



Hochschule für Angewandte Wissenschaften Hamburg  
*Hamburg University of Applied Sciences*

DEPARTMENT INFORMATION

## *Masterarbeit*

### **Welches Weltbild vermitteln Suchmaschinen? Untersuchung der Gewichtung inhaltlicher Aspekte von Google- und Bing-Ergebnissen in Deutschland und den USA zu aktuellen internationalen Themen**

*vorgelegt von*

**Markus Günther**

Studiengang Information, Medien, Bibliothek

erster Prüfer: Prof. Dr. Dirk Lewandowski

zweiter Prüfer: Prof. Dr. Steffen Burkhardt

Hamburg, August 2015



## **Abstract**

Vor dem Hintergrund stets vorhandener Suchmaschinenverzerrungen – allen Unterschieden eines tatsächlichen Suchergebnisses zum idealen – geht diese Arbeit der Frage nach, ob sich die von Google und Bing vermittelten Bilder aktueller internationaler Themen in Deutschland und den USA hinsichtlich Vollständigkeit, Abdeckung und Gewichtung der jeweiligen inhaltlichen Aspekte voneinander unterscheiden.

Um dies zu untersuchen, wurde eine Methode aus Ansätzen der empirischen Sozialwissenschaften, nämlich der Inhaltsanalyse, sowie der Informationswissenschaft, nämlich Retrievaltests, entwickelt und angewandt.

Die theoretischen Hintergründe zu Suchmaschinen als Kommunikationsmedien im Kontext des Internets bzw. Webs und Suchmaschinenverzerrungen werden in einem Literaturteil erläutert.

Für die Beantwortung der Forschungsfragen wurden am 28. April 2015 jeweils die Top-10 textlichen Ergebnisse von Google und Bing (jeweils unter .de- und .com-Domain) bzgl. 14 aus Tageszeitungen beider Regionen identifizierten Themen erfasst, und in Hinsicht auf das Vorhandensein sowie die Ausführlichkeit der jeweiligen inhaltlichen Aspekte bewertet. Von den 14 Themen wurden 9 ausgewertet.

Die Ergebnisse der empirischen Untersuchung zeigen, dass Google und Bing in Deutschland und den USA keine vollständigen Bilder aktueller internationaler Themen vermitteln, wobei es Unterschiede hinsichtlich der inhaltlichen Aspekte gibt; dass sie auf den ersten Trefferpositionen nicht die drei wichtigsten inhaltlichen Aspekte abdecken; und dass es bei der Gewichtung der inhaltlichen Aspekte keine signifikanten Unterschiede gibt.

Allerdings erfahren diese Ergebnisse diverse Einschränkungen und haben somit praktisch keine Aussagekraft. Dennoch kann die vorliegende Arbeit mit ihrer entwickelten Methode als Ausgangspunkt für weitere Forschung bzgl. inhaltlicher Suchmaschinenverzerrungen dienen.

**Keywords:** Suchmaschinen, Suchmaschinenverzerrungen, search engine bias, Google, Bing, Universal Search, Inhaltsanalyse, Retrievaltests

# Inhaltsverzeichnis

<b>1. Einleitung</b> .....	<b>S. 1</b>
1.1 Motivation .....	S. 2
1.2 Aufbau der Arbeit .....	S. 2
<b>2. Theoretischer Hintergrund</b> .....	<b>S. 3</b>
2.1 Kommunikation und Medien .....	S. 4
2.2 Massenkommunikation und Massenmedien .....	S. 8
2.3 Das Internet bzw. Web und Suchmaschinen als Medien .....	S. 19
2.4 Das Mediensystem Deutschlands .....	S. 25
2.5 Das Mediensystem der USA .....	S. 30
<b>3. Suchmaschinen</b> .....	<b>S. 40</b>
3.1 Suchmaschinen als technische Systeme .....	S. 42
3.2 Spezialsuchmaschinen .....	S. 48
<b>4. Stand der Forschung</b> .....	<b>S. 53</b>
4.1 Suchmaschinenverzerrungen .....	S. 53
4.2 Verwandte Forschung .....	S. 60
<b>5. Empirische Untersuchung</b> .....	<b>S. 69</b>
5.1 Forschungsfragen und -hypothesen .....	S. 70
5.2 Methodik .....	S. 71
5.2.1 Theorie .....	S. 71
5.2.2 Praxis .....	S. 76
5.3 Ergebnisse .....	S. 93
5.4 Diskussion, Einschränkungen und weitere Forschung .....	S. 104
<b>6. Fazit</b> .....	<b>S. 108</b>
<b>Literaturverzeichnis</b> .....	<b>S. 110</b>
<b>Anhang</b> .....	<b>S. 127</b>
A Beigabe: Inhalt der CD .....	S. 127
<b>Eidesstattliche Erklärung</b> .....	<b>S. 128</b>

## Abbildungsverzeichnis

Abb. 1: Medienevolution .....	S. 6
Abb. 2: Feldschema der Massenkommunikation (Maletzke) .....	S. 10
Abb. 3: Internet – ISO/OSI-Schichtenmodell .....	S. 20
Abb. 4: Das Mediensystem Deutschlands .....	S. 27
Abb. 5: Suchmaschinen-Referenzmodell .....	S. 43
Abb. 6: Universal Search – Prinzip .....	S. 51
Abb. 7: Universal Search – Google und Bing (April 2015) .....	S. 52
Abb. 8: Inhaltsanalyse – standardisierter Untersuchungsablauf (Früh) .....	S. 72
Abb. 9: Thema 7 – Häufigkeit der Gewichtungen nach Suchmaschinen .....	S. 88
Abb. 10: Thema 7 – Häufigkeit der Gewichtungen nach Regionen .....	S. 88
Abb. 11: Datenanalyse-Arten .....	S. 89
Abb. 12: Thema 7 – Mann-Whitney-Test Regionen – SPSS-Ergebnis .....	S. 91
Abb. 13: Thema 7 –Kruskal-Wallis-Test Suchmaschinen – SPSS-Ergebnis .....	S. 92
Abb. 14: Vollständige Bilder nach Suchmaschinen .....	S. 98
Abb. 15: Vollständige Bilder nach Regionen (Durchschnitte) .....	S. 98
Abb. 16: Abdeckung der drei wichtigsten Aspekte auf Position 1 nach Suchmaschinen .....	S. 99
Abb. 17: Abdeckung der drei wichtigsten Aspekte auf Position 1 nach Regionen (Durchschnitte) .....	S. 99
Abb. 18: Abdeckung der 75 Prozent wichtigsten Aspekte auf den ersten drei Positionen nach Suchmaschinen .....	S. 100
Abb. 19: Abdeckung der 75 Prozent wichtigsten Aspekte auf den ersten drei Positionen nach Regionen (Durchschnitte) .....	S. 100
Abb. 20: Häufigkeit der Gewichtungen nach Suchmaschinen .....	S. 101
Abb. 21: Häufigkeit der Gewichtungen nach Regionen .....	S. 101
Abb. 22: Mann-Whitney-Test Regionen – SPSS-Ergebnis .....	S. 102
Abb. 23: Kruskal-Wallis-Test Suchmaschinen – SPSS-Ergebnis .....	S. 102

## Tabellenverzeichnis

Tab. 1: Medienevolution – Zeitraffer .....	S. 6
Tab. 2: Online-Kommunikation nach Morris & Ogan (Beck) .....	S. 23
Tab. 3: Internet-Kommunikation nach Rössler (Beck) .....	S. 24
Tab. 4: Suchmaschinenverzerrungen google.de und bing.de (Thema 7) .....	S. 54
Tab. 5: Gemeinsame Themen deutscher und amerikanischer Zeitungen .....	S. 77
Tab. 6: Zeitungsthemen-Suchanfragen .....	S. 78
Tab. 7: Thema 7 – inhaltliche Aspekte .....	S. 82
Tab. 8: Thema 7 – Vorkommen und Ausführlichkeit der inhaltlichen Aspekte .....	S. 84
Tab. 9: Thema 7 – Identifizierung der jeweils drei wichtigsten Aspekte .....	S. 85
Tab. 10: Thema 7 – Wichtige Aspekte mit der Häufigkeit ihres Auftretens .....	S. 85
Tab. 11: Thema 7 – Abdeckung der wichtigen Aspekte .....	S. 86
Tab. 12: Thema 7 – Anzahlen abgedeckter wichtiger Aspekte (kumuliert) .....	S. 86
Tab. 13: Thema 7 – Abdeckung der drei wichtigsten Aspekte auf Trefferposition 1 .....	S. 87
Tab. 14: Thema 7 – Abdeckung der 75 Prozent wichtigster Treffer auf den ersten drei Positionen .....	S. 87
Tab. 15: Thema 7 – Häufigkeiten der Gewichtungen nach Suchmaschinen und Regionen .....	S. 88
Tab. 16: Häufigkeiten der Gewichtungen nach Suchmaschinen und Regionen .....	S. 100
Tab. 17: Mann-Whitney- und Kruskal-Wallis-Tests Einzelthemen – Ergebnisse .....	S. 103
Tab. 18: Post-Hoc-Tests Thema 2, 6 und 14 – Ergebnisse .....	S. 103

## Abkürzungsverzeichnis

ABC	American Broadcasting Company
AEUV	Vertrag über die Arbeitsweise der Europäischen Union
afp	Agence France Press
AOL	America Online
AP	Associated Press
ARD	Arbeitsgemeinschaft der öffentlich-rechtlichen Rundfunkanstalten der Bundesrepublik Deutschland
BPjM	Bundesprüfstelle für jugendgefährdende Medien
CBS	Columbia Broadcasting System
CDA	Communication Decency Act
CNN	Cable News Network
DEFA	Deutsche Film AG
DLR	DeutschlandRadio
dpa	Deutsche Presseagentur
dpad	Deutscher Auslandsdepeschendienst
DTV	Digital Television
EU	Europäische Union
FAZ	Frankfurter Allgemeine Zeitung
FCC	Federal Communications Commission
FOX	Fox Broadcasting Company
FSF	Freiwillige Selbstkontrolle Fernsehen e.V.
FSK	Freiwillige Selbstkontrolle der Filmwirtschaft GmbH
FSM	Freiwillige Selbstkontrolle Multimedia-Dienstleister e.V.
GG	Grundgesetz
GWB	Gesetz gegen Wettbewerbsbeschränkungen
HBO	Home Box Office
HDTV	High Definition Television
HTML	Hypertext Markup Language
ISO	Internationale Organisation für Normung
MOO	Multi User Dungeon/Dimension/Domain, object oriented
MTV	Music Television
MUD	Multi User Dungeon/Dimension/Domain
NBC	National Broadcasting Company

NPR	National Public Radio
OSI	Open System Interconnection
PBS	Public Broadcasting System
RAT	Relevance Assessment Tool
rtr	Thomson Reuters
SERP	Search engine result page
SPSS	urspr. Statistical Package for the Social Sciences
TREC	Text Retrieval Conference
UFA	ehem. Universum Film AG, heute UFA GmbH
UKA	analoge Ultrakurzwellen
UP	United Press International
URL	Uniform Resource Locator
USA	United States of America
W3	World Wide Web
WSJ	Wall Street Journal
WWW	World Wide Web
ZDF	Zweites Deutsches Fernsehen

*[Search engines] are like windows onto the web—and, like windows, tend to be largely unnoticed because our gaze focuses on what is visible through them. With windows, however, it is easy to detect when they are cloudy or distorted. With search engines, however, it is much more difficult to tell when they are providing distorted or incomplete pictures.*

– Lawrence Hinman, *Esse est indicato in Google: Ethical and Political Issues in Search Engines*

## **1. Einleitung**

In diesem Kapitel wird das Thema dieser Arbeit eingeführt, sowie ihre Motivation und ihr Aufbau erläutert.

Vorweg sei erwähnt, dass es allgemeine, *Spezial-*, *Hybrid-* und *Metasuchmaschinen* gibt (vgl. Lewandowski 2015, S. 16f). Für diese Arbeit sind allgemeine und Spezialsuchmaschinen von Interesse (s. Kap. 3). Wenn hier von ‚Suchmaschinen‘ die Rede ist, sind allgemeine Suchmaschinen gemeint – ansonsten wird spezifiziert.

Suchmaschinen sind jene Form des Zugangs zu Informationen im *World Wide Web* (kurz *Web*), die sich weitgehend durchgesetzt hat (vgl. Lewandowski 2015, S. 8). Sie haben eine enorme gesellschaftliche Bedeutung (vgl. ebd., S. 2) und sind der beliebteste Dienst des Internets – von 82 Prozent aller deutschen Internetnutzer werden sie mindestens wöchentlich genutzt (vgl. van Eimeren & Frees 2014, S. 386f; 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 & Frees 2014, S. 386f) – dabei sind Suchmaschinen für 42 Prozent der mit Abstand wichtigste Zugangsweg (vgl. Hölig & Hasebrink 2014, S. 535). Bereits 2003 wurde eingeschätzt, dass das Internet neben den klassischen Massenmedien Print, Hörfunk und Fernsehen „seine Funktion als viertes tagesaktuelles Medium gefunden hat“ (van Eimeren et al. 2003, S. 358). Im Zeitraum 2009 bis 2013 hat das Internet insgesamt insbes. zulasten von Fernsehen und Tageszeitung an Bedeutung gewonnen (vgl. TNS Infratest 2013, S. 8f, 30f): Online-Nachrichten ersetzen für manchen Nutzer den Kauf von Tageszeitungen (vgl. van Eimeren et al. 2003, S. 348).

Allerdings sind die Ergebnisse von Suchmaschinen nicht ‚neutral‘, sondern immer in irgendeiner Weise verzerrt – das liegt schon in der Funktionsweise von Suchmaschinen begründet (vgl. Weber

2011, S. 277). Man spricht hierbei von *Suchmaschinenverzerrungen* (*search engine bias*). Unter diesen Begriff fallen alle Unterschiede eines tatsächlichen Suchergebnisses zum idealen, das aus sämtlichen potenziell relevanten Dokumenten in einer allein nach objektiven Kriterien erfolgten und damit richtigen Reihung besteht (vgl. Lewandowski 2015, S. 272).

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

Um dies zu untersuchen, wurde für die vorliegende Arbeit eine Methode aus Ansätzen der empirischen Sozialwissenschaften und der Informationswissenschaft, genauer Inhaltsanalyse und Retrievaltests, entwickelt und angewandt.

## **1.1 Motivation**

In diesem Abschnitt wird die Motivation der vorliegenden Arbeit erläutert.

Die Motivation zu dieser Arbeit ergibt sich grundlegend aus den bereits im vorigen Kapitel geschilderten Umständen: Dadurch, dass Suchmaschinen zunehmend genutzt werden, um sich über das aktuelle Tagesgeschehen zu informieren, kommt ihnen eine besondere Verantwortung zu, der allerdings grundsätzlich die stets vorhandenen Suchmaschinenverzerrungen im Weg stehen. Zudem sind (wie in Kap. 4.2 zu sehen sein wird) unter den vielfältigen Suchmaschinenverzerrungen die inhaltlichen bislang vergleichsweise wenig erforscht – u.a. sicherlich deshalb, weil es sich dabei um eine Überschneidung der empirischen Sozialwissenschaften und der Informationswissenschaft handelt. Daher ist es reizvoll, zu diesem Thema eine Methode als möglichen Ausgangspunkt für weitere Forschung zu entwickeln und erste Ergebnisse zu erzielen.

## **1.2 Aufbau der Arbeit**

In dieser Sektion wird der Aufbau dieser Arbeit dargelegt.

Die vorliegende Arbeit gliedert sich grundlegend in drei Teilbereiche: den theoretischen Hintergrund, Suchmaschinen und den Stand der Forschung, und die empirische Untersuchung.

Im ersten Teilbereich (Kap. 2) wird sich Massenkommunikation und Massenmedien über Kommunikation und Medien genähert, bevor versucht wird, das Internet bzw. Web und Suchmaschinen als Medien zu verorten. Anschließend wird ein Blick auf die Mediensysteme Deutschlands und der USA geworfen.

Der zweite Teilbereich (Kap. 3 und 4) dreht sich um Suchmaschinen und Suchmaschinenverzerrungen, und es werden kurz einige Arbeiten zu diesem Thema vorgestellt.

Im dritten Teilbereich (Kap. 5) schließlich geht es um die empirische Untersuchung: Dabei werden die Forschungsfragen und -hypothesen dargelegt, es wird die Methodik erläutert, und abschließend werden die Ergebnisse präsentiert und diskutiert. Das zugehörige Material und die Einzelauswertungen der Untersuchung finden sich auf der beiliegenden CD.

Am Ende jedes Teilbereichs folgt ein Fazit, das jeweils die wichtigsten Punkte noch einmal zusammenfasst.

## **2. Theoretischer Hintergrund**

In dieser Arbeit geht es um Suchmaschinen. Eine Suchmaschine ist ein Dienst des Internets, und beide sind auf ihre Weise Kommunikationsmedien, und zwar vergleichsweise junge – es geht also grundlegend um Kommunikation. Dies ist die Perspektive dieses Kapitels.

Nach Merten (2007, S. 14f) unterscheidet die Kommunikationswissenschaft Kommunikationsformen in einer Dimension in *informell* (mündlich, personal, *face to face*) oder *medial*. Dabei wird unterstellt, dass Kommunikation und Massenkommunikation grundsätzlich die gleiche Struktur aufweisen resp. nach den gleichen Gesetzen funktionieren. Für eine Analyse von Kommunikation muss also bei der einfachsten Form, der informellen, angesetzt werden. So wird in diesem Kapitel zuerst auf Kommunikation und Medien eingegangen, um über Massenkommunikation und Massenmedien schließlich zum Internet bzw. Web und Suchmaschinen zu gelangen.

Anschließend werden die Mediensysteme Deutschlands und der USA, der beiden in der empirischen Untersuchung berücksichtigten Regionen, betrachtet.

Ziel dieses Kapitels ist dabei nicht, alle diese Themenbereiche erschöpfend zu behandeln und den aktuellen Stand der Forschung zu vermitteln, sondern es soll den theoretischen Unterbau für die in der restlichen Arbeit eher technischen Betrachtungen und die empirische Untersuchung darstellen.

Kap. 2.5 schließt mit einem Zwischenfazit, das die wichtigsten Punkte des ersten Teilbereichs noch einmal zusammenfasst.

## 2.1 Kommunikation und Medien

In diesem Abschnitt wird erläutert, was Kommunikation ist und wie sie funktioniert. Dabei kommt Medien eine besondere Rolle zu.

Beim Versuch, den Begriff (informeller) Kommunikation zu fassen bzw. zu definieren, stößt man auf erhebliche theoretische und methodische Schwierigkeiten. Merten führt in diesem Zusammenhang die Eigenschaften Profanität, Universalität, Flüchtigkeit, Relationalität und Unvermeidbarkeit an (vgl. 2007, S. 15ff). Zusammenfassend spricht er von

flüchtige[n] Prozesse[n], von denen man nur dann eine Reproduktion, einen Abdruck erzeugen kann, wenn man erneut Kommunikation initiiert, relationale[n] Prozesse[n], bei denen man keine Substanzen, keine Elemente dingfest machen kann, Phänomene[n], bei denen die Regel nicht durch die Ausnahme befestigt werden kann (ebd., S. 17).

Letztlich plädiert er sogar für die Verwendung einer *transklassischen Wissenschaft* (vgl. Günther 1963, S. 14), da die Analyse des Kommunikationsprozesses mittels des Kausalprinzips unbefriedigend bleibt (vgl. Merten 2007, S. 18f).

Davon abgesehen sieht Merten Kommunikation am ehesten konsensfähig mit dem Begriff der (sozialen) Interaktion (vgl. 1977, S. 88f; auch z.B. Jarren & Bonfadelli 2001, S. 19f; Pürer 2003, S. 59f; Burkart 2002, S. 30). Damit ist ein wechselseitiges Geschehen zwischen zwei oder mehr Personen gemeint, „welches mit einer Kontaktaufnahme (aufgrund von / oder verbunden mit wechselseitiger Wahrnehmung) beginnt und zu (Re-)Aktionen der im Kontakt stehenden Lebewesen führt“ (Burkart 2002, S. 30 nach Burghardt 1972, S. 42). Demnach kann Kommunikation als eine „spezifische Form der sozialen Interaktion“ verstanden werden (Graumann 1972, S. 1110; Burkart 2002, S. 30). Merten präzisiert diesen Ansatz über sechs Kommunikationseigenschaften (vgl. 1977, S. 74-88):

- Reziprozität (Wechselseitigkeit)
- Intentionalität
- Anwesenheit (gegenseitige Wahrnehmbarkeit)
- Sprachlichkeit (Symbole: Zeichen, die für etwas (Gemeintes) stehen (vgl. Burkart 2002, S. 63); Stichwort *Symbolischer Interaktionismus* (nach Mead 1934))
- Wirkung
- Reflexivität (Nicht-Kausalität; zeitlich, sachlich und sozial)

Von diesen sei noch einmal die Sprachlichkeit herausgegriffen.

In einem allgemeinen Verständnis definiert Maletzke Kommunikation als „Bedeutungsvermittlung zwischen Lebewesen“ (1963, S. 18). Nach Burkart (2002, S. 35f) impliziert hierbei der Begriff ‚Vermittlung‘, dass Kommunikation immer eine Instanz, fachspezifisch ein Medium, z.B. Sprache, braucht, durch das Bedeutungsinhalte Gestalt annehmen. Demnach scheinen Medien auf den ersten Blick schlicht Transportmittel für die auszutauschenden Bedeutungsinhalte zu sein. Diese *Container-Metapher* (Krippendorff 1994, S. 86) ist allerdings überholt; Kommunikation wird heute als „Aktualisierung von Sinn“ (Luhmann 1971, S. 32) begriffen, d.h. dass (im besten Fall ähnliche) Bedeutungsinhalte im Bewusstsein der jeweiligen Kommunikationspartner wachgerufen werden. Dennoch bedarf „alle Kommunikation [...] eines Mittels oder Mediums, durch das hindurch eine Nachricht [...] aufgenommen wird“ (Graumann 1972, S. 1182). Demnach steht ‚Medium‘ sowohl für personale Vermittlungsinstanzen als auch für technische Hilfsmittel zur Übertragung einer Botschaft.

In einer weit verbreiteten Differenzierung gliedert Pross die mediale Vielfalt menschlicher Kommunikation in drei Gruppen (vgl. 1972, S. 127f):

- primäre Medien: Medien des menschlichen Elementarkontaktes, z.B. Sprache, Mimik, Gestik – es ist kein Gerät zwischen Sender und Empfänger geschaltet
- sekundäre Medien: erfordern auf der Produktionsseite ein Gerät, z.B. Print
- tertiäre Medien: erfordern technische Sender und Empfänger, z.B. Telefon, elektronische Massenmedien

Geht man (abgesehen von den Medien des menschlichen Elementarkontaktes) von Schrift als dem ersten Medium aus und rafft die Zeit ab ihrer Entwicklung bis ungefähr heute in einer einzigen Stunde zusammen, zeigt sich, dass mehr als 55 Minuten lang nichts passierte. Dann allerdings überschlug sich die Entwicklung: Die Zeitung wurde erfunden, dann der Telegraf, das Telefon, das Radio, das Fernsehen usw. bis hin zum Internet (und tatsächlich darüber hinaus) – alles innerhalb von knapp fünf Minuten. Dies ist in Tab. 1 dargestellt.

Tab. 1: Medienevolution – Zeitraffer (Merten 2007, S. 20)

Medium	Jahr der Erfindung	Stundenzeit
Schrift	3000 v. Chr.	00'00"
Presse	1609	55'24"
Elektrischer Telegraph	1809	57'48"
Telefon	1872	58'33"
Film	1895	58'51"
Radio	1918	59'07"
Fernsehen	1931	59'17"
Tonband	1951	59'31"
Satellitenfernsehen	1971	59'46"
Fax	1979	59'52"
Btx	1980	59'53"
CD	1983	59'55"
CD ROM	1988	59'57"
Internet (WWW)	1992	59'59"

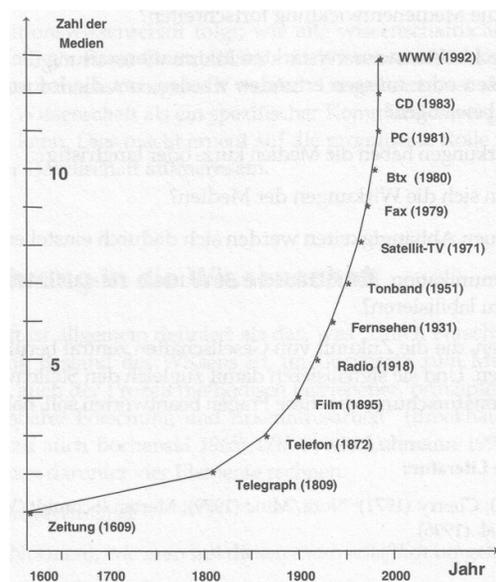


Abb. 1: Medienevolution (Merten 2007, S. 21)

Abb. 1 zeigt die Zahl der Medien in Abhängigkeit vom Jahr ihrer jeweiligen Entstehung anhand eines Graphen, der nahezu perfekt einer Evolutionskurve folgt. Merten (2007, S. 21) spricht hierbei von einem ‚Megatrend zur Mediengesellschaft‘, der das Bewusstsein des Menschen mit einem exponentiell gestiegenen Informationsangebot konfrontiert und ihn zu immer stärkerer Selektion zwingt.

Überdies fördert(e) die technische Entwicklung Ende des 20. / Anfang des 21. Jahrhunderts (die *Digitalisierung*), die *Medienkonvergenz*, d.h. die Verschmelzung von Telekommunikation, Computer, Unterhaltungselektronik und Rundfunk, wofür heute der Begriff *Multimedia* steht (vgl. Wilke 2009, S. 329; Pürer 2003, S. 266; Jarren & Bonfadelli 2001, S. 468; Dreier 2009, S. 257, 259). Als Folge dieser Entwicklung können Medien heute immer weniger über Eigenschaften bestimmter Geräte definiert werden als eher über die Dienste, für die man sie in Anspruch nimmt. Die Inanspruchnahme erfolgt zunehmend im Rahmen computerbasierter Vernetzungen, z.B. via Internet (vgl. Burkart 2002, S. 37f). Eine Bezeichnung für diese (neuen) Medien lautet in Anlehnung an Pross:

- quartäre Medien (Burkart 2002, S. 37f; Schulz 2009, S. 171): auch ‚digitale‘ oder ‚Online‘-Medien; die traditionelle Sender-Empfänger-Beziehung wird vielfach aufgeweicht, z.B. Online-Zeitungen, Suchmaschinen, E-Mail

Nach Burkart (vgl. 2002, S. 38f) verhelfen also Medien Mitteilungen dazu, in Erscheinung zu treten, und bestimmen auch die Form, in der dies geschieht (z.B. gesprochen oder geschrieben); sie stellen

unterschiedliche Rahmen bereit, innerhalb derer dann jeweils bestimmte Ausdrucksformen als Zeichen, z.B. Symbole, fungieren können. Das ist es auch, worauf McLuhan mit seinem bekannten Ausspruch „The medium is the message“ (1997, S. 7) aufbaut, wenn er sagt, dass nicht ausschlaggebend ist, was wir mitteilen (Inhalt), sondern wie (Form, die durch das Medium geprägt wird) (dabei geht es ihm allerdings mehr um die gesellschaftlich-kulturellen Wirkungsweisen als um die individuellen Wirkungen und Folgen von Kommunikation und Medien) (vgl. Beck 2013, S. 82).

Es existiert (noch) keine für die gesamte (Kommunikations-)Wissenschaftlergemeinschaft verbindliche Definition des Begriffs ‚Medium‘ (vgl. Burkart 2002, S. 40; Merten 2007, S. 81; Beck 2013, S. 14), doch eine vielzitierte Nominaldefinition von Saxer aus Publizistik-Perspektive lautet: „Medien sind komplexe institutionalisierte Systeme um organisierte Kommunikationskanäle von spezifischem Leistungsvermögen“ (1998, S. 54 urspr. in 1980, S. 532). Faulstich ergänzt: „... mit gesellschaftlicher Dominanz“ (2004b, S. 27). Hierbei werden also vier Bedeutungsdimensionen unterschieden:

1. ‚komplexe institutionalisierte Systeme‘: Medienkommunikation resultiert aus „Herstellungs-, Bereitstellungs- und Empfangsprozessen“ (Saxer 1998, S. 55), daher bilden Medien mehr oder weniger komplexe soziale Gebilde (z.B. Organisationen, bspw. eine Lokalzeitung oder ein großer Fernsehsender). Institutionalisiert sind sie in dem Sinne, dass sie allgemein bekannt sind und von vielen Menschen genutzt und akzeptiert werden (vgl. Faulstich 2004a, S. 12).
2. ‚organisierte Kommunikationskanäle‘: Nicht jeder Kommunikationskanal ist ein Medium in publizistischem Sinne; so wird man z.B. Fernsehkanäle nicht als Medien, sondern eher als Übertragungstechniken, als eine Art kommunikative Infrastruktur begreifen, mit der das Medium Fernsehen gesendet/empfangen werden kann (vgl. Burkart 2002, S. 44f). In diesem Zusammenhang erscheint es sinnvoll, Medien nach *erster* und *zweiter Ordnung* zu unterscheiden (vgl. Joerges & Braun 1994, S. 19; Kubicek et al. 1997, S. 32ff, 218ff): Medien erster Ordnung wären dann technische Möglichkeiten der Vermittlung, des Speicherns und/oder Abrufens von Mitteilungen wie z.B. auch das Internet (vgl. Beck 2013, S. 86; Beck 2012, S. 261f) oder der Computer. Medien zweiter Ordnung wären die Vermittlungstechniken, die von institutionalisierten Kommunikatoren zur Herstellung und Verbreitung von Inhalten benutzt werden.
3. ‚spezifisches Leistungsvermögen‘: Dies meint die Besonderheit (etwa die technischen Möglichkeiten) jedes Mediums im Verhältnis zu den anderen; die Bandbreite seiner charakteristischen Merkmale (vgl. Burkart 2002, S. 42; Faulstich 2004a, S. 12).

4. ‚gesellschaftliche Dominanz‘: „Jedes Medium unterliegt einem geschichtlichen Wandel und insofern einer Veränderung seiner Bedeutung“ (Faulstich 2004a, S. 12). Dominanz bedeutet hier Relevanz für die Gesellschaft.

Abschließend lassen sich vier Faktoren abstrahieren, die jedem Kommunikationsprozess zugrundeliegen (vgl. Burkart 2002, S. 63f):

- Jemand, der etwas mitteilen will, fachspezifisch *Kommunikator*: die Quelle ausgesendeter Botschaften; er produziert die Mitteilung i.d.R. und adressiert sie an jemanden
- Die Aussage/Botschaft (die mitzuteilenden Bedeutungsinhalte) selbst
- Ein Medium (eine Instanz, mit deren Hilfe der mitzuteilende Inhalt ‚transportabel‘ wird)
- Jemand, an den die Botschaft gerichtet ist, fachspezifisch *Rezipient*: an ihn ist die ausgesendete Botschaft adressiert; er empfängt oder konsumiert sie

## 2.2 Massenkommunikation und Massenmedien

In dieser Sektion wird erläutert, was Massenkommunikation und Massenmedien sind, und was es für Ansätze gibt, ihre Rolle innerhalb der Gesellschaft zu begreifen. Dabei wird als ein Beispiel insbes. auf die soziologische Systemtheorie nach Luhmann eingegangen.

Allgemein kann Massenkommunikation als eine spezifische Form öffentlicher Kommunikation verstanden werden, die durch Massenmedien vermittelt wird; klassisch sind dies die Bereiche Presse, also Zeitung und Zeitschrift, und Rundfunk, also Hörfunk und Fernsehen, außerdem Buch und Film (Kino, Video) (vgl. Beck 2013, S. 129; Pürer 2003, S. 74f; Jarren & Bonfadelli 2001, S. 22; Beck 2012, S. 338).

Der Begriff stammt vom englischen *mass communication* und etablierte sich in den 1960er Jahren (davor war *Publizistik* für an die Öffentlichkeit gerichtete Aussagen und Botschaften üblich) (vgl. Pürer 2003, S. 74).

Der Term ‚Masse‘ beinhaltet fachsprachlich keine massenpsychologischen oder kulturkritischen Assoziationen – er meint lediglich, dass sich die vermittelten Aussagen an eine Vielzahl von Menschen richten, die man angemessener als *Publikum* bezeichnet (vgl. ebd., S. 76f; Burkart 2002, S. 167ff). Maletzke (1963, S. 28f) hat hierfür den Begriff *dispersedes Publikum* geprägt, womit einzelne oder auch kleine Gruppen von Menschen gemeint sind, die sich räumlich und ggf. zeitlich getrennt einem gemeinsamen Gegenstand, nämlich den Aussagen der Massenkommunikation, zuwenden (im Unterschied zum *Präsenzpublikum*, das räumlich versammelt und denselben situativen Bedingungen

ausgesetzt ist, und dessen Interessen, Sinne und Erwartungen i.d.R. gleichgerichtet sind (z.B. im Kino)).

Der Term ‚Kommunikation‘ suggeriert, die Rezipienten könnten mit dem Kommunikator ‚kommunizieren‘ – das ist hier allerdings nur sehr eingeschränkt möglich. Massenkommunikation ist nicht an eine Person gerichtet, sondern eher an einen breiten Querschnitt einer Bevölkerung (vgl. Pürer 2003, S. 77f) – man kann sie daher grundsätzlich als öffentlich verstehen (vgl. Burkart 2002, S. 171). Hinzu kommt eine räumliche Distanz (z.B. bei Live-Sendungen) oder sogar eine raum-zeitliche Trennung von Rezipienten und Kommunikator (z.B. beim Zeitunglesen) (vgl. ebd., S. 170). Weiter kommt es bei Massenkommunikation i.d.R. zu einer Polarisierung der kommunikativen Rollen: „Es fehlt [weitestgehend – Erg. Pürer 2003, S. 78] der – für die direkte zwischenmenschliche Begegnung so typische – Rollentausch zwischen den Kommunikationspartnern“ (Burkart 2002, S. 170). Nach Pürer (2003, S. 78) schließt klassische Massenkommunikation *Feedback* (Rückkopplungen) zwar nicht grundsätzlich aus, doch erfolgt dies i.d.R. über Leserbriefe, Anrufe u.Ä., und wirkt sich, wenn überhaupt, erst mit Verzögerung auf das Kommunikationsverhalten aus – die beteiligten Kommunikationspartner tauschen also nicht grundsätzlich ihre Rollen. Damit ist Massenkommunikation in erster Linie Übertragung und nur selten Austausch.

Zusammengefasst finden sich diese Punkte in der im deutschen Sprachraum am weitesten verbreiteten Begriffsdefinition (vgl. Pürer 2003, S. 80) von Maletzke:

Unter Massenkommunikation verstehen wir jene Form der Kommunikation, bei der Aussagen öffentlich (also ohne begrenzte und personell definierte Empfängerschaft) durch technische Verbreitungsmittel (Medien) indirekt (also bei räumlicher oder zeitlicher oder raumzeitlicher Distanz zwischen den Kommunikationspartnern) und einseitig (also ohne Rollenwechsel zwischen Aussagendem und Aufnehmendem) an ein disperses Publikum (im soeben erläuterten Sinne [s.o.]) vermittelt werden (1963, S. 32).

Maletzke entwickelte auch das *Feldschema der Massenkommunikation*, ein Massenkommunikationsmodell, das „bis heute ein weithin anerkannter Versuch [ist], alle wesentlichen Faktoren der ‚Massenkommunikation‘ zusammenzufassen“ (Beck 2013, S. 134; vgl. Rau 2013, S. 87). Es stellt Abb. 2 dar.

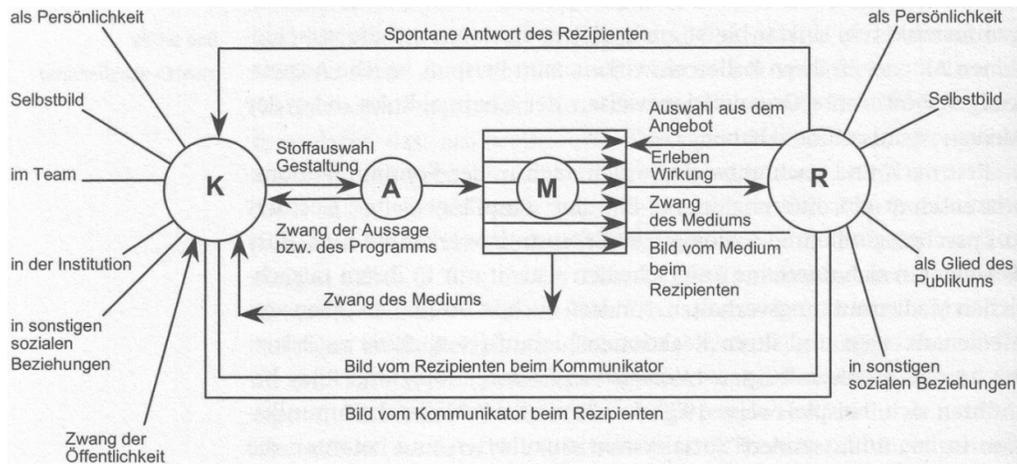


Abb. 2: Feldschema der Massenkommunikation (Maletzke) (1963, S. 41)

Das Besondere an diesem Modell ist nach Merten (2006, S. 199ff), dass erstmals die kognitiven Referenzen des Symbolischen Interaktionismus' (nach Mead 1934) – *Kommunikanten* (erstmalig Kommunizierende) bringen Einstellungen, Vorstellungen, Erfahrungen und insbes. Erwartungen an den/die Kommunikationspartner mit – eingearbeitet sind, und dass die Öffentlichkeit mitgedacht wird und als eine weitere soziale Rahmung zur Realitätsnähe beiträgt. Außerdem verweisen zwei Begrifflichkeiten auf etwas (damals) grundlegend Neues, nämlich *Selektivität* (das Auswahlverhalten) aller Kommunikation und *Selbstreferenz*. Selektivität bedeutet, dass der Rezipient ein Informationsangebot niemals 1:1 übernimmt, sondern stets selektive Operationen (Interpretation, Deutung etc.) durchführt. Selbstreferenz meint, dass der Kommunikationsprozess eine zirkuläre Struktur aufweist. In Abb. 2 stellt bspw. die ‚Spontane Antwort des Rezipienten‘ ein Feedback dar, aus dem der Kommunikator laufend lernt, wie seine Kommunikation wirkt. Das Bild vom Rezipienten beim Kommunikator ist eine andere Form des Feedbacks. Außerdem gibt es neben dem Fremd- auch ein Selbstbild.

