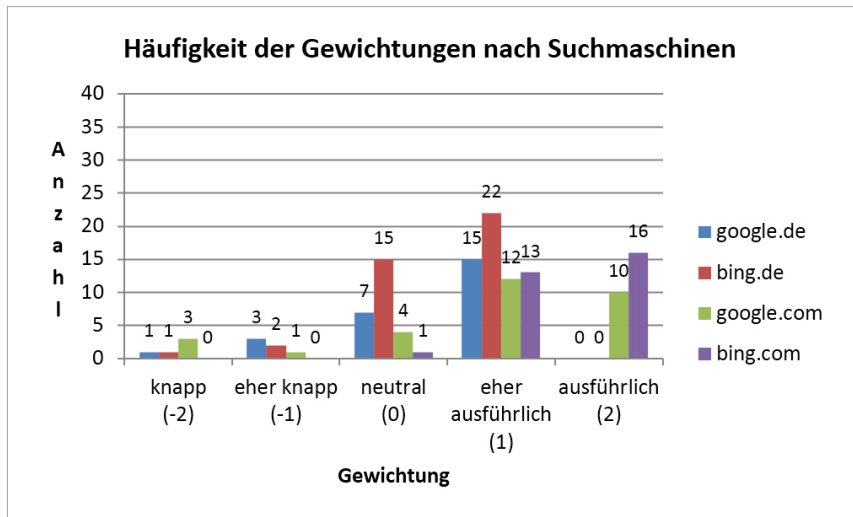


Abbildung 1: Thema 7 – Häufigkeit der Gewichtungen nach Suchmaschinen

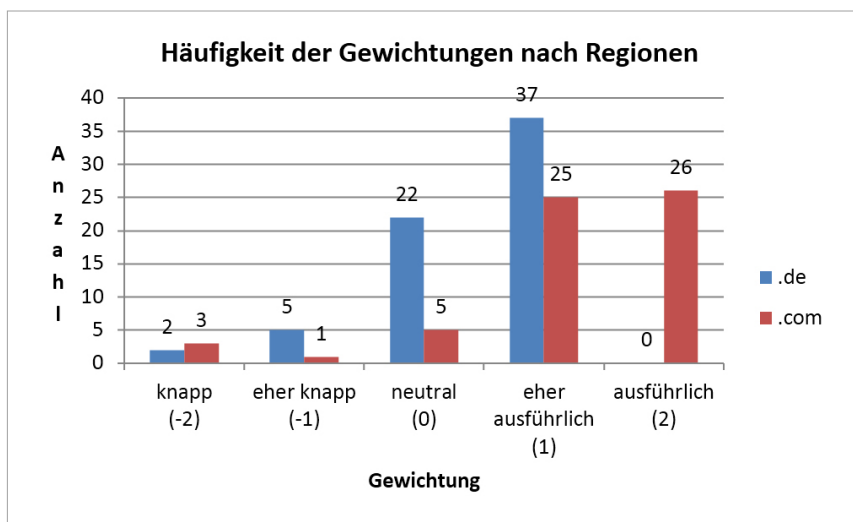


Abbildung 2: Thema 7 – Häufigkeit der Gewichtungen nach Regionen

und den USA gibt es in Hinsicht auf die wichtigen inhaltlichen Aspekte keine Unterschiede in puncto Vollständigkeit.

Als nächstes wurde die Abdeckung der drei wichtigsten Aspekte jeweils auf der ersten Trefferposition der Suchmaschinen untersucht. Google- und bing.de deckten so die drei wichtigsten Aspekte ab, google.com die ersten beiden und bing.com keinen.

An dieser Stelle kann für Thema 7 Hypothese 2.1 widerlegt werden: Auf Trefferposition 1 werden von Google und Bing in Deutschland und den USA nicht die drei wichtigsten inhaltlichen Aspekte aktueller internationaler Themen abgedeckt.

Danach wurde die Abdeckung von 75 Prozent der wichtigsten Aspekte, in diesem Fall 3,75, also vier (es

wurde jeweils wie üblich gerundet), auf den ersten drei Trefferpositionen untersucht. Es gelang so allen Suchmaschinen, die vier wichtigsten Aspekte abzudecken. An dieser Stelle kann für Thema 7 Hypothese 2.2 nicht widerlegt werden: Bei Berücksichtigung der ersten drei Trefferpositionen werden von Google und Bing in Deutschland und den USA mindestens 75 Prozent der wichtigsten inhaltlichen Aspekte aktueller internationaler Themen abgedeckt.

Somit kann für Thema 7 die Forschungsfrage zur Abdeckung nicht eindeutig beantwortet werden: Google und Bing in Deutschland und den USA decken teilweise bereits auf den ersten Trefferpositionen die wichtigsten inhaltlichen Aspekte aktueller internationaler Themen ab.

Zu guter Letzt wurden anhand Tabelle 3 die Häufigkeiten der inhaltlichen Gewichtungen nach Suchmaschinen und Regionen gezählt. Die Ergebnisse sind als Diagramme in Abbildung 1 nach Suchmaschinen und in Abbildung 2 nach Regionen dargestellt.

Die Summen der Gewichtungen pro Suchmaschine, und damit pro Region, können sich voneinander unterscheiden, weil nicht alle Suchergebnisse immer alle inhaltlichen Aspekte behandelten und es irrelevante Treffer gab, die keine Gewichtungen erhielten. Hier behandeln die amerikanischen Suchmaschinenergebnisse die inhaltlichen Aspekte des Themas tendenziell etwas ausführlicher als die deutschen.