Doch es gibt es eine Vielzahl theoretischer Ansätze zur Massenkommunikation aus unterschiedlichen Perspektiven und mit unterschiedlichen Ansprüchen. Einen Eindruck vermittelt eine Gliederung von Merten (vgl. 2007, S. 146-183):

- Publizistikwissenschaftliche Ansätze: zunächst im Rahmen der Zeitungswissenschaft
- Informationstheoretische Ansätze: nach syntaktischer, semantischer oder pragmatischer Definition von ‚Information‘
- Nachrichtentheorie: wann ein Ereignis zu einer Nachricht taugt (12 Nachrichtenfaktoren)

- Funktionalistische Ansätze: machen sich an den Leistungen, Aufgaben und Folgen von Massenkommunikation fest
- Politische Theorien der Massenmedien: *molekulare* (selbstbezogene) und *molare* (integrierte) Ansätze; fragen nach dem Wesen von Medien bzw. nach ihrer Funktion für größere Systeme (nicht zu verwechseln mit den u.g. umfassenderen und wirklichkeitsbezogenen Ansätze der systemischen und kritischen Theorien)
- Kritische Medientheorie: molar; die Analyse und Erklärung von Massenkommunikation findet dialektisch vor dem Hintergrund einer marxistischen Theorie der gesellschaftlichen Entwicklung statt; die Produktion von Medienaussagen wird im Rahmen kapitalistischer Verwertungszusammenhänge begriffen. Die wohl prominenteste Weiterentwicklung ist die *Theorie des kommunikativen Handelns* von Habermas.
- Materialistische Medientheorie: geht wie die kritische Medientheorie von einem marxistischen, historisch-dialektischen Ansatz aus, ist aber radikaler. Die unterschiedlichen Ansätze thematisieren vor allem den Zusammenhang zwischen Aussagenproduktion, ökonomischen Verhältnissen und Bewusstsein der Rezipienten, und setzen die Produktion von Kommunikationsinhalten mit kapitalistischer Warenproduktion gleich.
- Systemische Theorie der Massenkommunikation: unterschiedliche Ansätze entstanden aus der Systemtheorie selbst und lassen sich in zwei Dimensionen unterscheiden: den analytischen Grad, mit dem der Systembegriff verwendet wird, und die Analyse-Ebene, auf der eine Theorie der Massenkommunikation (bzw. ein Kommunikationssystem) angesiedelt ist. Analytischer Grad:
  - nominale Ebene: Massenkommunikation wird bspw. als Kommunikationssystem begriffen, „das für die Gesellschaft die Funktion der Information, Meinungsbildung und Unterhaltung erfüllt, Teil des gesamten Kommunikationssystems und Teilsystem der Gesellschaft (soziokulturelles System) ist“ (Merten 2007, S. 173).
  - deskriptive Ebene: Ansätze stützen sich z.B. auf die kybernetische Systemtheorie und die funktional-strukturelle Systemtheorie von Luhmann (1970) und betten z.B. das System der Massenkommunikation in ein Mediensystem ein, das als Subsystem der Gesellschaft verstanden wird.
  - analytische (theoretisch fundierte) Ebene: Entwürfe, die den Systembegriff nicht nur in seiner Relation zur Umwelt begreifen, sondern die diesbzgl. Mechanismen untersuchen. Prominent ist Luhmanns *funktional-strukturelle Systemtheorie* (Luhmann 1996).

Analyse-Ebene:

- gesamtgesellschaftliche Ebene: Herstellung von Konsens durch laufende Verbreitung gleicher Inhalte bzw. Themen
- Ebene der Teilsysteme: verbindet kommunikativ alle Teilsysteme mit allen Teilsystemen
- Ebene der Organisation: Untersuchung des Handelns kommunikativer Einrichtungen (z.B. von Redaktionen)
- Weitere Theorien der Massenkommunikation: entstehen während der Entwicklung der Kommunikationswissenschaft und greifen einzelne Aspekte neu auf (z.B. Zeit und Geschwindigkeit)
- Verknüpfung informeller und medialer Kommunikation: die erste systematische wurde 1940 von Lazarsfeld (et al. 1944) sozusagen zufällig entdeckt, als er die Wirkung der Massenmedien messen wollte und herausfand, dass die *Face-to-face*-Kommunikation wichtiger ist (*Two-step-flow*-Modell: nicht die Massenmedien üben die Wirkung aus, sondern bestimmte Personen, die *Opinion leaders*. Später fand eine Weiterentwicklung zum *Multi-step-flow*-Modell statt.)

Weitverbreitet von diesen Ansätzen ist u.a. die umfangreiche und verzweigte *Theorie des kommunikativen Handelns* von Habermas (Kritische Medientheorie) mit ihren Wurzeln in der *Kritischen Theorie der Frankfurter Schule*. Seinem Ansatz stellte sich später Luhmann mit seiner *funktional-strukturellen Systemtheorie* (Systemische Theorie) entgegen, und „beide Sozialphilosophen entwickelten ihre Theorien in den nächsten Jahrzehnten weiter, beide rieben sich aneinander und fanden so auch die Motivation, ihre jeweiligen Theorien genauer auszuarbeiten“ (Horster 1999, S. 33). An dieser Stelle sei daher noch einmal Luhmanns abstrakterer Ansatz herausgegriffen, der auch bei einer nur kurzen Betrachtung zeigt, als wie gewichtig und weitreichend die gesellschaftliche Rolle von Massenkommunikation und -medien gesehen werden kann.

Im Folgenden werden also Massenmedien in der Sichtweise der soziologischen Systemtheorie nach Luhmann als soziales System vorgestellt. Systemtheoretische Begriffe und Zusammenhänge werden dabei allerdings nicht immer bzw. bis ins Letzte verfolgt – das würde eine Betrachtung sozialer Systeme und der soziologischen Systemtheorie nach sich ziehen, die hier weder erfolgen kann noch soll. Es werden daher eher grob einige wichtige Aspekte herausgegriffen, die einen Eindruck dieser Sichtweise vermitteln sollen.

Die allgemeine Systemtheorie ist ein weitverzweigter, heterogener, interdisziplinärer Rahmen, der den Begriff ‚System‘ als Grundkonzept führt (vgl. Kneer & Nassehi 1993, S. 17, 21). Ganz allgemein

geht es um Elemente und ihre Relationen einer Ganzheit (vgl. Krieger 1998, S.12; Kneer & Nassehi 1993, S. 18, 21).

Die soziologische Systemtheorie hat den Anspruch einer Universaltheorie, d.h. sie behandelt alles Soziale (und ist somit auch selbstreferenziell) (vgl. Luhmann 2002, S. 9f; Kneer & Nassehi 1993, S. 33, 37).

Als ihre wichtigsten Vertreter gelten Parsons, „entscheidender soziologischer Lehrer und Gewährsmann Luhmanns“ (ebd., S. 29), und Luhmann.

Es sei noch angemerkt, dass Merten in seiner Gliederung (s.o.) mit dem Term ‚funktional-strukturelle‘ Systemtheorie explizit Luhmanns Frühwerk anspricht, das in Abgrenzung zu Parsons *strukturell-funktionaler Systemtheorie* entstand, doch wird in der folgenden Betrachtung auch das Spätwerk (u.a. das Prinzip der *Autopoiesis*) berücksichtigt, um einen Gesamteindruck zu vermitteln.

In der soziologischen Systemtheorie nach Luhmann bedeutet der Begriff *funktionale Differenzierung*, dass sich innerhalb eines sozialen Systems einzelne Teilsysteme herausbilden, die jeweils eine bestimmte Funktion für das Gesamtsystem erfüllen und daher auch *Funktionssysteme* genannt werden (vgl. Luhmann 1990, S. 479). Die Massenmedien sieht Luhmann als eins der Funktionssysteme der modernen Gesellschaft, das wie alle anderen (z.B. Wirtschaft, Politik und Recht) *operativ geschlossen* und *autopoietisch autonom* ist (vgl. Luhmann 2004, S. 21f).

Das Kunstwort ‚Autopoiesis‘ (aus griech. ‚autos‘ = selbst, und ‚poiein‘ = machen/hervorbringen (vgl. Kneer & Nassehi 1993, S. 48; Krieger 1998, S. 36)) stammt von den Biologen und Neurophysiologen Maturana und Varela (1982), die es zur Bestimmung lebender, sich selbst erzeugender und erhaltender Systeme verwendeten. Nach Kneer und Nassehi (1993, S. 56) sind autopoietische Systeme organisational geschlossen und damit autonom. Zugleich sind lebende Systeme materiell und energetisch offen. D.h. dass autopoietische Systeme hinsichtlich ihrer Komponenten keinen In- oder Output haben; sie können durch Umwelteinflüsse nicht determiniert, höchstens irritiert werden. Die konkreten Systemzustände werden somit nicht von der Umwelt, sondern vom System selbst bestimmt. Z.B. hat das Nervensystem keinen unmittelbaren Zugang zur Welt, folglich liefern Kognitions- und Wahrnehmungsprozesse kein Bild der Wirklichkeit an sich, sondern fertigen systeminterne *Konstruktionen* an.

Nach Krieger (1998, S. 36) wird auf der Ebene autopoietischer Systeme der Begriff ‚operative Geschlossenheit‘ im nicht-trivialen Sinne als *informationelle Geschlossenheit* verwendet. D.h. es gibt keinen Informationsaustausch zwischen System und Umwelt, also werden Informationen systemintern erzeugt und nicht von der Umwelt abgelesen.

Man kann nach Luhmann (2004, S. 13ff) die Realität der Massenmedien in zweierlei Hinsicht verstehen: Als den Verbreitungsprozess der Inhalte und die ihn umrankenden Kommunikationen, also eine Sequenz von Operationen, und als das, was für sie oder durch sie für andere als Realität erscheint, also eine Sequenz von Beobachtungen, genauer von beobachtenden Operationen. Für diese zweite Verstehensmöglichkeit muss die Einstellung eines Beobachters zweiter Ordnung eingenommen werden; eines Beobachters von Beobachtern.

In dem beobachteten Massenmedien-System findet eine Realitätsverdopplung statt; es kommuniziert, und zwar über etwas anderes oder über sich selbst. Das System kann also zwischen *Selbst-* und *Fremdreferenz* unterscheiden, es ist sogar dazu genötigt.

Fragen, inwieweit das, was die Medien berichten, stimmt, oder wie die Medien die Realität verzerren, begegnet Luhmann auf erkenntnistheoretischer Ebene mit dem *operativen Konstruktivismus*: Kognitive Systeme sind gar nicht in der Lage, zwischen Bedingungen der Existenz von Realobjekten und ihrer Erkenntnis zu unterscheiden, weil sie keinen erkenntnisunabhängigen Zugang zu Realobjekten haben. Massenmedien als operativ geschlossenes und autopoietisch autonomes System müssen folglich Realität konstruieren. Daher lautet für Luhmann die Frage vielmehr, *wie* sie das tun.

Nach Luhmann (2004, S. 33-36) liegt der Grund für die Ausdifferenzierung des Massenmedien-Systems in der Erfindung von Verbreitungstechnologien, die eine Interaktion unter Anwesenden nicht nur ersparen, sondern für die eigenen Kommunikationen der Massenmedien wirksam ausschließen (z.B. des Buchdrucks); stattdessen orientiert sich das System an der Unterscheidung von Selbst- und Fremdreferenz. Damit ist allerdings noch nicht gesagt, wie das Selbst das Selbst bestimmt bzw. wie die Anschlussfähigkeit von Operationen im System erkannt und die Differenz von System und Umwelt laufend reproduziert wird. Dies geschieht durch einen binären Code, der unter Ausschließung dritter Möglichkeiten einen positiven und einen negativen Wert fixiert. Der positive bezeichnet die im System gegebene Anschlussfähigkeit der Operationen, der negative dient nur der Reflexion der Bedingungen dafür. Der Code des Massenmedien-Systems ist die Unterscheidung von Information (positiv) und Nichtinformation (negativ).

Nach Luhmann (ebd., S. 41-44) liegt die wichtigste Besonderheit dieses Codes in seinem Verhältnis zur Zeit: Informationen lassen sich nicht wiederholen; sobald sie Ereignis werden, werden sie Nichtinformation. Das System führt daher ständig den eigenen Output, nämlich Bekanntheit von Sachverhalten, auf der Negativseite des Codes als Nichtinformation wieder in das System ein. Es veraltet sich also selbst und zwingt sich so, ständig für neue Informationen zu sorgen. So bewirken Massenmedien gesellschaftsweite soziale *Redundanz*, also den unmittelbar anschließenden Bedarf nach neuen Informationen. Luhmann spricht in diesem Zusammenhang von einem „geradezu

neurotische[n] Zwang in Wirtschaft, Politik, Wissenschaft und Kunst, etwas Neues bieten zu müssen (obwohl niemand weiß, woher die Neuheit des Neuen kommt und wie groß der Vorrat ist)“ (ebd., S. 44).

Als wichtigste interne Struktur des Massenmedien-Systems differenziert Luhmann (2004, S. 50f) die drei Programmbereiche Nachrichten/Berichte, Werbung und Unterhaltung, wobei jeder in verschiedenen Ausführungen den Code Information – Nichtinformation nutzt.

Nach Luhmann (2004, S. 53-57) ist der Programmbereich Nachrichten und Berichte am deutlichsten als Erarbeitung/Verarbeitung von Informationen erkennbar. Hier ist es zu einer Professionalisierung in Form des Journalismus' gekommen, und bei den angebotenen Informationen wird vorausgesetzt, dass sie zutreffen, wahr sind. Das System kann und muss sich selbst Beschränkungen auferlegen, und so finden sich bzgl. Nachrichten typisch folgende Selektoren (ebd., S. 58-71): Neuheit; Konflikte; Quantitäten; lokaler Bezug; Normverstöße, insbes. in Verbindung mit moralischen Bewertungen und/oder handelnden Personen; Aktualität und Meinungsäußerungen. Diese werden durch Routinen der Organisationen verstärkt, die die Selektion durchführen.

Von den Nachrichten sollte man nach Luhmann (ebd., S. 72f) die von Tagesereignissen unabhängigen Berichte unterscheiden. Diese informieren über die Kontexte etwaiger Neuigkeiten, ihr Neuigkeitswert liegt nicht in der Zeit, sondern ergibt sich aus dem vermuteten Wissensstand des Publikums. Auch hierbei geht es um Informationen mit Wahrheitsanspruch.

Durch die elektronische Speicherung und das Verfügbarhalten publizierter Nachrichten lässt sich schon länger eine Verflüssigung der Differenz von Nachrichten und Berichten beobachten.

Nach Luhmann (ebd., S. 74ff) findet durch Nachrichten und Berichte eine Identitätsgewinnung statt, und mit ständigen Identitätserneuerungen wird das soziale *Gedächtnis* gefüllt. Damit ist kein Speicher gemeint, sondern ein laufendes Diskriminieren zwischen Vergessen und Erinnern; Gedächtnis konstruiert Redundanz mit fortgesetzter Offenheit für Aktuelles, mit ständig erneuerter Irritabilität.

Werbung deklariert nach Luhmann (2004, S. 85f) ihre Motive und verdeckt häufig ihre Mittel; sie sucht zu manipulieren, arbeitet unaufrichtig und setzt voraus, dass das vorausgesetzt wird. Es geht nicht mehr darum, angebotene Objekte zu beschreiben, sondern man wirbt „mit komplexer eingreifenden Mitteln, die die zur Kritik neigende kognitive Sphäre umgehen“ (ebd., S. 85f). Statt der eigenen Motive werden die des Umworbenen unkenntlich gemacht; dieser erkennt, dass es sich um Werbung handelt, aber nicht, wie er beeinflusst wird. Ihm wird Entscheidungsfreiheit suggeriert, und so kann er wollen, was er eigentlich nicht wollte.

Nach Luhmann (ebd., S. 87f) ist eine der Werbetechniken die Vereinnahmung des Gegenmotivs, z.B. durch die Tendenz zur schönen Form, die für sich steht, oder paradoxen Sprachgebrauch, bspw. ‚Sparen durch Geldausgeben‘. Eine andere Technik ist die Paradoxierung der Motivlage, etwa wenn in Bildarrangements das beworbene Objekt in den Hintergrund gerückt ist, oder bei zeitlichen Sequenzen erst am Ende herauskommt, was beworben wird. Diese Vertauschungen muten dem Umworbenen Anstrengungen zu, die das Erinnern fördern.

Bei Werbung, die nicht als solche wahrgenommen wird, liegt die Paradoxie darin, dass bewusste Entscheidungen unbewusst getroffen werden.

Nach Luhmann (ebd., S. 89ff) besteht eine der wichtigsten latenten Funktionen der Werbung darin, Geschmack zu substituieren. So z.B. in Hinsicht auf Mode, mit der zu gehen „fast wie ein Zwang“ (ebd., S. 90) gilt; hier kann die Werbung Motive voraussetzen und in der Form von Informationen einfach anstoßen. Dann kann sie einen Faktor in der Erzeugung des Umschlagtempos darstellen, und sogar planungs- und produktionstechnische Prozesse betreffen.

Nach Luhmann (ebd., S. 94) hat Werbung (und damit das Massenmedien-System) die Funktion, ein Verhältnis von Redundanz und *Varietät* in der Alltagskultur zu stabilisieren. Redundanz wird durch Verkäufe erzeugt, Varietät dadurch, dass man die Produkte am Markt unterscheiden können muss. Daher wird die Illusion geschaffen, dasselbe sei etwas Neues.

Bei Unterhaltung findet nach Luhmann (2004, S. 96-99) eine Realitätsverdopplung statt; Massenmedien stellen dann einen Rahmen dar, der eine Welt freisetzt, in der eine eigene fiktionale Realität gilt. In der fiktiven Geschichte muss der Rezipient allerdings durch bekannte Details in die Lage versetzt werden, schnell ein passendes, auf die Geschichte zugeschnittenes Gedächtnis zu bilden.

Wie immer bei operativer Schließung entstehen nach Luhmann (ebd., S. 108f) durch Ausdifferenzierung zunächst Möglichkeitsüberschüsse. Daher unterscheiden sich Unterhaltungsformen danach, wie diese reduziert werden. Das Grundmuster dafür ist die Erzählung, die sich zu einem beträchtlichen Formenreichtum differenziert hat. Eins der wenigen funktionalen Äquivalente sind Wettbewerbe aller Art (Quizsendungen, Sportveranstaltungen etc.).

Durch die Bezugnahme auf beim Rezipienten bereits vorhandenes Wissen hat Unterhaltung einen diesbzgl. Verstärkereffekt. Sie nutzt vorhandenes Wissen allerdings nur, um sich davon abzuheben. Unterhaltung zielt auf die Aktivierung von selbst Erlebtem, Erhofftem, Befürchtetem oder Vergessenem; „Unterhaltung re-imprägniert das, was man ohnehin schon ist“ (ebd., S. 109).

Nach Luhmann (ebd., S. 111f) hat sich neben die am Roman gewonnene Form erzählender Unterhaltung durch das Fernsehen die Gattung der persönlichen Erfahrungsberichte gesetzt. Anscheinend liegt das Interesse an solchen Sendungen darin, eine glaubwürdige, aber nicht konsenspflichtige Realität geboten zu bekommen.

Unterhaltung bietet nach Luhmann (ebd., S. 115f) einen Sonderbeitrag zur allgemeinen Erzeugung von Realität: Durch Selbstbeobachtung, genauer durch Beobachten des eigenen Beobachtens, kann sich der Rezipient als Individuum identifizieren bzw. seine Identität finden; er kann sich in der dargestellten Welt verorten. Damit regelt die Unterhaltung auch auf der Seite der Subjekte die Inklusion und Exklusion; Inklusion aller mit Ausnahme derer, die sich „per Abstinenz (oft arroganter Abstinenz) an ein Selbst gewöhnt haben, daß [sic] [auf Unterhaltung] nicht angewiesen ist und sich eben dadurch definiert“ (ebd., S. 116).

Insgesamt liegt der Beitrag aller drei Programmbereiche nach Luhmann (2004, S. 120f, 149, 173) darin, Voraussetzungen für weitere Kommunikation zu schaffen, die nicht eigens mitkommuniziert werden müssen. Die Massenmedien stellen ein Hintergrundwissen bereit und schreiben es fort, sie erzeugen ein Gedächtnis für das moderne Gesellschaftssystem, das für jede Kommunikation bestimmte Realitätsannahmen als bekannt voraussetzt. Gleichzeitig spannen die Massenmedien im Prozess der Erarbeitung von Informationen einen Horizont selbsterzeugter Ungewissheit auf, der durch immer weitere Informationen bedient werden muss.

Letztlich sieht Luhmann die Funktion der Massenmedien „im Dirigieren der Selbstbeobachtung des Gesellschaftssystems“ (ebd., S. 118).

Systemtheoretische Ansätze wie der Luhmanns haben oft gemein, dass sie Leistungen der Massenmedien auf einer sehr hohen Abstraktionsebene fassen (vgl. Schulz 2007, S. 28). Mehr oder weniger normative Funktions- und Leistungskataloge von Massenmedien (in Demokratien) finden sich etwa bei Pürer (2003, S. 422-430) oder Burkart (2002, S. 378-412) (vgl. Beiler 2013, S. 46).

Nach Beiler (2013, S. 46f) sehen Wildenmann und Kaltefleiter (1965) die Informationsfunktion als „ursprünglichste Funktion der Massenmedien“ (ebd., S. 15) an. Diese entzieht sich einer eindeutigen Zuordnung zu anderen Funktionsbereichen (vgl. Burkart 2002, S. 402), da sie die Basis und Voraussetzung für alle anderen Leistungen der Massenmedien darstellt.

Die zentralen gesellschaftlichen Leistungen der Massenmedien basieren auf ihren politischen und sozialen Funktionen. Eine wesentliche Funktion ist die Herstellung von Öffentlichkeit, indem Massenmedien Informationen veröffentlichen und zugänglich machen (vgl. Ronneberger 1974, S. 199; Burkart 2002, S. 391). Die am politischen Prozess Beteiligten machen über die Massenmedien ihre Programme, Absichten, Forderungen und Ziele bekannt und treten so miteinander in Verbindung. Die Massenmedien sollen dabei der Vielfalt der vorhandenen Interessen und Meinungen zur Artikulation verhelfen (vgl. Starkulla 1963). Dies dient letztlich der Mitwirkung an der Meinungs- und Willensbildung, die Saxer als Korrelationsleistung fasst (vgl. 1974).

Weiter kommt den Massenmedien die Funktion der politischen Sozialisation und Integration zu (vgl. Ronneberger 1971, S. 50; Ronneberger 1974, S. 201; Pürer 2003, S. 83). Es werden die politischen Rollen transparent gemacht, um dem Einzelnen eine aktive Teilnahme am politischen Geschehen zu ermöglichen, und gesamtgesellschaftliche Interessen aufgezeigt. Damit eng verbunden ist die politische Bildungsfunktion: Massenmedien leisten einen Beitrag zur „Heranbildung von sich am politischen Prozess beteiligenden Staatsbürgern“ (Ronneberger 1974, S. 204), um diese letztlich zur Meinungsbildung zu befähigen. Wesentlich ist auch die Kritik- und Kontrollfunktion (vgl. Burkart 2002, S. 395ff): Massenmedien geben den Mitgliedern einer Gesellschaft die Möglichkeit zur Kritik an den Machträgern, etwa indem sie Opposition und Interessengruppen öffentliches Gehör verschaffen. Außerdem kritisieren sie selbst Staat, Gesellschaft oder Organisationen. In gewissem Maße bringt dies Kontrolle über die kritisierten Zustände, ohne dass die Massenmedien weitere Sanktionsmöglichkeiten besitzen.

Pürer (2003, S. 84) sieht Massenkommunikation als gesamtgesellschaftliches Phänomen sozio-politisch, sozioökonomisch und soziokulturell eingebunden. Dies sei kurz erläutert.

Die politischen Rahmenbedingungen von Massenkommunikation sind nach Pürer (2003, S. 84f) primär in den rechtlichen Grundlagen zu sehen, die ihre Basis bilden. Besonders wichtig in pluralistischen Systemen ist hier das Grundrecht auf Informations-, Meinungs- und Pressefreiheit. Weiter zielen rechtliche Regelungen darauf ab, mithilfe von Aufsichtsorganen Meinungsvielfalt durch publizistischen und ökonomischen Medienwettbewerb zu gewährleisten. Zu nennen sind ebenfalls alle beobachtbaren Formen der Einflussnahme auf Journalismus und Massenmedien. – In Kap. 2.4 werden die politischen Rahmenbedingungen von Massenkommunikation in Deutschland noch etwas näher betrachtet.

Die wirtschaftlichen Gegebenheiten und ökonomischen Zwänge von Massenkommunikation sind nach Pürer (2003, S. 85) vorrangig in den marktwirtschaftlichen Bedingungen zu sehen, denen auch die Massenmedien als Kultur- und Wirtschaftsgüter unterliegen. Die Massenmedien müssen sich auf dem Publikums- und dem Werbemarkt behaupten – daraus resultieren unterschiedliche Erlösquellen und Finanzierungsformen, etwa Vertriebs- und Anzeigenerlöse (Printmedien) oder Gebühren- und Werbefinanzierung (Rundfunk). Generell macht ihre Abhängigkeit von Werbeerlösen die Massenmedien konjunkturanfällig und führt zu internationalen Monopol- und Konzernbildungen. – Auch die wirtschaftlichen Gegebenheiten von Massenkommunikation in Deutschland werden in Kap. 2.4 etwas näher beleuchtet.

Die sozialen und kulturellen Dimensionen von Massenkommunikation sieht Pürer (2003, S. 85f) als ein „sehr unterschiedlich strukturiertes und diskutiertes Feld“ (ebd., S. 85), bei dem es allgemein um die komplexen Wechselwirkungen zwischen Gesellschaft, Medien und Kultur geht. Es stehen

dabei zwei Thesen im Widerstreit: Laut *Reflexionsthese* spiegeln die Massenmedien die in einer Gesellschaft dominanten Wertvorstellungen und Leitmotive lediglich wider, während laut *Kontrollthese* massenmediale Inhalte kulturelle Trends schaffen und prägen, und der Wertewandel auf die Medien zurückzuführen ist. Damit ist die schwierige Frage nach Wirkungen bzw. Folgen von Massenkommunikation angesprochen. Wird dabei zwischen individuellen und sozialen Wirkungen unterschieden, muss bedacht werden, dass beide Bereiche nicht trennscharf voneinander abgrenzbar sind. Unter individuellen Wirkungen von Massenkommunikation versteht man solche, die Kognition, Emotion und Habitus einer Person betreffen. Soziale Wirkungen meinen die in der Gesellschaft beobachtbaren Erscheinungen und Folgen von Massenkommunikation. Diskutiert werden hier z.B. Fragen der politischen Beeinflussung, die Problematik gewaltdarstellender Inhalte, oder auch, welches Abbild der Realität die Massenmedien vermitteln.

### **2.3 Das Internet bzw. Web und Suchmaschinen als Medien**

In diesem Kapitel wird versucht, das Internet bzw. Web und Suchmaschinen als Medien zu verorten, bzw. es werden diesbzgl. Schwierigkeiten aufgezeigt. Außen vor gelassen wird dabei eine Betrachtung des Computers als Medium, der aber natürlich die Schnittstelle zum Internet und damit zum Web und Suchmaschinen darstellt (s. ggf. etwa Beck 2006).

Eine umfassende Behandlung kann und soll an dieser Stelle nicht erfolgen, es werden daher einige grundlegendere Punkte angesprochen.

Die Geschichte und Entwicklung des Internets fasst Merten knapp so zusammen:

Das Internet – 1969 zunächst als digitales Netzwerk des US-Verteidigungsministeriums entwickelt – wurde 1986 als Netzwerk zur Verbindung von vier amerikanischen Universitäten erweitert und erfuhr seine weltweite Verbreitung vor allem ab 1989 durch die Installation des World Wide Web (WWW) am *Europäischen Zentrum für Teilchenphysik* (CERN) in Genf, die 1992 für die Öffentlichkeit frei gegeben wurde. (2007, S. 320 nach Runkehl et al. 1998, S. 9ff)

Einen detaillierteren Blick auf die Struktur des Internets ermöglicht die Anwendung des durch die *Internationale Organisation für Normung* (ISO) standardisierten *Open System Interconnection* (OSI)-Schichtenmodells, eines Referenzmodells für Kommunikationssysteme. Dies ist in Abb. 3 dargestellt.

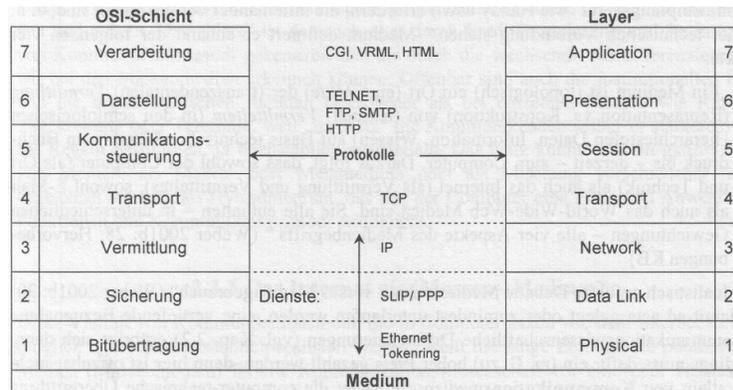


Abb. 3: Internet – ISO/OSI-Schichtenmodell (Beck 2006, S. 19 nach Neumahr 2000, S. 74)

Nach Elektronik Kompendium 2015 besteht das OSI-Schichtenmodell allgemein aus sieben Schichten, die innerhalb der Kommunikation zweier Systeme bestimmte Aufgaben erfüllen, wofür jeweils Funktionen und Protokolle definiert werden. Die Schichten 1 bis 4 sind transport-, 5 bis 7 anwendungsorientiert. Die Kommunikation oder der Datenfluss durchläuft alle Schichten mindestens zweimal; einmal beim Sender und einmal beim Empfänger.

Beck erläutert das Modell in Hinsicht auf das Internet (Abb. 3):

Die ersten beiden Schichten stellen [...] materielle Netze (z.B. Ethernet, Token Ring und Telekommunikationsnetze) dar. Die dritte oder Netzvermittlungsschicht ist beim Internet durch das Internet Protocol (IP) gekennzeichnet, die vierte (oder Transportschicht) durch das Transmission Control Protocol (TCP). In den Schichten fünf und sechs, den Steuerungs- und Darstellungsschichten, finden sich verschiedene weitere Protokolle, die dienstspezifisch sind: FTP (File Transfer Protocol), TELNET (Terminal Emulation Remote Login), SMTP (Simple Mail Transfer Protocol) und HTTP (Hypertext Transfer Protocol). Die Anwendungsschicht (Schicht 7) beschreibt die Software-Sprachen: CGI (Common Gateway Interface), VRML (Virtual Reality Markup Language) oder HTML (Hypertext Markup Language), die schließlich die technische Basis für die verschiedenen Dienste wie Archie, Gopher, WorldWideWeb, E-Mail usw. komplettieren. (2006, S. 18)

Bereits hier wird auch das Web (auch WWW oder W3) als Dienst des Internets eingeführt. Dies soll an dieser Stelle kurz und schon leicht in Hinsicht auf Suchmaschinen aufgegriffen werden: Nach Lewandowski (2015, S. 20, 31ff) ist das Web ein Bestandteil des Internets und besteht aus *Dokumenten*. Es ist multimedial, seine Inhalte können neben der Textform auch als bspw. Bilder, Audio oder Video vorliegen. ‚Dokumente‘ meint daher schlicht die Aufzeichnungen von Informationen, unabhängig von ihrem Format. Die Web-Dokumente sind gewöhnlich in HTML verfasst, besitzen eindeutige Adressen (*Uniform Resource Locators* (URLs)) und können über (*Hyper*-)Links miteinander verknüpft werden. Einzelne Dokumente sind innerhalb von *Websites* abgelegt und werden *Webseiten* genannt. Eine Website ist also ein in sich geschlossenes Angebot im Web – die Abgrenzung erfolgt über die *Domain* oder *Subdomains* bzw. Verzeichnisse (z.B. *tagesschau.de* bzw. *tagesschau.de/inland*) – und eine Webseite ist ein i.d.R. aus Text und zugehörigen medialen

Elementen zusammengesetztes Dokument (die englische Unterscheidung *website* – *webpage* ist griffiger).

Nach Beck (2014, S. 1) stellen das Internet und Online-Kommunikation eine Herausforderung für Publizistik-, Kommunikations- und Medienwissenschaft dar, denn bislang hatte man es lediglich mit (vermeintlich) distinkten Massenmedien zu tun (vgl. auch Merten 2007, S. 320; Burkart 2002, S. 368). Der Telekommunikationssektor lag weit abseits davon. Doch mit dem Aufkommen des Internets stellte sich auch die Frage nach der Konvergenz der Kommunikationsmedien bzw. der Multimedialität der Netzkommunikation (s. Kap. 2.1), denn im Internet fand nun interpersonale und publizistische, private, organisationsinterne und öffentliche Kommunikation statt.

Merten nennt folgende Punkte, die eine Einordnung des Internets erschweren (2007, S. 320f):

- Es ist sowohl zur Rezeption als auch zur Kommunikation nutzbar und hebt damit auf spezifische Weise die Rollentrennung von Kommunikator und Rezipient auf.
- Es fungiert sowohl als Medium als auch als Metamedium.
- Sein Radius ist unbegrenzt, so dass alles elektronisch erfasste Wissen, über das die Weltgesellschaft verfügt – sowohl aktuelles als auch kumuliertes Wissen, sowohl persönliches als auch gesellschaftliches Wissen – *on demand* vorgehalten bzw. jederzeit ergänzt werden kann.
- Der Zielgruppenbezug dieses Mediums ist uneinholbar, weil die Selektivität des Zugriffs einerseits auf den individuellen Fall zugeschnitten werden kann (und dies vom Individuum selbst über E-Mail) und auf der anderen Seite beliebig viele Rezipienten beliebig oft auf eine Adresse zugreifen können.
- Durch seine potentiell unbeschränkte Zugänglichkeit macht das Internet die Entstehung einer panthematischen *Weltöffentlichkeit* von allen für alle möglich.
- Seine Geschwindigkeit bei der Nachrichtenbereitstellung erhöht sich ständig.
- Es ist theoretisch von allen Punkten der Erde für alle Punkte der Erde rund um die Uhr verfügbar.

(Weitere Punkte finden sich z.B. bei Berghaus 1994.)

Für Merten (2007, S. 321) ist das Internet in Hinsicht auf die Massenmedien in erster Linie ein *Metamedium*, weil viele Massenmedien mit einem Eintrag, also einer Website, in ihm vertreten sind; z.B. Online-Angebote großer Zeitungen, die mit aktuellen Informationen bestückt werden (und gegenüber den ursprünglichen Medien viele Vorteile haben, z.B. unbeschränkter Platz und Archivierung von Informationen (vgl. Neuberger & Tonnemacher 1999, S. 33ff)).

Doch es gibt diverse Einordnungen, z.B. einfach als (neues) ‚Medium‘ (vgl. Faulstich 2004a, S. 147) bzw. „kein Medium im klassischen Sinn“ (Pürer 2003, S. 267), als „Massenmedium“ (z.B. Morris & Ogan 1996) bzw. „kein Massenmedium“ (z.B. Schmitt-Walter 2004, S. 16; Laurin 1997, S. 20) (zu dieser Diskussion s. Neuberger 2001, S. 24-30; Höflich 2003, S. 76-80), wie gerade gesehen als Metamedium, als (integrierendes) „Hybridmedium“ (Höflich 2003, S. 75f; Dobal & Werner 1997, S. 108) usw.

Gibt es einen roten Faden in dieser Konfusion, so ist es folgender: U.a. in Rückgriff auf Abb. 3 schreibt Beck:

Betrachtet man also das Internet aus kommunikationswissenschaftlicher Sicht und legt einen differenzierten Medienbegriff zugrunde, dann erscheint das **Internet nicht als Kommunikationsmedium**, das sich mit Hörfunk, Fernsehen, Zeitung oder auch dem Telefon vergleichen ließe. Das Internet lässt sich **zutreffender als technische Infrastruktur oder Medium erster Ordnung** beschreiben. Dies gilt insbesondere für die Schichten 1 bis 4 (TCP/IP), während sich in den darüber liegenden Schichten (den einzelnen Diensten) zumindest die Konturen distinkter Kommunikationsmedien abzuzeichnen beginnen. (2006, S. 19, Hervorh. i. O.; vgl. auch 2014, S. 4f)

Oder in Hinsicht auf die Internet-Dienste:

Führt man sich die semiotische, institutionelle und organisatorische Heterogenität von internet-basierten Kommunikationsformen wie E-Mail, Chat oder WorldWideWeb vor Augen, dann wird deutlich, dass „das Internet“ als Medium erster Ordnung oder eine technische Plattform fungiert, auf der ein – technisch zunehmend auf WWW-Interfaces integriertes – Bündel von Medien zweiter Ordnung aufbaut (2014, S. 4).

Beck (als einer von vielen) greift also auf die Einteilung von Medien in erste und zweite Ordnung (s. Kap. 2.1) zurück, und fasst das Internet seinem Wesen nach als physikalische Infrastruktur für Medien, die in seinem kommunikativen Gebrauch entstehen, (auch) von Beck *Kommunikationsmodi* genannt – darunter das Web, das eher den publizistischen Medien als den Telekommunikationsmedien gleicht (vgl. Beck 2006, S. 28; auch Wirth & Schweiger 1999, S. 47; Weinreich 1998, S. 132). Diese kommunikationswissenschaftliche Einordnung des Internets bzw. Webs scheint am meisten Konsens zu finden (vgl. z.B. Pürer 2003, S. 267f; Lang & Bekavac 2004, S. 433ff; Rußmann 2007, S. 23; Beck 2012, S. 261ff).

Ein grundlegender Versuch, Online-Kommunikation zu systematisieren, stammt von Morris und Ogan (1996). Sie schlagen eine Zeitdimension (synchron vs. asynchron) und eine Sozialdimension (Konfiguration; Zahl der Kommunikationspartner und Struktur der Kommunikationsprozesse) als Ordnungskriterien vor. Daraus ergibt sich eine Matrix, die in einer von Beck aktualisierten Fassung in Tab. 2 dargestellt ist (Dienste, die heute keine Rolle mehr spielen, z.B. *Gopher*, wurden entfernt, dafür neuere Dienste bzw. Anwendungen aufgenommen).

Tab. 2: Online-Kommunikation nach Morris & Ogan (Beck) (2014, S. 8 nach Morris & Ogan 1996)

Sozialdimension Konfiguration	Zeitdimension	
	synchron	asynchron
<i>One-to-one</i>	MUD , MOO , IRC , Instant Messaging, Onlinespiele	E-Mail, WWW, FTP
<i>One-to-few</i>	MUD, MOO, IRC, Instant Messaging, Onlinespiele	Mailinglist, Blog, Microblog , Social Network Services
<i>One-to-many</i>	MUD, MOO, IRC, Onlinespiele	WWW, FTP, Mailinglist, Blog, Microblog, Social Network Services
<i>Many-to-one</i>		WWW, FTP
<i>Many-to-many</i>	Onlinespiele	Usenet , Bulletin Board System

Nach Beck (2014, S. 7) erlaubt diese Methode nur eine grobe Verordnung. Es zeigt sich, dass auf diese Weise keine trennscharfe Zuordnung von Kommunikationsformen möglich ist; die Übergänge zwischen synchronen und asynchronen Kommunikationsformen erweisen sich im Gebrauch der Medien als fließend: Z.B. findet E-Mail-Kommunikation z.T. synchron als Frage-Antwort-Dialog statt, während ältere (*Multi User Dungeon/Dimension/Domain* (MUD); *MUD, object oriented* (MOO)) und neuere Formen netzbasierter Online-Spiele oder sogar *Social-Media*-Angebote (etwa *Facebook*) auf gespeicherte, also asynchrone *Kommunikate* (Profile, Timelines, Posts usw.) zurückgreifen. Die zeitliche Einordnung ganzer Dienste ist also problematisch. Und auch die soziale mittels der Konfiguration ist nicht eindeutig: In Tab. 2 werden einige Online-Kommunikationsmodi mehrfach zugeordnet. Letztlich wird die Konfiguration mehr durch konkrete sozialisierte Handlungen und institutionalisierte Nutzungsweisen bestimmt als durch das medientechnische Potenzial. Besonders auffällig ist dies beim Web, in dem sich private Homepages oder Blogs finden, die sich nur an einen Kommunikationspartner richten, dabei aber auch – zumindest theoretisch – für die *Weltöffentlichkeit* rezipierbar sein können. Private, intime Homepages, gruppen- und gemeinschaftsbezogene Websites und professionelle Sites aktueller publizistischer Medien mit dem Anspruch, zur Meinungsbildung beizutragen, existieren nebeneinander. (Eine Zuordnung des Internets oder auch nur des Webs als Massenmedium verbietet sich folglich.)

Ein anderer Ansatz zur Systematisierung von Online-Kommunikation stammt von Rössler (1998). Dabei sollen als Definitions- und Beschreibungskriterien nicht technische, sondern „eher für den Gebrauch der Modi (und damit ihrem medialen Charakter) relevante Dimensionen“ (ebd., S. 29) dienen. Genannt werden Individualität, Interaktivität und Medialität – wobei die Definition von Individualität, die nur als Differenz zur klassischen Massenkommunikation verstanden wird, unklar bleibt. Dieser Ansatz ist in Tab. 3 dargestellt.

Tab. 3: Internet-Kommunikation nach Rössler (Beck) (2006, S. 24 nach Rössler 1998, S. 36)

	Individualität	Interaktivität	Medialität
World Wide Web	-	-	++
E-Mail	++	+	-
Usenet-Newsgroups	+	-	-
Chat (IRC)	++	++	--
MU*s	++	++	++

Nach Beck (2006, S. 24) führt die Verwendung des Kriteriums Interaktivität zu vorschnellen Schlüssen, und die Internet-Dienste werden letztlich nach ihren Potenzialen und nicht anhand des konkreten Umgangs des einzelnen Nutzers beurteilt (wie Rössler selbstkritisch anmerkt). Insofern ist die Zuweisung der Ausprägungen in den einzelnen Beschreibungsdimensionen diskussionswürdig.

Sowohl Morris' und Ogans als auch Rösslers Ansatz zeigen nach Beck (2014, S. 8), dass pauschale medientheoretische Einordnungen problematisch sind, „weil die konkreten Gebrauchsweisen und die sich hieraus im Laufe der Zeit entwickelnden Gebrauchsregeln entscheidend die Institutionen-dimension von Onlinemedien prägen“ (Beck 2014, S. 8).

Und Beck schließt: „Das Internet' als Bündel bzw. Verbund von Medien kann rahmenanalytisch als ‚Hybridmedium‘ beschrieben werden, bei dem sich verschiedene Kommunikationsmodi auf spezifische Weise im Mediengebrauch (ver-)mischen“ (2006, S. 24).

Suchmaschinen nun gehören nach den bisherigen Erörterungen wie das Web zu den Medien zweiter Ordnung, die auf dem Internet aufbauen. Nach Lewandowski (2015, S. 3, 8, 13, 16f, 31) sind sie neben *Web-Verzeichnissen* (auch *-Kataloge*), *Social-bookmarking-* und *Frage-Antwort-Diensten* (vgl. Griesbaum et al. 2009) jene Art des Zugangs zu Daten im Web, die sich durchgesetzt hat.

Ihre Aufgabe ist also die Vermittlung zwischen Inhalten des Webs und den Nutzern. Dazu legen Suchmaschinen grundsätzlich mittels *Crawlern* eine Kopie des Webs an, bereiten diese zu einem *Index* auf und durchsuchen diesen anhand von Suchanfragen der Nutzer nach passenden Inhalten. (Insbes. auf die Funktionsweise von Suchmaschinen wird noch ausführlich in Kap. 3.1 eingegangen.)

Laut Beck (s.o.) gleicht das Web eher den publizistischen Medien als denen der Telekommunikation – in Form von Dokumenten werden Inhalte erzeugt wie z.B. als Artikel in Zeitungen (Unterschiede bestehen u.a. darin, durch wen und wie dies geschieht). Hinsichtlich Suchmaschinen lässt sich dies umkehren: Diese gleichen eher den Telekommunikationsmedien als den publizistischen, denn durch sie werden (i.d.R.) keine Inhalte erzeugt, sondern es wird auf sie verwiesen. Verglichen werden könnte dies vielleicht mit Fernsehen und Teletext: Das Fernsehen (Web) kann ohne Weiteres genutzt werden – Teletext (eine Suchmaschine) aber bietet Informationen darüber, was wann wo läuft (was es wo gibt).

Das ist es auch, was sich in der Literatur am häufigsten zu Suchmaschinen findet: dass sie eine „Schnittstelle“ zwischen Nutzer und Anbieter (Machill et al. 2003, S. 9; vgl. auch Lewandowski 2015, S. 31), ein „Tor zum Netz“ (Bündnis 90 / Die Grünen 2005), „Fenster zum Web“ (vgl. Hinman 2005, S. 21) darstellen, „die zentrale Zugangsfunktion“ zu Daten im Internet bzw. Web (LfM 2005; vgl. auch Lewandowski 2015, S. 13) und eine „Orientierungs-“ (Neuberger 2005, S. 5) und „Navigationsfunktion“ (vgl. Beck 2012, S. 271) haben, „Selektions- und Vermittlungsinstanzen“ (Machill et al. 2003, S. 27), „Informationsintermediäre“ (Paal 2012, S. 19) sind. (Oft wird in solchen Kontexten anhand des Begriffs ‚Gatekeeper‘ die Macht von Suchmaschinen kritisiert – dazu in Kap. 3 mehr.)

Aus informationswissenschaftlicher Sicht sind Suchmaschinen Werkzeuge des *Web Information Retrievals*, also dem Suchen und Finden von Informationen im Web (vgl. Stock 2007, S. 2, 369-436; Ceri et al. 2013, S. 3f; s. auch Lewandowski 2005).

Nach Beck (2012, S. 172) kommt (u.a.) Suchmaschinen in Anbetracht des großen und intransparenten Gesamtangebotes des Webs eine „Schlüsselrolle“ (ebd., S. 172) zu, weil sie als Startpunkte dienen und Orientierung bieten sollen. Ihre Orientierungs- und Navigationsdienstleistung besteht dabei darin, „mithilfe aufwändiger und intransparenter Suchalgorithmen auf elektronischem Wege mit Werbung vermischte Trefferlisten zusammenstellen“ (ebd., S. 172) (denn Suchmaschinen finanzieren sich in aller Regel über Werbung (vgl. Lewandowski 2015, S. 4, 152)). Weiter können sie Zugriff auf kompilierte Fremdinhalte bieten, wie z.B. bei *Google News*, einer Spezialsuchmaschine für Nachrichten (s. Kap 3.2), wobei Auszüge der (Online-)Presse zu einem Angebot kombiniert werden, ohne dass dafür eine Vergütung etwa der dahinterstehenden Verlage oder Journalisten erfolgen würde. Folglich klassifiziert Beck Suchmaschinen als Metamedien (vgl. auch Winkler 1997 (zur Kritik s. Röhle 2010, S. 35f); Beiler 2005, S. 165f).

## **2.4 Das Mediensystem Deutschlands**

In diesem Abschnitt werden zuerst einige allgemeinere Zusammenhänge des deutschen Mediensystems aufgezeigt, bevor kurz auf die einzelnen Teilsysteme eingegangen wird. Dieses Kapitel orientiert sich relativ stark an Beck 2012, da diese Quelle integrativ aus publizistikwissenschaftlicher Perspektive analysiert und noch recht aktuell ist. Als ebenfalls ergiebig, wenn auch etwas älter, hat sich Dreier 2009 erwiesen.

Auf allzu detaillierte Informationen wird aus Platzgründen verzichtet.

Historisch ist die Struktur des deutschen Mediensystems von der Presse-, Film- und Rundfunk-Neuordnung durch die Alliierten nach dem Zweiten Weltkrieg und dem Beitritt der neuen

Bundesländer nach der Wiedervereinigung 1989 geprägt (vgl. Beck 2012, S. 333ff; auch Thomaß 2013, S. 221f; Dreier 2009, S. 257).

Nach Beck (2012, S. 335) werden im *Grundgesetz* (GG) der Bundesrepublik Deutschland die Kommunikationsfreiheiten, also Meinungs-, Informations- und Medienfreiheit, als umfangreiche Menschenrechte garantiert (vgl. auch Weischenberg 2004, S. 130-138):

Jeder hat das Recht, seine Meinung in Wort, Schrift und Bild frei zu äußern und zu verbreiten und sich aus allgemein zugänglichen Quellen ungehindert zu unterrichten. Die Pressefreiheit und die Freiheit der Berichterstattung durch Rundfunk und Film werden gewährleistet. Eine Zensur findet nicht statt (GG Art. 5 Abs. 1).

Für das Mediensystem sind besonders die Kommunikationsgrundrechte der Medieninstitutionen und der sich hieraus ergebende Schutz vor staatlicher Zensur oder Lenkung relevant.

Nach Beck (2012, S. 335ff) wird die Pressefreiheit im weiten Sinne als Medienfreiheit verstanden, die neben Rundfunk und Film auch die publizistischen Online-Medien umfasst. Medien tragen zur Meinungsbildung bei und haben deshalb eine besondere verfassungsrechtliche Stellung inne. Der Staat darf keine Medienorganisationen besitzen.

Die meisten Medien sind wie gewinnorientiert arbeitende Unternehmen nach privatem Recht organisiert (vgl. auch Weischenberg 2004, S. 148f), hinzu kommen öffentlich-rechtliche Organisationen im Rundfunk- und Telemediensektor sowie Non-Profit-Anbieter auf den anderen Medienfeldern (z.B. konfessionelle, Parteien- und Verbandspresse). Außerdem begünstigen technologische Innovationen sowie gesunkene technische und wirtschaftliche Barrieren die Entstehung von Medienressourcenschwacher zivilgesellschaftlicher Akteure und Plattformen für Publikationen durch Laien.

Die zentrale Institution der deutschen Medien ist der Markt mit seiner Angebot-und-Nachfragedynamik. Da eine rein kommerzielle Institutionalisierung hinsichtlich der zentralen öffentlichen, *meritorischen* Funktion (meritorische Güter werden aufgrund zu geringer Nachfrage durch den Staat bereitgestellt (nach Musgrave 1957, S. 341)) jedoch zu Marktversagen führt, wird eine öffentliche, aber nicht-staatliche Rundfunkorganisation mit einem entsprechend staatsfernen Programm betrieben.

Um die beiden Hauptziele – flächendeckende mediale Versorgung und strukturelle Vielfaltssicherung – zu erreichen, werden Medien und Medienvertrieb steuerlich begünstigt (Buch und Presse) und z.T. explizit gefördert (Film und Fernsehen). Außerdem werden diese Medienmärkte besonderen Wettbewerbsregelungen unterworfen, z.B. gilt im Buchsektor eine Preisbindung. Da die kapitalistische Produktions- und Wirtschaftsweise allerdings regelmäßig zur Marktkonzentration führt,

bestehen für die Presse und den Rundfunk strenge kartell- und medienrechtliche Konzentrationsregeln.

Paradoxerweise hat der Staat für eine möglichst staatsferne Medienregulierung zu sorgen (vgl. auch Weischenberg 2004, S. 141). Tatsächlich wird diese weitgehend nicht-staatlichen Regulierungsakteuren überlassen (z.B. *Freiwilligen Selbstkontrollen für Kinofilm, Fernsehen und Multimedia* (FSK, FSF, FSM)).

Die Regulierungskompetenzen unterliegen einem mehrgliedrigen System: Den Bundesländern kommen die Kompetenzen für den Rundfunk und die Presse zu, beim Film und den Online-Medien gibt es geteilte, und der Bund besitzt die Kompetenz für die Telekommunikation. Alle deutschen politischen Akteure sind überdies an die Richtlinien, besonders die rundfunk- und online-politischen Vorgaben, der *Europäischen Union* (EU) gebunden. Zur Wahrung einheitlicher Rechtsverhältnisse koordinieren die Bundesländer ihre Medienpolitik in Staatsverträgen, gegenüber der EU fungiert der Bund als Sachwalter der Länder.

Nach Beck (2012, S. 338) besteht das deutsche Mediensystem aus verschiedenen Teilsystemen: den Printmedien, also Buch und Presse (Zeitung und Zeitschrift); dem Film (Kino und Video); dem Rundfunk (Fernsehen und Hörfunk); und den publizistischen Online-Medien. Sie alle bauen auf einer gemeinsamen normativen, technischen und infrastrukturellen Grundlage auf, die für die Beschaffung (Agenturen), die Verbreitung (Telekommunikation), und auch die Finanzierung (Werbung) wichtig ist. Dies ist in Abb. 4 dargestellt.

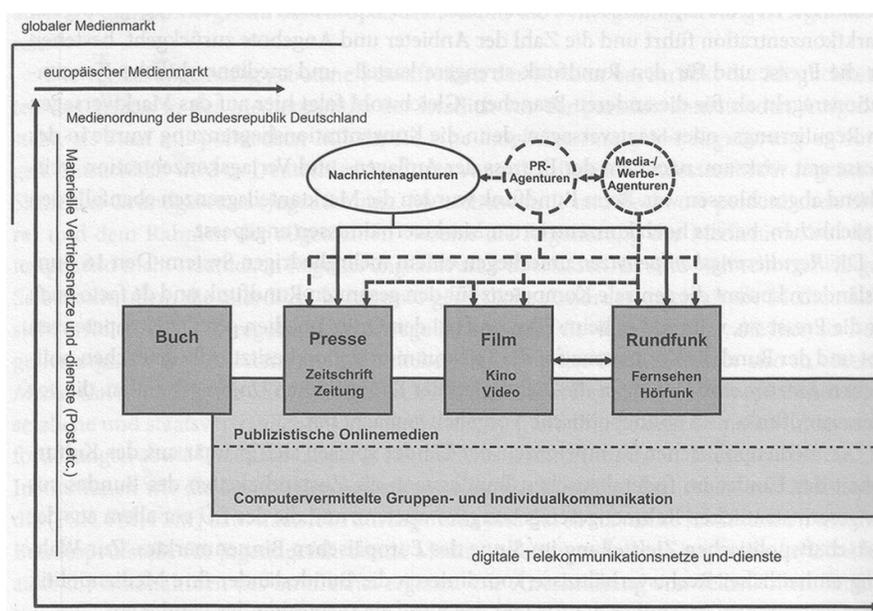


Abb. 4: Das Mediensystem Deutschlands (Beck 2012, S. 338)

Nach Beck (2012, S. 339) greifen publizistische Medienorganisationen auf Nachrichten-, PR-, Media- und Werbeagenturen zurück.

Für Medien mit aktueller Berichterstattung sind Nachrichtenagenturen wie die *Deutsche Presseagentur* (dpa) und der *Deutsche Auslandsdepeschendienst* (dpad), Weltagenturen wie die *Associated Press* (AP), *Thomson Reuters* (rtr) und die *Agence France Press* (afp), sowie Spezialagenturen (z.B. für Sport und Wirtschaft) wichtig. Die Agenturen sind staatsfern organisiert. Ihnen kommt am Anfang des Nachrichtenprozesses eine Schlüsselstellung für die Prüfung (Verifikation), Auswahl (Selektion) und Einordnung (Ressortierung) der Nachrichten zu.

Da sich die meisten deutschen Publikumsmedien zum Großteil über Werbung finanzieren, spielt der Werbemarkt innerhalb des deutschen Mediensystems eine große Rolle. Werbe- und Mediaagenturen kümmern sich um die Werbekommunikation, z.B. bei der Planung und Schaltung von Anzeigen.

Journalisten sind auf die Zusammenarbeit mit Pressesprechern und PR-Agenturen angewiesen.

Das Buch hat in Deutschland nach Beck (2012, S. 337, 340f) eine lange Tradition und genießt hohes Ansehen. Zwischen Autoren und Lesern nehmen Verlagsunternehmen eine zentrale medienökonomische Rolle bei der Produktion und Vermarktung ein. Die Buchkommunikation wird durch einen reduzierten Umsatzsteuersatz indirekt subventioniert und unterliegt einer gesetzlichen Buchpreisbindung. Deutschland verfügt über ein im internationalen Vergleich außergewöhnliches Buchangebot, eins der leistungsfähigsten Vertriebssysteme und ein differenziertes Netz öffentlicher Bibliotheken.

Nach Beck (2012, S. 335, 337, 342ff) genießt die Presse in Deutschland aufgrund ihrer wichtigen öffentlichen Aufgabe rechtliche und steuerliche Privilegien, gleichzeitig unterliegt sie strengen Wettbewerbsregeln (Konzentrationskontrolle).

Das Geschäftsmodell der meisten bedeutenden Pressetitel koppelt den Leser- bzw. Käufermarkt mit dem Werbemarkt – publizistischer Inhalt wird dabei in erster Linie produziert, um eine hohe Publikumsreichweite zu erzielen oder ein bestimmtes Zielpublikum zu erreichen. So entsteht eine Abhängigkeit der Presse von der werbetreibenden Wirtschaft, was negative Auswirkungen auf die Unabhängigkeit der Berichterstattung und die innere Pressefreiheit haben kann. Es existieren jedoch u.a. die *Publizistischen Grundsätze* (Pressekodex).

Tageszeitungen, Wochenzeitungen, Publikumszeitschriften sowie Spezial- und Fachzeitschriften werden in Deutschland oft per Abonnement bezogen. Weiter gibt es acht lokale bzw. regionale Boulevardzeitungen (z.B. *Bild*), rund zehn überregional verbreitete Qualitätszeitungen (z.B. *FAZ*, *Süddeutsche* und *Frankfurter Rundschau*), drei bundesweite Wirtschafts- und Finanzzeitungen (z.B.

*Handelsblatt*), und parteinahe Tageszeitungen (z.B. *Neues Deutschland*) (vgl. auch Weischenberg 2004, S. 141ff).

Der Zeitungs- und Zeitschriftenmarkt ist von einem Wettbewerb privater Verlagsunternehmen geprägt. Bei den Publikumszeitschriften dominieren die vier großen Zeitschriftenverlage *Burda*, *Bauer*, *Springer* und *Bertelsmann* (*Gruener+Jahr*; im November 2014 komplett übernommen (vgl. Bertelsmann 2014)).

Durch die wirtschaftliche Abhängigkeit der Presse sowie strukturelle Herausforderungen aufgrund veränderter Mediennutzungsmuster und von Online-Angeboten (s.u.), erweist sich die Presse als krisenanfällig.

Ihre Regulierung erfolgt durch die jeweiligen Landespressegesetze der 16 Bundesländer. Der Bund besitzt lediglich eine Rahmengesetzgebungskompetenz, die er aber de facto nicht ausübt.

Der Film stand in Deutschland nach Beck (2012, S. 337, 344ff) von Anfang an unter starkem politischem Einfluss – er wurde lange zensiert und staatlich gelenkt, sowie als Propagandamittel genutzt.

In der DDR wurde ab 1946 am Standort der *Universum Film AG* (UFA) (heute *UFA GmbH*) die staatliche *Deutsche Film AG* (DEFA) aufgebaut. Nach dem zweiten Weltkrieg entstand eine starke Marktdominanz der US-Majorfilmkonzerne. Aufgrund hohen Kapitalaufwands, großer Produktions- und Marktrisiken, sowie eines im internationalen Vergleich begrenzten Binnenmarkts wird die deutsche Filmproduktion durch den Bund, die Länder, die EU und den *Europarat* gefördert.

Nach Beck (2012, S. 334, 336f, 342ff) ist der Rundfunk in Deutschland föderalistisch organisiert und soll als Medium und Faktor der Meinungsbildung dienen. Der verfassungsrechtliche Auftrag des öffentlich-rechtlichen Rundfunks lautet, allen Bürgern ein *innenpluralistisches* (inhaltlich vielfältiges) und sozial integrierend wirkendes Informations-, Bildungs- und Unterhaltungsprogramm zu bieten, sowie zu beraten. Entsprechend hoch sind die Anforderungen. Ein Problem sind Versuche seitens Landes- und Bundespolitikern, den Rundfunk für ihre Machtinteressen zu instrumentalisieren (vgl. auch Weischenberg 2004, S. 140).

Die Anstalten der *Arbeitsgemeinschaft der öffentlich-rechtlichen Rundfunkanstalten der Bundesrepublik Deutschland* (ARD) bieten insgesamt 54 Hörfunkprogramme und 20 öffentlich-rechtliche Fernsehprogramme an. Die öffentlich-rechtlichen Landesrundfunkanstalten der ARD sowie das *Zweite Deutsche Fernsehen* (ZDF) und *DeutschlandRadio* (DLR) werden durch eine staatliche Beitragspflicht (*Rundfunkbeitrag*: [www.rundfunkbeitrag.de](http://www.rundfunkbeitrag.de)) finanziert (vgl. auch Dreier 2009, S. 262).

Seit Mitte der 1980er Jahre trat an die Stelle des öffentlich-rechtlichen Monopols das *duale System*; seitdem wird Rundfunk auch von kommerziellen Unternehmen (und außerdem als nicht-kommerzieller Bürgerfunk bzw. *Offener Kanal*) betrieben.

Fernsehen wird in Deutschland meist noch analog über Kabelverteilnetze verbreitet, außerdem digital per Satellit, terrestrisch und selten via Internet. Hörfunk wird über *analoge Ultrakurzwelle* (UKA), und selten als Webradio, per Kabel oder Satellit gesendet.

2013 waren 84 Prozent der Einwohner Deutschlands Internetnutzer, 34,6 Prozent verfügten über einen Breitband-Internetanschluss, und 44,7 Prozent nutzten einen mobilen Breitband-Internetvertrag (Destatis 2015a).

Publizistische Online-Medien werden in Deutschland nach Beck (2012, S. 283, 348f) von professionellen Angeboten von Medienorganisationen aus dem Presse- und Rundfunksektor dominiert. Hinzu kommen Marktakteure aus dem Kommunikationssektor, die den technischen Zugang realisieren, und zahlreiche neue Akteure, die vielfältige Online-Dienstleistungen erbringen, z.B. die Bereitstellung von Plattformen für Blogs und Videos, Social Network Services, Suchmaschinen und Bezahlsysteme. Dabei herrscht eine starke Marktkonzentration.

Die Zahlungsbereitschaft für publizistische (Qualitäts-)Inhalte ist beschränkt, und so wird in Deutschland bzgl. tragfähiger Geschäftsmodelle immer noch experimentiert (z.B. mit *Paid Content* im Verlagsbereich). Der Werbemarkt ist zwar gewachsen, für eine originäre Finanzierung jedoch nicht ausreichend. Oft verfolgen Online-Medien-Anbieter daher Kostensenkungsstrategien (z.B. werden Online-Redaktionen nur mit geringen Ressourcen ausgestattet und greifen auf Inhalte des eigenen Unternehmens oder anderer Medien zurück (*Content-Syndication*)).

Die Regulierung der Online-Medien basiert auf dem *Staatsvertrag für Rundfunk und Telemedien*, dem *Jugendmedienschutz-Staatsvertrag* und dem *Telemediengesetz* des Bundes, die im Einklang mit der *Richtlinie über Audiovisuelle Mediendienste* der EU stehen (vgl. auch Dreier 2009, S. 266). Eine große Rolle spielen auch Institutionen der Selbstkontrolle.

## **2.5 Das Mediensystem der USA**

Nach Kleinsteuber hat sich kein Mediensystem als vergleichbar richtungsweisend für den Rest der Welt erwiesen wie das der USA: „Dort wurde erstmals konsequent das Prinzip der Pressefreiheit konstitutionell verankert, dort entwickelte sich die Presse zuerst zu einem echten Massenmedium, dort entstand das kommerzielle Modell von Radio und Fernsehen, dort wurde auch das Zeitalter der Online-Medien eingeläutet“ (2009, S. 1211). Und noch immer werden neue Medientrends in den USA begonnen und in anderen Teilen der Welt adaptiert (auch wenn nach Turnstall (2008) die Impulse

weniger werden, neue Informationstechniken zunehmend aus Fernost kommen, die US-Programmexporte zurückgehen und sich Medien *glokalisieren*, also in einem globalen Umfeld lokaler werden (vgl. Kleinsteuber & Kutz 2008, S. 33)).

In dieser Sektion wird das Mediensystem der USA nach ein paar grundlegenden Informationen in Hinsicht auf die Bereiche Print, Rundfunk und Online umrissen. Kleinsteuber (2009, S. 1211) bemängelt, dass verlässliche Angaben zu Medienindikatoren kaum erhältlich sind, und so sind auch die hier verwendeten Hauptquellen bereits etwas älter.

Ein Zwischenfazit am Ende der Sektion fasst die wichtigsten Punkte des ersten Teilbereichs dieser Arbeit, dem theoretischen Hintergrund, nochmals zusammen.

Nach Kleinsteuber (2009, S. 1211f) gibt es angesichts der Verankerung der Pressefreiheit im ersten Zusatzartikel zur Verfassung, dem *First Amendment* von 1791, mit dem die Kommunikationsfreiheit in den Grundrechtskatalog der USA aufgenommen wurde und der die wichtigste normative Basis des Mediensystems darstellt (vgl. Weischenberg 2004, S. 162), keine spezielle Pressegesetzgebung des Bundes. Allerdings finden sich für die Medien relevante Rechtssätze in anderen Bundesgesetzen (z.B. Staatsschutz und Strafrechtsdelikte) und in Gesetzen der Einzelstaaten.

Um die starke Pressefreiheit gegen andere Rechtsgüter abzuwägen, wurden oft Gerichte bemüht, besonders der *Supreme Court*. Die hervorgehobene Bedeutung der Presse wird auch durch das Bild der *Fourth Branch of Government* deutlich; die Presse soll als vierte Gewalt (neben Exekutive, Legislative und Judikative) rechtlich privilegiert sein, dafür aber eine besondere Wächterrolle gegenüber der Politik wahrnehmen. Dies wird durch eine lange Tradition des investigativen Journalismus' unterstrichen.

Der *Freedom of Information Act* (zuerst 1966) garantiert allen Bürgern freien Zugang zu öffentlichen Dokumenten – der Staat kann die Herausgabe nur in bestimmten und festgelegten Fällen verweigern.

Natürlich können auch Journalisten in den USA mit dem Gesetz in Konflikt geraten, doch weit typischer ist ein Agieren innerhalb selbstgewählter Grenzen, die von einer ethischen Professionalität bestimmt werden.

Für Kleinsteuber zeichnet sich das Mediensystem der USA durch zwei zentrale Charakteristika aus (2007a, S. 377):

- Es entwickelte sich unter den Bedingungen weitgehender Staatsfreiheit.
- Alle wesentlichen Medien sind privatwirtschaftlich verfasst und arbeiten gewinnorientiert.

Damit stellen die USA einen marktwirtschaftlich-kapitalistischen Prototyp unter den Mediensystemen der westlichen Welt dar (vgl. Kleinsteuber 2007a, S. 377).

Für Weischenberg (2004, S. 167) stellt als Folge dieser Entwicklungen das zentrale Merkmal des Mediensystems Konzentration dar, und zwar in dreierlei Hinsicht: die Akkumulation von Medienbesitz (ökonomische Konzentration), eine unzureichende Repräsentation der sozialen Schichten im Journalismus (Status-Konzentration), und eine ungenügende Verteilung der Medieninstitutionen auf der publizistischen Landkarte (geografische Konzentration). Die geografische Konzentration bedeutet, dass die Massenmedien von den Medienzentren an der Ostküste dominiert werden, aus der Status-Konzentration resultiert mangelnde Vielfalt der Berichterstattung, und die ökonomische Konzentration führt schließlich zur Bildung von lokalen (Presse-)Monopolen sowie Zeitungsketten, und von Konglomeraten (Medienunternehmen mit Beteiligungen in diversen Medienbereichen und darüber hinaus) als vorherrschender Konzentrationsform.

Nach Kleinsteuber (2009, S. 1212ff; 2007b, S. 258f) wies die Pressegeschichte der USA dem Land über lange Zeiträume eine Avantgarde-Rolle zu. Bereits in der Kolonialzeit bestanden Zeitungen mit erheblichen Freiheiten, in der Phase von Revolution und Unabhängigkeitskrieg war die Presse bereits weit entwickelt. Ab den 1830er Jahren entstand mit Frühformen einer Boulevardpresse (*One-Penny-Press*) die moderne Massenpresse. Innovationen im Pressewesen wurden i.d.R. früher als in Europa eingeführt, z.B. die Gründung einer Nachrichtenagentur auf Telegrafien-Basis, AP (Vorläufer seit 1848), oder der Einsatz fotografischer Abbildungen und Comics (ab 1889). Vor dem ersten Weltkrieg wurde *investigative reporting* zur Kontrolle der Regierenden fester Bestandteil des journalistischen Selbstverständnisses (Presse als vierte Gewalt). Hearst (1863-1951) war der erste Zeitungsmogul, machte mit Sensationen Geschäfte und nutzte seine Zeitung zur Förderung eigener politischer Ambitionen. Sein Widersacher Pulitzer (1847-1911) betonte die ethischen Wurzeln des eigenen Gewerbes, gab Anstoß zur Gründung der *Columbia School of Journalism* (Modell für viele weitere Journalistenschulen), und stiftete den nach ihm benannten Preis.

Noch heute ist die US-Presse größtenteils lokal konstituiert (fast alle führenden Zeitungen tragen ihren Erscheinungsort im Titel). In der Vergangenheit hatte dies in Hinsicht auf die Größe des Landes v.a. technische Gründe – erst in den 1970er Jahren wagte das *Wall Street Journal* (via Satellitenübertragung) den zeitgleichen Druck in verschiedenen Landesteilen, und erst 1981 wurde mit *USA Today* eine betont nationale Zeitung konzipiert, die sofort ein Erfolg wurde und noch heute die auflagenstärkste Tageszeitung ist (vgl. Statista 2015b). Dennoch treten die international bekannten Zeitungen wie die *New York Times* oder die *Washington Post* außerhalb der Zentren weniger durch ihre Präsenz als mehr durch Teil-Nachdrucke in Erscheinung (*syndicated news*). Die Dezentralität der

Zeitungslandschaft wirkt sich auch auf die Auflagenhöhen aus; kein Presseorgan erreicht z.B. die der deutschen *Bild*.

Trotz einer Vielzahl unabhängig erscheinender Zeitungen sind die Märkte nach jahrzehntelangen Prozessen von Verdrängung und Fusion hochkonzentriert. Ein Großteil der Zeitungen zählt zu nationalen oder regionalen Ketten wie *Gannett*, *Knight-Ridder*, *Newhouse* oder *Scripps-Howard*.

Die heutige Zeitungslandschaft der USA ist von Stagnation geprägt: Wesentliche Medienfunktionen sind (besonders im Boulevard-Bereich) an das Fernsehen übergegangen – waren früher für die Zeitungsstruktur Verlegerfamilien typisch, sind heute vielfach Hedgefonds und Investmentgesellschaften in die inzwischen börsennotierten Verlage eingestiegen und haben deren Geschäfte übernommen – die Auflagen sinken, die Werbeeinnahmen stagnieren, die Konkurrenz aus dem Internet steigt, Redaktionen werden verkleinert, Renditen sinken. Als ein Beispiel des Wandels sei der Verkauf von 33 Lokalzeitungen in 2013 durch Medienmogul Murdoch erwähnt, womit dieser sein Unternehmen *News Corporation* stärker auf Online-Angebote ausrichten wollte (vgl. Spiegel Online 2013).

Hinzu kommt, dass alle nationalen und internationalen Informationen von sehr wenigen Quellen stammen, insbes. von den beiden Nachrichtenagenturen AP und *United Press International* (UPI), von denen 99 Prozent der täglich berichtenden Nachrichtenmedien ihre Informationen beziehen. U.a. dies führt zu einer hohen Homogenität bei der aktuellen Berichterstattung.

Nach Kleinsteuber (2009, S. 1214f; 2007b, S. 259f) sind auch Zeitschriften in der Mediengeschichte der USA bereits früh zu finden – sie begannen 1732 als Almanache. Im 19. Jahrhundert bestand bereits eine vielfältige Zeitschriftenlandschaft, die bald durch Illustrationen und später Fotos bereichert wurde. Einige Publikationen sind seit weit mehr als 100 Jahren am Markt, wie z.B. das *Harper's Magazine* (1850 als *Harper's Monthly* gegründet). Eine beachtliche Tradition haben politische Wochenmagazine wie das seit 1923 erscheinende *Time Magazine* (Vorbild des *Spiegels*) und dessen stärkster Konkurrent *Newsweek*.

Heute ist die Zeitschriftenlandschaft hoch diversifiziert, wobei unterhaltungs- und serviceorientierte Medien am auflagenstärksten sind, doch auch politische Magazine haben ihren Platz. Seit vielen Jahren ist *Reader's Digest* die meistverkaufte Publikation (in den USA und in vielen Übersetzungen auch weltweit; allerdings meldete die Zeitschrift in den USA 2009 und 2013 Insolvenz an (vgl. Tagesspiegel 2013)). Anders als die Tagespresse sind Zeitschriften am nationalen Markt orientiert, weswegen viele sehr hohe Auflagen verzeichnen. Einige erfolgreiche amerikanische Zeitschriftenformate erreichten auch den deutschen Markt, z.B. *Playboy*, *Cosmopolitan* und *National Geographic*.

Nach Kleinsteuber (2009, S. 1215-1222; 2007b, S. 260f) entstanden die ersten elektronischen Medien in den USA nach dem Ersten Weltkrieg. Damals war (anders als in Europa) nicht der Staat beteiligt, sondern Amateure und Konstrukteure experimentierten mit der neuen Funktechnik, und Unternehmen folgten. Der Hörfunk entstand 1919/20 (einige Jahre vor Deutschland), und als man feststellte, dass Werbung den Stationsbetrieb finanzieren kann, begann ein explosives Wachstum. Dieses mündete schließlich in einen aggressiven Verdrängungswettbewerb der vielen neuen Sender, sodass ab 1927 staatliche Regulierung eingreifen musste. 1934 wurde die *Federal Communications Commission* (FCC) gegründet, die heute für den gesamten Radio- und Fernsbereich und auch die Telekommunikation zuständig ist. Sie erteilt Sendelizenzen auf der rechtlichen Annahme, dass die Frequenzen (*public airwaves*) der Allgemeinheit gehören und für ihre Nutzung eine zeitlich begrenzte Autorisierung zu erwerben ist. Damit wurde gleichzeitig entschieden, dass die unbegrenzte Freiheit, die für die Presse in der Verfassung verankert wurde, für den elektronischen Bereich so nicht gelten kann (vgl. Creech 2013).

Auf der Grundlage dieses Lizenzsystems entfaltete sich ein kommerzielles Radio- und später Fernsehsystem, bei dem die Finanzierung des Programms über Werbeeinblendungen erfolgte. Es basiert auf der Logik des kommerziellen Rundfunks, in dem letztlich Hörer- oder Zuschauerschaften an die Werbewirtschaft verkauft werden. Eine Folge war, dass staatlich begründete und gebührenfinanzierte Anbieter mit einem gesetzlichen Programmauftrag wie im *Public-service*-Sektor Europas niemals eine Chance hatten, weil sie als Verstoß gegen das Gebot der Medienfreiheit interpretiert wurden. Der kommerzielle Typ des Sendens war bereits seit Ende der 1920er Jahre fest etabliert und breitete sich früh aus – in Europa kam er nach einigen Ausnahmen (z.B. Großbritannien) erst in den 1980er Jahren an und markiert hier seitdem die zweite Säule des dualen Systems (s. voriges Kap.).

In den USA bildeten sich aus Kostengründen bereits früh *networks*; dabei verbinden sich örtlich lizenzierte Stationen mit nationalen Netzwerken, die wesentliche Teile des Programms bereitstellen – die Werbeeinnahmen werden geteilt. Lange dominierten die *Big Three Network*-Gesellschaften *National Broadcasting Company* (NBC), *Columbia Broadcasting System* (CBS) und *American Broadcasting Company* (ABC) die TV-Versorgung, bis 1986 die *Fox Broadcasting Company* (FOX; unter Kontrolle von Murdochs *News Corp*) dazustieß und das Trio zu den *Big Four* erweiterte. Bei dieser Art TV-Vollprogramm steht Unterhaltung im Vordergrund, da nur diese ein optimales Umfeld für Werbung bietet und große Publika anzieht.

Parallel zu den in den 1970er Jahren neuen Übertragungswegen Kabel und Satellit begann die Expansion von Pay-TV, bei dem werbefreie Programme gegen ein Abonnement angeboten werden. Populär ist hier der Anbieter *Home Box Office* (HBO; zu *Time Warner* gehörend). Inzwischen entwickelte sich auch ein verzweigtes System an ‚freien‘ und Pay-Spartenkanälen, von denen inzwischen etliche auch in Europa angekommen sind, z.B. *Cable News Network* (CNN; Nachrichten),

MTV (vormals *Music Television*; ehem. Musik, heute eher TV-Shows für Jugendliche) und *Discovery Channel* (Dokumentationen).

In kleinem Umfang gibt es in den USA auch öffentliche Anbieter wie *National Public Radio* (NPR) und *Public Broadcasting System* (PBS), die sich um kommerziell benachteiligte Bereiche wie Kultur und Bildung kümmern. Sie werden meist von Universitäten oder zivilgesellschaftlichen Trägern betrieben und finanzieren sich durch Spenden und staatliche Zuweisungen. Die Bezeichnung ‚*public*‘ bezieht sich also auf meist privat betriebene Sender, die einem *Non-profit*-Auftrag folgen.

1996 wurde durch den kontrovers debattierten *Telecommunications Act* der FCC das Telekommunikationsrecht grundlegend überarbeitet und der Wettbewerb verstärkt.

Seit 1998 wird digitales *High Definition Television* (HDTV) angeboten.

Der Umstieg auf digitales Fernsehen (*Digital Television* (DTV) und HDTV) verlief/verläuft nicht ohne Schwierigkeiten – 2009 sollte ein ‚harter‘ Wechsel erfolgen, 2011 ordnete die FCC an, dass letzte analoge *low-power television stations* ihren Betrieb bis zum 01. September 2015 einstellen sollen (vgl. FCC 2011).

Im Laufe der Zeit sind die Medienmärkte immer stärker unter die Kontrolle von immer weniger Unternehmen gekommen. Mit den großen Medienkonzernen können weltweit nur noch *Sony* (Japan), *Bertelsmann* und evtl. *Vivendi SA* (Frankreich) konkurrieren (vgl. IFM 2014).

Auch das Internet entstand in den USA, und Anfang der 1990er Jahre wurde es zum Massenphänomen (s. Kap. 2.1 und 2.3). 1993 proklamierte Vizepräsident Al Gore die Errichtung einer *National Information Infrastructure* (auch *Information Superhighway*), die alle Einrichtungen und Haushalte über vernetzte und interaktive Computer umfassen sollte. Über sie sollten neben Computerdiensten auch Medienangebote laufen. In den 1990ern verbreiteten sich Online-Dienste, *America Online* (AOL) wurde zum größten Anbieter, jedoch hat dieser Ansatz seine Bedeutung verloren.