Um auf signifikante Unterschiede zwischen den Regionen und Suchmaschinen zu prüfen, wurden via SPSS der *Mann-Whitney-Test* bzw. der *Kruskal-Wallis-Test* durchgeführt.

Das Ergebnis des Mann-Whitney-Tests für die Regionen liegt mit 0,841 (exakte Signifikanz) bzw. 0,834 (asymptotische) deutlich über dem üblichen Signifi-

kanzniveau von  $p = 0,05$ ; die Regionen unterscheiden sich in ihren inhaltlichen Gewichtungen also nicht signifikant voneinander. Das Ergebnis des Kruskal-Wallis-Tests für die Suchmaschinen liegt mit 0,925 ebenfalls deutlich über  $p = 0,05$ ; auch die Suchmaschinen unterscheiden sich somit in ihren inhaltlichen Gewichtungen nicht signifikant voneinander.

An dieser Stelle kann für Thema 7 die Forschungsfrage zur Gewichtung beantwortet werden.

Hypothese 3 kann nicht widerlegt werden: Bei der Gewichtung der wichtigen inhaltlichen Aspekte durch Google und Bing vermittelter Bilder aktueller internationaler Themen in Deutschland und den USA gibt es keine signifikanten Unterschiede.

Somit kann Forschungsfrage 3 verneint werden: Bei den durch Google und Bing vermittelten Bildern aktueller internationaler Themen in Deutschland und den USA gibt es keine signifikanten Unterschiede bei der Gewichtung der wichtigen inhaltlichen Aspekte.

## 4 Ergebnisse der empirischen Untersuchung

Vor der Präsentation des zu den Forschungsfragen und -hypothesen (vgl. Abschnitt 2) korrespondierenden Gesamtergebnisses der empirischen Untersuchung ist zu sagen, dass die Themen 1, 5, 8, 10 und 11 (s. Abschnitt 3.2) nicht ausgewertet wurden, etwa, weil es zu wenig relevante Ergebnisse gab oder das Thema zu divers war. Von den 14 Themen wurden somit neun ausgewertet.

### 4.1 Vollständigkeit inhaltlicher Aspekte

Abbildung 3 zeigt die Anzahl vollständig vermittelter Bilder nach Suchmaschinen, Abbildung 4 nach Regionen im Durchschnitt.

Die höchste erreichte Anzahl vollständig vermittelter Bilder erzielen google- und bing.de mit sechs von neun Fällen – die amerikanischen Suchmaschinen schneiden schlechter ab. Dies wird auch bei der Ansicht nach Regionen deutlich: Die deutschen Suchmaschinen vermitteln durchschnittlich sechs von neun Bildern vollständig, die amerikanischen lediglich 3,5.

Damit kann Hypothese 1.1 widerlegt werden: Google und Bing vermitteln in Deutschland und den USA in Hinsicht auf die wichtigen inhaltlichen Aspekte in mindestens 75 Prozent der Fälle keine voll-

ständigen Bilder aktueller internationaler Themen. Somit kann Forschungsfrage 1.1 verneint werden: Google und Bing in Deutschland und den USA vermitteln in Hinsicht auf die wichtigen inhaltlichen Aspekte keine vollständigen Bilder aktueller internationaler Themen.

Hypothese 1.2 kann ebenfalls widerlegt werden: Trotz der ähnlichen Suchmaschinen und Kultur- und Sprachräume gibt es bei den durch Google und Bing vermittelten Bildern aktueller internationaler Themen in Deutschland und den USA in Hinsicht auf die wichtigen inhaltlichen Aspekte in puncto Vollständigkeit in mindestens 75 Prozent der Fälle Unterschiede. Somit kann Forschungsfrage 1.2 bestätigt werden: Bei den durch Google und Bing vermittelten Bildern aktueller internationaler Themen in Deutschland und den USA gibt es in Hinsicht auf die wichtigen inhaltlichen Aspekte Unterschiede in puncto Vollständigkeit.

### 4.2 Abdeckung inhaltlicher Aspekte

Abbildung 5 zeigt die Anzahl von Fällen mit den auf Position 1 abgedeckten drei wichtigsten inhaltlichen Aspekten nach Suchmaschinen, Abbildung 6 nach Regionen im Durchschnitt.

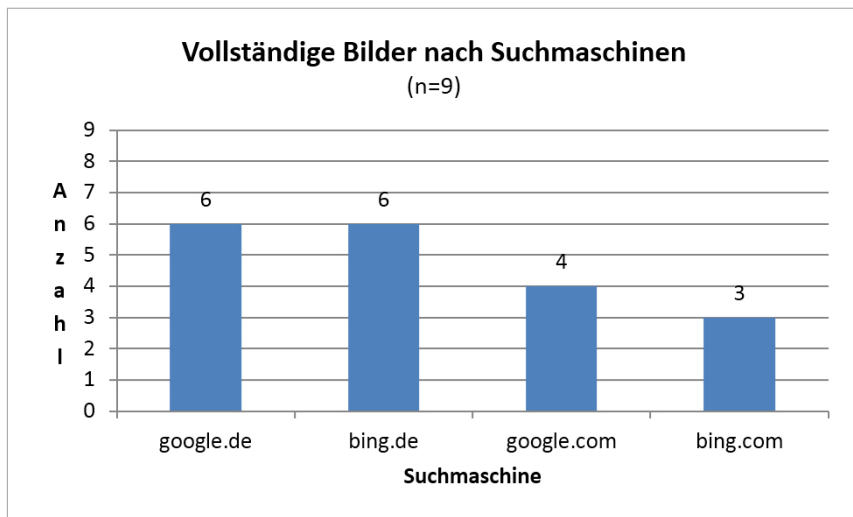


Abbildung 3: Vollständige Bilder nach Suchmaschinen

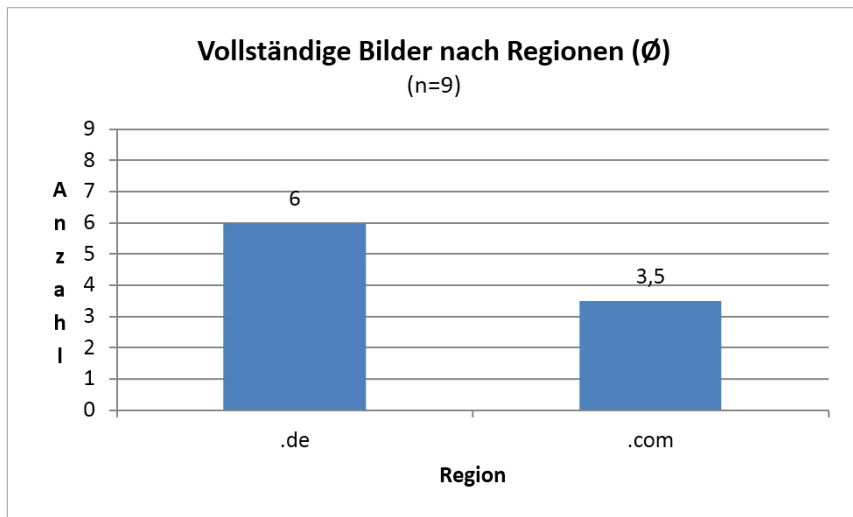


Abbildung 4: Vollständige Bilder nach Regionen (Durchschnitte)

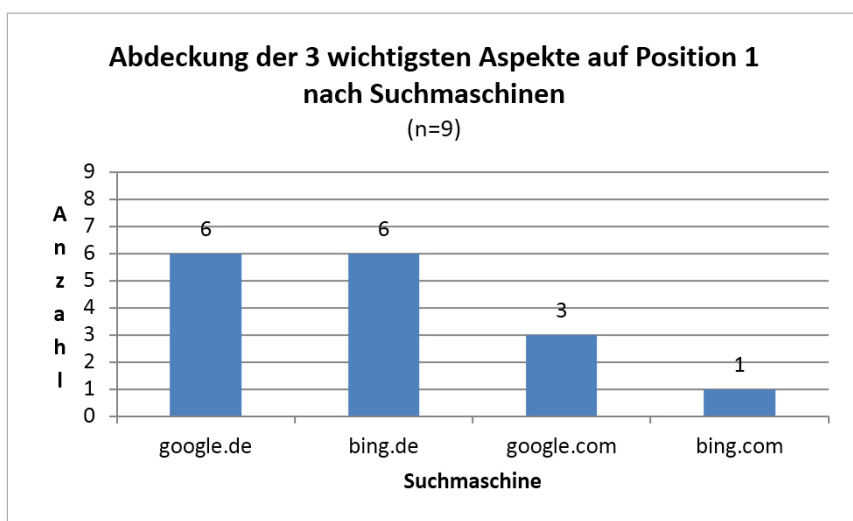


Abbildung 5: Abdeckung der drei wichtigsten Aspekte auf Position 1 nach Suchmaschinen

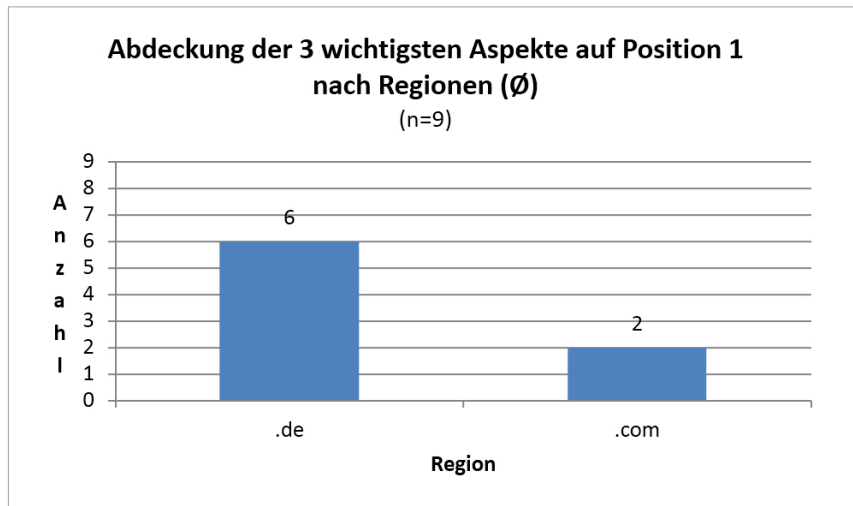


Abbildung 6: Abdeckung der drei wichtigsten Aspekte auf Position 1 nach Regionen (Durchschnitte)