In den USA hat sich noch kein eigenes Internetrecht entwickelt – einige Regelungen finden sich in Gesetzen wie dem *Telecommunications Act*. Der Versuch, in diesem Kontext inhaltliche Qualitätsstandards für das Web festzuschreiben (*Communication Decency Act*; CDA), scheiterte vor dem *Supreme Court*. Zeitweise Versuche der Regierung, Zugang in die Computernetze zum Abhören des Datenverkehrs (*Clipper chip*) zu erzwingen, blieben erfolglos. Inzwischen halten jedoch mit dem Argument der Terroristenabwehr nach dem 11. September 2001 zunehmend staatliche Überwachungsrechte in Bezug auf das Internet Eingang in die Gesetzgebung, insbes. über den *Patriot Act* (2001). Die Industrie und große Teile der Internetgemeinde favorisieren dagegen gemeinsame Absprachen, die mit technischen Mitteln (z.B. Filterprogrammen) oder auch Ächtungen durchgesetzt werden sollen. Insgesamt ist das Datenschutzniveau geringer als in Europa; z.B. ist es üblich, dass

private Firmen systematisch Daten über Privatleute sammeln und an Großkunden verkaufen (*data mining*). Allerdings kommen aus den USA auch innovative Ansätze, eine *Internet Governance* zu erproben.

Nahezu alle bedeutsamen Medien verfügen über eigene Internet-Portale – als Top-Adressen gelten jene überregionaler Medien (z.B. *Cable News Network*; CNN). Die Nachfrage nach solchen Angeboten steigt, doch die Verweildauer bleibt meist kurz.

2010 erließ die FCC den *National Broadband Plan*, um den Internetzugang in den USA zu verbessern.

2013 waren 84,2 Prozent der US-Einwohner Internetnutzer, 28,5 Prozent verfügten über einen Breitband-Internetanschluss, und 92,8 Prozent nutzten einen mobilen Breitband-Internetvertrag (Destatis 2015b).

Auch die Idee der Konvergenz von elektronischen und digitalen Medien (s. Kap 2.1 und 2.3) stammt aus den USA; mit der Fusion von AOL und *Time Warner* zum Medienunternehmen *AOL Time Warner* wurde die Strategie verfolgt, die gesamte Wertschöpfungskette aller erdenklichen Medieninhalte (Film und TV, Musik und Print, Online) von der Produktion von Inhalten (z.B. *Warner Bros*) bis zu deren Verteilung (Kabelnetze, Sender, *broad band*), unter einem Unternehmen zu kontrollieren. Jedoch wurde ihm das Platzen der Dotcom-Blase zum Verhängnis. Dagegen orientierte sich z.B. Murdochs *News Corp*, eigentlich fokussiert auf Satellitenfernsehen, neu, verkaufte 2006 seine Anteile an *DirecTV* und setzt seitdem auf das Internet. Schon lange verbinden sich ‚alte‘ und ‚neue‘ Medienunternehmen, um gemeinsam die digitale Welt zu erschließen, und nur selten gelingt es Neugründungen wie damals Google, in die boomenden Märkte einzudringen und selbst zum großen Player zu werden – üblicher sind Aufkäufe oder Kooperationen bekannter Akteure. Besonders der *Web 2.0*-Markt mit seinem *user-generated content* geriet ins Visier der Investoren. Google steigert seinen Werbeumsatz stetig: von 23 Mrd. \$ 2009 auf 59 Mrd. \$ 2014 (Statista 2015a). 2008 versuchte *Microsoft*, *Yahoo* zu übernehmen, was schließlich 2009 zu einer Kooperation führte. Murdoch hatte bereits 2006 das soziale Netzwerk *MySpace* erworben (und stieß es 2011 wieder ab). Die Konturen zwischen alten und neuen Medien verschwimmen, was einerseits die Akteure (Unternehmen der Medien- und Internetbranche) und andererseits die Nutzungsgewohnheiten der Konsumenten betrifft.

Im Folgenden werden die wichtigsten Punkte des ersten Teilbereichs dieser Arbeit noch einmal zusammengefasst, außerdem werden Deutschland bzw. Europa und die USA in Hinsicht auf Suchmaschinen kurz rechtsvergleichend betrachtet.

Zuerst wurde der theoretische Hintergrund zu Suchmaschinen als Kommunikationsmedien beleuchtet. Dabei wurde sich klassisch Massenkommunikation und Massenmedien über Kommunikation und Medien genähert.

Kommunikation ist eine spezifische Form sozialer Interaktion (vgl. Graumann 1972, S. 1110; Burkart 2002, S. 30), die der Bedeutungsvermittlung zwischen Lebewesen dient (vgl. Maletzke 1963, S. 18). Dazu braucht es immer ein Medium, das die Nachricht aufnimmt – Medien können sowohl personale Vermittlungsinstanzen als auch technische Hilfsmittel sein (vgl. Graumann 1972, S. 1182). Oft werden sie in die Gruppen ‚primäre Medien‘ (menschliche Elementarkontakte), ‚sekundäre Medien‘ (erfordern auf Produktionsseite ein Gerät) und ‚tertiäre Medien‘ (erfordern auf Produktions- und Empfängerseite ein Gerät) eingeteilt (vgl. Pross 1972, S. 127f). Durch die Digitalisierung und die Medienkonvergenz (Multimedia) entstanden digitale oder Online-Medien (darunter auch Suchmaschinen), die manchmal als ‚quartäre Medien‘ bezeichnet werden (vgl. Burkart 2002, S. 37f, Schulz 2009, S. 171). Medien wurden definiert als „komplexe institutionalisierte Systeme um organisierte Kommunikationskanäle von spezifischem Leistungsvermögen mit gesellschaftlicher Dominanz“ (vgl. Saxer 1998, S. 54 ursprünglich in 1980, S. 532; Faulstich 2004b, S. 27). Vier Faktoren, die jedem Kommunikationsprozess zugrundeliegen, sind der Kommunikator, die Aussage/Botschaft, ein Medium und der Rezipient (vgl. Burkart 2002, S. 63f).

Massenkommunikation wurde als Form von Kommunikation definiert, „bei der Aussagen öffentlich [...] durch technische Verbreitungsmittel [...] indirekt [...] und einseitig [...] an ein disperses Publikum [...] vermittelt werden“ (Maletzke 1963, S. 32). Es gibt eine Vielzahl theoretischer Ansätze zur Massenkommunikation – als einer von ihnen wurde die soziologische Systemtheorie nach Luhmann betrachtet. Nach ihr stellen Massenmedien ein soziales System dar, das operativ geschlossen und autopoietisch autonom ist (vgl. Luhmann 2004, S. 21f), und dessen interne Struktur sich in die drei Programmbereiche Nachrichten/Berichte, Werbung und Unterhaltung differenzieren lässt (vgl. ebd., S. 50f). Demnach liegt die Funktion der Massenmedien letztlich im Dirigieren der Selbstbeobachtung des Gesellschaftssystems (vgl. ebd., S. 118). Allgemein ist Massenkommunikation soziopolitisch, sozioökonomisch und soziokulturell eingebunden, und letztlich auch Bestandteil der Sozialisation (vgl. Pürer 2003, S. 84).

Als nächstes wurde versucht, Suchmaschinen als Medien über das Internet (und das Web) zu verorten.

Mit dem Aufkommen des Internets stellte sich auch die Frage nach der Konvergenz der Kommunikationsmedien bzw. der Multimedialität der Netzkommunikation, und daher stellen das Internet und Online-Kommunikation eine Herausforderung für Publizistik-, Kommunikations- und Medienwissenschaft dar (vgl. auch Merten 2007, S. 320; Burkart 2002, S. 368). Aus den daraus

resultierenden vielfältigen Einordnungsversuchen scheint sich das Verständnis des Internets als technische Infrastruktur oder Medium erster Ordnung (vgl. Joerges & Braun 1994, S. 19; Kubicek et al. 1997, S. 32ff, 218ff) durchgesetzt zu haben (vgl. z.B. Pürer 2003, S. 267f; Lang & Bekavac 2005, S. 433ff; Rußmann 2007, S. 23; Beck 2012, S. 261ff). Das Web als ein Dienst des Internets ist demnach ein Medium zweiter Ordnung (vgl. Beck 2006, S. 28; auch Wirth & Schweiger 1999, S. 47; Weinreich 1998, S. 132) – genauso wie Suchmaschinen in ihrer Funktion als Werkzeuge des Web Information Retrievals (vgl. Stock 2007, S. 2, 369-436; Ceri et al. 2013, S. 3f; s. auch Lewandowski 2005). Beck klassifiziert Suchmaschinen weiter als Metamedien (vgl. 2012, S. 172; auch Winkler 1997 (zur Kritik s. Röhle 2010, S. 35f); Beiler 2005, S. 165f).

Anschließend wurde ein Blick auf die Mediensysteme Deutschlands und der USA geworfen.

Dabei ist das deutsche Mediensystem von der Neuordnung durch die Alliierten nach dem Zweiten Weltkrieg und dem Beitritt der neuen Bundesländer nach der Wiedervereinigung 1989 geprägt (vgl. Beck 2012, S. 333ff). Wichtigste normative Basis ist GG Art. 5 Abs. 1, in dem die Kommunikationsfreiheiten als umfangreiche Menschenrechte garantiert werden. Das Mediensystem besteht aus den Teilsystemen Printmedien, also Buch und Presse (Zeitung und Zeitschrift); Film (Kino und Video); Rundfunk (Fernsehen und Hörfunk); und publizistische Online-Medien (vgl. ebd., S. 338). Sie alle bauen auf einer gemeinsamen normativen, technischen und infrastrukturellen Grundlage auf, die für die Beschaffung (Agenturen), die Verbreitung (Telekommunikation), und auch die Finanzierung (Werbung) wichtig ist.

Die Presse ist von der Werbewirtschaft abhängig und dadurch krisenanfällig (vgl. ebd., S. 343).

Das von den Alliierten geprägte *Public-service*-Modell des Rundfunks hat sich (wie in nahezu allen europäischen Staaten) zu einem dualen System (öffentlich-rechtlich / privat, gewinnorientiert) entwickelt (vgl. ebd., S. 346). Weltweit einmalig ist die föderalistische Art der Aufsicht durch die Landesmedienanstalten (vgl. Kleinsteuber 2007b, S. 264f).

Publizistische Online-Medien werden von professionellen Angeboten von Medienorganisationen aus dem Presse- und Rundfunksektor dominiert (vgl. Beck 2012, S. 348f). Die Zahlungsbereitschaft für publizistische (Qualitäts-)Inhalte ist beschränkt, und so wird in Deutschland bzgl. tragfähiger Geschäftsmodelle immer noch experimentiert.

Das Mediensystem der USA stellt einen marktwirtschaftlich-kapitalistischen Prototyp unter den Systemen der westlichen Welt dar, und kein anderes hat sich für den Rest der Welt als vergleichbar richtungsweisend erwiesen (vgl. Kleinsteuber 2009, S. 1211). Besonders für die europäischen Mediensysteme sind die USA eine Region der *Temporanz*: dort begannen Entwicklungen, die später übernommen oder eigenständig entwickelt wurden (vgl. Kleinsteuber 2007b, S. 264). Auf der

Schattenseite ist das Mediensystem der USA durchweg von hoher Konzentration geprägt (vgl. Weischenberg 2004, S. 167).

Wichtigste normative Basis ist das *First Amendment*, das die Pressefreiheit in der Verfassung verankert (vgl. ebd., S. 162). Die besondere Stellung der Pressefreiheit wird auch durch das Bild von der Presse als vierter Gewalt deutlich, die gegenüber der Politik eine Wächterrolle einnehmen soll (vgl. Kleinsteuber 2009, S. 1211f). Dies wird durch eine lange Tradition des investigativen Journalismus' unterstrichen.

Durch die Größe des Landes ist die US-Presse noch heute größtenteils lokal konstituiert.

Die FCC ist heute für den gesamten Radio- und Fernsehbereich sowie die Telekommunikation zuständig und vergibt Sendelizenzen.

Charakteristisch für den Rundfunk sind die *Big Four Network*-Gesellschaften NBC, CBS, ABC und FOX.

In Bezug auf das Internet halten seit dem 11. September 2001 zunehmend staatliche Überwachungsrechte Eingang in die Gesetzgebung. Insgesamt ist das Datenschutzniveau geringer als in Europa. Nahezu alle bedeutsamen Medien verfügen über eigene Internet-Portale.

Der Medienkonvergenz versuchen Unternehmen insbes. durch Aufkäufe oder Kooperationen zu begegnen.

Nimmt man nun bzgl. der beiden Regionen in Hinsicht auf Suchmaschinen eine rechtsvergleichende Perspektive ein, so überformen nach Paal (2012, S. 62-65) die Rechtssetzungsakte der EU vielfach die nationale Rechtsordnung, während die USA bereits eine längere Historie der Auseinandersetzung mit Suchmaschinensachverhalten aufweisen.

Auf europäischer Ebene kommt zur Eingrenzung der Marktmacht bei Vorliegen eines missbräuchlichen Verhaltens insbes. Art. 102 des *Vertrags über die Arbeitsweise der Europäischen Union* (AEUV) in Betracht – die Bestimmung entspricht weitgehend der Regelung im deutschen Recht (§§ 19, 20 *Gesetz gegen Wettbewerbsbeschränkungen* (GWB)). Die EU-Kommission untersucht seit 2010 mögliche Kartellrechtsverstöße durch Google (vgl. EU 2010), mit dem aktuellen Zwischenergebnis, das Unternehmen bevorzuge systematisch seinen Dienst *Google Shopping* (vgl. EU 2015). Es regt sich in mehreren EU-Mitgliedsstaaten Widerstand gegenüber dem Marktverhalten von Google – so entschied etwa ein Mailänder Gericht in Italien, dass Google teilweise für die Darstellung fremder Inhalte in den Ergebnissen verantwortlich sei, und in Frankreich attestierte die Wettbewerbsbehörde (*Autorité de la concurrence*) dem Unternehmen eine klar marktbeherrschende Stellung auf dem Markt für mit Suchmaschinen verknüpfte Online-Anzeigen.

Nach US-amerikanischem Recht greift zur Eingrenzung von Marktmacht von Suchmaschinenbetreibern insbes. der *Sherman Act* als Kernelement des Kartellrechts: Section 2 regelt ein Verbot der

Monopolisierung durch wettbewerbswidriges Verhalten. Unter Berufung auf diese Bestimmung erfolgte eine Beschwerde über Google bzgl. der Bevorzugung eigener Dienste von *Foundem* an die FCC. 2011 erfolgte eine Anhörung vor dem Justizausschuss des US-Senats mit dem Ziel, eine freiwillige Verhaltensänderung Googles zu bewirken. Der damalige Vorstandschef Googles Schmidt erklärte, Google wolle zwar eine bessere Informationspolitik betreiben, manipulierte seine Ergebnisse aber in keinem Fall. V.a. die Frage nach der Feststellung eines wettbewerbswidrigen Verhaltens nach Section 2 Sherman Act wirft Probleme auf. Würde man allerdings einen Kartellverstoß annehmen, bestünde in den USA – anders als nach deutschem und EU-Kartellrecht – die Möglichkeit, eine durch einen Verstoß erlangte Monopolmacht durch eine Zerschlagung bzw. Entflechtung aufzuheben.

Mit diesem Hintergrundwissen geht es im zweiten Teilbereich dieser Arbeit auf technischerer Ebene um Suchmaschinen.

### **3. Suchmaschinen**

In diesem Kapitel wird zuerst kurz auf die gesellschaftliche Bedeutung von Suchmaschinen und ihre Bezeichnung als ‚Gatekeeper‘ eingegangen, anschließend werden sie als technische Systeme betrachtet und es wird ein Blick auf Spezialsuchmaschinen geworfen.

Nach Lewandowski (2015, S. 2) haben Suchmaschinen eine immense gesellschaftliche Bedeutung. Erklärt werden kann dies zum einen durch ihre massenhafte Nutzung, zum anderen durch die Reihung und Darstellung der Suchergebnisse. Beides soll im Folgenden kurz vertieft werden.

Wie bereits in der Einleitung (Kap. 1) angesprochen, sind nach Lewandowski (2015, S. 2ff) Suchmaschinen mittlerweile der beliebteste Dienst des Internets – damit haben sie E-Mail abgelöst, die lange diesen Rang innehatte – sie werden von der klaren Mehrheit aller deutschen Internetnutzer mindestens wöchentlich genutzt (vgl. van Eimeren & Frees 2014, S. 386f; auch AGOF 2015, S. 11). In anderen Ländern, etwa den USA, ist es ähnlich (vgl. Purcell et al. 2012, S. 3, 5-8). Jeden Monat werden weltweit mehr als 175 Mrd. Anfragen an Suchmaschinen gestellt – mehr als 5,8 Mrd. pro Tag bzw. mehr als 2 Mio. pro Sekunde.

Auch der Online-Werbemarkt trägt zur Bedeutung von Suchmaschinen bei: Mit Internet-Werbung werden in Deutschland jährlich mittlerweile mehr als 7 Mrd. Euro umgesetzt – davon entfallen rund 35 Prozent, also mehr als 2,5 Mrd. Euro, auf Textanzeigen in Suchmaschinen (vgl. BVDW 2013, S. 5, 7). Für Unternehmen sind Suchmaschinen also bedeutende Werbepattformen. Wichtig ist für sie aber auch, dass sie von den Nutzern in den regulären Suchergebnissen gefunden werden. Um diese

Wahrscheinlichkeit zu erhöhen, wird mittels Optimierung der eigenen Website (*Suchmaschinenoptimierung* – dazu mehr im folgenden Kapitel) um die ersten Trefferpositionen in Suchmaschinen gekämpft – denn beim typischen Blickverlaufsmuster bzgl. Listen werden vor allem die ersten Elemente betrachtet (*Goldenes Dreieck / golden triangle*; vgl. Lewandowski 2015, S. 131).

Suchmaschinen werden oft als ‚Gatekeeper‘ des Internets bzw. Webs bezeichnet und kritisiert (vgl. z.B. Machill & Beiler 2008; Stark 2014). Dabei wird auf die publizistik- und kommunikationswissenschaftliche Kommunikatorforschung Bezug genommen, was im Folgenden kurz erläutert wird.

Nach Beck (2013, S. 179f), Jarren und Bonfadelli (2001, S. 264), Pürer (2003, S. 128f) und Burkart (2002, S. 276f) stammt das Konzept des *Gatekeepers*, des Torwächters oder Schleusenwärters, aus der Sozialpsychologie: Lewin entwickelte es 1947, um die Schlüsselpositionen einzelner Entscheidungsträger in sozialen Gruppen zu verdeutlichen (Lewin 1947); konkret ging es um das Einkaufsverhalten von Hausfrauen am Beispiel der Auswahl von Lebensmitteln (was kommt in den Einkaufswagen, was nicht). 1950 übertrug White das Konzept in einer Untersuchung grob vereinfacht auf den Prozess der Nachrichtenauswahl (White 1950); es ging darum, welche Nachrichten vom Redakteur einer Zeitung, Mr. Gates genannt, für die Veröffentlichung ausgewählt werden, und welche nicht. Es zeigte sich, dass Mr. Gates 90 Prozent des Materials nicht verwendete und den Lesern letztlich nur das anbot, was er selbst für die Wahrheit und die ‚Realität‘ hielt. Bei der Nachrichtenauswahl findet also ein Gatekeeping in erheblichem Umfang und z.T. in mehreren Stufen statt (Reporter, Agentur, Redakteur etc.). Wurden ursprünglich die persönlichen Ansichten und Einstellungen des selektierenden Kommunikators als ausschlaggebend angenommen, zeigte sich in weiterführenden Studien, dass auch Aspekte etwa sozialpsychologischer (der Gatekeeper als Träger einer Berufsrolle) und soziologischer Art (Strukturen und Funktionen einer Gesamtedaktion) eine Rolle spielen. Doch soll an dieser Stelle kein Abriss der Gatekeeper-Forschung erfolgen (für eine ausführlichere Darstellung s. etwa Weischenberg 2004, S. 317-327).

Mit der Analogie als Gatekeeper werden Suchmaschinen also zu Instanzen gemacht, die über die Weitervermittlung von Informationen an ein Massenpublikum entscheiden (vgl. Röhle 2010, S. 30; Granka 2010, S. 365). Hier gibt Lewandowski (2015, S. 180) zu bedenken, dass Suchmaschinen für jede Suchanfrage neue Suchergebnisseiten produzieren, also kein einheitliches Massenprogramm wie die Massenmedien bieten – außerdem wird nicht jede Suchergebnisseite einzeln von Menschen zusammengestellt. Oder etwas anders ausgedrückt: „Search engines universally apply an algorithm, whereas traditional news media makes case-by-case decisions“ (Granka 2010, S. 365).

Röhle schreibt, die Bezeichnung von Suchmaschinen als Gatekeeper erscheint sinnvoll, „wenn dadurch die empirische oder theoretische Fokussierung auf den Selektionsaspekt und dessen Effekte angezeigt werden soll“ (2010, S. 30). Bei näherer Betrachtung der erfolgten Weiterentwicklungen

und Differenzierungen des Gatekeeper-Konzepts wird jedoch deutlich, dass die Vielfalt von Einflussfaktoren auf die Selektionskriterien bei seiner Verwendung hinsichtlich Suchmaschinen keine Rolle spielt – dadurch erscheint der Suchmaschinen-Gatekeeper viel stärker als Ausgangspunkt denn als Resultat spezifischer Machtverhältnisse, und „die gesellschaftlichen Verhandlungen, die in die Technikentwicklungen eingeflossen sind, werden ausgeblendet“ (ebd., S. 31).

### 3.1 Suchmaschinen als technische Systeme

In diesem Abschnitt wird die Funktionsweise von Suchmaschinen erläutert, und es wird kurz auf Suchmaschinenoptimierung und Personalisierung eingegangen.

Eine Definition von Suchmaschinen lautet:

Eine Suchmaschine (auch: Web-Suchmaschine; Universalsuchmaschine) ist ein Computersystem, das verteilte Inhalte aus dem *World Wide Web* mittels *Crawling* erfasst und über eine Benutzerschnittstelle durchsuchbar macht, wobei die Ergebnisse in einer nach systemseitig angenommener Relevanz geordneten Darstellung aufgeführt werden (Lewandowski 2013, S. 495; Hervorh. i.O.).

Die Elemente dieser Definition seien kurz erläutert (vgl. Lewandowski 2015, S. 29f):

- ‚Computersystem‘: Die Suchmaschine wird von mehreren verbundenen Rechnern mit unterschiedlichen Funktionen gebildet.
- ‚verteilte Inhalte aus dem World Wide Web‘: Suchmaschinen durchsuchen nur das Web, nicht das ganze Internet (was z.B. E-Mails einschließen würde). Die Inhalte des Webs liegen in Form von Dokumenten auf Webservern vor. Erst durch die Verlinkungen entsteht ein Netz, das für Nutzer navigier- und für Suchmaschinen erfassbar ist.
- ‚Crawling‘: Das Erfassen von Web-Inhalten durch die Traversierung von Links.
- ‚Benutzerschnittstelle‘: Durch sie können Nutzer mittels Suchanfragen auf Suchmaschinen-Inhalte in Form von Suchergebnissen zugreifen.
- ‚systemseitig angenommene Relevanz‘: Es gibt keine einzig richtige Anordnung von Relevanz; für unterschiedliche Nutzer können unterschiedliche Dokumente relevant sein.
- ‚geordnete Darstellung‘: Die Ergebnispräsentation muss nicht in einfacher Listenform stattfinden, sondern kann komplexer sein.

Der grundlegende Aufbau von Suchmaschinen setzt sich zusammen aus dem *Crawler* (auch *Spider* oder *Robot*), der *Datenbasis (local store)*, dem *Indexer* und dem *Searcher*. Die prinzipielle Funktionsweise sieht so aus, dass der Crawler eine Kopie des Webs anlegt, der Indexer diese zu einem Index

aufbereitet, und der Searcher diesen anhand von Suchanfragen der Nutzer nach passenden Inhalten durchsucht. Dies ist in Abb. 5 dargestellt und wird im Folgenden näher erläutert.

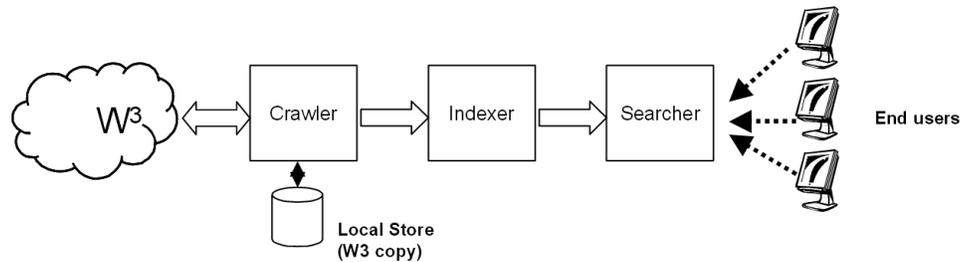


Abb. 5: Suchmaschinen-Referenzmodell (Risvik & Michelsen 2002, S. 290)

Nach Lewandowski (2015, S. 34ff, 37-40) indexiert der Crawler eine Ausgangsmenge bekannter Webseiten (*seed set*), folgt den Links auf ihnen, indexiert die so neu gefundenen Dokumente, folgt den Links auf wiederum ihnen usw. – so sollen möglichst alle im Web vorhandenen Dokumente gefunden werden. Dies ist jedoch nicht möglich. Erst einmal ist nur ein Teil des Webs über Links zugänglich: das sog. *Oberflächenweb (surface web)*, das aus vielen Milliarden Dokumenten besteht. Daneben gibt es jedoch das sog. *Deep Web (tiefes Web, auch Unsichtbares/Invisible Web; vgl. Sherman & Price 2000)*, das aus Inhalten besteht, die nicht gecrawlt werden können, weil sie nicht verlinkt, durch Passwortabfragen gesichert, oder dynamisch sind. Zur Größe des Deep Webs gibt es keine verlässlichen Angaben (vgl. Lewandowski 2015, S. 235). Weiter kann jedoch auch das Oberflächenweb aufgrund seiner Struktur nicht komplett gecrawlt werden, da es z.B. Bereiche enthält, von denen zwar verlinkt, auf die aber nicht zurückverlinkt wird (*Fliegenmodell / bow-tiemodel; Broder et al. 2000*).

Dennoch ist Crawling die wichtigste Methode zum Auffinden von Inhalten durch Suchmaschinen und die einzige, mit der tatsächlich weite Teile des Webs erfasst werden können.

Das Crawling muss kontinuierlich stattfinden, da ständig neue Dokumente zum Web hinzukommen und ‚alte‘ gelöscht oder verändert werden. Suchmaschinen priorisieren die zu besuchenden Dokumente v.a. nach Popularität (Verlinkungsstruktur und Zugriffe) und Aktualisierungsintervall (z.B. werden Nachrichtenwebsites häufiger besucht).

Neben dem Umstand, dass es für Suchmaschinen unmöglich ist, das komplette Web zu erfassen, ist das womöglich auch nicht wünschenswert; z.B. in Hinsicht auf *Dubletten (duplicate content)*, also mehrfach vorhandene Inhalte (bspw. Wikipedia-Artikel-Kopien auf verschiedenen Websites), oder *Spam*, also irrelevante Inhalte meist werblicher Natur. Auch gibt es Inhalte, deren Aufnahme zumindest strittig ist und die z.B. je nach Land verboten oder erlaubt sind (bspw. Holocaustleugnungen in Deutschland bzw. den USA), oder zur Unterdrückung in Suchmaschinen beantragte Inhalte, durch die sich Personen geschädigt fühlen (*Recht auf Vergessen*).

Nach Lewandowski (2015, S. 41-44) wird beim Crawling erst einmal alles erfasst, was gefunden wird. Suchmaschinen können jedoch von Websitebetreibern über Metadaten-Befehle gesteuert werden, die sich in einzelnen Dokumenten oder in der *robots.txt*-Datei, die mehrere Dokumente (bis hin zur ganzen Website) abdecken kann, befinden. Diese beiden Konventionen werden zusammen oft als *Robots Exclusion Standard* bezeichnet. Sie sind jedoch nicht verbindlich, sondern eine Übereinkunft zwischen den großen Suchmaschinenbetreibern.

Suchmaschinen schließen selbst Inhalte aufgrund technischer Beschränkungen und Spam-Identifikation aus, es erfolgt aber auch ein ‚Eingreifen von Hand‘ seitens der Betreiber: Dabei kann es sich vom Beugen vor staatlicher Zensur (etwa in China), bis zur Jugendschutz-Filterung oder dem Nachgehen von Urheberrechts-Beschwerden handeln.

Aus all diesen Gründen können bei der Erfassung von Web-Inhalten Unterschiede zwischen Suchmaschinen auftreten (obwohl womöglich derselbe Web-Bereich gecrawlt wurde).

Die Aufgabe des Indexers ist nach Lewandowski (2015, S. 48ff, 52f, 54f), die vom Crawler gelieferten Dokumente so aufzubereiten, dass sie in der Suche effizient verarbeitet werden können. Dabei werden Verfahren angewandt, die bei allen *Information-Retrieval*-Systemen zum Einsatz kommen. Das *Syntaxanalyse-System (parsing module)* zerlegt die gefundenen Dokumente in indexierbare Einheiten (Wörter, Wortstämme oder *N-Gramme*) und verzeichnet deren Vorkommen innerhalb des Dokuments. So wird ein invertierter Index erstellt, der für jede indexierte Einheit die Dokumente verzeichnet, in denen sie vorkommt. Dieses Prinzip ähnelt also dem von Buchregistern (wobei dort nur die für den Inhalt sinntragenden Begriffe aufgeführt werden). Invertierte Indices ermöglichen einen schnellen Zugriff, da nicht jedes Dokument nach den Suchbegriffen durchsucht, sondern nur abgeglichen werden muss, in welchen sie vorkommen (denn es wird ja der Index der Suchmaschine durchsucht, nicht das Web selbst). Außerdem werden invertierte Indices für das Ranking benötigt; so können den Indexierungs-Einheiten textstatistische Strukturinformationen beigefügt werden, z.B. dass ein Wort in der Überschrift eines Dokuments vorkommt und es sich daher für einen nach diesem Wort suchenden Nutzer um ein besonders relevantes Dokument handeln dürfte.

Bei der Indexierung anderer Dokumententypen als Text, etwa Bildern, Audiodateien und Videos, werden aus den betreffenden Dateien selbst (noch) nur basale Informationen gewonnen; wichtiger sind hier Kontextinformationen wie Metadaten oder Umgebungstexte.

Beim Aufbau von Indices gibt es v.a. aufgrund der Masse der zu indexierenden Dokumente technische Schwierigkeiten, außerdem muss auch das Indexieren kontinuierlich stattfinden, und so wird in der Praxis nicht mit einem zentralen Index, sondern mit einem verteilten System von Indices gearbeitet.

Suchmaschinen indexieren unterschiedlich und weisen ergo unterschiedliche Indices auf (wieder: obwohl womöglich derselbe Web-Bereich gecrawlt wurde).

Die Aufgaben des Searchers bestehen darin, Suchanfragen zu ‚verstehen‘ und relevante Suchergebnisse zu ranken (vgl. Lewandowski 2015, S. 58).

Nach Lewandowski (ebd., S. 69ff) werden Suchanfragen gemeinhin nach Broder (2002) in *informations-*, *navigations-* oder *transaktionsorientiert* unterschieden. Hinter informationsorientierten Anfragen steht ein problemorientierter Informationsbedarf; der Nutzer möchte sich über ein Thema informieren und liest daher mehrere Dokumente (z.B. ‚organspende‘). Mit navigationsorientierten Anfragen soll eine Seite (wieder)gefunden werden, die der Nutzer kennt oder von der er annimmt, dass sie existiert (bspw. ‚ebay‘). Transaktionsorientierte Anfragen zielen auf Websites ab, auf denen eine Transaktion stattfinden soll, etwa ein Kauf, ein Download oder eine Datenbankrecherche (z.B. ‚firefox download‘).

Zu dieser eher groben Kategorisierung kommt nach Lewandowski (2015, S. 59f, auch 72ff, 74-78), dass die meisten Nutzer sich bei der Formulierung ihres Suchziels keine allzu große Mühe geben (vgl. Machill et al. 2003, S. 169; Stark et al. 2014, S. 440); Suchanfragen sind oft kurz und unspezifisch, und deshalb lässt sich oft nicht aus ihnen herauslesen, was die Nutzer mit ihnen erreichen wollen. Mit diesem Problem beschäftigt sich die *Anfrageinterpretation (query understanding)*: Dabei werden die Suchanfragen mit Kontextinformationen angereichert, und zwar über den aktuell suchenden Nutzer, oder auch über frühere Nutzer, die die jeweilige Suche bereits ausgeführt haben.

Man kann einen Suchprozess als *Session* sehen; als Abfolge von Suchanfragen und Dokumentensichtungen innerhalb einer bestimmten Zeitspanne zu einem bestimmten Thema. Befindet sich der aktuelle Nutzer in einer Session, lassen sich aus seinen vorigen Suchanfragen, anschließenden Klicks auf Suchergebnisse und dortigen Verweildauern Rückschlüsse auf seine aktuellen Interessen ziehen (sucht er z.B. nach ‚golf‘ und hat vorher eindeutige Anfragen zu Autos eingegeben, wäre es nicht sinnvoll, ihm Ergebnisse zur Sportart auszugeben). Solche Sessionanalysen können auch in großem Maßstab zu früheren Nutzern durchgeführt werden.

Man unterscheidet zwischen impliziter und expliziter Anfrageinterpretation; erstere findet unbemerkt statt (z.B. auf den Standort des Nutzers ausgerichtete Suchergebnisse bei Google), bei letzterer wird der Nutzer über die Interpretation aufgeklärt (zu sehen z.B. bei der Suchmaschine *Wolfram Alpha*: [www.wolframalpha.com](http://www.wolframalpha.com)).

Bei der anderen Aufgabe des Searchers, dem Ranking, geht es nach Lewandowski (2015, S. 89f) darum, die zu einer Suchanfrage gefundenen Ergebnisse in eine sinnvolle Reihenfolge von absteigender Relevanz zu bringen. Über die Bestimmung des Begriffs ‚Relevanz‘ wird jedoch seit Langem diskutiert (vgl. Mizzaro 1997; Saracevic 2007a, b), und es bleibt festzuhalten, dass es

zumindest für informationsorientierte Anfragen keine objektive Relevanzbestimmung gibt: Was für den einen Nutzer relevant ist, kann für den anderen nutzlos sein; sucht z.B. ein Laie nach einem bestimmten Thema, sind für ihn andere Dokumente relevant als für einen Experten. Es gibt also nicht ‚das‘ richtige Ranking, sondern lediglich verschiedene algorithmische Sichtweisen auf die erfassten Web-Inhalte, die letztlich auf von Menschen getroffenen Annahmen beruhen.

Das Ranking wird von einer Vielzahl von Faktoren bestimmt; Google bspw. nennt hier die Zahl 200 (Google 2015a). Die genauen Verfahren und ihr Zusammenspiel halten Suchmaschinenbetreiber jedoch geheim, bildet das Ranking doch den Kern jeder Suchmaschine (vgl. auch Granka 2010, S. 366).

Nach Lewandowski haben sich bzgl. Rankingfaktoren sechs Bereiche herausgebildet (2015, S. 91f, 99, 122):

- **textspezifische Faktoren**  
Mit ihnen wird abgeglichen, ob und ggf. wo Wörter der Suchanfrage in den zu durchsuchenden Dokumenten vorkommen (z.B. in Überschriften).
- **Popularität**  
Die Popularität von Dokumenten wird v.a. durch ihre Verlinkung gemessen (*linktopologische Verfahren*; bspw. Googles *PageRank*-Algorithmus, der im Grunde anhand von *Inlinks* (Links, die auf eine Website verweisen) die Wahrscheinlichkeit angibt, mit der ein Nutzer, der sich nur über Links durchs Web bewegt, auf diese Website treffen würde (vgl. ebd., S. 101f)), aber auch durch das Klickverhalten der Nutzer (*nutzungsstatistische Verfahren*).
- **Aktualität**  
Die Aktualität von Dokumenten wird als Ausgleichsfaktor zu linktopologischen Verfahren (Popularität) herangezogen, die tendenziell eher ältere Dokumente bevorzugen.
- **Lokalität**  
Durch sie werden der Standort des Nutzers berücksichtigt und Dokumente (bzw. beschriebene Inhalte) aus der Nähe bevorzugt. Lokalität hat v.a. im Kontext der Suche mit mobilen Endgeräten an Bedeutung gewonnen.
- **Personalisierung**  
Eine Sonderform nutzungsstatistischer Verfahren (Popularität): Der Nutzer wird als Individuum gesehen und soll maßgeschneiderte Ergebnisse erhalten. Dazu werden v.a. seine früheren Suchanfragen herangezogen.
- **technische Rankingfaktoren**

Grundlegende technische Eigenschaften von Websites bzw. Servern, die für das Ranking genutzt werden; z.B. die Geschwindigkeit und Fehlerlosigkeit, mit der eine Seite geladen wird.

Am wichtigsten von diesen Bereichen ist die Textstatistik – die anderen Gruppen dienen darauf aufbauend v.a. der Qualitätsbewertung der gefundenen Dokumente.

(Es finden sich auch andere Einteilungen, z.B. in *On-Page*-, *On-Site*-, *Link*- und *Nutzerfaktoren* bei Griesbaum et al. 2009.)

Viele Betreiber optimieren ihre Websites, um im Suchergebnis möglichst hoch gerankt zu werden – man spricht dann von Suchmaschinenoptimierung. Im Folgenden wird ein Blick auf diesen Bereich geworfen und abschließend noch einmal der Rankingbereich ‚Personalisierung‘ herausgegriffen.

Unter Suchmaschinenoptimierung zählen „alle Maßnahmen, die darauf abzielen, die Position von Webseiten für bestimmte Anfragen in den Suchmaschinen-Ergebnislisten zu verbessern“ (Erlhofer 2013, S. 716). Zu diesem Zweck wird sich das (unvollständige) Wissen über die Rankingfaktoren von Suchmaschinen (s.o.) zunutze gemacht (vgl. Lewandowski 2015, S. 165). Diese werden bei der Suchmaschinenoptimierung klassisch in *On-Page*- und *Off-Page-Faktoren* unterteilt (vgl. ebd., S. 166-172; Erlhofer 2013, S. 25f, 411-456, 457-530):

*On-Page-Faktoren* sind auf der eigenen Website beeinflussbar. Dabei geht es um den Inhalt (*content*) (z.B. Verbesserung der Qualität, Ausrichtung auf passende Suchwörter/-phrasen, Sicherstellen der Aktualität), HTML (sinnvolle *Tag*-Nutzung, bspw. für Überschriften) und Website-Architektur (*site architecture*) (technischer Aufbau und technische Struktur; z.B. Vermeidung von *Flash* und mehrfach vorkommenden Inhalten (*duplicate content*), Reduzierung der Ladezeiten und Optimierung der Darstellung für mobile Endgeräte).