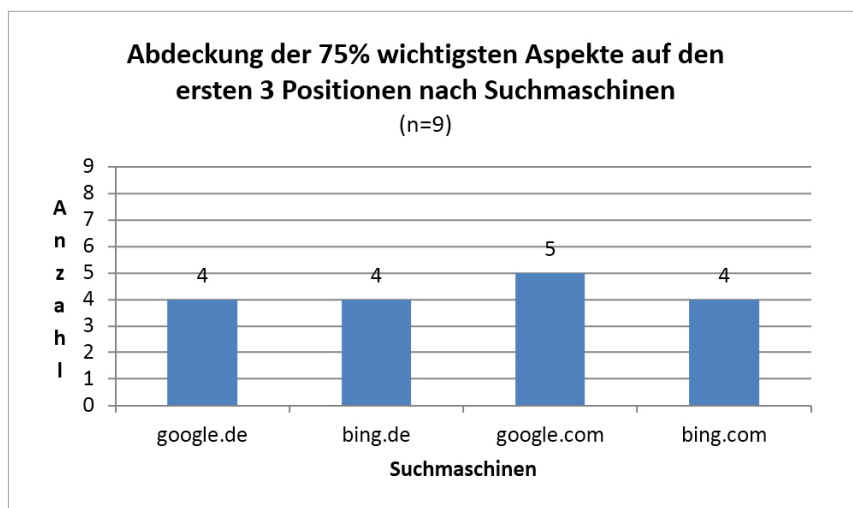


Abbildung 7: Abdeckung der 75 Prozent wichtigsten Aspekte auf den ersten drei Positionen nach Suchmaschinen

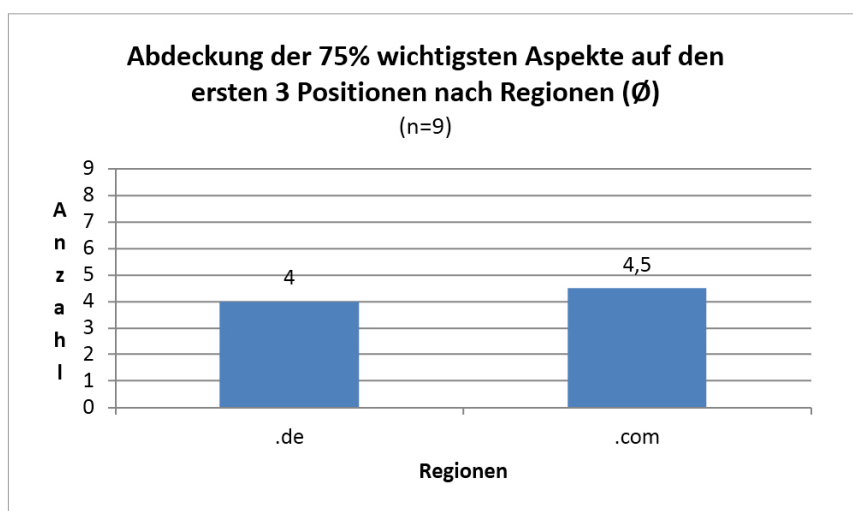


Abbildung 8: Abdeckung der 75 Prozent wichtigsten Aspekte auf den ersten drei Positionen nach Regionen (Durchschnitte)

Die höchste erreichte Anzahl der auf Trefferposition 1 abgedeckten drei wichtigsten inhaltlichen Aspekte erzielen wieder google- und bing.de mit sechs – die amerikanischen Suchmaschinen schneiden erneut schlechter ab (bing.com deutlich mit nur einem Fall). Dies wird auch bei der Ansicht nach Regionen deutlich: Die deutschen Suchmaschinen decken in durchschnittlich sechs von neun Fällen die drei wichtigsten inhaltlichen Aspekte auf Trefferposition 1 ab, die amerikanischen in lediglich zwei. Damit kann Hypothese 2.1 widerlegt werden: Auf Trefferposition 1 werden von Google und Bing in Deutschland und den USA nicht die drei wichtigsten inhaltlichen Aspekte aktueller internationaler Themen abgedeckt.

Abbildung 7 zeigt die Anzahl von Fällen mit den auf den ersten drei Positionen abgedeckten 75 Prozent wichtigsten inhaltlichen Aspekten nach Suchmaschinen, Abbildung 8 nach Regionen im Durchschnitt.

Es zeigt sich, dass drei Suchmaschinen mit vier solcher Fälle recht ähnlich abschneiden – nur bei google.com waren es fünf. Dieses Bild zeigt sich auch bei der Ansicht nach Regionen. Damit kann auch Hypothese 2.2 widerlegt werden: Bei Berücksichtigung der ersten drei Trefferpositionen werden von Google und Bing in Deutschland und den USA nicht mindestens 75 Prozent der wichtigsten inhaltlichen Aspekte aktueller internationaler Themen abgedeckt.

Somit kann Forschungsfrage 2 verneint werden: Von Google und Bing in Deutschland und den USA werden nicht bereits auf den ersten Trefferpositionen die drei wichtigsten inhaltlichen Aspekte aktueller internationaler Themen abgedeckt.

### 4.3 Gewichtung inhaltlicher Aspekte

Abbildung 9 zeigt die Häufigkeiten der inhaltlichen Gewichtungen nach Suchmaschinen, Abbildung 10 nach Regionen.

Die Summen der Gewichtungen der inhaltlichen Aspekte reichen bei den Suchmaschinen von 200

(bing.com) bis 307 (bing.de). Insgesamt wurden bei den deutschen Suchmaschinen mit 587 mehr Aspekte als bei den amerikanischen mit 459 gewichtet – eine Differenz von 128.

Es zeigt sich bezüglich aller Suchmaschinen eine ähnliche Verteilung der Gewichtungen: Es gibt einige knappe Aspekte, einige wenige eher knappe, etwas mehr neutrale, deutlich mehr und am meisten eher ausführliche und wieder weniger ausführliche. Dieses Bild zeigt sich auch bei der Ansicht nach Regionen: Hier gibt es kaum größere Unterschiede, besonders, wenn man die größere deutsche Grundgesamtheit bedenkt: Es fällt lediglich auf, dass die inhaltlichen Aspekte in den deutschen Ergebnissen deutlich öfter eher ausführlich waren.

Um auf signifikante Unterschiede zwischen den Regionen und Suchmaschinen zu prüfen, wurden mittels SPSS der Mann-Whitney-Test bzw. der Kruskal-Wallis-Test durchgeführt. Das Ergebnis des Mann-Whitney-Tests für die Regionen liegt mit 0,69 (exakte Signifikanz) bzw. 0,602 (asymptotische) deutlich über dem üblichen Signifikanzniveau von  $p = 0,05$ ; die Regionen unterscheiden sich in ihren inhaltlichen Gewichtungen also nicht signifikant voneinander. Das Ergebnis des Kruskal-Wallis-Tests für die Suchmaschinen liegt mit 0,473 ebenfalls deutlich über  $p = 0,05$ ; auch die Suchmaschinen unterscheiden sich somit in ihren inhaltlichen Gewichtungen nicht signifikant voneinander.

Damit kann Hypothese 3 nicht widerlegt werden: Durch die ähnlichen Suchmaschinen und Kultur- und Sprachräume gibt es bei der Gewichtung der wichtigen inhaltlichen Aspekte durch Google und Bing vermittelter Bilder aktueller internationaler Themen in Deutschland und den USA keine signifikanten Unterschiede. Somit kann Forschungsfrage 3 verneint werden: Bei den durch Google und Bing vermittelten Bildern aktueller internationaler Themen in Deutschland und den USA gibt es bei der Gewichtung der wichtigen inhaltlichen Aspekte keine signifikanten Unterschiede.

## 5 Diskussion, Einschränkungen und weitere Forschung

Die Ergebnisse der Untersuchung zeichnen ein negatives Bild: Google und Bing vermitteln in Deutschland und den USA keine vollständigen Bilder aktueller internationaler Themen (F1.1, vgl. Abschnitt 2), wobei es Unterschiede hinsichtlich der wichtigen Aspekte gibt (F1.2); sie decken auf den ersten Trefferpositio-

nen nicht die drei wichtigsten inhaltlichen Aspekte ab (F2); und bei der Gewichtung der wichtigen inhaltlichen Aspekte gibt es keine signifikanten Unterschiede (F3). Demnach sind tatsächlich unterschiedlich wirkende inhaltliche Suchmaschinenverzerrungen vorhanden.