*Off-Page-Faktoren* liegen nicht im direkten Einflussbereich der Website. Grundsätzlich geht es darum, deren Popularität zu erhöhen bzw. für die Suchmaschinen höher erscheinen zu lassen. Der wichtigste Bereich ist hier das *Linkbuilding* (Linkaufbau; Sammeln von Inlinks) – weitere Maßnahmen zielen ebenfalls auf das Vertrauen der Suchmaschinen ab (etwa durch Shares und weitere Faktoren), und betonen z.B. die Social-Media-Relevanz (durch *Influencer*; wichtige Personen in sozialen Netzwerken, Anzahlen von Shares und Bewertungen).

Nach Lewandowski (2015, S. 172f) haben Suchmaschinenbetreiber ein ambivalentes Verhältnis zur Suchmaschinenoptimierung: Einerseits müssen sie sich vor übermäßiger Optimierung schützen, um eine gewisse Qualität der Suchergebnisse wahren zu können, andererseits profitieren sie von gut zu verarbeitenden Websites. Daher gibt es für die Suchmaschinenoptimierung Regeln (s. z.B. Google 2015c). Gehen Maßnahmen mit diesen konform, spricht man in Anlehnung an die Hackerkultur von

*White-Hat*-, andernfalls von *Black-Hat*-Optimierung (Praktiken wie z.B. *Keyword Stuffing* oder *Cloaking*) (vgl. auch Erlhofer 2013, S. 717, 708). Entdecken Suchmaschinen Black-Hat-Optimierungen, werden die betreffenden Websites in Form von Ranking-Verlusten oder sogar Entfernungen aus dem Index abgestraft.

„Mit Personalisierung wird allgemein das Anpassen eines Objekts an die Bedürfnisse eines Subjekts bezeichnet“ (Riemer & Brüggemann 2009, S. 153). Nach Lewandowski (2015, S. 118ff) erfolgt die Personalisierung von Suchergebnissen wie bei der Anfrageninterpretation aus im- und expliziten Nutzerverhaltensdaten. Jedem Nutzer werden auf Basis seines Profils unterschiedliche Ergebnisse bzw. eine unterschiedliche Ergebnisreihung ausgegeben. (Es ist ein Missverständnis, dass dabei Ergebnisse unterschlagen würden – diese werden lediglich auf niedrigeren Rankingpositionen ausgegeben, sodass sie i.d.R. nicht mehr wahrgenommen werden.) Damit soll eine bessere Qualität der Ergebnisse für den jeweiligen Nutzer erreicht werden (s. obiges Beispiel mit der ‚golf‘-Anfrage). Google gibt seit 2009 weltweit personalisierte Suchergebnisse aus (vgl. Google 2009).

Die Suchergebnis-Personalisierung wird häufig kritisch gesehen; durch sie würden Persönlichkeitsprofile erstellt, die personalisierten Ergebnisse würden v.a. die eigene Meinung bestätigen und andere unterdrücken, und zufällige Entdeckungen (*serendipity*) wären nicht mehr möglich bzw. unwahrscheinlicher. Bei seiner Kritik durch sein Buch *Filter Bubble* berücksichtigt Pariser (2012), dass Nutzer Medien seit jeher nach ihrem Geschmack ausgewählt hätten, allerdings würde durch die Personalisierung von Suchergebnissen (und sozialen Netzwerken) eine neue Dimension erreicht: So wäre die Filterblase auf jeden Nutzer individuell angepasst, bleibe für den Nutzer unsichtbar, und der Nutzer könne sich nicht für oder gegen die Personalisierung entscheiden – sie wird einfach durchgeführt.

Es ist allerdings unstrittig, dass die Suchergebnis-Personalisierung zu einer besseren Ergebnisqualität führen kann, da eben die zwar für die Masse, aber nicht für den einzelnen Nutzer, relevanten Ergebnisse ausgefiltert werden können.

Durch die Diskussion über die Filter Bubble hat sich die Debatte von einer Kritik an Suchmaschinen bzgl. massenkonformen, oberflächlichen Ergebnissen hin zu einer Kritik an der (zu starken) Anpassung an den einzelnen Nutzer verlagert (vgl. auch Jürgens et al. 2014, S. 108f). (Für eine ausführlichere Darstellung von Personalisierung s. etwa ebd., S. 104-113.)

### **3.2 Spezielsuchmaschinen**

In dieser Sektion geht es um Spezielsuchmaschinen, wobei der Fokus auf ihrer Einbindung in die *Universal Search* liegt.

Suchmaschinen sind Beschränkungen aus vier Problembereichen unterworfen, die z.T. bereits im vorigen Kapitel angeklungen sind (Lewandowski 2008, S. 54ff):

- Technische Probleme  
Das Web kann nicht vollständig erfasst werden; überdies werden beim Crawling Prioritäten nach ‚allgemeiner Bedeutung‘ von Dokumenten und nicht nach Bedürfnissen zu bestimmten Themen bzw. bestimmter Nutzergruppen ausgerichtet. Dadurch sind Suchmaschinen in spezifischen Bereichen oft kaum vollständig und aktuell.
- Finanzielle Hürden  
Neben den technischen Problemen wird die vollständige und aktuelle Erfassung der Web-Inhalte durch finanzielle Restriktionen i.d.R. noch weiter eingeschränkt. Suchmaschinenbetreiber müssen zwischen dem Anspruch einer möglichst vollständigen Erfassung und dem Einsatz finanzieller Ressourcen abwägen; das kann bedeuten, dass eine Erschließung nur bis zu einer bestimmten Tiefe erfolgt.
- Ausrichtung an einem einzigen Nutzermodell  
Suchmaschinen, die sich an eine allgemeine Nutzerschaft wenden, müssen immer Rücksicht auf die Ansprüche und Bedürfnisse des angenommenen ‚Durchschnittsnutzers‘ nehmen. Sie bieten z.B. keine themenspezifischen Einschränkungsmöglichkeiten bei der Anfrageformulierung, in den erweiterten Suchformularen oder auf den Suchergebnisseiten.
- Probleme der Erschließung  
Suchmaschinen behandeln alle Dokumente in ihrem Index, und damit auch die Dokumentrepräsentationen, gleich. Spezifische Informationen, die z.T. bereits in strukturierter Form in Dokumenten aus bestimmten Quellen vorliegen, gehen bei der Erschließung verloren.

Diesen Problemen versuchen Spezialsuchmaschinen zu begegnen. Eine Definition lautet: „Spezialsuchmaschinen [...] sind solche, die sich thematisch oder anhand formaler Dokumentenmerkmale (Bsp. Dateityp) beschränken“ (Lewandowski 2008, S. 57).

Die Vorteile solch einer Beschränkung sind, dass die ausgesuchten Inhalte vollständiger und tiefer erschlossen, und die Suchoberfläche sowie das Ergebnisranking an jeweilige Nutzergruppen angepasst werden können (vgl. ebd., S. 56).

Es gibt verschiedene Typen von Spezialsuchmaschinen, die sich durch unterschiedliche Charakteristika auszeichnen (vgl. Lewandowski 2015, S. 242f):

- Fokussiertes Crawling (*focused crawling*)  
Hierbei werden ausgewählte Bereiche des Oberflächenwebs erfasst. Die Quellen, die erschlossen werden sollen, werden von Hand ausgewählt (*white list*). So können Inhalte vollständiger und ggf. tiefer erschlossen werden.
- Bestimmter Dokumenttyp  
Ein zu erfassender Dokumenttyp kann technisch bestimmt sein (z.B. durch Bilder-Dateiformate) oder auch anhand eines Genres (bspw. Nachrichtentexte in Nachrichtensuchmaschinen). (Die Unterscheidung zum fokussierten Crawling ist nicht trennscharf.)
- Hybridsuchmaschinen  
Sie verbinden ausgewählte Inhalte des Deep-Webs mit solchen des Oberflächenwebs (z.B. *Google Scholar*).
- Archivsuchmaschinen  
Ein Sonderfall: Sie machen Inhalte aus dem Web in der jeweiligen Version dauerhaft verfügbar (bspw. *Wayback Machine*).

Ein bekanntes Beispiel einer Spezialsuchmaschine ist nach Lewandowski (2015, S. 244f) die Nachrichtensuchmaschine Google News. Nachrichtensuchmaschinen funktionieren ähnlich wie allgemeine Suchmaschinen, weisen aber einige Besonderheiten auf (vgl. Machill et al. 2005, S. 108f; Beiler 2008, S. 293ff). Google News erfasst per fokussiertem Crawling Inhalte von mehr als 700 manuell zusammengestellten deutschen Nachrichten-Quellen (vgl. Google 2015b). Durch diese Beschränkung können die Quellen in recht kurzen Abständen abgefragt werden. Vorteile sind qualitativ insgesamt hochwertigere Dokumente (durch die selektive Quellen-Auswahl) und deren höhere Aktualität. Außerdem werden die Rankingfaktoren (s. voriges Kapitel) anders gewichtet, wobei es v.a. um ein geeignetes Verhältnis von inhaltlicher Passung (Textstatistik) und Aktualität geht, was letztlich zu relevanteren Ergebnissen führt.

Dieses Beispiel zeigt außerdem, dass Spezialsuchmaschinen nicht komplett eigenständig sein müssen, sondern in allgemeinen Suchmaschinen integriert sein können. Tatsächlich bieten die großen Suchmaschinen heute i.d.R. gesonderte Kollektionen für besondere Inhalts-Arten (z.B. Bilder) oder bestimmte -Typen (bspw. wissenschaftliche Texte) an (vgl. Lewandowski 2015, S. 45). Zwar stellt sich eine solche Suchmaschine für den Nutzer einheitlich dar, doch besteht sie i.d.R. aus der *horizontalen* Suchmaschine, die das Web in seiner Breite abdecken soll und den Web-Index aufbaut, und zusätzlich mehreren *vertikalen*, eben Spezialsuchmaschinen, die in bestimmten Bereichen in die Tiefe gehen und mit speziellen Crawlern eigene (kleinere) Kollektionen aufbauen (vgl. ebd., S. 239, 47).

Diese speziellen Kollektionen konnten (und können immer noch) über Tabs erreicht werden, was von den Nutzern jedoch kaum wahrgenommen und genutzt wurde (man spricht in diesem Zusammenhang von der *Tab blindness*; Sullivan 2003). Das führte zu der Idee, den regulären Suchergebnissen aus dem Web-Index automatisch Ergebnisse aus den speziellen Kollektionen hinzuzufügen; das Prinzip der *Universal Search* – es ist in Abb. 6 dargestellt.

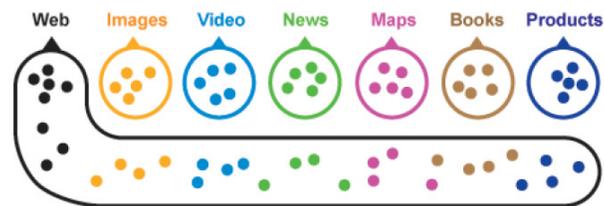


Abb. 6: Universal Search – Prinzip (Elliance 2007)

Bei einer normalen Suche werden also neben dem Web-Index auch die speziellen Kollektionen durchsucht und passende Treffer in das Ergebnis der Web-Suche eingebunden.

Nach Lewandowski (2015, S. 136ff) wird bei der Zusammenführung von regulären Ergebnissen aus dem Web-Index (auch *organische* Ergebnisse genannt) und solchen der Universal Search (sowie Werbung) kein gemeinsames Ranking durchgeführt, sondern innerhalb jeder Kollektion findet ein eigenes Ranking statt, und es wird lediglich festgelegt, wo die Universal-Search-Ergebnisse auf der Suchergebnisseite (*search engine result page* (SERP)) platziert werden. Die Darstellung dieser Ergebnisse findet in Containern oberhalb oder inmitten der organischen Ergebnisse statt und erfolgt je nach algorithmisch eingeschätzter Bedeutung des Ergebnistyps.

Suchmaschinen treffen bei der Zusammenstellung der SERPs also zwei Entscheidungen: ob bzw. welche Universal-Search-Container für die Suchanfrage relevant sind, und auf welchen Positionen sie angezeigt werden.

Im Laufe der Zeit kamen immer neue Kollektionen hinzu. Heute sind die wichtigsten Nachrichten (*News*), Bilder, Videos und lokale Ergebnisse.

Mit dem Prinzip der Universal Search haben Suchmaschinenanbieter die Gleichbehandlung aller Dokumente aufgegeben. Können aufgrund der gleichen Rankingfaktoren alle organischen Ergebnisse in die Trefferlisten gelangen, sind die Kriterien bei Universal-Search-Ergebnissen komplexer. Erst einmal muss ein Dokument bzw. seine Quelle gewisse Anforderungen erfüllen, um überhaupt in die Kollektion aufgenommen zu werden (denn das kann, wie bei Google News, händisch geschehen). Weiter werden Universal-Search-Ergebnisse unterschiedlich dargestellt, während organische Ergebnisse weitgehend gleich gestaltet sind; so werden z.B. Bilder als Thumbnails oder Videos durch Vorschaubilder dargestellt, was die Blicke der Nutzer auf den SERPs lenkt. Schließlich sind

Suchmaschinenbetreiber mittlerweile nicht mehr nur Vermittler, sondern selbst Inhalte-Anbieter, wodurch Konkurrenz entsteht und Suchmaschinen auf eigene Inhalte verweisen, was problematisch sein kann. Dies soll an dieser Stelle jedoch nicht vertieft werden.

Abb. 7 zeigt die SERPs von Google und Microsofts Bing zur Anfrage ‚star wars‘. Hervorgehoben sind dabei die in diesem Fall innerhalb der organischen Ergebnisse platzierten Universal-Search-Container: bei Google News und Bilder, bei Bing zusätzlich Videos. Ebenfalls hervorgehoben sind die Tabs, über die man direkt zu den jeweiligen Kollektionen gelangen kann. Dorthin kann auch über einen Link von jedem Universal-Search-Container aus gewechselt werden (z.B. „Weitere Nachrichten für star wars“).

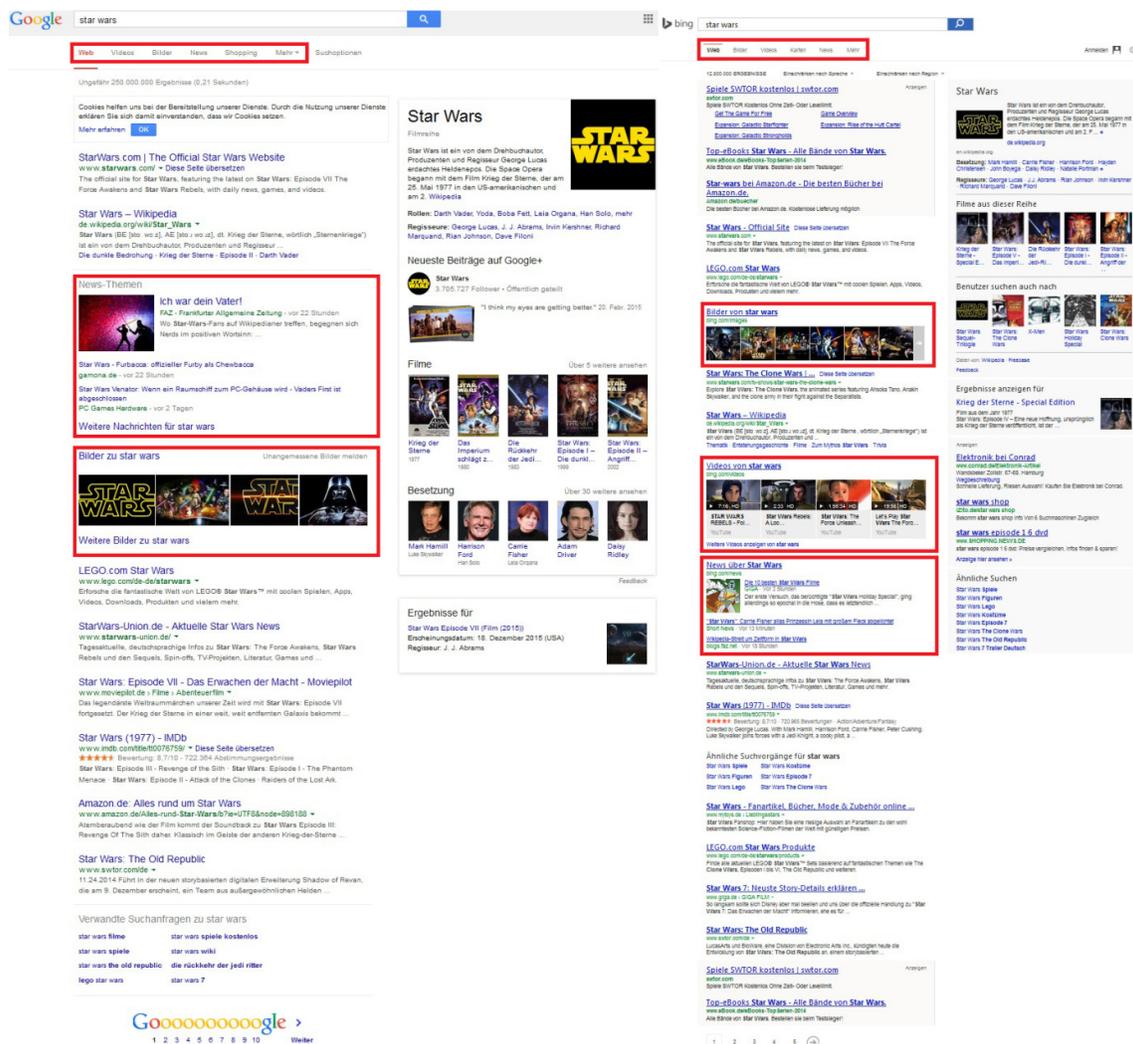


Abb. 7: Universal Search – Google und Bing (April 2015)

Die Universal Search war ein großer Schritt weiter weg von den berühmt-berüchtigten *ten blue links* (früher bestanden SERPs i.d.R. aus einer Liste von zehn schlichten Ergebnissen) hin zu einer vielfäl-

tigeren Ergebnisauswahl und -darstellung (vgl. Lewandowski 2015, S. 128f). Letztere hat großen Einfluss auf das Nutzerverhalten, weil durch sie das typische Blickverlaufsmuster (Goldenes Dreieck / *golden triangle*; s. Kap. 3) aufgebrochen wird, was wiederum Auswirkungen z.B. auf die Wahrnehmung von Werbung hat. (Zu den Auswirkungen der Universal Search auf das *User Interface* und die Nutzerführung s. Quirnbach 2009.) Nicht zuletzt entstehen durch die Universal Search viele Wege zu einem Dokument, was den Nutzern entgegenkommt (Lewandowski 2015, S. 254f).

## 4. Stand der Forschung

In diesem Kapitel geht es um das Kernthema dieser Arbeit: Suchmaschinenverzerrungen. Zuerst werden diese grundlegend behandelt, anschließend werden kurz neun Forschungsarbeiten zu ihnen vorgestellt.

### 4.1 Suchmaschinenverzerrungen

In diesem Abschnitt wird erläutert, was Suchmaschinenverzerrungen sind, wo ihre Ursachen liegen, und was gegen sie unternommen werden kann.

Das ideale Ergebnis einer Suchmaschine zu einer Anfrage besteht 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 (vgl. Lewandowski 2015, S. 272 – für weitere Definitionen s. etwa Weber 2011, S. 276f).

Hierzu folgt in Anlehnung an Weber (2011, S. 273ff) ein kleines Beispiel anhand eines Material-Ausschnitts der empirischen Untersuchung dieser Arbeit (Kap. 5): Tab. 4 zeigt die ersten zehn Ergebnisse von google.de und bing.de zur Suchanfrage ‚israel luftangriff syrien grenze‘ (die ersten drei sind jeweils Universal-Search-News-Treffer). Der Luftangriff Israels an der syrischen Grenze am 27. April 2015 war Einzelthema Nr. 7 der empirischen Untersuchung, an dem in Kap. 5.2.2 auch ihre praktische Methodik erläutert wird. (Die beiden SERPs finden sich im Original auf der beiliegenden CD.) Die Suchen wurden am 28. April 2015 unter Umgehung von Personalisierungen (s. Kap. 3.1), etwa durch Suchhistorie oder Standort, durchgeführt – diesbzgl. sind die Ergebnisse also ‚rein‘.

Tab. 4: Suchmaschinenverzerrungen google.de und bing.de (Thema 7)

	Google	Bing
1	<p><b>Golanhöhen: Israel fliegt Luftangriff auf mutmaßliche Terroristen</b> ZEIT ONLINE - vor 2 Tagen Das israelische Militär hat mehrere Männer an der Grenze zu Syrien attackiert. Sie sollen nach Armeeingaben einen Anschlag geplant haben.</p>	<p><b>Israel führt Luftschlag an Syriens Grenze</b> Deutsche Welle · Vor 1 Tag Die israelische Armee hat an der Grenze zu Syrien ... Tiefe Feindschaft Israel und Syrien sind erbitterte Feinde.</p>
2	<p><b>Konflikt zwischen Israel und Syrien Luftangriff auf Golanhöhen</b> taz.de - vor 1 Tag</p>	<p><b>Golanhöhen: Israelische Kampffjets fliegen Angriffe in syrischem Grenzgebiet</b> Spiegel Online · Vor 16 Stunden</p>
3	<p><b>Israel: Vier mutmaßliche Terroristen in Dolanhöhen an syrischer Grenze getötet</b> STERN - vor 1 Tag</p>	<p><b>Nahost-Konflikt Terrorverdächtige an syrischer Grenze gestorben</b> N24 · Vor 1 Tag</p>
4	<p><b>Konflikt zwischen Israel und Syrien: Luftangriff auf Golanhöhen - taz.de</b> www.taz.de/!158932/ vor 1 Tag ... JERUSALEM ap   Das israelische Militär hat an der Grenze zu Syrien einen Luftangriff auf mutmaßliche Extremisten geflogen. „Eine Gruppe ...</p>	<p><b>Golanhöhen: Israel fliegt Luftangriff auf mutmaßliche ...</b> www.zeit.de › Politik › Ausland Das israelische Militär hat an der Grenze zu Syrien einen Luftangriff auf mutmaßliche Attentäter geflogen. "Eine Gruppe bewaffneter Terroristen" habe ...</p>
5	<p><b>Israel: Vier mutmaßliche Terroristen in Dolanhöhen an syrischer ...</b> www.stern.de/.../israel-vier-mutmassliche-terroristen-in-dolanhoehen-an-syrischer-grenze-getoetet-2189984.html vor 1 Tag ... Mutmaßliche Terroristen getötet: Israel fliegt Luftangriff an syrischer Grenze. 27. ... Israelische Soldaten haben an der Grenze zwischen Syrien und Israel eine ... an der syrischen Grenze zu platzieren", hieß es aus israelischen ...</p>	<p><b>Israel: Vier verdächtige Terroristen an syrischer Grenze ...</b> www.spiegel.de › Politik › Ausland › Golanhöhen · Vor 1 Tag Bei einem Luftangriff des israelischen Militärs sind in den Golanhöhen im Grenzgebiet zwischen Israel und Syrien vier mutmaßliche Terroristen getötet ...</p>
6	<p><b>Luftangriff auf den Golan-Höhen : Israel tötet iranischen General in ...</b> www.tagesschau.de/ausland/israel-iran-syrien-101.html 19. Jan. 2015 ... Golan-Höhen an der Grenze von Israel und Syrien ... Schon in der Vergangenheit war Israel vereinzelt Luftangriffe in Syrien geflogen, unter ...</p>	<p><b>Israel führt Luftschlag an Syriens Grenze   Aktuell ...</b> www.dw.de/israel-führt-luftschlag-an-syriens-grenze/a-18409967 Israel Israel führt Luftschlag an Syriens Grenze. Die israelische Armee hat an der Grenze zu Syrien eine Gruppe mutmaßlicher Terroristen getötet.</p>
7	<p><b>Staatsfernsehen meldet Luftangriffe: Israel attackiert angeblich Ziele ...</b> www.tagesschau.de/ausland/israel-syrien-103.html 7. Dez. 2014 ... Golan-Höhen an der Grenze von Israel und Syrien ... Syrien meldet israelischen Luftangriff, K. Senz, ARD Kairo   audio · Weltatlas   Syrien ...</p>	<p><b>Israel fliegt Angriffe an syrischer Grenze</b> www.welt.de › Panorama › Leute Vor 1 Tag Israel tötet mutmaßliche Terroristen an syrischer Grenze Die israelische Armee hat an der Grenze zu Syrien eine Gruppe mutmaßlicher Terroristen getötet.</p>
8	<p><b>Terroristen an der israelisch-syrischen Grenze getötet   Israel ...</b> www.israel-nachrichten.org/archive/17401 vor 1 Tag ... Vier Terroristen die versuchen eine Bombe an Israels Grenze zu Syrien zu platzieren, wurden in einem israelischen Luftangriff getötet.</p>	<p><b>Israel: Vier mutmaßliche Terroristen auf Golanhöhen an ...</b> www.stern.de/.../israel-vier...syrischer-grenze-getoetet-2189984.html · Vor 1 Tag Israelische Soldaten haben an der Grenze zwischen Syrien und Israel eine Gruppe bewaffneter Männer getötet - es soll sich um Terroristen gehandelt haben</p>
9	<p><b>Nach Luftangriff: Spannungen an Grenze Israels mit Syrien und ...</b> www.faz.net/.../nach-luftangriff-spannungen-an-grenze-israels-mit-syrien-und-libanon-verschaerft-13395138.html 28. Jan. 2015 ... Die israelische Luftwaffe hat zwei syrische Militärstellungen angegriffen. ... Nach Luftangriff: Spannungen an Grenze Israels mit Syrien und ...</p>	<p><b>Nahost-Konflikt: Terrorverdächtige an syrischer Grenze ...</b> www.n24.de › Start › Nachrichten › PolitikVor 1 Tag Israel hat nach eigenen Angaben einen Anschlag durch Extremisten vereitelt. An der syrischen Grenze griff die Luftwaffe des Landes mehrere mutmaßliche ...</p>
10	<p><b>Luftangriffen nahe Damaskus: Israel sperrt Luftraum an der Grenze ...</b></p>	<p><b>Israel führt Luftschlag an Syriens Grenze - Worldnews.com</b></p>

<p>www.focus.de/.../luftangriffen-nahe-damaskus-israel-sperret-luftraum-an-der-grenze-zu-syrien_aid_981264.html          Bewertung: 5 - 1 Abstimmungsergebnis          6. Mai 2013 ... Nach dem israelischen Luftangriff nahe Damaskus hat Syrien mit Vergeltung gedroht. Zwar rechnet Israel nicht mit einem Gegenangriff des ...</p>	<p>article.wn.com/.../04/27/Israel_fuehrt_Luftschlag_an_Syriens_Grenze          27.04.2015 · DPA Israelischer Soldat an der Grenze zwischen Syrien und Israel ... der Stadt an der syrisch-türkischen Grenze allein mit Luftangriffen.</p>
--	--

Es zeigt sich, dass sich die Ergebnislisten erheblich voneinander unterscheiden. Gemein sind beiden Suchmaschinen nur zwei Treffer: Das *Zeit*-Ergebnis bei Google auf Platz 1 und bei Bing auf 4, und der *Stern*-Treffer bei Google auf Platz 3 und 5 und bei Bing auf 8. Allerdings doppelten sich Universal-Search-News-Treffer mit organischen Ergebnissen derselben Suchmaschine: Google 2 und 4, sowie 3 und 5, Bing 1 und 10, 2 und 5, sowie 3 und 9. An den wenigen Überschneidungen wird deutlich, dass das Ranking hier sehr unterschiedlich ist (auch zeigt das Beispiel, wie sich die Konsultation einer weiteren Suchmaschine lohnen kann), die Doppelungen der Universal-Search-News-Treffer zeigen Mängel in den Ergebnispräsentationen. Beides sind nach obiger Definition Suchmaschinenverzerrungen.

Nach Lewandowski (2015, S. 272) wird in diesem Kontext oft der Wunsch nach ‚neutralen‘ Suchmaschinen laut, doch kann es schon aufgrund der Funktionsweise von Suchmaschinen in Hinsicht auf das Ranking (und früher noch das Crawling) (s. Kap. 3.1) keine unverzerrten Ergebnisse geben (vgl. auch Weber 2011, S. 277).

So wäre es nach Weber (2011, S. 277) zumindest hilfreich, den Grad der Verzerrung qualitativ festzustellen, oder wenigstens die Ursachen zu kennen, um Gegenstrategien entwickeln zu können.

Arbeiten, die sich mit dem Abschätzen der Verzerrungen beschäftigen, werden in Kap. 4.2 vorgestellt, daher wird dieser Punkt hier nicht weiter behandelt.

Die Ursachen von Verzerrungen ordnet Weber (ebd., S. 278) in drei Kategorien ein: Implementierung, Anbieterverhalten und Nutzerverhalten. Dies wird im Folgenden erläutert.

Implementierung bezieht sich nach Weber (2011, S. 278f) auf die Indexierungs- und Suchalgorithmen von Suchmaschinen (s. Kap 3.1), auf rechtliche Vorgaben und auf mögliche Einflussnahmen seitens Suchmaschinenbetreibern.

Ein Verfahren zur Komplexitätsreduktion bei der Reaktion auf Suchanfragen ist das *Stemming*, bei dem anstatt jeder Wortform jeweils nur der Wortstamm indexiert wird. Insbes. natürlichsprachige Suchanfragen werden also auf ihre Wortstämme reduziert, nach denen dann gesucht wird. So kann der Umfang des Indexes reduziert werden und alle Wortformen, z.B. Deklinationen und Konjugationen, werden abgedeckt. Sollen verschiedene Sprachen berücksichtigt werden, müssen die jeweiligen sprachlichen Eigenheiten beachtet und als Algorithmen formuliert werden. Die Implementierung solcher Algorithmen entscheidet, ob das *Stemming* die Suchergebnisse verbessert oder

verzerrt. Uyar (2009a) zeigt am Beispiel von Google, dass die dort verwendeten Algorithmen großen Einfluss darauf haben, welche Inhalte gefunden und in welchem Ranking sie ausgegeben werden (allerdings scheint es, dass Google jede einzelne Wortform indexiert, nicht nur den Wortstamm). Es besteht also eine Wechselbeziehung zwischen der Unterstützung natürlichsprachiger Suchanfragen und der Ergebnislistenverzerrung. Und es ist denkbar, dass in Suchmaschinen weitere Mechanismen dieser Art mit ähnlichen Wirkungen implementiert sind.

Eine andere Ursache der Ergebnislistenverzerrung sind rechtliche Vorgaben – nicht nur in China werden Suchmaschinen reguliert und die Ausgabe bestimmter Inhalte verhindert. Nach FSM 2015 haben in Deutschland 2005 die bekanntesten Suchmaschinen (darunter Google) unter dem Dach der FSM die *Selbstkontrolle Suchmaschinen* gegründet, „um sich gemeinsam über einheitliche Standards zur Gewährleistung von Verbraucher- und Jugendschutz bei gleichzeitiger Wahrung der Meinungsfreiheit und der Vermeidung von Zensur zu verständigen“ (FSM 2015). Sie verpflichteten sich, URLs, die von der *Bundesprüfstelle für jugendgefährdende Medien* (BPjM) auf den Index jugendgefährdender Medien gesetzt wurden, nicht mehr auszugeben.

Für diese Regulierungen muss eine entsprechende Infrastruktur geschaffen werden, die auch zu anderen Zwecken genutzt werden kann: So könnten Suchmaschinenbetreiber über die Gestaltung ihrer Algorithmen Angebote von Konkurrenzunternehmen verdrängen oder gar ausschließen (s. Edelman 2010; Edelman & Lockwood 2011). Bspw. untersucht die EU-Kommission seit 2010 mögliche Kartellrechtsverstöße durch Google (s. EU 2010), und kam 2015 zu dem vorläufigen Ergebnis, dass das Unternehmen „seinen eigenen (derzeit ‚Google Shopping‘ genannten) Preisvergleichsdienst auf seinen allgemeinen Suchergebnisseiten systematisch bevorzugt, indem es beispielsweise Google Shopping auf dem Bildschirm sichtbarer platziert“ (EU 2015).

Unter dem Verhalten von Inhalte-Anbietern als Ursache von Verzerrungen versteht Weber die Suchmaschinenoptimierung. Diese wurde bereits in Kap. 3.1 behandelt.

Nach Weber (2011, S. 279) nehmen viele Suchmaschinen-Nutzer jeweils nur die erste Seite der Ergebnisliste wahr (vgl. Wirth et al. 2007, pdf S. 8f) – die übrigen Fundstellen werden weitgehend ignoriert. Daher ist es für Unternehmen, die ja vom Verkauf von Produkten oder Dienstleistungen leben, entscheidend, unter den Suchergebnissen möglichst weit oben zu stehen. Und tatsächlich ist nach Lewandowski (2015, S. 164) im Bereich Produkte und Dienstleistungen kaum noch eine Lücke zu finden, in der nicht bereits optimiert wurde – z.B. bei Suchanfragen in den Bereichen Versicherungsvergleiche und Kredite, wo ein Neukunde viel wert ist und oft Vermittlungsprämien gezahlt werden.

Mit Nutzerverhalten als Ursache von Verzerrungen meint Weber (2011, S. 280f), dass Nutzer die Möglichkeiten von Suchmaschinen oft nicht ausreizen, indem sie z.B. nur in einer Sprache suchen

(bspw. Deutsch, obwohl der gesuchte Gegenstand womöglich aus dem englischsprachigen Raum stammt), oder sich nur einer Standard-Suchmaschine wie Google bedienen (anstatt etwa Spezial-suchmaschinen (s. Kap. 3.2)). Die entsprechenden Ergebnisse sind hochgradig verzerrt, da sie jeweils nur eine kleine Auswahl der im (Deep) Web (s. Kap. 3.1) vorhandenen Inhalte darstellen.

Zu guter Letzt nennt Weber (2011, S. 281) noch Kombinationen von Implementierung, Anbieter- und Nutzerverhalten als Ursache von Verzerrungen. Ein Beispiel hierfür ist die Inselbildung von Inhalten im Web: Sie kommt zustande, indem Inhalte-Anbieter Suchmaschinen-Crawler anweisen, ihre Webseiten nicht zu indexieren (s. Kap. 3.1) (weil sie womöglich ihr geistiges Eigentum schützen wollen), und Suchmaschinenbetreiber dies respektieren. Das führt dazu, dass Inhalte zum Deep Web gehören.

Nach Jürgens et al. (2014, S. 108ff) hat die (im Beispiel am Anfang dieses Kapitels ausgeblendete) Personalisierung potenziell drastische Auswirkungen auf Suchmaschinenverzerrungen. Personalisierung wurde bereits in Kap. 3.1 behandelt.

Im ersten Moment erscheint sie als kleine Modifikation des ohnehin komplexen und schwer zugänglichen Auswahl- und Sortierverhaltens von Suchmaschinen, doch „mit ihr wandelt sich das Wirkungspotenzial von Search Engine Bias von einer uniformen Beeinflussung zu einer – zumindest potenziell – differenziert und individuell wirkenden Kraft“ (ebd., S. 108f). Im Grunde verlagert sich so die Quelle der Verzerrungen von der Suchmaschine auf den Nutzer. Daraus ergeben sich zwei neue Wirkungen:

1. Personalisierung kann zu einem individuellen ‚rich-get-richer-Effekt‘ führen, der die ohnehin ausgeprägten Interessen des Nutzers weiter in den Vordergrund rückt. Dieser Effekt verstärkt sich selbst, denn mit der Zeit basieren die gesammelten Nutzerdaten auf bereits personalisierten Suchergebnissen. Dies kann die Vielfalt der Suchergebnis-Inhalte immer weiter reduzieren (*overspecialization* – s. Adomavicius & Tuzhilin 2005, S. 737).
2. Wenn Suchmaschinen Nutzern unterschiedliche Ergebnisse zu denselben Suchbegriffen liefern, löst das möglicherweise die ehemals geteilte Wahrnehmung von Themen auf und führt im Extremfall langfristig zu einer Fragmentierung der Öffentlichkeit (*Filter Bubble*; Pariser 2012, vgl. auch Webster & Ksiazek 2012, Napoli 2011).

Dieses dystopische Szenario erscheint aus dem Blickwinkel der Praxis jedoch nicht sonderlich realistisch, merken Jürgens et al. (2014, S. 111) an.

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

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

Würden Nutzer viele Suchmaschinen verwenden, hätte zwar jede spezifische Verzerrungen, doch würden sich diese gewissermaßen gegenseitig ausgleichen (außerdem wären systematische Verzerrungen im direkten Vergleich unterschiedlicher Suchmaschinen leichter zu entdecken) (vgl. Lewandowski 2015, S. 273).

Bereits im Jahr 2000 wiesen Inirona & Nissenbaum (2000, S. 169f, 175-178, 181f) 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 (demokratische Möglichkeiten; öffentlicher Nutzen, Verbreitung von Macht und Zugang zu Informationen), 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.

10 Jahre später schaut Granka (2010, S. 364f) auf Inironas 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 Inironas 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 (2010, S. 365f, 368f) 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 der Algorithmen, sodass die Nutzer sehen könnten, falls eine Suchmaschine von ihren organischen abweicht, um Ergebnisse einzubinden, von denen die Suchmaschinenbetreiber finanziell profitieren.

Nach Granka (2010, S. 369ff) gingen Introna und Nissenbaum bei ihrer Kritik an einem Suchmarkt bzgl. der Auswahl von Suchmaschinen von einem zu simplen Nutzerverhalten aus – z.B. Eyetracking zeigt, dass Nutzer sich diesbzgl. unterschiedlich verhalten. Das fördert die Entwicklung eines stabilen Online-Informationsmarktes, v.a. eines, bei dem die besten Suchmaschinen am meisten besucht werden. Wenn sich Suchmaschinenbetreiber also ethisch am Nutzerverhalten orientieren, bedarf es keiner Regulierung – nur wenn sie ihre Macht missbrauchen, besteht die Gefahr einer Marktverschlechterung.

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 spezifischeren).

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.

Generell kritisiert Granka (ebd., S. 371f) die vielfache Verwendung von aggregierten Skalen in der Suchmaschinenforschung, insbes. bzgl. der inhaltlichen Vielfalt (*Diversity*; deren Definition auch noch Probleme bereitet). So wird nur der Massengeschmack auf der Makroebene betrachtet und es

werden kaum brauchbare Ergebnisse erzielt. Nützlicher wäre es womöglich, Suchergebnisse für einzelne Anfragen, also auf der Mikroebene, zu betrachten.

Nach Weber (2011, S. 282f) gibt es einige Gegenstrategien bzgl. Suchmaschinenverzerrungen. Zuerst einmal sollten Alternativen genutzt werden (s. als Beispiel Tab. 4). Dies betrifft nicht nur die Konkurrenz zwischen Google und Bing, den beiden einzigen relevanten Wettbewerbern (von länderspezifischen Ausnahmen wie z.B. *Baidu* in China abgesehen), sondern insbes. für die Nutzung bereichsspezifischer Suchmaschinen bspw. von Bibliotheken, Verlagen und anderen Informationsanbietern – durch sie ist eine wenigstens teilweise Erschließung des Deep Webs möglich.

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 anzuschauen.

Viele Nutzer neigen dazu, nur der eigenen Meinung entsprechende Inhalte wahrzunehmen und damit *kognitive Dissonanz* zu vermeiden (vgl. z.B. Weber & 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.

Alle diese Strategien können helfen, Suchmaschinenverzerrungen zu umgehen, ändern jedoch nichts an der eigentlichen Problematik.

## **4.2 Verwandte Forschung**

Im vorigen Abschnitt wurde beleuchtet, was Suchmaschinenverzerrungen sind, wie sie entstehen, und was Nutzer gegen sie tun können. In dieser Sektion liegt der Fokus darauf, wie sich Suchmaschinenverzerrungen messen lassen – diesbzgl. werden chronologisch einige bedeutende Arbeiten aus dem Zeitraum 2002 bis 2013 vorgestellt. Dies geschieht in stark verkürzter Weise, und die Ergebnisse der älteren sind inzwischen freilich veraltet, doch geht es darum, einen Eindruck der Vielfältigkeit von Suchmaschinenverzerrungen sowie Ansätzen, sie nachzuweisen, und damit der Weite dieses Felds zu vermitteln.

Ein Zwischenfazit am Ende der Sektion fasst die wichtigsten Punkte des zweiten Teilbereichs dieser Arbeit, Suchmaschinen und Stand der Forschung, nochmals zusammen.

Nach Mowshowitz und Kawaguchi (2002a, S. 142; b, S. 57) sind Suchmaschinenverzerrungen unangemessener Ein- oder Ausschluss von Ergebnissen bzw. unangemessene Gewichtungen – die Schwierigkeit dabei ist, ‚unangemessen‘ zu definieren. Weil dazu irgendeine Form von Norm benötigt

wird (vgl. Mowshowitz & Kawaguchi 2002a, S. 145), schlugen sie vor, die Gesamtleistung einer Gruppe ähnlicher Suchmaschinen als Norm zu definieren und dann die Leistungen einzelner Suchmaschinen mit dieser zu vergleichen.

Konkret entwickelten sie eine Art Metasuchmaschine (diese haben keinen eigenen Datenbestand, sondern führen Ergebnisse von anderen Suchmaschinen zusammen (vgl. Lewandowski 2015, S. 21f)), mit der sie jeweils Ergebnis-URLs der abgedeckten Suchmaschinen zu Suchanfragen erfassen und untersuchen konnten.

Sie deckten Themenbereiche ab, indem sie jeweils mehrere verwandte Suchanfragen stellten, z.B. zu ‚home refrigerators‘ zusätzlich ‚fridges‘, ‚iceboxes‘, ‚refrigerators‘ etc. Dieses Pooling stellte eine bessere Basis für die späteren Messungen der Abweichungen dar (vgl. Mowshowitz & Kawaguchi 2002b, S. 57). Alle erfassten URLs bildeten dann (beliebig nach Domains trunkiert) in ihrer Einzigartigkeit, Häufigkeit und ggf. Position die Grundlage für statistische Berechnungen mit URLs einzelner Suchmaschinen (vgl. Mowshowitz & Kawaguchi 2002a, S. 144-151).

Die Autoren fanden heraus, dass Suchmaschinenverzerrungen weder von den Themen abhängig zu sein schienen (was sie ursprünglich angenommen hatten (vgl. ebd., S. 154)), noch von den ausgewählten Suchtermen (vgl. Mowshowitz und Kawaguchi 2002b, S. 59f).

Vaughan und Thelwall (2004) widmeten sich Verzerrungen durch die länderspezifische Abdeckung von Suchmaschinen. In ihrer Studie untersuchten sie die Abdeckung dreier Suchmaschinen (Google und die inzwischen eingestellten *AltaVista* und *AlltheWeb*) bzgl. vierer Länder (die USA, China, Taiwan und Singapur). Ihre Fragestellung lautete, ob die Abdeckung bedeutender Suchmaschinen in nationaler oder linguistischer Hinsicht verzerrt ist, und wenn ja, was die Gründe dafür sind (vgl. ebd., pdf S. 2).

Um dies zu beantworten, nahmen sie Stichproben kommerzieller Websites aus jedem der vier Länder, stellten die Größen dieser Sites durch einen angepassten Crawler anhand der Anzahl ihrer Webseiten fest, suchten die Websites mit den Suchmaschinen (bei Google etwa mit dem ‚site:‘-Operator), und verglichen die Ergebnisse mit den Crawler-Daten. Zusätzlich betrachteten sie jeweils die Inlinks (vgl. ebd., pdf S. 3, 7).

Die Auswertung bestand in der Berechnung des Prozentsatzes der abgedeckten Websites und des Abdeckungsverhältnisses der Suchmaschinen bzgl. der Länder (Anzahl abgedeckter Webseiten / Anzahl gecrawlter Webseiten) (vgl. ebd., pdf S. 3, 7).

Vaughan und Thelwall fanden heraus, dass es bzgl. des Prozentsatzes der abgedeckten Websites einen starken Unterschied zwischen den USA und den anderen Ländern gab (bei den USA lag er deutlich höher). Auch bzgl. des Abdeckungsverhältnisses gab es signifikante Unterschiede zwischen den Ländern und den Suchmaschinen (deutlich mehr US-Websites wurden abgedeckt).

Als mögliche technische Ursachen für die Verzerrungen wurden die verschiedenen Sprachen (Englisch und Chinesisch) und die Sichtbarkeit der Websites (gemessen an den Inlinks) betrachtet (vgl. ebd., pdf S. 10). Während die Ergebnisse jedoch nicht auf den sprachlichen Faktor zurückgeführt werden konnten, spielte die Sichtbarkeit der Websites tatsächlich eine Rolle: Je mehr Inlinks es gab, desto besser waren die Sites i.d.R. abgedeckt (vgl. ebd., pdf S. 10f).

Cho und Roy (2004) untersuchten die Auswirkungen von Suchmaschinen auf die Popularitätsentwicklung von Websites. Wenn Suchmaschinen auf den ersten Trefferpositionen gerade sehr populäre Websites ausgeben, werden dadurch neue Sites, die noch nicht so bekannt sind, benachteiligt? Erlangen so populäre Websites immer mehr Popularität, während neue ignoriert werden? Um diese Fragen zu beantworten, führten die Autoren eine experimentelle Studie durch und untersuchten anschließend im Vergleich mit einem suchmaschinenfreien Szenario (in dem sich Nutzer nur durch das Verfolgen von Links durchs Web bewegen), wie viel länger es im Suchmaschinenkontext für eine neue Website dauert, eine große Zahl Nutzer anzuziehen (vgl. ebd., pdf S. 1).

Im Rahmen der experimentellen Studie wurden über mehrere Monate hinweg zweimal einige Websites heruntergeladen, und anhand der Inlinks die PageRanks berechnet und verglichen (vgl. ebd., pdf S. 2f). Als Ergebnis zeigte sich, dass die bereits populären Websites tatsächlich noch populärer geworden waren, der ‚rich-get-richer‘-Effekt also tatsächlich vorhanden war (vgl. ebd., pdf S. 3f).

Um zu eruieren, wie viel davon auf Suchmaschinenverzerrungen zurückzuführen war, d.h. um zu sehen, wie es mit der Popularität von Websites bestellt wäre, wenn Suchmaschinen sie *nicht* nach ihrer aktuellen Popularität ranken würden, entwarfen Cho und Roy zwei Nutzermodelle: das *Random-surfer model*, bei dem Nutzer keine Suchmaschinen benutzten, und das *Search-dominant model*, bei dem Nutzer immer von einer Suchmaschine aus ins Web starteten (vgl. ebd., pdf S. 2ff). Das Ergebnis einiger statistischer Untersuchungen mit diesen Modellen (vgl. ebd., pdf S. 6ff) war, dass neue Websites im Suchmaschinenkontext ein Vielfaches länger brauchten, um populär zu werden, als im suchmaschinenfreien (es wird konkret von 66-mal so lange gesprochen (vgl. ebd., pdf S. 6)). Demnach hätten Suchmaschinen eine besorgniserregende Wirkung auf das Entdecken neuer Websites durch Nutzer (vgl. ebd., pdf S. 1).

Joachims et al. (2007) untersuchten die Zuverlässigkeit impliziten Feedbacks aus Klickdaten und Suchanfragen-Reformulierungen in der Websuche. Dazu analysierten sie mittels einer eyetracking-unterstützten Studie die Interaktion von ca. 30 Studierenden mit Google-Suchergebnisseiten, und wie ihr Verhalten als Relevanzbewertungen interpretiert werden konnte (vgl. ebd., pdf S. 2, 5).

Dabei gab es drei experimentelle Szenarien: Die Teilnehmer erhielten jeweils entweder die Top-10-Treffer im originalen Ranking, mit vertauschtem ersten und zweiten Ergebnis, oder in umgekehrter Reihenfolge (vgl. ebd., pdf S. 2, 5).

Joachims et al. fanden heraus, dass die Teilnehmer vernünftige Entscheidungen trafen und ihre Klicks Relevanzbewertungen widerspiegeln – dass die Klickentscheidungen aber auf mindestens zwei Weisen verzerrt waren (vgl. ebd., pdf S. 2). Erstens führte eine Vertrauensverzerrung (*trust bias*) zu mehr Klicks auf von Google hoch gerankte Links, sogar wenn diese Trefferbeschreibungen (*Snippets*) weniger relevant waren als andere, die die Teilnehmer gesehen hatten (vgl. ebd., pdf S. 13f). Zweitens gab es eine Kontextqualitäts-Verzerrung (*quality-of-context bias*); die Klickentscheidungen der Teilnehmer wurden nicht nur durch die Relevanz des geklickten Links, sondern auch durch die Gesamtqualität der anderen Snippets beeinflusst (vgl. ebd., pdf S. 14).

Beim Vergleich mit expliziten Relevanzbewertungen, die Joachims et al. ebenfalls im Rahmen der Studie durchführen ließen, kam heraus, dass implizites Feedback aus Klicks sowohl innerhalb eines Ergebnis-Sets als auch zwischen mehreren bei Suchanfragen-Ketten (durch Reformulierungen) unter Anwendung verschiedener Strategien zur relativen Interpretation gut mit diesen übereinstimmte (vgl. ebd., pdf S. 22ff).

Die ‚Vertrauensverzerrung‘ – dass Nutzer bevorzugt hoch gerankte Ergebnisse anklicken – wird auch *Position-Bias*-Effekt genannt (vgl. Yue et al. 2010, pdf S. 1) und ist mittlerweile einer der am meisten erforschten Verzerrungseffekte (vgl. Jeong et al. 2012, pdf S. 1).

Brown et al. (2008) verglichen die Ergebnisse von Suchmaschinen mit einer traditionellen, bibliotheksbasierten Suche. Hierzu trugen sie zunächst in einer mehrjährigen Bibliotheksrecherche alle Erwähnungen zehn fossiler Amphibien aus dem Perm in Zeitschriften und anderen Periodika zusammen, die sie finden konnten – dies betraf den Zeitraum 1895 bis 2006 (vgl. ebd., S. 279, 281). Das Ergebnis dieser Suche stellte die Norm dar. Dann suchten sie die fossilen Amphibien mit drei Suchmaschinen (darunter *Google Scholar*), drei bibliografischen Datenbanken und einem aggregierten Zeitschriften-Archiv, und verglichen die Anzahlen der ausgegebenen Erwähnungen mit dem Bibliotheksrecherche-Ergebnis (vgl. ebd., S. 281).

Es kam heraus, dass die sieben Suchmaschinen zusammengenommen weniger als die Hälfte der in der Bibliotheksrecherche gefundenen Erwähnungen abdeckten (41 Prozent; einzeln betrachtet variierten die Abdeckungen von 4 bis 23 Prozent) (vgl. ebd., S. 282). Auch bei einer Betrachtung der Abdeckungen über den berücksichtigten Zeitraum im Detail, der fünf Zeitschriften mit den meisten Erwähnungen, der fünf produktivsten Autoren und der nicht-englischen Erwähnungen zeigten sich bei den Suchmaschinen Ineffizienz und Verzerrungen (vgl. ebd., S. 279, 283-286).

So empfahlen Brown et al., einige der besten Suchmaschinen lediglich als Einstieg für eine traditionelle Bibliotheks-Recherche zu nutzen (vgl. ebd., S. 287).

Alle bedeutenden Suchmaschinen geben mit den Suchergebnisseiten Trefferanzahl-Angaben zur jeweiligen Suchanfrage aus, die etwa bei Einwort-Anfragen bis in den Milliardenbereich gehen können (vgl. Uyar 2009b, S. 470, 472). Uyar (2009b) untersuchte die Genauigkeit dieser Angaben bei Google, Yahoo und Microsofts *Live Search* (heute Bing). Die Forschungsfragen lauteten dabei, wie genau die Trefferanzahl-Angaben für Einwort- und Mehrwort-Anfragen sind, und ob die Genauigkeit durch ein Monitoring über einige Tage hinweg möglicherweise steigt.

Bei der Untersuchung wurden von den drei Suchmaschinen zu Suchanfragen über ein *Java*-Programm die Ergebnisseiten inkl. der Trefferanzahl-Angaben abgerufen und erfasst. Die tatsächliche Trefferanzahl wurde über die letzte Ergebnisseite festgestellt – dann wurden die angegebene und die tatsächliche Trefferanzahl verglichen (vgl. ebd., S. 472). (Die drei Suchmaschinen gaben standardmäßig lediglich die ersten 1.000 Ergebnisse zu Suchanfragen aus, sodass Anfragen aus Zahlen verwendet wurden, bei denen die Trefferanzahlen niedriger lagen (vgl. ebd., S. 470).)

Bzgl. der Tests zu Einwort-Suchanfragen kam heraus, dass alle drei Suchmaschinen kein exaktes Zählen von Ergebnissen durchführten, sondern Schätzungen ausgaben, die sich in ihren Algorithmen stark voneinander unterschieden (vgl. ebd., S. 473, 479). Am besten schnitt Google ab.

Bei den Tests zu Mehrwort-Suchanfragen (solche aus zwei bzw. fünf Zahlen) zeigte sich, dass sich bei allen drei Suchmaschinen die Genauigkeit der Trefferanzahl-Angaben um fast die Hälfte reduzierte (vgl. ebd., S. 477, 479): „With the increase in the number of query words, the error in estimation increases and the number of accurate estimations decreases“ (ebd., S. 479).

Bzgl. der Tests zur Konsistenz der Trefferanzahl-Angaben über einige Tage hinweg kam heraus, dass sich die Schätzungen aller Suchmaschinen mit der Zeit nicht der tatsächlichen Trefferzahl annäherten, sondern i.d.R. sehr ähnlich blieben (vgl. ebd., S. 478f).

Yue et al. (2010) beschäftigten sich über den *Position-Bias*-Effekt (s.o.) hinaus mit der Attraktivität von Suchergebnissen als potenzielle Quelle für Präsentationsverzerrungen. Genauer untersuchten sie den Einfluss gefetteter Wörter innerhalb der Titel und Trefferbeschreibungen (die also mit der Suchanfrage übereinstimmen) auf das Klickverhalten von Nutzern (vgl. ebd., pdf S. 1). Dafür überprüften sie Klickdaten auf den *Position-Bias*-Effekt hin und kontrollierten die Qualität der Suchergebnisse durch manuelle Bewertungen (vgl. ebd., pdf S. 1).

Da Nutzer Suchergebnisse typischerweise in der gerankten Reihenfolge durchgehen, zeugt ein Klick auf ein höher geranktes Ergebnis nicht unbedingt von (angenommener) Relevanz – wird allerdings ein niedriger geranktes angeklickt, kann davon ausgegangen werden, dass es dem höheren

vorgezogen wird (vgl. ebd., pdf S. 2). Auf dieser Basis verwendeten Yue et al. den *FairPairs*-Algorithmus (vorgestellt von Radlinski & Joachims 2006), um rankingbenachbarte Suchergebnis-Paare in eins von zwei Schemata zu gliedern (1 & 2, 3 & 4, 5 & 6 usw. oder 2 & 3, 4 & 5, 6 & 7 usw.) und dann in originaler Reihenfolge oder getauscht Nutzern zu präsentieren (vgl. Yue et al. 2010, pdf S. 2). So sammelten sie *FairPairs*-Daten aus Google-Traffic, hunderttausende Klicks – dabei wurden mehr obere als untere Ergebnisse angeklickt (vgl. ebd., pdf S. 3).

Zu einer Stichprobe der *FairPairs*-Daten wurden manuelle Relevanzbewertungen vorgenommen – dabei wurden zu jedem ausgewählten *FairPair* fünf Bewertern die Suchanfrage und zwei konkurrierende URLs vorgelegt, und sie mussten sich für eine entscheiden, nachdem sie beide Ergebnis-Webseiten begutachtet hatten (vgl. ebd., pdf S. 3). Wie bei den Klicks wurden auch hierbei mehr obere als untere Ergebnisse bevorzugt (vgl. ebd., pdf S. 3).

Die Klickdatenanalyse und die manuellen Bewertungen wurden mittels eines *logistic regression model* zusammengebracht (vgl. ebd., pdf S. 4f), mit dem die Wahrscheinlichkeit eines Klicks auf das ursprünglich höher gerankte Ergebnis eines *FairPairs* anhand von vier Faktoren bestimmt werden konnte: die Titelfettungs-Unterschiede, die Trefferbeschreibungsfettungs-Unterschiede, ob das *FairPair* getauscht wurde und die relative Qualität aus der manuellen Bewertung (vgl. ebd., pdf S. 4).

Yue et al. fanden heraus, dass Fettungen in Titeln eindeutig einen positiven Effekt auf Klicks hatten (bzgl. Trefferbeschreibungen war dieser nicht signifikant) und somit die Attraktivität von Suchergebnissen zu Präsentationsverzerrungen führen kann (vgl. ebd., pdf S. 5, 7).

Sozusagen ergänzend zu Yue et al. untersuchten leong et al. (2010) die URLs von Suchergebnissen, genauer die Domains. Dabei führten sie eine Reihe von Experimenten durch, darunter ein Pepsi/Coke-artiges, bei dem Relevanz-Beurteilern zu einer Suchanfrage von zwei Treffern entweder nur die Beschreibungen, die Beschreibungen und die tatsächlichen URLs, oder die Beschreibungen mit ausgetauschten URLs vorgelegt wurden (vgl. ebd., pdf S. 1f, 5f).

Es zeigte sich, dass Nutzer Webseiten als relevanter ansahen und bevorzugten, wenn sie von einer bestimmten Domain stammten, dass Domain-Verzerrungen sowohl hinsichtlich Klickverhalten als auch expliziter Beurteilungen vorlag, und dass Domains Nutzerpräferenzen in etwa einem Viertel der Fälle umkehren konnten, unabhängig von Ranking und Relevanz (vgl. ebd., pdf S. 1, 9f).

Nach leong et al. haben Domain-Verzerrungen Auswirkungen auf Klickvorhersagen (Klickmodelle), das Lernen von Relevanz durch Klicks, und existierende Ansätze zur Klassifizierung von Suchanfragen (vgl. ebd., pdf S. 9f).

White (2013) beschäftigte sich mit Überzeugungen und Verzerrungen innerhalb von Websuche: Mit Fokus auf Ja/Nein-Fragen aus dem Gesundheitsbereich demonstrierte er u.a. anhand einer

retrospektiven Befragung, einer manuellen Antwortsuche in Trefferbeschreibungen (Titel, Beschreibungen und URLs), und einer großangelegten Analyse von Suchverhaltens-Logdaten Bings die Ausmaße von Verzerrungen in Suchverhalten und Suchergebnissen (vgl. ebd., pdf S. 1, 4f).

Die Befragung von ca. 200 Microsoft-Mitarbeitern nach Erfahrungen bzgl. Websuchen mit Ja/Nein-Fragen zeigte v.a., dass diese bei ihrer ursprünglichen Überzeugung blieben, bestätigende Informationen bevorzugten, und die Wahrscheinlichkeit, sich bei Unentschlossenheit für eine positive Antwort zu entscheiden, fast doppelt so hoch war wie die für eine negative (vgl. ebd., pdf S. 2f, 4).

Die Antwortsuche in Trefferbeschriftungen wurde anhand von Logdaten von Bing zu knapp 700 gesundheitlichen Ja/Nein-Fragen durch *Crowdsourcing*-Beurteiler jeweils zu den Top-10-Ergebnissen durchgeführt – sie ergab, dass über die Hälfte der Trefferbeschriftungen keine Antworten enthielten, und bei den restlichen die Wahrscheinlichkeit einer positiven Antwort drei- bis viermal so hoch war wie die einer negativen (vgl. ebd., pdf S. 5f). Weiter stellte White fest, dass Treffer mit positiven Antworten ungeachtet der Wahrheit an den Anfang der Ergebnisliste und über solche mit negativen gerankt wurden (vgl. ebd., pdf S. 5f).

Bei der Analyse von Suchverhaltens-Logdaten von Bing wurde jeweils der erste Klick von ca. 500 gesundheitlichen Ja/Nein-Fragen betrachtet (vgl. ebd., pdf S. 7). Als Ergebnis (vgl. ebd., pdf S. 7ff) tendierten Nutzer zu positiven Antworten und übersprangen deutlich lieber Nein-Antworten, um auf Ja-Antworten zu klicken, als umgekehrt. Weiter zeigte sich, dass die Nutzer lediglich zu etwa der Hälfte der Fragen die korrekte Antwort gefunden hatten.

White schloss, die besorgniserregendste Erkenntnis seiner Arbeit sei vielleicht, dass die Kombination von Suchmaschinen- und Nutzerverzerrungen Nutzer sich in ungefähr der Hälfte der Fälle für eine falsche Antwort entscheiden lässt (vgl. ebd., pdf S. 9).

White vertiefte und erweiterte seine Forschung gemeinsam mit Horvitz in White & Horvitz 2015.

Es bleibt anzumerken, dass jede dieser Methoden, Suchmaschinenverzerrungen zu messen, ihre Vor- und Nachteile hat und die Ergebnisse selbst verzerrt sein können – daher sind diesbzgl. Aussagen mit Vorsicht zu genießen (vgl. Weber 2011, S. 278).

Im Folgenden werden die wichtigsten Punkte des zweiten Teilbereichs dieser Arbeit noch einmal zusammengefasst.

Zuerst wurde sich Suchmaschinen kurz gesellschaftlich und ausführlicher technisch gewidmet (Kap. 3 und 3.1).

Suchmaschinen haben durch ihre massenhafte Nutzung und durch die Reihung und Darstellung der Suchergebnisse eine enorme gesellschaftliche Bedeutung (vgl. Lewandowski 2015, S. 2). Oft werden sie daher in Anlehnung an die publizistik- und kommunikationswissenschaftliche Kommunikationsforschung als ‚Gatekeeper‘ des Internets bzw. Webs bezeichnet und kritisiert (vgl. z.B. Machill & Beiler 2008; Stark 2014). Allerdings können Suchmaschinen auf diese Weise schlecht mit Massenmedien verglichen werden (vgl. Lewandowski 2015, S. 180; Granka 2010, S. 365; Röhle 2010, S. 30f).

Eine Suchmaschine wurde als Computersystem definiert, „das verteilte Inhalte aus dem Web mittels Crawling erfasst und über eine Benutzerschnittstelle durchsuchbar macht, wobei die Ergebnisse in einer nach systemseitig angenommener Relevanz geordneten Darstellung aufgeführt werden“ (Lewandowski 2013, S. 495). Suchmaschinen bestehen aus dem Crawler, der Webseiten indiziert (und dabei eine lokale Kopie des Webs, den Index der Suchmaschine, anlegt), dem Indexer, der die gefundenen Dokumente so aufbereitet, dass sie in der Suche effizient verarbeitet werden können, und dem Searcher, der Suchanfragen ‚verstehet‘ und relevante Suchergebnisse rankt (vgl. Risvik & Michelsen 2002, S. 290; Lewandowski 2015, S. 37, 48, 58).

Bzgl. Rankingfaktoren haben sich sechs Bereiche herausgebildet, von denen die Textstatistik, mit der abgeglichen wird, ob und ggf. wo Wörter der Suchanfrage in den zu durchsuchenden Dokumenten vorkommen, am wichtigsten ist (vgl. Lewandowski 2015, S. 91ff).

Unter Suchmaschinenoptimierung fallen „alle Maßnahmen, die darauf abzielen, die Position von Webseiten für bestimmte Anfragen in den Suchmaschinen-Ergebnislisten zu verbessern“ (Erlhofer 2013, S. 716). Dabei werden die Rankingfaktoren klassisch in On-Page-Faktoren, die auf der eigenen Seite beeinflussbar sind, und Off-Page-Faktoren, die nicht im direkten Einflussbereich der Seite liegen, unterteilt (vgl. Lewandowski 2015, S. 166; Erlhofer 2013, S. 25f). Für die Suchmaschinenoptimierung gibt es Regeln (s. z.B. Google 2015c).

Personalisierung meint im Suchmaschinenkontext die individuelle Anpassung der Suchergebnisse an den Nutzer aus im- und expliziten Nutzerverhaltensdaten zum Zweck der Qualitätsverbesserung (vgl. Lewandowski 2015, S. 118ff). Die Suchergebnis-Personalisierung wird häufig kritisch gesehen; populär ist die These der Filter Bubble von Pariser (2012), nach der jeder Nutzer unbemerkt in seiner eigenen ‚Filterblase‘ lebt und sich nicht für oder gegen die Personalisierung entscheiden kann.

Anschließend wurden Spezialsuchmaschinen, insbes. im Kontext der Universal Search, betrachtet (Kap. 3.2).

Spezialsuchmaschinen wurden definiert als „solche, die sich thematisch oder anhand formaler Dokumentenmerkmale (Bsp. Dateityp) beschränken“ (Lewandowski 2008, S. 57). Die Vorteile solcher einer Beschränkung sind, dass die ausgesuchten Inhalte vollständiger und tiefer erschlossen, und die Suchoberfläche sowie das Ergebnisranking an jeweilige Nutzergruppen angepasst werden können

(vgl. ebd., S. 56). Es gibt verschiedene Typen von Spezialsuchmaschinen, die sich durch unterschiedliche Charakteristika auszeichnen. Ein bekanntes Beispiel ist die Nachrichtensuchmaschine Google News (vgl. Lewandowski 2015, S. 244f).

Spezialsuchmaschinen sind heute oft in allgemeine Suchmaschinen integriert und dort für spezielle Inhalts-Arten oder -Typen zuständig, für die sie eigene (kleinere) Kollektionen aufbauen (vgl. ebd., S. 239, 47). Ergebnisse aus den Spezialsuchmaschinen werden dabei i.d.R. in das Ergebnis der Web-Suchmaschine eingestreut – dies ist das Prinzip der Universal Search. Die Universal Search führt zu einer vielfältigeren Ergebnisauswahl und -darstellung, was großen Einfluss auf das typische Blicksverlaufsmuster hat.

Als nächstes wurde sich Suchmaschinenverzerrungen gewidmet (Kap. 4.1).

Suchmaschinenverzerrungen wurden definiert als Unterschiede eines tatsächlichen Suchergebnisses zum Ideal aus sämtlichen potenziell relevanten Dokumenten in einer allein nach objektiven Kriterien erfolgten und damit richtigen Reihung (vgl. Lewandowski 2015, S. 272). Aufgrund der Funktionsweise von Suchmaschinen kann es keine unverzerrten Ergebnisse geben (vgl. Weber 2011, S. 277; Lewandowski 2015, S. 272).

Die Ursachen von Verzerrungen lassen sich in drei Kategorien einteilen (vgl. Weber 2011, S. 278): Implementierung, d.h. Indexierungs- und Suchalgorithmen von Suchmaschinen, rechtliche Vorgaben und mögliche Einflussnahmen seitens Suchmaschinenbetreibern (vgl. ebd., S. 278f); Anbieterverhalten, d.h. Suchmaschinenoptimierung (vgl. ebd., S. 279); und Nutzerverhalten, d.h. Nutzer reizen die Möglichkeiten von Suchmaschinen oft nicht aus (vgl. ebd., S. 280f). Schließlich nennt Weber (2011, S. 281) noch Kombinationen von Implementierung, Anbieter- und Nutzerverhalten als Ursache von Verzerrungen.

Mit der Personalisierung verlagert sich die Quelle der Verzerrungen von der Suchmaschine auf den Nutzer (vgl. Jürgens et al. 2014, S. 108f).

Zum Problem werden Suchmaschinenverzerrungen v.a. durch die Verbindung der Dominanz des Modells algorithmischer Web-Suchmaschinen im Web, der Dominanz Googles in diesem Bereich, und des eher restringierten Verhaltens der Suchmaschinennutzer (vgl. Lewandowski 2014, S. 233).

Gegenstrategien zu Suchmaschinenverzerrungen resultieren aus dem Bewusstsein über sie und führen zu entsprechendem Verhalten, wie z.B. weitere Seiten der Ergebnislisten anzuschauen, in weiteren Sprachen zu suchen, oder Alternativen zu nutzen (vgl. Weber 2011, S. 282f).

Schließlich wurden in aller Kürze neun Arbeiten aus dem Zeitraum 2002 bis 2013 vorgestellt, die sich auf vielfältige Weise mit Suchmaschinenverzerrungen beschäftigen. Dabei ging es u.a. um die Zuverlässigkeit impliziten Feedbacks aus Klickdaten und Suchanfragen-Reformulierungen (Joachims

et al. 2007); den Vergleich von Suchmaschinen mit einer traditionellen, bibliotheksbasierten Recherche (Brown et al. 2008); die Attraktivität von Suchergebnissen (Yue et al. 2010) bzw. der Domains (leong et al. 2012) als potenzielle Quelle für Präsentationsverzerrungen; und um Überzeugungen und Verzerrungen innerhalb von Websuche (White 2013).

Mit diesem Wissen geht es im folgenden dritten Teilbereich dieser Arbeit, der empirischen Untersuchung, um inhaltliche Suchmaschinenverzerrungen bei deutschen und amerikanischen Suchergebnissen.

## **5. Empirische Untersuchung**

Dieses Kapitel dreht sich um die empirische Untersuchung dieser Arbeit.

Wie bereits in der Einleitung (Kap. 1) und Kap. 3 angesprochen, haben Suchmaschinen eine enorme gesellschaftliche Bedeutung, sind der beliebteste Dienst des Internets und werden massenhaft genutzt, u.a., um aktuellen Nachrichten zu suchen (vgl. Lewandowski 2015, S. 2, 3; van Eimeren & Frees 2014, S. 386f; auch AGOF 2014, S. 11; Purcell et al. 2012, S. 3, 5-8; Hölzig & Hasebrink 2014, S. 535).

In Kap. 4.1 wurde allerdings gezeigt, dass Suchergebnisse niemals ‚neutral‘, sondern immer in irgendeiner Weise verzerrt sind (vgl. Weber 2011, S. 277; Lewandowski 2015, S. 272).

Daher stellt sich die Frage, was für Bilder Suchmaschinen eigentlich von aktuellen Themen vermitteln – genauer, ob und inwiefern mögliche inhaltliche Verzerrungen bestehen.

Dies wird anhand dieser empirischen Untersuchung vergleichend für Deutschland und die USA untersucht.

Hierzu wurden zuerst aus der aktuellen Tagespresse beider Regionen Themenüberschneidungen identifiziert. Zu diesen Themen wurden bei den berücksichtigten Suchmaschinen Google und Bing (jeweils unter .de- und .com-Domain) Anfragen gestellt und automatisiert jeweils die Top-10 textlichen Ergebnisse erfasst. Zu jedem Thema wurden aus den zugehörigen Ergebnissen aller Suchmaschinen die inhaltlichen Aussagen identifiziert, gesammelt, und zu inhaltlichen Pools, sog. ‚Aspekten‘, geclustert. Dann wurden diese Aspekte in den einzelnen Ergebnissen (so vorhanden) ausgemacht und hinsichtlich ihrer Ausführlichkeit betrachtet – so wurden auch die ‚wichtigen‘ (= die ausführlichsten) identifiziert. Anschließend wurden die einzelnen Suchmaschinen anhand der

wichtigen inhaltlichen Aspekte in Hinsicht auf Vollständigkeit, Abdeckung und Gewichtung betrachtet. Zu guter Letzt wurden so die deutschen mit den amerikanischen Suchmaschinen verglichen.

(Dieses Vorgehen wird noch ausführlicher in Kap. 5.2.2 beschrieben.)

Im Folgenden werden die Forschungsfragen und -hypothesen der empirischen Untersuchung dargestellt, die Methodik erläutert, die Ergebnisse präsentiert und diskutiert, und Einschränkungen sowie Anregungen zu weiterer Forschung aufgezeigt.

## **5.1 Forschungsfragen und -hypothesen**

In diesem Abschnitt werden die Forschungsfragen und -hypothesen der empirischen Untersuchung dargestellt. Sie beziehen sich auf drei Themenbereiche: die Vollständigkeit, Abdeckung und 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 ähnlichen Suchmaschinen (s. Kap 3.1) 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 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: 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.

## 5.2 Methodik

In dieser Sektion geht es um die Methodik der empirischen Untersuchung.

Im Folgenden wird zuerst auf ihren theoretischen Hintergrund eingegangen, bevor sie im Detail erläutert und anhand einer Beispiel-Auswertung demonstriert wird.

### 5.2.1 Theorie

In diesem Kapitel wird auf den theoretischen Hintergrund der Untersuchungsmethodik eingegangen. Diese ist im Grunde eine Kombination von Ansätzen der empirischen Sozialwissenschaften, nämlich Inhaltsanalyse, und der Informationswissenschaft, nämlich Retrievaltests. Diese werden im Folgenden etwas näher betrachtet.

Mit der inzwischen zur Geschichte gehörenden Definition des Inhaltsanalyseklassikers Berelson (vgl. Mayring 2010, S. 11; s. Berelson 1952, S. 18) als Grundlage definiert Früh 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 Mayring (2010, S. 13) will Inhaltsanalyse zusammenfassend

- Kommunikation analysieren.
- fixierte [d.h. protokollierte, festgehaltene] Kommunikation analysieren.
- dabei systematisch vorgehen.
- dabei also regelgeleitet vorgehen.
- dabei auch theoriegeleitet vorgehen.
- das Ziel verfolgen, Rückschlüsse auf bestimmte Aspekte der Kommunikation [z.B. Aussagen über den Sender (bspw. dessen Ansichten), über Wirkungen beim Empfänger o.Ä.] zu ziehen.

Nach Früh besteht der Ablauf einer Inhaltsanalyse standardisiert aus vier Phasen: der Planungs-, Entwicklungs-, Test- und Anwendungsphase. Dies ist in Abb. 8 dargestellt und wird im Folgenden etwas näher erläutert.

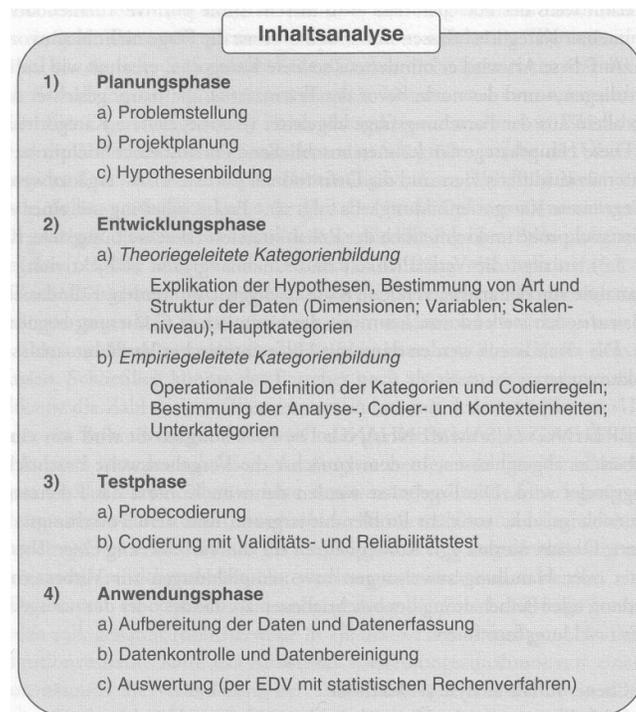


Abb. 8: Inhaltsanalyse - standardisierter Untersuchungsablauf (Früh) (2011, S. 102)

(Für ein anderes Ablaufmodell s. etwa Mayring 2010, S. 60.)

Nach Früh (2011, S. 147-153) wird in der Planungsphase zunächst das Untersuchungsziel bestimmt, d.h. eine Fragestellung erarbeitet. Dann werden theoriegeleitet Hypothesen gebildet, die durch empiriegeleitetes Vorgehen ergänzt werden sollten. Auf dieser Basis wird die Grundgesamtheit des Untersuchungsmaterials (z.B. Zeitungen) bestimmt und ggf. eine Stichprobe gezogen.

In der Entwicklungsphase findet nach Früh (2011, S. 153-163) die Bildung von Kategorien statt. Diese geben an, auf welche Merkmale der untersuchten Mitteilungen sich die Analyse beziehen soll, und welche konkreten Texteinheiten als ähnlich betrachtet werden (vgl. auch Mayring 2010, S. 49f). Aus den Hypothesen abgeleitete theoretische Konstrukte stellen als Hauptkategorien den Kern des Kategoriensystems dar.

Bei der empiriegeleiteten Kategorienbildung werden die Hauptkategorien in Unterkategorien ausdifferenziert und ggf. ergänzt. Außerdem werden die Kategorien operational definiert, sodass klar wird, welche Merkmale und Einheiten des Textmaterials von den Codierern in welche Codierungen überführt werden. Um sicherzustellen, dass alle für das Forschungsziel relevanten Aspekte und Bedeutungen erkannt und einbezogen werden können, muss eine repräsentative Textstichprobe aus dem Untersuchungsmaterial gezogen werden, die einer systematischen qualitativen Bearbeitung aus vier Schritten unterzogen wird:

1. Selektion/Reduktion: Textpassagen, die in Zusammenhang mit der Forschungsfrage stehen, werden extrahiert, redundante und inhaltsleere Stellen ausgeblendet.
2. Bündelung: Die extrahierten Textstellen werden nach inhaltlichen Gemeinsamkeiten auf einer jeweils einheitlichen Abstraktionsebene gruppiert. Auftauchende, relevant erscheinende Untergruppen werden separiert.
3. Generalisierbarkeit/Abstraktion/Bezeichnung: Den gebündelten Textpassagen werden Labels zugewiesen, die den abstrahierten gemeinsamen Bedeutungsgehalt bezeichnen.
4. Rückbezug auf Theorie: Überprüfung, ob die als relevant betrachteten Textpassagen den in den Hypothesen enthaltenen theoretischen Konstrukten (die Hauptkategorien) zugeordnet werden können. Wenn ja, sind die gefundenen relevanten Ausdifferenzierungen als Unterkategorien zu übernehmen. Bei überschüssigem Textmaterial muss geprüft werden, ob daraus weitere Hypothesen generiert werden können, die dann wieder denselben Prozess durchlaufen.

Nach Früh (2011, S. 163-198) werden in der Testphase Codierungen zur Probe und zu Reliabilitäts- und Validitätstests durchgeführt (vgl. auch Mayring 2010, S. 51f, 116-122).