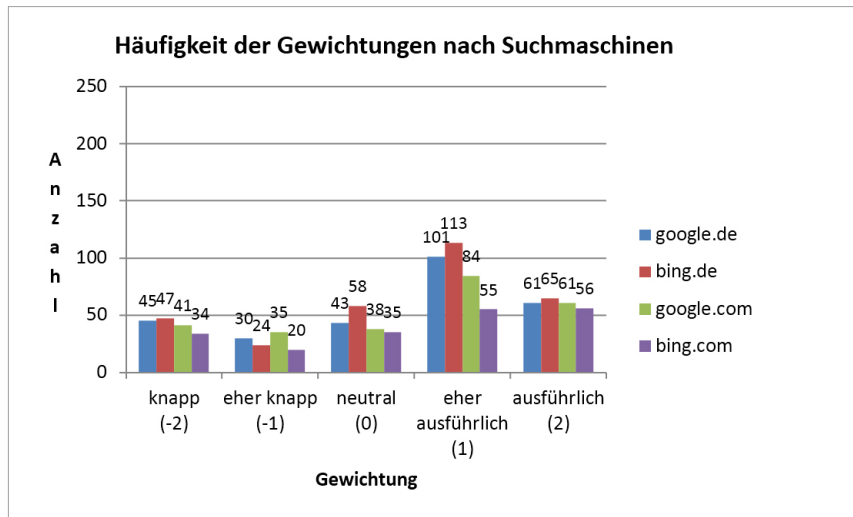


Abbildung 9: Häufigkeit der Gewichtungen nach Suchmaschinen

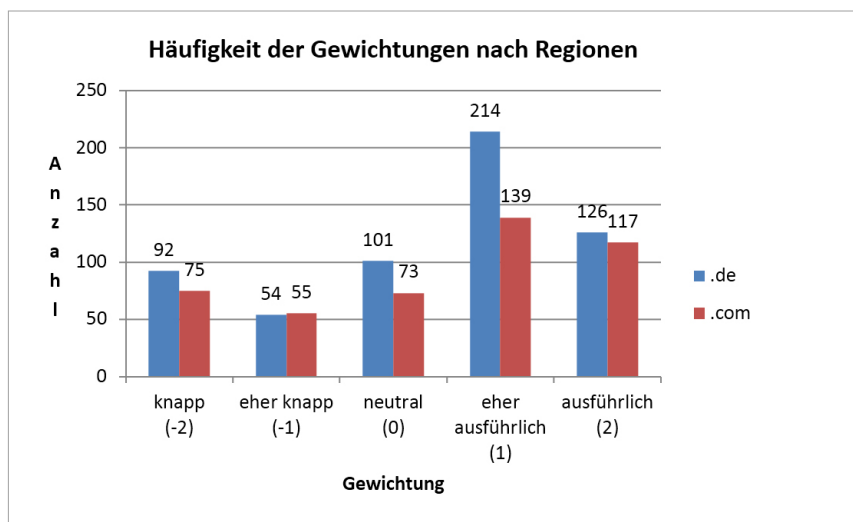


Abbildung 10: Häufigkeit der Gewichtungen nach Regionen

Allerdings ist die Aussagekraft dieser Ergebnisse eingeschränkt. Die Methodik der Untersuchung baut auf Inhaltsanalyse und Retrievaltests auf, d.h. dass Kritik an diesen Methoden auch auf diese Methodik zutreffen kann. In Hinsicht auf Retrievaltests ist dies etwa beim zugrundeliegenden Nutzermodell der Fall, bei der Beschränkung auf organische und Universal-Search-News-Treffer und bei der geringen Anzahl der Suchanfragen (14 Themen, von denen neun ausgewertet wurden, sind zu wenig, um valide Ergebnisse zu produzieren) (vgl. Lewandowski 2015, S. 220 f.).

In diesem Zusammenhang hat sich bzgl. der Zeitungsauswahl (vgl. Abschnitt 2) gezeigt, dass die beiden Boulevardblätter – *Bild* und *USA Today* –

kaum Nutzen gebracht haben. Statt diesen sollten das nächste Mal weitere auflagenstarke Qualitätszeitungen einbezogen werden.

Bezüglich der Suchanfragen hat sich gezeigt, dass diese z.T. zu allgemein gewählt waren. Dies führte zu allzu diversen bzw. irrelevanten Suchergebnissen. Auch durch Ergebnisse, die aufgrund der Suchmaschinen irrelevant waren, variierte die Anzahl der auswertbaren Suchergebnisse bei den Einzelauswertungen z.T. erheblich.

Weitere Einschränkungen ergeben sich aufgrund von Problemen beim Scraping; so wurden organische Ergebnisse, die vorher als Universal-Search-News-Treffer erfasst worden waren, ausgelassen, oder Universal-Search-News-Treffer nicht oder un-

vollständig erfasst. Dies wurde allerdings bei der Auswertung durch das Abgleichen der SERPs mit den Listen der gescrapten Ergebnisse so weit wie möglich korrigiert. Zudem konnten bei der Auswertung mögliche Eigenarten des Bearbeiters nicht ausgeglichen werden – Interrater-Reliabilität ist nicht vorhanden. Das nächste Mal sollten mindestens zwei Bearbeiter zusammenarbeiten. Aufgrund all dieser Punkte ist die Aussagekraft der Untersuchungsergebnisse als eingeschränkt zu betrachten.

Dennoch hat sich die Untersuchung gelohnt, liegt jetzt doch ein Versuch des Nachweises überregionaler inhaltlicher Suchmaschinenverzerrungen vor, der weiterer Forschung in diesem Bereich als Ausgangspunkt dienen kann.

Es wäre interessant, die Untersuchung unter Berücksichtigung der verbessernden Anmerkungen

zu wiederholen und so zu validieren Ergebnissen zu gelangen. Ebenfalls interessant wäre es, einen Ansatz zur automatisierten Inhaltserschließung von Suchergebnissen zu entwickeln. Dazu gibt es diverse *Text-mining*-, *Text-analysis*- und weitere Tools. Allerdings ist die automatisierte Inhaltserschließung sehr komplex und technisiert, manche basalen Probleme sind noch ungelöst. Ein weiterer Problem- punkt besteht darin, dass diese Tools meist aus dem englischsprachigen Raum stammen und oft keine deutschen Dokumente verarbeiten können. Trotzdem wäre die Entwicklung eines automatisierten Textanalyse-Ansatzes eine dankbare Aufgabe, da der manuelle Aufwand sehr hoch ist. Interessant wäre in diesem Zusammenhang wohl auch ein Vergleich zwischen einem manuellen und einem automatisierten Verfahren.