Um das Kategoriensystem zu testen, operational exakter zu definieren und ggf. zu modifizieren, wird eine weitere Teilstichprobe gezogen und anhand des Kategoriensystems vom Forscher und versierten Codierern verschlüsselt, wobei jede Zuordnung besprochen wird.

Durch eine Codierschulung werden Codierer, die auch die endgültige Codierung durchführen, mit der vom Forscher gemeinten und im Codebuch formulierten Interpretationsweise vertraut gemacht. Sie endet mit einer quantitativen Gültigkeits- und Verlässlichkeitskontrolle. Nach dem Training

entwirft der Forscher einen Codierbogen, um den Codiervorgang stärker zu formalisieren und die Daten für eine Weiterverarbeitung via Computer vorzubereiten.

Am Ende der Codierschulung wird ein Reliabilitätstest durchgeführt, an dem auch der Forscher selbst teilnimmt. Dabei wird die Übereinstimmung mehrerer Codierer am selben Textmaterial (*Intercoder-Reliabilität*), bei langer Projektdauer zusätzlich mit zeitlichem Abstand (*Intracoder-Reliabilität*) gemessen. Zum Abschluss wird noch eine Validitätsprüfung durchgeführt, um zu überprüfen, ob die Codierungen (also die produzierten Daten) den in der Forschungsfrage anvisierten Bedeutungsgehalt (das zu messende theoretische Konstrukt) tatsächlich treffen.

In der Anwendungsphase wird nach Früh (2011, S. 198-207) das Textmaterial systematisch oder ggf. nach Zufallskriterien auf die Codierer verteilt, damit evtl. Codierer-Verzerrungen überall gleich auftreten und sich so neutralisieren.

Wenn anfangs trotz Trainingsphase und Reliabilitätstest noch Probleme und Zweifelsfälle auftreten, sollten diese sofort diskutiert werden – zusätzliche Regelungen sind vom Forscher sofort schriftlich festzuhalten und für alle Codierer zu vervielfältigen. Bzgl. des Kategoriensystems kann erstmals auftretenden Problemen problemlos mit einer Modifizierung begegnet werden – handelt es sich um einen Sachverhalt, der schon häufiger vorgekommen ist, ist der betroffene Teil des bereits codierten Materials anhand der neuen Regelung nachzucodieren.

Für die Auswertung werden die Codes von den Codierbögen durch manuelle Eingabe oder Scannen auf maschinenlesbare Datenträger übertragen.

Bei der Auswertung sollen die erhobenen Daten mit Hilfe statistischer Auswertungsmethoden (anwendbar durch Statistik- und Analyse-Software wie z.B. der Marke SPSS (urspr. *Statistical Package for the Social Sciences*)) derart weiterverarbeitet werden, dass Schlussfolgerungen im Sinne der Hypothesen möglich sind. Der Auswertungsplan hat sich daher eng an den Hypothesen und dem Erkenntnisinteresse der Forschungsfrage zu orientieren. Neben der Frage, welche Methode unter den gegebenen Bedingungen überhaupt angemessen bzw. zulässig ist, muss der Forscher oft auch entscheiden, welches Verfahren am ökonomischsten, präzisesten und informationsreichsten ist.

Nach Lewandowski (2015, S. 215-221) (ausführlicher s. Lewandowski 2011) sind Retrievaltests eine Methode mit langer Tradition, die *Retrievaleffektivität* 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 – da dafür also alle vorhandenen Dokumente bekannt sein und hinsichtlich ihrer Relevanz bewertet werden müssen, ist die Bestimmung des Recalls 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 & Voorhees 2006)) bekannt ist: Eine bestimmte Anzahl 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 Trefferpositionen 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. Die von Hawking et al. genannten fünf Kriterien für einen validen Suchmaschinentest beziehen sich auf die Abbildung realer Informationsbedürfnisse, die Mitteilung des Informationsbedürfnisses (falls ein Informationsvermittler eingesetzt wird), die genügend große Zahl von Testanfragen, die Auswahl der wichtigsten Suchmaschinen, und den sorgfältigen Aufbau der Untersuchung sowie eine ebensolche Durchführung.

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.

Wünschenswert ist dabei natürlich, möglichst viele Suchanfragen zu verwenden. Allerdings treten in der Praxis oft Probleme auf, genügend Juroren zu finden, und so hat sich mittlerweile ein Minimum von 50 Suchanfragen etabliert. Die Anfragen-Anzahl hängt freilich auch vom Zweck des Tests ab. Sollen viele Themenbereiche und ein verschiedenartiges Suchverhalten abgedeckt werden, muss die Zahl entsprechend erhöht werden.

Vom Zweck des Tests hängt auch die Auswahl der Suchanfragen ab. Sollen Aussagen über die Trefferqualität der Suchmaschinen allgemein getroffen werden, sollten die Anfragen möglichst breit gewählt und auch die Verteilung ihrer Länge beachtet werden.

Um die Juroren bei der Bewertung zu unterstützen, empfiehlt es sich evtl., die Suchanfragen mit Beschreibungen und ggf. expliziten Bewertungsinformationen anzureichern.

Bei der Festlegung der Anzahl der Suchergebnisse ist das typische Verhalten der betreffenden Nutzergruppe zu berücksichtigen. Die gängigen Untersuchungen beschränken sich i.d.R. auf die ersten zehn Ergebnisse – mit einer höheren Trefferanzahl steigt der Umfang des Tests entsprechend.

Die Anzahl der benötigten Juroren richtet sich natürlich nach der Anzahl der zu untersuchenden Suchanfragen. Üblicherweise bewertet jeder Juror die Treffer aller Suchmaschinen zu einer Suchanfrage. I.d.R. wird jeder Treffer nur von einer Person beurteilt – da allerdings davon auszugehen ist, dass auch innerhalb einer einzelnen Gruppe die *Interrater-Reliabilität* nur ein gewisses Maß erreicht (vgl. Schaer et al. 2010, pdf S. 5f), sollten die Treffer von mindestens zwei Personen bewertet werden.

Weitere Faktoren, die die Testergebnisse verfälschen können, sind Markeneffekte bei Treffern mit nicht verschleierte Herkunft (vgl. Jansen et al. 2007a, b; Bailey et al. 2007) und Lerneffekte bei sichtbarer ursprünglicher Reihung der Treffer (vgl. Bar-Ilan et al. 2009). Auch sollten die Treffer unterschiedlicher Suchmaschinen gemischt werden. Dubletten sollten den Juroren nur einmal (pro Suchanfrage) vorgelegt werden, um einheitliche Urteile zu erhalten.

Bei der Auswertung des Tests werden die Relevanzurteile zusammengeführt und ausgewertet – dabei können neben der Precision auch weitere bzw. andere Kennzahlen verwendet werden.

Durch die Kombination von Inhaltsanalyse und Retrievaltests ist eine ganzheitlichere Sicht auf Suchmaschinen und ihre Ergebnisse möglich, als die Methoden für sich alleine bieten. Dies wird im folgenden Kapitel deutlich werden.

## **5.2.2 Praxis**

In diesem Abschnitt wird u.a. anhand einer kompletten Einzelauswertung und korrespondierend zu den Forschungsfragen und -hypothesen (Kap. 5.1) die praktische Methodik der empirischen Untersuchung erläutert.

Sinn und Zweck der Untersuchung war es, einen Weg zu finden, potenzielle inhaltliche Suchmaschinenverzerrungen bzgl. aktueller internationaler Themen festzustellen. Die Grundidee

war, solche Themen aus der ‚analogen Welt‘ anhand von aktuellen Zeitungen zu identifizieren und in der ‚digitalen‘ in Form von Suchergebnissen 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 2015e bzw. Statista 2015b) – von einem bestimmten Tag, Montag den 28. April 2015, in Hinsicht auf gemeinsame Themen gesichtet. Auswahlkriterium war, dass ein Thema mindestens in einer deutschen und einer amerikanischen Zeitung behandelt werden musste.

So wurden 14 aktuelle internationale Themen identifiziert, darunter etwa Varoufakis‘ ‚Entmachtung‘ nach dem Riga-Gipfel, das Erdbeben in Nepal, Hillary Clinton und Spendengelder, und die erneuten Kämpfe in der Ostukraine bei Mariupol. Tab. 5 zeigt die Themen und ihr jeweiliges Vorkommen in den Zeitungen.

Tab. 5: Gemeinsame Themen deutscher und amerikanischer Zeitungen

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

Es sollten Suchmaschinenergebnisse aus Deutschland und den USA verglichen werden, also musste es die zu untersuchenden Suchmaschinen in einer oder beiden Regionen geben. Außerdem sollten sie bedeutend sein, um dem Informations-Massenkonsum möglichst nahe zu kommen, und außerdem unabhängig (unbedeutendere Suchmaschinen geben i.d.R. in Kooperation mit einer bedeutenden deren Ergebnisse aus – dann lässt sich besser gleich die Quelle untersuchen).

Nach Statista (2015d bzw. c) wird sowohl in Deutschland als auch den USA der Suchmaschinenmarkt deutlich von Google dominiert – in Deutschland mit einem aktuellen Marktanteil von ca. 95 Prozent, in den USA mit knapp 65. Stärkster Konkurrent ist in beiden Regionen Bing, die zwar keine mit Google vergleichbaren Werte erreicht, doch mit ihrem Partner Yahoo, der Bing-Ergebnisse

ausgibt (vgl. Computerwoche 2015), in Deutschland ca. 4 Prozent ausmacht, in den USA gut 20. Andere Suchmaschinen sind unbedeutend und i.d.R. abhängig. So fiel die Wahl der zu untersuchenden Suchmaschinen auf google.de und bing.de, sowie google.com und bing.com.

Zu den 14 Themen aus den Zeitungen wurden jeweils eine deutsche und eine englische Suchanfrage formuliert. Diese sind in Tab. 6 dargestellt.

Tab. 6: Zeitungsthemen-Suchanfragen

Nr.	Anfrage google.de/bing.de	Anfrage google.com/bing.com
1	deutsche bank einsparmaßnahmen	deutsche bank saving measures
2	varoufakis entmachtung	varoufakis disempowerment
3	erdbeben nepal	earthquake nepal
4	hillary clinton geld	hillary clinton money
5	englands rolle in der welt	britain's role in the world
6	burundi proteste nkurunziza	burundi protests nkurunziza
7	israel luftangriff syrien grenze	israel airstrike syria border
8	usa japan verteidigungsabkommen	us japan defense agreement
9	lee wan-koo rücktritt	lee wan-koo resignation
10	40 jahre vietnamkrieg	40 years vietnam war
11	afrika migranten	africa migrants
12	klimawandel hitze regen	climate change heat rain
13	mylan teva übernahme	mylan teva takeover
14	ostukraine kämpfe mariupol	east ukraine fights mariupol

Anhand dieser Anfragen wurden noch am Tag der Zeitungssichtung mithilfe des an der *HAW Hamburg* entwickelten *Relevance Assessment Tools* (RAT) die jeweils ersten zehn textlichen Ergebnisse der vier berücksichtigten Suchmaschinen unter Umgehung von Personalisierungen (s. Kap. 3.1), etwa durch Suchhistorie oder Standort, erfasst.

Das RAT ist nach Lewandowski und Sünkler (2012, S. 237f) ein Werkzeug zur Durchführung von Retrievaltests (s. voriges Kap.) und zur weitgehend automatisierten Erfassung, Aufbereitung und Auswertung der Daten. Es besteht aus folgenden Komponenten:

- Suchmaschinenscraper
- Backend für die Verwaltung und die Testdesigns
- Frontend zur Durchführung der Tests (Relevanzbewertungen durch Probanden)
- Auswertung

Da die vorliegende Untersuchung nicht mit Probanden und manuell durchgeführt wurde, ist hier lediglich der *Suchmaschinenscraper* von Interesse.

Allgemein ist ein *Screenscraper* (auch *Wrapper*) eine Anwendung, die Informationen aus Texten extrahiert und speichert. Bei Webseiten wird dies z.B. in zwei Schritten durchgeführt: Aufruf der Seite und Extraktion, Aufbereitung und Speicherung der Informationen.

Der Suchmaschinenscraper des RATs simuliert den Aufruf einer Suchmaschine und das Abschicken einer Suche durch einen Webbrowser, und kann dann Ergebnisseiten (SERPs) und Ergebnisse selbst erfassen und speichern (für die technischen Details s. ebd., S. 240f).

Bei der Untersuchung wurden jeweils die SERPs als lokale Kopien erfasst, und die ersten zehn textlichen Ergebnisse für die Anfragen aus Tab. 6 als Links in *Excel*-Listen sowie ebenfalls als lokale Kopien (für dieses Material s. beiliegende CD). ‚Textliche Ergebnisse‘ bedeutet hier organische Ergebnisse und evtl. vorhandene Universal-Search-News-Treffer (s. Kap. 3.2). Dies wurde so gehandhabt, weil die 14 Themen aus den Zeitungen natürlich aktuell waren und Nachrichtencharakter hatten, sodass von einer Ähnlichkeit dessen in der ‚digitalen Welt‘ auszugehen, und es realistischer war, auch Universal-Search-News-Treffer zu berücksichtigen, da sie Suchmaschinennutzern ja so im Suchergebnis präsentiert werden.

Leider kam es zu Problemen beim Scraping, sodass etwa organische Ergebnisse, die vorher als Universal-Search-News-Treffer erfasst worden waren, ausgelassen wurden, oder Universal-Search-News-Treffer nicht oder unvollständig erfasst wurden. Diese Fehler wurden bei der Auswertung durch das Abgleichen der SERPs mit den Listen der gescrapten Ergebnisse so weit wie möglich korrigiert.

Ab hier soll die Auswertung zu Thema 7 – dem Luftangriff Israels an der syrischen Grenze – als Leitfaden dienen. Das Thema wurde ausgewählt, weil es recht kompakt ist, die Auswertung unproblematisch war, sodass sich gut auf die Methodik konzentriert werden kann, und sich dabei zwei interessante Umstände zeigten.

Zu Thema 7 sind folgende Anmerkungen bzgl. des Scrapings zu machen:

- Bei google.de wurden das zweite und fünfte organische Ergebnis nicht mit erfasst, es handelte sich jedoch um Dubletten bereits gescrapter Universal-Search-News-Treffer. (Die Auswertungsergebnisse der Universal-Search-Treffer wurden also für die beiden fehlenden organischen Ergebnisse übernommen.)
- Bei bing.de wurden die drei Universal-Search-News-Treffer am Anfang der Ergebnisliste nicht mit erfasst, sie wurden jedoch als organische Dubletten gescrapet. (Die Auswertungsergebnisse der organischen Ergebnisse wurden also für die fehlenden Universal-Search-News-Treffer übernommen.)

- Bei google.com wurden das zweite und dritte organische Ergebnis nicht mit erfasst, es handelte sich jedoch um Dubletten bereits erfasster Universal-Search-News-Treffer. (S. erster Punkt.)
- Bei bing.com wurden die drei Universal-Search-News-Treffer am Anfang der Ergebnisliste nicht mit erfasst, eins davon wurde als organische Dublette gescrapet. (Die beiden auch nicht als organische Dubletten erfassten Universal-Search-News-Treffer konnten nicht ersetzt werden – sie wurden aus der Auswertung ausgeblendet, die nachfolgenden Ergebnisse rückten auf.)

Manche Ergebnisse wurden für die Auswertungen als irrelevant eingestuft. Dabei handelte es sich um solche, die

- mit dem Themenfeld zu tun hatten, aber nicht aktuell waren und sich um frühere Geschehnisse drehten (z.B. ging es im ersten google.de-Treffer zu ‚Deutsche Bank Einsparmaßnahmen‘ zwar um dieses Thema, er war allerdings bereits über zehn Jahre alt - <http://www.n-tv.de/archiv/Ackermann-in-der-Kritik-article107815.html>),
- womöglich mit dem Themenfeld zu tun hatten und aktuell waren, aber andere Inhalte behandelten (bspw. beschäftigte sich der dritte google.de-Treffer zum Thema ‚Hillary Clinton und Spendengelder‘ ausschließlich mit Clintons Frisuren im Laufe der Zeit - <http://www.welt.de/icon/article140137308/Hillary-Clintons-Frisur-kann-die-Wahl-entscheiden.html>),
- womöglich mit dem Themenfeld zu tun hatten und aktuell waren, aber entweder zu wenige Informationen beinhalteten, wie etwa Dossiers, aus denen man die eigentlichen Texte erst herausuchen und anklicken musste (z.B. der siebte google.de-Treffer zum ‚Hillary Clinton und Spendengelder‘-Thema - [http://www.sueddeutsche.de/thema/Hillary\\_Clinton](http://www.sueddeutsche.de/thema/Hillary_Clinton)), oder die Informationen nicht als Text anboten (bspw. stellte das Ergebnis auf Position 2 und 5 von google.de zum ‚Proteste in Burundi‘-Thema lediglich ein Video (mit zu knapper Textinformation) dar - <http://www.welt.de/videos/article140173907/Blutige-Protteste-gegen-Praesident-Pierre-Nkurunziza.html>),
- schlicht nichts mit dem gesuchten Thema zu tun hatten (z.B. führte der sechste bing.com-Treffer des Varoufakis-Themas zur Webseite eines Verlags, auf der das Buch ‚Intellectuals and African Development – Pretension and Resistance in African Politics‘ präsentiert wurde - <http://zedbooks.co.uk/node/21406>),
- sich ohne Registrierung/Anmeldung nicht (ganz) lesen ließen (bspw. Treffer 5 und 9 von bing.com zum Erdbeben in Nepal - <http://www.wsj.com/articles/how-the-nepal-earthquake-happened-like-clockwork-1430044358>).

(Die Auswahl der Beispiele erfolgte willkürlich und enthält keine Aussage zur Retrievaleffektivität der Suchmaschinen.)

Ergebnisse, bei denen man erst Werbung, ein Abo-Angebot o.Ä. abwarten und/oder wegklicken musste, um zum Inhalt zu gelangen, wurden nicht als irrelevant eingestuft, auch nicht solche, bei denen der Inhalt nur angeteasert wurde und man z.B. auf ‚Den vollständigen Artikel lesen ...‘ klicken musste. Verteilten sich die Texte über mehrere Seiten, wurde natürlich geblättert, um die Inhalte vollständig zu erfassen.

Dem bei Retrievaltests klassischen Problem der Bestimmung des Recalls (s. voriges Kap. – es war natürlich unmöglich, in Hinsicht auf die 14 Themen jeweils an *alle vorhandenen* relevanten Dokumente der vier Suchmaschinen zu gelangen), wurde in dieser Untersuchung so begegnet, dass zu jedem Thema aus den inhaltlichen Aussagen der erfassten relevanten Ergebnissen der vier Suchmaschinen ein Pool aus mehreren ‚Aspekten‘ gebildet wurde. Dieser bildete jeweils die Grundgesamtheit, gegen den später die Retrievaleffektivität der einzelnen Suchmaschinen gespiegelt wurde. (Diese Pooling-Methode ähnelt im Ansatz also dem Vorgehen von Mowshowitz und Kawaguchi in ihren Arbeiten über Suchmaschinenverzerrungen (s. Kap. 4.2). – Natürlich ist dies ein Behelf, der von der Realität abweicht, doch galten so für alle 14 Themen zumindest die gleichen Bedingungen, sodass die Leistungen der Suchmaschinen im Rahmen dieser Untersuchung verglichen werden konnten.)

Die Bildung der inhaltlichen Pools sah so aus, dass zu jedem Thema alle zugehörigen erfassten Suchergebnisse gesichtet und die darin enthaltenen inhaltlichen Aussagen gesammelt wurden. Da die Ergebnisse und damit die Suchmaschinen letztlich ja bzgl. 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 bei den Themen von vier (Thema 9 – Lees Rücktritt, 12 – Klimawandel und 14 – erneute Kämpfe in der Ostukraine) bis neun (3 – Erdbeben in Nepal und 4 – Hillary Clinton und Spendengelder). Die Gruppierung der Aspekte bildete natürlich gleichzeitig ihre ausführlichste Version. Zusätzlich wurde jeweils eine knappe Version definiert, die lediglich aus den grundlegenden inhaltlichen Aussagen bestand. So wurden jeweils Ausführlichkeitsdimensionen aufgespannt, anhand derer dann die einzelnen Ergebnisse hinsichtlich des Vorhandenseins und der Ausführlichkeit der inhaltlichen Aspekte beurteilt werden konnten.

Tab. 7 zeigt so die fünf inhaltlichen Aspekte von Thema 7 – dem Luftangriff Israels an der syrischen Grenze.

Tab. 7: Thema 7 – inhaltliche Aspekte

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
IV	Hintergrund: Verhältnis von Israel & Syrien, Rolle der Golanhöhen	Israel & Syrien sind Feinde, Sie streiten seit Jahrzehnten um die Golanhöhen; In Syrien herrscht seit Jahren Bürgerkrieg; Israel versucht sich rauszuhalten, reagiert jedoch ab und zu mit Vergeltungsangriffen	Israel & Syrien gelten als erbitterte Feinde & streiten sich seit fast 50 Jahren um die Golanhöhen; Israel hatte 1967 im Sechstagekrieg Teile des 1150+ km <sup>2</sup> großen syr. Plateaus Golan erobert/besetzt & (Ende) 1981 annektiert, was international nicht anerkannt wird; Ein Teil blieb unter syr. Kontrolle; Offiziell befinden sich beide Länder weiter im Kriegszustand; Der Großteil wird heute von Rebellengruppen kontrolliert, darunter die Al-Nusra-Front (syr. Ableger des Terrornetzwerk Al-Quaida); Im Gebiet leben 41.000 Bewohner - Juden, Drusen & Alawiten; Bis zum Ausbruch des Bürgerkriegs in Syrien war es auf den Golanhöhen relativ ruhig, bewaffnet dürfen sich in dem Gebiet lediglich die Soldaten der Uno-Mission Undof aufhalten; In Syrien herrscht seit Jahren / 2011 Bürgerkrieg, in dem gemäßigte Rebellen, aber auch Dschihadisten gegen die Armee von Syriens Staatschef Baschar al-Assad kämpfen, die von der Hisbollah unterstützt wird; In dem von Isr. besetzten Teil der Golanhöhen schlagen regelmäßig aus Syr. abgeschossene Geschosse ein; Häufig ist unklar, ob Isr. mit Absicht beschossen wurde; Israel versucht zwar, sich aus dem Bürgerkrieg herauszuhalten, reagiert jedoch auf gelegentlichen Beschuss mit Vergeltungsangriffen; Isr. droht immer wieder, die Lieferung hoch entwickelter Waffen an die schiit. Hisbollah im Libanon notfalls auch mit milit. mitteln zu stoppen, die isr. Luftwaffe soll mehrfach iran. Raketen ins Visier genommen haben; Trotz der frostigen Beziehungen hatte die Assad-Dynastie / Baschar al-Assad sowie sein Vater Hafiz in den vergangenen 40 Jahren dafür gesorgt, dass es in der Grenzregion überwiegend ruhig blieb; Isr. befürchtet, dass Syrien nach einem möglichen Sturz Assads in die Hände von

			Extremisten der Terrormiliz IS oder mit Al-Kaida verbündeten Gruppen fallen könnte, die dann Isr. angreifen könnten
V	Der Luftangriff im Kontext von Auseinander-setzungen	An der Stelle des Luftangriffs und im weiteren Umkreis kam es schon zu verschiedenen Kampfhandlungen; Es wird eingeschätzt, dass Israel durch Attacken gezielt syr. Waffenlieferungen an die Hisbollah zu verhindern versucht	Im März 2014 waren an derselben Stelle wie beim mutmaßlichen Angriff am Sonntag 4 isr. Soldaten verwundet worden, als ihr Jeep über eine Mine fuhr; Im August schoss Isr. eine Drohne ab, die im selben Luftraum wie am Sonntag von Syrien aus gekommen war; Im September 2014 hatte die isr. Luftwaffe einen syr. Kampfjet über den Golanhöhen abgeschossen; Mitte Januar hatte Israel bei einem Luftangriff auf den syrisch kontrollierten Teil der Golanhöhen 6 Hisbollah-Kämpfer & 1 iran. General getötet; Die Hisbollah griff knapp 2 Wochen später einen isr. Militärkonvoi an & tötete 2 Soldaten; Israel reagierte mit Artilleriebeschuss auf Hisbollah-Stellungen im Libanon; Der Angriff am Sonntag kam 2 Tage nachdem arab. Medien berichtet hatten, dass isr. Luftangriffe auf Stützpunkte der Hisbollah & der syr. Armee nördl. von Damaskus stattgefunden hätten, die Ziele seien syr. Armeekonvois gewesen, von denen man glaubte, sie führten strategische Waffen & Langstreckenraketen für die Hisbollah im Norden des Libanon mit sich; Das isr. Militär antwortete nicht auf die Behauptungen; Am Montagmorgen berichteten arab. Medien über einen erneuten Luftangriff in derselben Gegend auf Waffenlager der syr. Armee & der Hisbollah; Isr. dementierte das; Isr. wird jedoch seit 2 Jahren verdächtigt, durch Luftangriffe gezielt syr. Waffenlieferungen an die Hisbollah zu verhindern; Isr. milit. Einschätzungen zufolge stellt für die isr. Sicherheit immer noch die Hisbollah die größte Bedrohung dar

Für die Inhaltliche-Aspekte-Tabellen aller Themen s. beiliegende CD.

Anhand der inhaltlichen Aspekte in ausführlicher und knapper Definition wurden die einzelnen Ergebnisse der Suchmaschinen dann mithilfe einer Fünfer-Likert-Skala beurteilt. Diese sah so aus:

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

Bei der Beurteilung der Ergebnisse wurde auch die Länge der Texte berücksichtigt – ein Aspekt konnte also auch mit einer 2 bewertet werden, wenn er in einem kurzen Text nicht so ausführlich wie in einem langen behandelt wurde, aber eben verhältnismäßig viel Raum einnahm (d.h. auch die Vollständigkeit der inhaltlichen Aussagen war kein K.o.-Kriterium für die Beurteilung (außerdem schon gar nicht ihre Richtigkeit)).

Tab. 8 zeigt die Auswertung der jeweils zehn ersten textlichen Ergebnisse der vier Suchmaschinen zu Thema 7 hinsichtlich des Vorkommens und der Ausführlichkeit der inhaltlichen Aspekte (s. Tab. 7).

Für die Auswertungen aller Themen s. beiliegende CD.

Die Farben markieren gleiche Inhalte – graue Ergebnisse sind irrelevant (Gründe s.o.).

Tab. 8: Thema 7 – Vorkommen und Ausführlichkeit der inhaltlichen Aspekte

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

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

Z.B. deckte also das erste google.com-Ergebnis drei der fünf Aspekte ab: I, den Luftangriff, eher ausführlich; II, Reaktionen zum Luftangriff von israelischer Seite, neutral; und IV, Hintergrund: Verhältnis von Israel & Syrien, Rolle der Golanhöhen, eher knapp. Aspekt III, Reaktionen zum Luftangriff von syrischer Seite, und V, der Luftangriff im Kontext von Auseinandersetzungen, wurden in diesem Text nicht behandelt.

Es treten bei dieser Auswertung zwei interessante Umstände zutage: Erstens gaben alle Suchmaschinen bei diesem Thema Universal-Search-News-Treffer am Anfang der Ergebnisliste aus und es zeigte sich, dass diese oft anschließend nochmals als organische Ergebnisse ausgegeben wurden (bei bing.com ist nur das erste Ergebnis ein Universal-Search-News-Treffer, s.o.). Das ist schlecht, weil es Nutzer Zeit kostet, Mängel in der Ergebnispräsentation offenbart und wertvollen sichtbaren Platz mit Dubletten verschwendet, anstatt dass weitere relevante Ergebnisse ausgegeben würden.

Zweitens fallen bei den amerikanischen Suchmaschinen die vielen rot markierten Ergebnisse auf – das war tatsächlich immer der gleiche Inhalt. Zwar waren die URLs und Titel unterschiedlich, doch es handelte sich immer um den gleichen Text (von AP). Das ist schlecht, weil so etwa bei bing.com von zehn Ergebnissen fünf den gleichen Inhalt haben, was Nutzer, die sich über das Thema informieren wollen, frustrieren dürfte, und wieder die Chance vertan wird, an diesen Stellen weitere relevante Ergebnisse auszugeben.

Dass Universal-Search-News-Treffer 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 (was bei

Thema 7 mit insgesamt vier Aspekten freilich nicht so effektiv ist wie etwa bei den Themen mit neun).

Tab. 9 zeigt jeweils die drei wichtigsten (also ausführlichsten) Aspekte mit ihrer Gewichtung der Suchmaschinenergebnisse bzgl. Thema 7.

Tab. 9: Thema 7 – Identifizierung der jeweils drei wichtigsten Aspekte

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

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

Die drei wichtigsten (und einzigen) Aspekte des ersten google.com-Ergebnisses waren also I mit der Gewichtung ‚eher ausführlich‘, II mit ‚neutral‘, und IV mit ‚eher knapp‘.

Anschließend wurde das Vorkommen dieser wichtigen Aspekte in allen Ergebnissen gezählt. Das Resultat ist in Tab. 10 dargestellt.

Tab. 10: Thema 7 – Wichtige Aspekte mit der Häufigkeit ihres Auftretens

I	21
II	19
V	17
III	16
IV	15

Es zeigt sich, dass in diesem Fall alle Aspekte mindestens einmal (hier: 15-mal) zu den wichtigsten drei Aspekten der Ergebnisse gehörten und daher alle zu den wichtigen Aspekten zählten. Dies war bei den Themen mit überschaubaren Anzahlen von Aspekten öfter so.

Als nächstes wurde untersucht, ob und wie die wichtigen Aspekte von den Suchmaschinen innerhalb der ersten zehn textlichen Ergebnisse abgedeckt wurden. Dies ist in Tab. 11 dargestellt.

Tab. 11: Thema 7 – Abdeckung der wichtigen Aspekte

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

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

Es wird deutlich, dass in diesem Fall die wichtigen Aspekte von allen Suchmaschinen abgedeckt wurden. Dies zeigt auch Tab. 12.

Tab. 12: Thema 7 – Anzahlen abgedeckter wichtiger Aspekte (kumuliert)

google.de	5
bing.de	5
google.com	5
bing.com	5
.de	5
.com	5

Das war die Ausnahme – i.d.R. deckten die Suchmaschinen, manchmal sogar beide einer Region, einen oder mehrere der wichtigen Aspekte nicht ab.

An dieser Stelle können für Thema 7 die beiden Forschungsfragen zur Vollständigkeit (s. Kap 5.1) beantwortet werden.

Hypothese 1.1 kann für Thema 7 nicht widerlegt werden: Google und Bing vermitteln in Deutschland und den USA in Hinsicht auf die wichtigen inhaltlichen Aspekte vollständige Bilder aktueller internationaler Themen.

Damit kann Forschungsfrage 1.1 für Thema 7 bestätigt werden: Google und Bing vermitteln in Deutschland und den USA in Hinsicht auf die wichtigen inhaltlichen Aspekte vollständige Bilder aktueller internationaler Themen.

Hypothese 1.2 kann für Thema 7 ebenfalls nicht widerlegt werden: Durch die 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 keine Unterschiede.

Damit kann Forschungsfrage 1.2 für Thema 7 verneint 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 keine Unterschiede in puncto Vollständigkeit.

Als nächstes wurde die Abdeckung der drei wichtigsten Aspekte (s. Tab. 10) jeweils auf der ersten Trefferposition der Suchmaschinen untersucht. Dies ist in Tab. 13 dargestellt.

Tab. 13: Thema 7 – Abdeckung der drei wichtigsten Aspekte auf Trefferposition 1

	g.de	b.de	g.com	b.com
<b>Aspekt</b>	<b>1</b>	<b>1</b>	<b>1</b>	<b>1</b>
<b>I</b>	x	x	x	
<b>II</b>	x	x	x	
<b>V</b>	x	x		

Damit kann an dieser Stelle 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 (die amerikanischen Suchmaschinen decken Aspekt V nicht ab, bing.com außerdem auch I und II nicht).

Dann 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. Dies ist in Tab. 14 dargestellt.

Tab. 14: Thema 7 – Abdeckung der 75 Prozent wichtigster Treffer auf den ersten drei Positionen

	google.de			bing.de			google.com			bing.com		
<b>Aspekt</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>1</b>	<b>2</b>	<b>3</b>
<b>I</b>	x	x	x	x	x	x	x		x			x
<b>II</b>	x	x	x	x	x	x	x		x			x
<b>V</b>	x	x	x			x			x			x
<b>III</b>	x	x		x		x			x			x

Damit 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.

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

Zu guter Letzt wurden anhand Tab. 8 die Häufigkeiten der inhaltlichen Gewichtungen nach Suchmaschinen und Regionen gezählt. Die Ergebnisse sind in Tab. 15 dargestellt.

Tab. 15: Thema 7 – Häufigkeiten der Gewichtungen nach Suchmaschinen und Regionen

	knapp	eher knapp	neutral	eher ausführlich	ausführlich	Summe
google.de	1	3	7	15	0	26
bing.de	1	2	15	22	0	40
google.com	3	1	4	12	10	30
bing.com	0	0	1	13	16	30
.de	2	5	22	37	0	66
.com	3	1	5	25	26	60

Diese Ergebnisse sind (bis auf die Summen) zur besseren Anschaulichkeit als Diagramme in Abb. 9 nach Suchmaschinen und in Abb. 10 nach Regionen dargestellt.

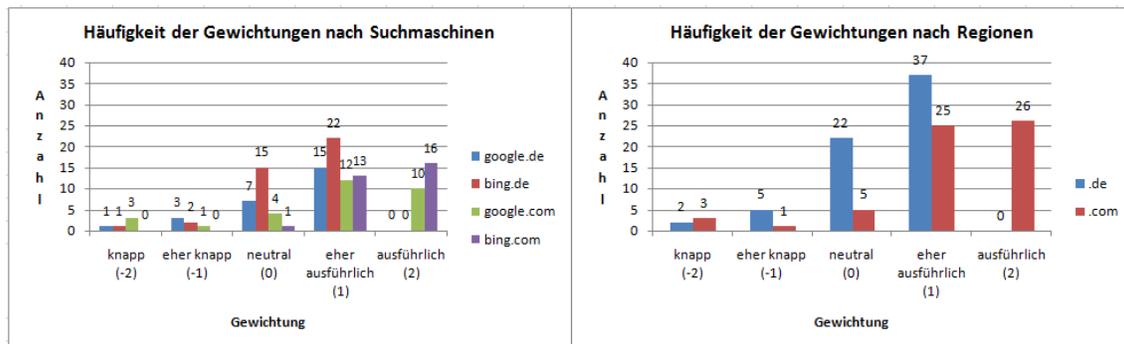


Abb. 9: Thema 7 – Häufigkeit der Gewichtungen nach Suchmaschinen

Abb. 10: Thema 7 – Häufigkeit der Gewichtungen nach Regionen

Die Summen der Gewichtungen pro Suchmaschine, und damit pro Region, können sich natürlich voneinander unterscheiden, weil nicht alle Suchergebnisse immer alle inhaltlichen Aspekte behandelten und es ja überdies irrelevante Treffer gab, die gar keine Gewichtungen erhielten. So wurden hier bei google.de am wenigsten, nämlich 26, inhaltliche Aspekte gewichtet, bei bing.de am meisten, nämlich 40. Hinsichtlich der Regionen wurden bei den deutschen Suchmaschinen insgesamt etwas mehr inhaltliche Aspekte (66) als bei den amerikanischen (60) gewichtet. Auf eine Angleichung der Daten wurde verzichtet, da die nachfolgenden statistischen Tests nicht mit aggregierten Daten funktionieren.

Es zeigt sich in diesem Fall, dass die amerikanischen Suchmaschinenergebnisse die inhaltlichen Aspekte des Themas tendenziell etwas ausführlicher behandeln als die deutschen.

Um auf signifikante Unterschiede zwischen den Regionen und Suchmaschinen zu prüfen, wurden anhand der Daten aus Tab. 15 der *Mann-Whitney-Test* bzw. der *Kruskal-Wallis-Test* durchgeführt. Dabei handelt es sich um nichtparametrische (auch verteilungsfreie oder annahmefreie) Tests – sie wurden als solche ausgewählt, weil die abhängige Variable, die Gewichtung der inhaltlichen Aspekte, ordinalskaliert, und eine Normalverteilung der Daten unwahrscheinlich ist (u.a. weil die Grundgesamtheiten der Messdaten eher klein sind) (vgl. Janssen & Laatz 2013, S. 601) – nichtparametrische Tests können allerdings auch für normalverteilte Daten angewandt werden (vgl. Keller 2012), sodass sich mit dieser Wahl auf der sicheren Seite befunden wird. Nichtparametrische Tests ordinalskaliert (oder metrischer) Testvariablen basieren meist auf Rangordnungen der Messdaten (vgl. Janssen & Laatz 2013, S. 602).

Nichtparametrische Tests unterscheiden sich durch die Anzahl der verwendeten Stichproben (ein, zwei oder mehr (allgemein  $k$ )), den Variablentyp, das Skalenniveau der Variablen und die Frage, ob die verwendeten Stichproben voneinander unabhängig sind oder nicht (vgl. ebd., S. 603). Abb. 11 zeigt in dieser Hinsicht die Verortung des Mann-Whitney- und des Kruskal-Wallis-Tests.

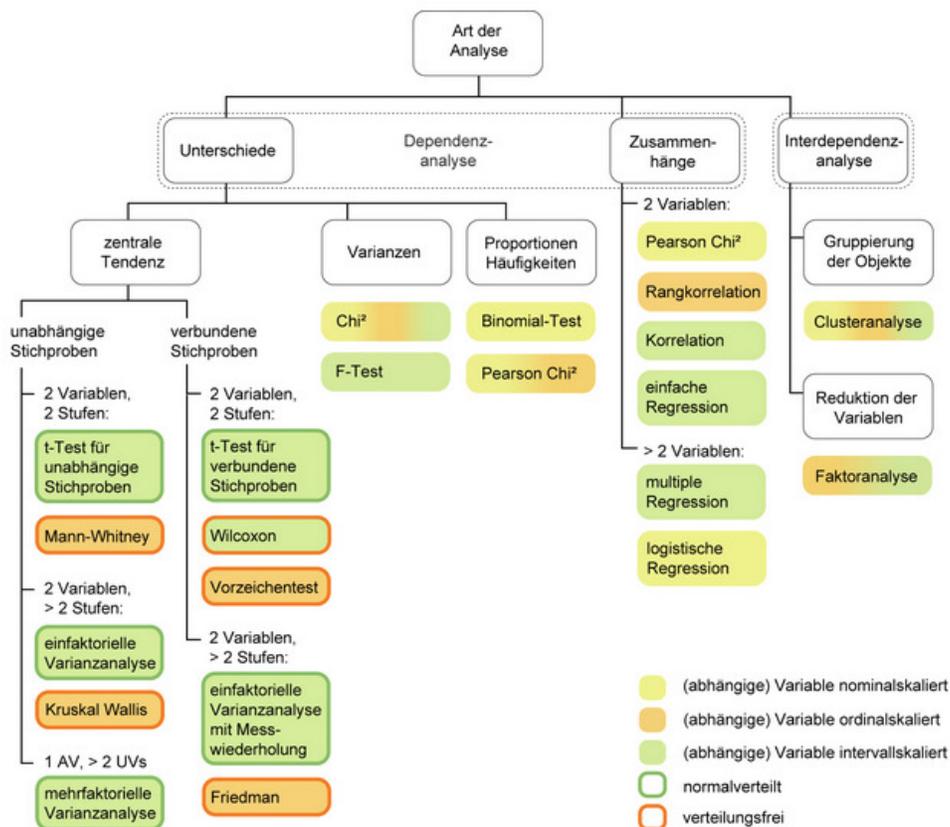


Abb. 11: Datenanalyse-Arten (UZH 2010a)

Bei der Prüfung, die hier vorgenommen werden sollte, geht es um ‚Unterschiede‘ und die ‚zentrale Tendenz‘ (also um Mittelwertsunterschiede (die abhängige Variable ist ja ordinalskaliert (vgl. UZH 2014a)) bei ‚unabhängigen Stichproben‘ (es handelt sich ja um Messwerte verschiedener Regionen bzw. Suchmaschinen, nicht um die derselben vor und nach einer Veränderung (vgl. Janssen & Laatz 2013, S. 603)) mit ‚2 Variablen und 2 Stufen‘ bzw. ‚2 Variablen > 2 Stufen‘ (die Variablen sind eine Gruppierungs- (also die Region bzw. Suchmaschine) und eine abhängige Variable (die inhaltliche Gewichtung), die Stufen sind hier die zwei Regionen bzw. vier Suchmaschinen (vgl. UZH 2014b)).