## Literatur

- AGOF e.V., Hrsg. (2015). *Internet Facts 2015-01*. [http://www.agof.de/download/Downloads\\_Internet\\_Facts/Downloads\\_Internet\\_Facts\\_ZIP/internet\\_facts\\_2015-01.zip?dbfe35](http://www.agof.de/download/Downloads_Internet_Facts/Downloads_Internet_Facts_ZIP/internet_facts_2015-01.zip?dbfe35) zuletzt abgerufen am 23. März 2016.
- Brown, Lauren; Dubois, Alain; Shepard, Donald (2008). Inefficiency and Bias of Search Engines in Retrieving References and Containing Scientific Names of Fossil Amphibians. In *Bulletin of Science, Technology & Society* 28(4), S. 279–288. [http://www.researchgate.net/publication/249747653\\_Inefficiency\\_and\\_Bias\\_of\\_Search\\_Engines\\_in\\_Retrieving\\_References\\_Containing\\_Scientific\\_Names\\_of\\_Fossil\\_Amphibians](http://www.researchgate.net/publication/249747653_Inefficiency_and_Bias_of_Search_Engines_in_Retrieving_References_Containing_Scientific_Names_of_Fossil_Amphibians) zuletzt abgerufen am 23. März 2016.
- Früh, Werner (2011). *Inhaltsanalyse*. 7. Aufl. München: Ölschläger.
- Gordon, Michael; Pathak, Praveen (1999). Finding information on the World Wide Web: the retrieval effectiveness of search engines. In *Information Processing & Management* 35, S. 141–180. [http://www.jasonmorrison.net/iakm/cited/Gordon\\_Pathak.pdf](http://www.jasonmorrison.net/iakm/cited/Gordon_Pathak.pdf) zuletzt abgerufen am 23. März 2016.
- Granka, Laura (2010). The Politics of Search: A Decade Retrospective. In *The Information Society* 26, S. 364–374. <http://www.australianscience.com.au/research/google/36914.pdf> zuletzt abgerufen am 23. März 2016.
- Harman, Donna; Voorhees, Ellen (2006). TREC: An overview. In *Annual Review of Information Science and Technology* 40, S. 113–155.
- Hawking, David; Craswell, Nick; Bailey, Peter; Griffiths, Kathy (2001). Measuring Search Engine Quality. In *Information Retrieval* 4, S. 33–59. [http://es.csiro.au/pubs/hawking\\_ir01.pdf](http://es.csiro.au/pubs/hawking_ir01.pdf) zuletzt abgerufen am 23. März 2016.
- Hölig, Sascha; Hasebrink, Uwe (2014). Nachrichtennutzung im Wandel: Neue Plattformen, Endgeräte und Zugänge: International vergleichende Befunde auf Basis des Reuter Institute Digital News Survey 2014. In *Media Perspektiven* 11, S. 530–538. [http://www.ard-werbung.de/fileadmin/user\\_upload/media-perspektiven/pdf/2014/11-2014\\_Hoelig\\_Hasebrink.pdf](http://www.ard-werbung.de/fileadmin/user_upload/media-perspektiven/pdf/2014/11-2014_Hoelig_Hasebrink.pdf) zuletzt abgerufen am 23. März 2016.
- leong, Samuel; Mishra, Nina; Sadikov, Eldar; Zhang, Li (2012). Domain Bias in Web Search. In *Proceedings of the fifth ACM international conference on Web search and data mining*. New York: ACM, S. 413–422. <http://research.microsoft.com/pubs/155941/domainbias.pdf> zuletzt abgerufen am 23. März 2016.
- Introna, Lucas; Nissenbaum, Helen (2000). Shaping the Web: Why the Politics of Search Engines Matters. In *The Information Society* 16, S. 169–185. <http://www.nyu.edu/projects/nissenbaum/papers/ShapingTheWeb.pdf> zuletzt abgerufen am 23. März 2016.

- Joachims, Thorsten; Granka, Laura; Pan, Bing; Hembrooke, Helene; Radlinski, Filip; Gy, Geri (2007). Evaluating the Accuracy of Implicit Feedback from Clicks and Query Reformulations in Web Search. In *ACM Transactions on Information Systems* 25(2), Article No. 7. [http://www.cs.cornell.edu/people/tj/publications/joachims\\_etal\\_07a.pdf](http://www.cs.cornell.edu/people/tj/publications/joachims_etal_07a.pdf) zuletzt abgerufen am 23. März 2016.
- Lewandowski, Dirk (2011). Evaluierung von Suchmaschinen. In *Handbuch Internet-Suchmaschinen 2: Neue Entwicklungen in der Web-Suche*. Hrsg. von Lewandowski, Dirk. Heidelberg: AKA, S. 203–228. <http://searchstudies.org/wp-content/uploads/2015/10/Evaluierung-von-Suchmaschinen-2011.pdf> zuletzt abgerufen am 23. März 2016.
- Lewandowski, Dirk (2014). Die Macht der Suchmaschinen und ihr Einfluss auf unsere Entscheidungen. In *Information – Wissenschaft & Praxis* 65(4/5), S. 231–238.
- Lewandowski, Dirk (2015). *Suchmaschinen verstehen*. Berlin und Heidelberg: Springer Vieweg.
- Lewandowski, Dirk; Sünkler, Sebastian (2012). Relevance Assessment Tool: Ein Werkzeug zum Design von Retrievaltests sowie zur weitgehend automatisierten Erfassung, Aufbereitung und Auswertung der Daten. In *Proceedings der 2. DGI-Konferenz: Social Media und Web Science – Das Web als Lebensraum*. Frankfurt a.M.: DGI, S. 237–249. <http://searchstudies.org/wp-content/uploads/2015/10/Relevance-Assessment-Tool-2012.pdf> zuletzt abgerufen am 23. März 2016.
- Purcell, Kristen; Brenner, Joanna; Rainie, Lee (2012). *Search Engine Use 2012*. [http://www.pewinternet.org/files/old-media/Files/Reports/2012/PIP\\_Search\\_Engine\\_Use\\_2012.pdf](http://www.pewinternet.org/files/old-media/Files/Reports/2012/PIP_Search_Engine_Use_2012.pdf) zuletzt abgerufen am 23. März 2016.
- Statista, Hrsg. (2015a). *Ranking der Zeitungen mit den höchsten verkauften Auflagen in den USA im Zeitraum von Oktober 2013 bis März 2014 und April 2014 bis September 2014*. Stand: 2015. <http://de.statista.com/statistik/daten/studie/362329/umfrage/zeitungen-mit-den-hoechsten-verkauften-auflagen-in-den-usa/> zuletzt abgerufen am 23. März 2016.
- Statista, Hrsg. (2015b). *Share of search queries handled by leading U.S. search engine providers as of April 2015*. Stand: 2015. <http://www.statista.com/statistics/267161/market-share-of-search-engines-in-the-united-states/> zuletzt abgerufen am 23. März 2016.
- Statista, Hrsg. (2015c). *Suchmaschinenverteilung in Deutschland im Jahr 2015 (sowie Vorjahresvergleich)*. Stand: 2015. <http://de.statista.com/statistik/daten/studie/167841/umfrage/marktanteile-ausgewaehlter-suchmaschinen-in-deutschland/> zuletzt abgerufen am 23. März 2016.
- Statista, Hrsg. (2015d). *Verkaufte Auflage der überregionalen Tageszeitungen in Deutschland im 1. Quartal 2015*. Stand: 2015. <http://de.statista.com/statistik/daten/studie/73448/umfrage/auflage-der-ueberregionalen-tageszeitungen/> zuletzt abgerufen am 23. März 2016.
- Tague-Sutcliffe, Jean (1992). The pragmatics of information retrieval experimentation, revisited. In *Information Processing & Management* 28, S. 467–490.
- TNS Infratest, Hrsg. (2013). *Relevanz der Medien für die Meinungsbildung: Empirische Grundlagen zum MedienVielfaltsMonitor der BLM*. [https://www.blm.de/files/pdf1/Studie\\_Relevanz\\_der\\_Medien\\_2013.pdf](https://www.blm.de/files/pdf1/Studie_Relevanz_der_Medien_2013.pdf) zuletzt abgerufen am 23. März 2016.
- van Eimeren, Birgit; Frees, Beate (2014). Ergebnisse der ARD/ZDF-Online-Studie: 79 Prozent der Deutschen online – Zuwachs bei mobiler Internetnutzung und Bewegtbild. In *Media Perspektiven* 45(7/8), S. 378–396. [http://www.ard-zdf-onlinestudie.de/fileadmin/Onlinestudie\\_2014/PDF/0708-2014\\_Eimeren\\_Frees.pdf](http://www.ard-zdf-onlinestudie.de/fileadmin/Onlinestudie_2014/PDF/0708-2014_Eimeren_Frees.pdf) zuletzt abgerufen am 23. März 2016.
- van Eimeren, Birgit; Gerhard, Heinz; Frees, Beate (2003). Internetverbreitung in Deutschland: Unerwartet hoher Zuwachs: ARD/ZDF-Online-Studie 2003. In *Media Perspektiven* (8), S. 338–358. [http://www.ard-zdf-onlinestudie.de/fileadmin/Onlinestudie\\_2003/Online03\\_Nutzung.pdf](http://www.ard-zdf-onlinestudie.de/fileadmin/Onlinestudie_2003/Online03_Nutzung.pdf) zuletzt abgerufen am 23. März 2016.
- Weber, Karsten (2011). Search Engine Bias. In *Handbuch Internet-Suchmaschinen 2: Neue Entwicklungen in der Web-Suche*. Hrsg. von Lewandowski, Dirk. Heidelberg: AKA, S. 265–285. [http://www.researchgate.net/publication/256473910\\_Search\\_Engine\\_Bias](http://www.researchgate.net/publication/256473910_Search_Engine_Bias) zuletzt abgerufen am 23. März 2016.
- Weber, Karsten; Drüeke, Ricarda (2010). Konvergente Medien, Fragmentierung der Öffentlichkeit und medienethische Anmerkungen. In *Zeitschrift für Kommunikationsökologie und Medienethik* 12(1), S. 68–72.



- White, Ryen (2013). Beliefs and Biases in Web Search. In *Proceedings of the 36th international ACM SIGIR conference on Research and development in information retrieval*. New York: ACM, S. 3–12. <http://research.microsoft.com/en-us/um/people/ryenw/papers/WhiteSIGIR2013.pdf> zuletzt abgerufen am 23. März 2016.
- White, Ryen; Horvitz, Eric (2015). Belief Dynamics and Biases in Web Search. In *ACM Transactions on Information Systems* ([In Press]). <http://research.microsoft.com/en-us/um/people/ryenw/papers/WhiteTOIS2015.pdf> zuletzt abgerufen am 23. März 2016.
- Yue, Yisong; Patel, Rajan; Roehrig, Hein (2010). Beyond position bias: examining result attractiveness as a source of presentation bias in clickthrough data. In *Proceedings of the 19th international conference on World wide web*. New York: ACM, S. 1011–1018. [http://www.yisongyue.com/publications/www2010\\_bias.pdf](http://www.yisongyue.com/publications/www2010_bias.pdf) zuletzt abgerufen am 23. März 2016.