Es wird deutlich, dass sich die beiden Tests lediglich durch die Anzahl ihrer Stufen unterscheiden; Mann-Whitney mit 2 und Kruskal-Wallis mit  $k$ , und tatsächlich ist letzterer in dieser Hinsicht eine Erweiterung des ersteren (vgl. UZH 2010b).

Nach UZH 2014b können mittels des Mann-Whitney- und des Kruskal-Wallis-Tests die Unterschiede der zentralen Tendenz einer mindestens ordinalskalierten Variablen zwischen zwei bzw.  $k$  unabhängigen Stichproben untersucht werden. Bei beiden Tests handelt es sich um Rangsummentests bzw. Rangtests – die Berechnung der Teststatistik basiert also auf dem Vergleich zweier Rangreihen.

Nach Janssen und Laatz (2013, S. 622) prüft der Mann-Whitney-Test auf Unterschiede hinsichtlich der zentralen Lage von Verteilungen. Voraussetzung ist, dass die getestete Variable mindestens ordinalskaliert ist. Bei dem Test werden nicht die Messwerte der Variablen, sondern deren Rangplätze zugrunde gelegt. Der Mann-Whitney-Test ist die Alternative zum parametrischen t-Test für den Vergleich von zwei Mittelwerten von Verteilungen.

„Der Kruskal-Wallis-Test eignet sich gut zur Prüfung auf eine unterschiedliche zentrale Lage von Verteilungen (auf unterschiedliche Mediane). Er entspricht einer einfaktoriellen Varianzanalyse für Rangziffern. (ebd., S. 635)“ (Für ausführlichere Erläuterungen der beiden Tests s. etwa Janssen & Laatz 2013, S. 622-626 bzw. 635ff.)

Das Ergebnis des Mann-Whitney-Tests für die Regionen via SPSS ist in Abb. 12 dargestellt.

**Descriptive Statistics**

	N	Mean	Std. Deviation	Minimum	Maximum
VAR00002	10	12,6000	13,45940	,00	37,00
VAR00001	10	1,5000	,52705	1,00	2,00

**Mann-Whitney Test**

**Ranks**

	VAR00001	N	Mean Rank	Sum of Ranks
VAR00002	1,00	5	5,30	26,50
	2,00	5	5,70	28,50
	Total	10		

**Test Statistics<sup>a</sup>**

	VAR00002
Mann-Whitney U	11,500
Wilcoxon W	26,500
Z	-,210
Asymp. Sig. (2-tailed)	,834
Exact Sig. [2*(1-tailed Sig.)]	,841 <sup>b</sup>

a. Grouping Variable: VAR00001

b. Not corrected for ties.

Abb. 12: Thema 7 – Mann-Whitney-Test Regionen – SPSS-Ergebnis

‚VAR00001‘ ist die Gruppierungsvariable (also ‚.de (‚1,00‘) oder ‚.com (‚2,00‘) und ‚VAR00002‘ die abhängige Variable (also die Anzahlen der inhaltlichen Gewichtungen).

Was sich in der ‚Ranks‘-Tabelle an den ähnlichen ‚Mean Ranks‘ abzeichnet, gilt auch für die Signifikanzen in Tabelle ‚Test Statistics‘ (bei kleineren Stichproben ( $n_1 < 10$  und  $n_2 < 10$ ) sollte die ‚Exact Significance‘, bei größeren die ‚Asymptotic‘ verwendet werden (UZH 2010c)): Hier liegen beide Werte mit 0,841 bzw. 0,834 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 via SPSS ist in Abb. 13 dargestellt.

	N	Mean	Std. Deviation	Minimum	Maximum
VAR00002	20	6,3000	6,88324	,00	22,00
VAR00001	20	2,5000	1,14708	1,00	4,00

### Kruskal-Wallis Test

	VAR00001	N	Mean Rank
VAR00002	1,00	5	10,00
	2,00	5	11,10
	3,00	5	11,60
	4,00	5	9,30
Total		20	

	VAR00002
Chi-Square	,474
df	3
Asymp. Sig.	,925

a. Kruskal Wallis Test  
b. Grouping Variable: VAR00001

Abb. 13: Thema 7 – Kruskal-Wallis-Test Suchmaschinen – SPSS-Ergebnis

Wieder ist ‚VAR00001‘ die Gruppierungsvariable (also google.de (‚1,00‘), bing.de (‚2,00‘), google.com (‚3,00‘) oder bing.com (‚4,00‘)) und ‚VAR00002‘ die abhängige Variable (also die Anzahlen der inhaltlichen Gewichtungen).

Was sich in der ‚Ranks‘-Tabelle an den ähnlichen ‚Mean Ranks‘ abzeichnet, gilt auch für die Signifikanz in Tabelle ‚Test Statistics‘: Hier liegt der Wert mit 0,925 ebenfalls deutlich über  $p = 0,05$ . Auch die Suchmaschinen unterscheiden sich in ihren inhaltlichen Gewichtungen also nicht signifikant voneinander.

Bei einem signifikanten Kruskal-Wallis-Test-Ergebnis sollte ein sog. *Post-Hoc-Test* durchgeführt werden, um zu prüfen, wo (also zwischen welchen Rangreihen (hier: Suchmaschinen)) die signifikanten Unterschiede auftreten (vgl. UZH 2010b).

Nach UZH 2010b werden bei diesem Post-Hoc-Test für alle oder einzelne Paarvergleiche Mann-Whitney-Tests durchgeführt. Allerdings muss dabei zur Vermeidung von Fehlern durch mehrfache Testung (*Alphafehler-Kumulierung* oder  $\alpha$ -Inflation) das Signifikanzniveau korrigiert werden. Dies kann bspw. durch die *Bonferroni-Methode* geschehen: Dabei ergibt sich das neue Signifikanzniveau aus dem alten geteilt durch die Anzahl der Paarvergleiche, die durchgeführt werden sollen.

Wäre das Kruskal-Wallis-Test-Ergebnis der Suchmaschinen für Thema 7 also signifikant gewesen, wären an dieser Stelle alle vier Suchmaschinen paarweise durch Mann-Whitney-Tests miteinander verglichen worden, und zwar unter dem neuen Signifikanzniveau  $p = 0,05 / 6 = 0,0083$ .

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

Hypothese 3 kann für Thema 7 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.

Damit kann Forschungsfrage 3 für Thema 7 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.

### 5.3 Ergebnisse

Nachdem im vorigen Abschnitt die praktische Methodik der empirischen Untersuchung erläutert wurde, wird in dieser Sektion das Gesamtergebnis der empirischen Untersuchung dargestellt.

Zunächst werden Bemerkungen zu Problemen beim Scraping und ggf. weitere Anmerkungen zu den einzelnen Themen gemacht. Diese können am Material auf der beiliegenden CD nachvollzogen werden.

Hinsichtlich Thema 1 – neuer Sparkurs der Deutschen Bank – wurde die Auswertung verworfen, da alle deutschen Ergebnisse irrelevant waren. Viele hatten schlicht andere Themen zum Inhalt, und einige drehten sich zwar um Einsparmaßnahmen der Deutschen Bank, waren aber veraltet. Interessanterweise gab es relevante amerikanische Ergebnisse, doch ein Vergleich deutscher und amerikanischer Suchmaschinen war so nicht möglich.

Zu Thema 2 – der Entmachtung Varoufakis‘ nach dem Riga-Gipfel:

- Bei bing.de wurden die drei Universal-Search-News-Treffer am Anfang der Ergebnisliste nicht mit erfasst, zwei davon wurden als organische Dubletten gescrapet.

Es zeigte sich, dass viele amerikanische Suchergebnisse irrelevant waren. Einige drehten sich zwar um Varoufakis, waren jedoch veraltet, und manche hatten gänzlich andere Themen zum Inhalt. Da jedoch ein paar relevante dabei und die deutschen alle relevant waren, wurde die Auswertung trotzdem durchgeführt.

Zu Thema 3 – Erdbeben in Nepal:

- Bei bing.de wurden die Universal-Search-News-Treffer am Anfang der Ergebnisliste nicht mit erfasst.

- Bei bing.com wurden die Universal-Search-News-Treffer am Anfang der Ergebnisliste nicht mit erfasst.

Zu Thema 4 – Kritik an Wahlkampfspenden Hillary Clintons:

- Bei google.de gab es nach den ersten zwei organischen Ergebnissen drei Universal-Search-News-Treffer, diese wurden als die ersten Ergebnisse erfasst.
- Bei bing.de gab es nach den ersten vier organischen Ergebnissen drei Universal-Search-News-Treffer, einer davon und das fünfte organische Ergebnis wurden ausgelassen, es handelte sich jedoch um Dubletten bereits erfasster Ergebnisse.
- Bei google.com wurden das zweite und dritte organische Ergebnis ausgelassen, es handelte sich jedoch um Dubletten bereits erfasster Universal-Search-News-Treffer.
- Bei bing.com wurden die drei Universal-Search-News-Treffer am Anfang der Ergebnisliste nicht mitgescrapet.

Die Suchanfragen („hillary clinton geld/money“) waren zu allgemein gewählt; so waren insbes. einige amerikanische Suchergebnisse irrelevant, weil sie sich etwa lediglich um das Sammeln Clintons von Geldern drehten, ohne auf die Kontroversen einzugehen. Die Auswertung wurde aufgrund der relevanten Ergebnisse dennoch durchgeführt.

Hinsichtlich Thema 5 – Großbritanniens Rolle hinsichtl. EU und Welt – wurde die Auswertung verworfen, da keins der deutschen Ergebnisse relevant war. Viele waren um mehrere Jahre veraltet, drehten sich um grundlegende Informationen zu England, beschäftigten sich mit Englands Rolle im Ersten Weltkrieg oder hatten völlig andere Themen zum Inhalt. Es gab relevante amerikanische Ergebnisse, doch ein Vergleich deutscher und amerikanischer Suchmaschinen war so wieder nicht möglich. (Außerdem wäre es durch die unterschiedliche Benennung der Region in den Suchanfragen ‚england‘ bzw. ‚britain‘ zu einer Bedeutungsverschiebung gekommen.)

Zu Thema 6 – Proteste in Burundi gegen Präsident Nkurunziza:

- Bei google.de wurden das zweite und sechste organische Ergebnis nicht mit erfasst, es handelte sich jedoch um Dubletten bereits gescrapter Universal-Search-News-Treffer.
- Bei bing.de wurden die drei Universal-Search-News-Treffer am Anfang der Ergebnisliste nicht mit erfasst, eins davon wurde als organische Dublette gescrapet.
- Bei google.com wurden das zweite und sechste organische Ergebnisse nicht mit erfasst, es handelte sich jedoch um Dubletten bereits erfasster Universal-Search-News-Treffer.
- Bei bing.com wurden die drei Universal-Search-News-Treffer am Anfang der Ergebnisliste nicht mit erfasst.

Von den amerikanischen Ergebnissen (insbes. von bing.com) waren einige bereits ein bis zwei Wochen alt. Da sie, auch wenn sie nicht die aktuellen Entwicklungen berücksichtigten, in Bezug auf das Thema relevant waren, wurden sie trotzdem ausgewertet.

Die Anmerkungen zu Thema 7 – Luftangriff Israels im syrischen Grenzgebiet – wurden bereits im vorigen Kapitel genannt.

Hinsichtlich Thema 8 – USA-Japan-Verteidigungsabkommen – wurde die Auswertung verworfen, da alle deutschen Ergebnissen irrelevant waren. Sie hatten oft europäische Bemühungen um ein Freihandelsabkommen mit Japan zum Inhalt, ein Verteidigungsabkommen der USA mit den Philippinen, oder zwar das USA-Japan-Verteidigungsabkommen, doch im Kontext von bereits ein Jahr alten Geschehnissen. Ein Vergleich deutscher und amerikanischer Suchmaschinen war so wieder nicht möglich.

Zu Thema 9 – Rücktritt Lee Wan-Koos:

- Bei google.de gab es nach den ersten zwei organischen Ergebnissen drei Universal-Search-News-Treffer, diese wurden als die ersten Ergebnisse erfasst.
- Bei bing.de wurden die drei Universal-Search-News-Treffer am Anfang der Ergebnisliste nicht mit erfasst, eins davon wurde als organische Dublette gescrapet.
- Bei google.com wurden das dritte und sechste organische Ergebnis nicht mit erfasst, es handelte sich jedoch um Dubletten bereits erfasster Universal-Search-News-Treffer.
- Bei bing.com wurden die drei Universal-Search-News-Treffer am Anfang der Ergebnisliste nicht mit erfasst, eins davon wurde als organische Dublette gescrapet. Zwischen dem dritten und vierten organischen Ergebnis wurde ein „Trends in sozialen Netzwerken zu lee wan-koo resignation“-Universal-Search-Container mit zwei Treffern (einer von Twitter, einer von Facebook) gescrapet. Diese wurden als normale Ergebnisse in die Auswertung mit einbezogen. Der Link hinter der „Trends in sozialen Netzwerken zu lee wan-koo resignation“-Überschrift war auch als Ergebnis erfasst worden, wurde als solches jedoch aus der Auswertung entfernt, da die Überschriften der News-Container so ja auch nicht erfasst wurden.

Bei diesem Thema kann man ähnlich wie bei Thema 7 sehen, wie der gleiche Inhalt unter unterschiedlichen URLs und Titeln auftaucht.

Hinsichtlich Thema 10 – der 40. Jährung des Vietnamkrieg-Endes – wurde die Auswertung verworfen, da die Ergebnisse zu divers waren. Themen der deutschen waren z.B. die Geschichte eines Vaters und

seines Sohnes, die angeblich während des Krieges in den Dschungel geflohen waren und erst jetzt entdeckt wurden, Spätfolgen des von US-Truppen verwendeten Herbizids ‚Agent Orange‘, Bücher zum Vietnamkrieg usw. Themen der amerikanischen Suchergebnisse waren z.B. auch Nachwirkungen ‚Agent Oranges‘, die Rolle Amerikas im Krieg, die aktuelle Lage in Vietnam, Erinnerungen usf. Angesichts dieser Vielfalt war es nicht möglich, für dieses Thema einen gemeinsamen Nenner aus ein paar Aspekten zu definieren.

Hinsichtlich Thema 11 – Entwicklungshilfe- und Migrantenproblematik Afrikas – wurde die Auswertung verworfen, da die Suchanfragen ‚afrika migranten‘ / ‚africa migrants‘ zu allgemein gewählt und die Ergebnisse daher zu divers waren. Deutsche behandelten z.B. die Einwanderungspolitik Europas, Flüchtlingshilfsorganisationen wie *Pro Asyl*, und allgemeine Informationen etwa zu Migration oder der ‚Great Migration‘. Amerikanische Ergebnisse drehten sich bspw. um den gefährlichen Weg afrikanischer Migranten nach Europa, um den Tod vieler im Mittelmeer, und um Angriffe auf afrikanische Migranten in Südafrika. Angesichts dieser Vielfalt war es wieder nicht möglich, für dieses Thema einen gemeinsamen Nenner aus ein paar Aspekten zu definieren.

Zu Thema 12 – Hitze und Regen durch Klimawandel:

- Bei google.de wurde das erste organische Ergebnis nicht mit erfasst, es handelte sich jedoch um eine Dublette eines bereits gescrapten Universal-Search-News-Treffer.
- Bei google.com gab es nach dem ersten organischen Ergebnis drei Universal-Search-News-Treffer – diese wurden als die ersten Ergebnisse gescrapet. Das vierte organische Ergebnis wurde nicht mit erfasst, es handelte sich jedoch um eine Dublette eines Universal-Search-News-Treffers.
- Bei bing.com wurden die drei Universal-Search-News-Treffer am Anfang der Ergebnisliste nicht mit erfasst, eins davon wurde als organische Dublette gescrapet.

Die Ergebnisse erwiesen sich wieder als ziemlich divers. In den Zeitungen wurde das Thema Klimawandel im Kontext einer neuen Schweizer Studie behandelt – dies wurde als Relevanzmerkmal für die Suchergebnisse übernommen. Das führte natürlich dazu, dass viele Ergebnisse irrelevant wurden (einige waren allerdings sowieso nicht aktuell) und es wäre besser gewesen, diesen Aspekt in die Suchanfragen (‚klimawandel hitze regen‘ bzw. ‚climate change heat rain‘) mit aufzunehmen, doch musste die Auswertung so wenigstens nicht ganz verworfen werden.

Zu Thema 13 – Übernahme Mylans durch Teva:

- Bei google.de wurden das erste und zweite organische Ergebnis nicht mit erfasst, es handelte sich jedoch um Dubletten bereits gescrapter Universal-Search-News-Treffer.

- Bei bing.de wurden die drei Universal-Search-News-Treffer am Anfang der Ergebnisliste nicht mit erfasst, eins davon wurde als organische Dublette gescrapet.
- Bei google.com wurde das dritte organische Ergebnis nicht mit erfasst, es handelte sich jedoch um eine Dublette eines bereits gescrapten Universal-Search-News-Treffers.
- Bei bing.com wurden die drei Universal-Search-News-Treffer am Anfang der Ergebnisliste nicht mit erfasst.

Zu Thema 14 – erneute Kämpfe in Mariupol:

- Bei google.de gab es nach den ersten drei organischen Ergebnissen drei Universal-Search-News-Treffer, diese wurden als die ersten Ergebnisse erfasst.
- Bei google.com gab es nach den ersten drei organischen Ergebnissen drei Universal-Search-News-Treffer, diese wurden als die ersten Ergebnisse erfasst.
- Bei bing.com wurden die drei Universal-Search-News-Treffer am Anfang der Ergebnisliste nicht mit erfasst.

Bei diesem Thema fiel es den Suchmaschinen schwer, relevante – was in diesem Fall hieß: aktuelle – Ergebnisse auszugeben. Das Problem war, dass ‚erneute Kämpfe in Mariupol‘ ein längerfristiges Thema ist (diese immer mal wieder aufflammenden Gefechte zwischen der ukrainischen Armee und prorussischen Aufständischen finden seit über einem Jahr statt), und dass das einzige Merkmal, was die hier gemeinten ‚erneuten Kämpfe‘ für die Suchmaschinen charakterisierte, eben Aktualität war. Das Scraping fand am Montag, dem 28.4.2015 statt, die gemeinten Kämpfe ereigneten sich direkt am Wochenende, am 26./27.4. Freilich hätte man die Suchanfragen ‚ostukraine kämpfe mariupol‘ / ‚east ukraine fights mariupol‘ präzisieren können, etwa durch den Zusatz ‚april 2015‘, andererseits sollte man als Nutzer davon ausgehen können, dass Suchmaschinen in Fällen wie diesem automatisch aktuelle Treffer nach oben ranken. Das hat hier jedoch mehr schlecht als recht funktioniert: Fast alle Ergebnisse drehten sich zwar um das Thema, doch konnten alle von vor dem Wochenende als irrelevant verworfen werden, was auch getan wurde. So lieferte Bing unter beiden Domains keine relevanten Ergebnisse, google.de nur drei und google.com zwei.

Trotz dieser mageren Ausbeute und dem Umstand, dass Bing ausfiel, wurde die Auswertung durchgeführt, da es offensichtlich möglich war, relevante Ergebnisse in beiden Regionen auszugeben, und es Bings Manko ist, dass bei ihr keins dabei war.

Insgesamt wurden von den 14 Themen also neun ausgewertet, wobei die Fehler beim Scraping zu einem großen Teil bei der Auswertung korrigiert werden konnten.

Im Folgenden werden die Ergebnisse der Auswertungen korrespondierend zu den Forschungsfragen und -hypothesen (s. Kap. 5.1) präsentiert.

Es wird mit den beiden Forschungsfragen zur Vollständigkeit der vermittelten Bilder begonnen. Abb. 14 zeigt die Anzahl vollständig vermittelter Bilder nach Suchmaschinen, Abb. 15 nach Regionen im Durchschnitt.

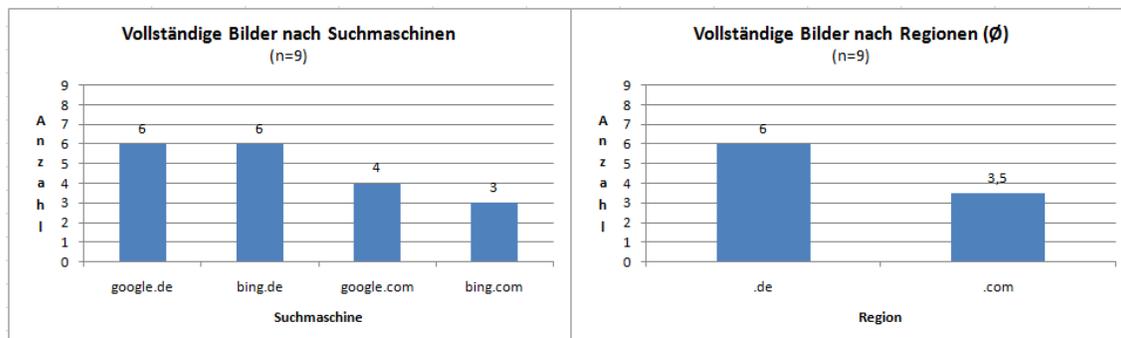


Abb. 14: Vollständige Bilder nach Suchmaschinen

Abb. 15: Vollständige Bilder nach Regionen (Durchschnitte)

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 vollständigen Bilder aktueller internationaler Themen.

Damit 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.

Damit 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.

Es wird mit der Forschungsfrage zur Abdeckung wichtiger inhaltliche Aspekte fortgefahren. Abb. 16 zeigt die Anzahl von Fällen mit den auf Position 1 abgedeckten drei wichtigsten inhaltlichen Aspekten nach Suchmaschinen, Abb. 17 nach Regionen im Durchschnitt.

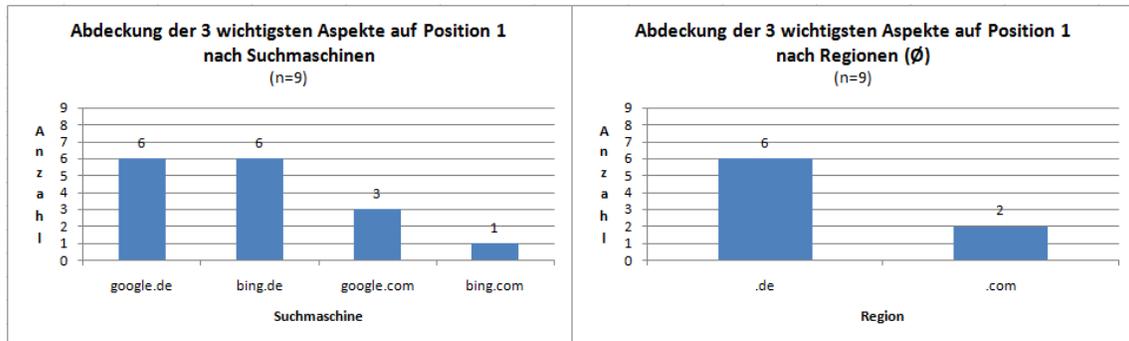


Abb. 16: Abdeckung der drei wichtigsten Aspekte auf Position 1 nach Suchmaschinen

Abb. 17: Abdeckung der drei wichtigsten Aspekte auf Position 1 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.

Abb. 18 zeigt die Anzahl von Fällen mit den auf den ersten drei Positionen abgedeckten 75 Prozent wichtigsten inhaltlichen Aspekten nach Suchmaschinen, Abb. 19 nach Regionen im Durchschnitt.

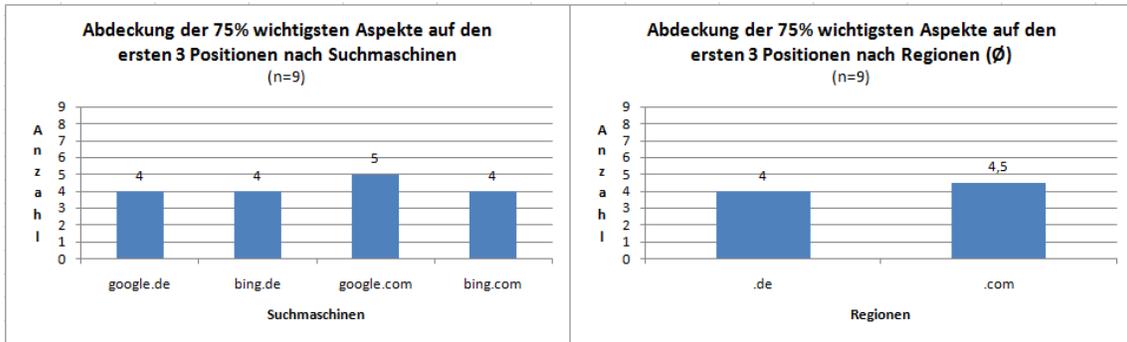


Abb. 18: Abdeckung der 75 Prozent wichtigsten Aspekte auf den ersten drei Positionen nach Suchmaschinen

Abb. 19: Abdeckung der 75 Prozent wichtigsten Aspekte auf den ersten drei Positionen nach Regionen (Durchschnitte)

Es zeigt sich, dass die Suchmaschinen mit etwa 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.

Damit 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.

Es wird mit der Forschungsfrage zur Gewichtung wichtiger inhaltlicher Aspekte abgeschlossen. Tab. 16 zeigt die Häufigkeiten der inhaltlichen Gewichtungen nach Suchmaschinen und Regionen.

Tab. 16: Häufigkeiten der Gewichtungen nach Suchmaschinen und Regionen

	knapp	eher knapp	neutral	eher ausführlich	ausführlich	Summe
google.de	45	30	43	101	61	280
bing.de	47	24	58	113	65	307
google.com	41	35	38	84	61	259
bing.com	34	20	35	55	56	200
.de	92	54	101	214	126	587
.com	75	55	73	139	117	459

Diese Ergebnisse sind (bis auf die Summen) zur besseren Anschaulichkeit als Diagramme in Abb. 20 nach Suchmaschinen und in Abb. 21 nach Regionen dargestellt.

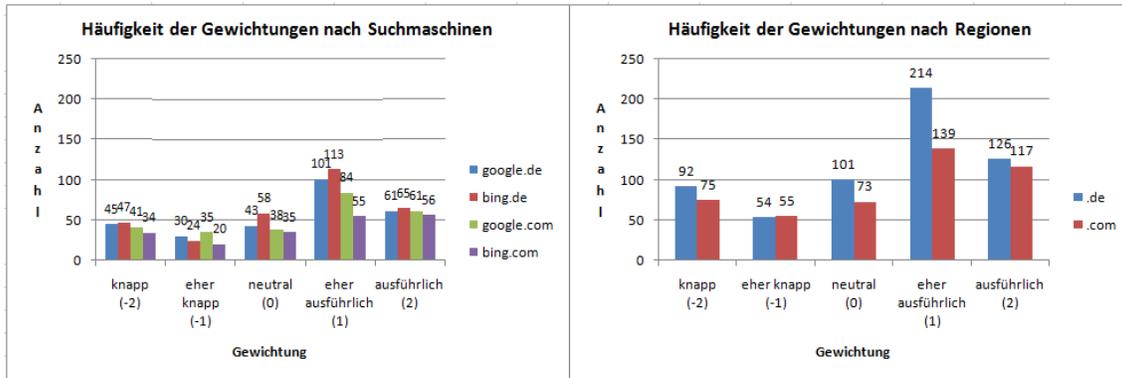


Abb. 20: Häufigkeit der Gewichtungen nach Suchmaschinen

Abb. 21: Häufigkeit der Gewichtungen 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 bzgl. aller Suchmaschinen eine ähnliche Verteilung der Gewichtungen: Es gibt einige knappe Aspekte, ein paar 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 weiter den Suchmaschinen zu prüfen, wurde anhand der Daten aus Tab. 16 via SPSS der Mann-Whitney- bzw. der Kruskal-Wallis-Test durchgeführt (s. voriges Kap.). Die Ergebnisse sind in Abb. 22 und 23 dargestellt.

Descriptive Statistics						Descriptive Statistics					
	N	Mean	Std. Deviation	Minimum	Maximum		N	Mean	Std. Deviation	Minimum	Maximum
VAR00002	10	104,6000	48,10914	54,00	214,00	VAR00002	20	52,3000	24,28146	20,00	113,00
VAR00001	10	1,5000	,52705	1,00	2,00	VAR00001	20	2,5000	1,14708	1,00	4,00

### Mann-Whitney Test

Ranks				
	VAR00001	N	Mean Rank	Sum of Ranks
VAR00002	1,00	5	6,00	30,00
	2,00	5	5,00	25,00
	Total	10		

Test Statistics <sup>a</sup>	
	VAR00002
Mann-Whitney U	10,000
Wilcoxon W	25,000
Z	-,522
Asymp. Sig. (2-tailed)	,602
Exact Sig. [2*(1-tailed Sig.)]	,690 <sup>b</sup>

a. Grouping Variable: VAR00001  
b. Not corrected for ties.

### Kruskal-Wallis Test

Ranks			
	VAR00001	N	Mean Rank
VAR00002	1,00	5	11,30
	2,00	5	12,80
	3,00	5	10,80
	4,00	5	7,10
	Total	20	

Test Statistics <sup>a,b</sup>	
	VAR00002
Chi-Square	2,515
df	3
Asymp. Sig.	,473

a. Kruskal Wallis Test  
b. Grouping Variable: VAR00001

Abb. 22: Mann-Whitney-Test  
Regionen – SPSS-Ergebnis

Abb. 23: Kruskal-Wallis-Test  
Suchmaschinen – SPSS-Ergebnis

Zum Mann-Whitney-Test bzgl. der Regionen (Abb. 22): ‚VAR00001‘ ist die Gruppierungsvariable (also ‚.de (,1,00‘) oder ‚.com (,2,00‘)) und ‚VAR00002‘ die abhängige Variable (also die Anzahlen der inhaltlichen Gewichtungen).

Was sich in der ‚Ranks‘-Tabelle an den ähnlichen ‚Mean Ranks‘ abzeichnet, gilt auch für die Signifikanzen in Tabelle ‚Test Statistics‘: Hier liegen beide Werte mit 0,602 bzw. 0,69 deutlich über dem üblichen Signifikanzniveau von 0,05. Die Regionen unterscheiden sich in ihren inhaltlichen Gewichtungen also nicht signifikant voneinander.

Zum Kruskal-Wallis-Test bzgl. der Suchmaschinen (Abb. 23): Wieder ist ‚VAR00001‘ die Gruppierungsvariable (also ‚google.de (,1,00‘), ‚bing.de (,2,00‘), ‚google.com (,3,00‘) oder ‚bing.com (,4,00‘)) und ‚VAR00002‘ die abhängige Variable (also die Anzahlen der inhaltlichen Gewichtungen).

Was sich in der ‚Ranks‘-Tabelle an den relativ ähnlichen ‚Mean Ranks‘ abzeichnet, gilt auch für die Signifikanz in Tabelle ‚Test Statistics‘: Hier liegt der Wert mit 0,473 ebenfalls deutlich über  $p = 0,05$ . Auch die Suchmaschinen unterscheiden sich in ihren inhaltlichen Gewichtungen also nicht signifikant voneinander.

Insgesamt zeigen sich bzgl. der inhaltlichen Gewichtungen der Suchergebnisse also keine signifikanten Unterschiede zwischen den Regionen, und auch nicht zwischen den Suchmaschinen. Damit kann im Prinzip Hypothese 3 und mit ihr Forschungsfrage 3 beantwortet werden – um allerdings zu prüfen, ob signifikante Unterschiede vielleicht in Hinsicht auf die einzelnen Themen

vorliegen, wurden noch für die neun Einzelthemen jeweils der Mann-Whitney-Test für die Regionen und der Kruskal-Wallis-Test für die Suchmaschinen durchgeführt. Die Ergebnisse sind zusammengefasst in Tab. 17 dargestellt (für die ausführlichen SPSS-Ergebnisse s. beiliegende CD).

Tab. 17: Mann-Whitney- und Kruskal-Wallis-Tests Einzelthemen – Ergebnisse

Nr.	Test	p
2	Mann-Whitney Regionen	0,21 bzw. 0,16
	Kruskal-Wallis Suchmaschinen	<b>0,006</b>
3	Mann-Whitney Regionen	0,6 bzw. 0,69
	Kruskal-Wallis Suchmaschinen	0,967
4	Mann-Whitney Regionen	0,138 bzw. 0,151
	Kruskal-Wallis Suchmaschinen	0,324
6	Mann-Whitney Regionen	0,071 bzw. 0,095
	Kruskal-Wallis Suchmaschinen	<b>0,006</b>
7	Mann-Whitney Regionen	0,834 bzw. 0,841
	Kruskal-Wallis Suchmaschinen	0,925
9	Mann-Whitney Regionen	0,753 bzw. 0,841
	Kruskal-Wallis Suchmaschinen	0,929
12	Mann-Whitney Regionen	1 bzw. 1
	Kruskal-Wallis Suchmaschinen	0,752
13	Mann-Whitney Regionen	0,829 bzw. 0,841
	Kruskal-Wallis Suchmaschinen	0,803
14	Mann-Whitney Regionen	0,281 bzw. 0,31
	Kruskal-Wallis Suchmaschinen	<b>0,014</b>

Es zeigt sich, dass in Hinsicht auf die Regionen bei keinem der neun Themen signifikante Unterschiede vorliegen, aber hinsichtlich der Suchmaschinen bei dreien, nämlich 2 – der Entmachtung Varoufakis’ –, 6 – Proteste in Burundi gegen Präsident Nkurunziza –, und 14 – erneute Kämpfe in Mariupol.

Um zu versuchen, diese signifikanten Unterschiede zu lokalisieren, wurden wie im vorigen Kap. erläutert Post-Hoc-Tests durchgeführt. Die Ergebnisse dieser paarweisen Mann-Whitney-Tests sind zusammengefasst in Tab. 18 dargestellt (für die ausführlichen SPSS-Ergebnisse s. beiliegende CD).

Tab. 18: Post-Hoc-Tests Thema 2, 6 und 14 – Ergebnisse

Neues Signifikanzniveau nach Bonferroni-Methode -> p = 0,05 / 6 = 0,0083				
		Nr.		
		2	6	14
google.com	google.de	0,221 bzw. 0,667	not enough valid cases	empty groups
google.com	bing.de	0,221 bzw. 0,667	not enough valid cases	empty groups
bing.com	google.de	empty groups	not enough valid cases	not enough valid cases
bing.com	bing.de	empty groups	not enough valid cases	not enough valid cases
google.com	bing.com	0,48 bzw. 0,667	not enough valid cases	empty groups
google.de	bing.de	not enough valid cases	not enough valid cases	empty groups

Aufgrund der ungenügenden Daten konnten die Tests oft nicht durchgeführt werden. Lediglich bei Thema 2 – der Entmachtung Varoufakis’ – treten Unterschiede bei der inhaltlichen Gewichtung zwischen google.com und den drei anderen Suchmaschinen auf, diese sind jedoch nicht signifikant.

Evtl. könnte hier mit anderen Post-Hoc-Methoden zu besseren Ergebnissen gelangt werden – dies soll an dieser Stelle jedoch nicht weiterverfolgt werden.

Hypothese 3 kann 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.

Damit 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.

#### **5.4 Diskussion, Einschränkungen und weitere Forschung**

In diesem Kapitel werden die in der vorigen Sektion präsentierten Ergebnisse diskutiert und in ihrer Aussagekraft eingeschränkt, und es werden Ansätze zu weiterer Forschung aufgezeigt.

Die Erkenntnisse der empirischen Untersuchung dieser Arbeit lauten also, dass Google und Bing in Deutschland und den USA keine vollständigen Bilder aktueller internationaler Themen vermitteln (F1.1, s. Kap. 5.1), wobei es Unterschiede hinsichtlich der wichtigen Aspekte gibt (F1.2); dass sie auf den ersten Trefferpositionen nicht die drei wichtigsten inhaltlichen Aspekte abdecken (F2); und dass es bei der Gewichtung der wichtigen inhaltlichen Aspekte keine signifikanten Unterschiede gibt (F3).

Das zeichnet ein relativ negatives Bild: Es sind definitiv unterschiedlich wirkende inhaltliche Suchmaschinenverzerrungen vorhanden. Allerdings ist die Aussagekraft dieser Ergebnisse stark eingeschränkt. Warum, wird im Folgenden durch einen kritischen Blick auf die Methodik erläutert.

Die Methodik der empirischen Untersuchung baut wie beschrieben grundlegend auf den Methoden der Inhaltsanalyse und von Retrievaltests auf (s. Kap. 5.2.1). D.h. dass Kritik an diesen Methoden auch auf diese Methodik zutreffen kann. In Hinsicht auf Retrievaltests ist dies etwa bei dem zugrundeliegenden Nutzermodell der Fall: Nach Lewandowski (2015, S. 220f) wird davon ausgegangen, dass ein Nutzer die Treffer nacheinander durchgeht, ohne sich von der Trefferpräsentation (sowohl die Trefferbeschreibungen aus Überschrift, Beschreibungstext und URL (vgl. ebd., S. 140-144) als auch Hervorhebungen z.B. durch das Universal-Search-Design (s. Kap. 3.2)) beeinflussen zu lassen, oder seine Suchanfrage zu verändern. Standard-Retrievaltests sind nicht fähig, den oftmals interaktiven und in mehreren Schritten verlaufenden Prozess der Recherche abzubilden. So kann kaum von einer realistischen Abbildung des Nutzerverhaltens gesprochen werden.

Davon abgesehen wurden bei dieser Untersuchung nicht alle Arten von Top-Ergebnissen erfasst, sondern nur textliche, also organische und Universal-Search-News-Treffer (s. Kap. 5.2.2). D.h. andere Universal-Search-Treffer wie z.B. Videos wurden nicht berücksichtigt.

Auch das etablierte Minimum von 50 Suchanfragen konnte hier nicht erreicht werden (wobei dann die Auswertung mit der angewandten Methodik vom Aufwand auch nicht machbar gewesen wäre), 14 Themen, von denen letztlich nur neun ausgewertet wurden, liegen weit darunter und sind schlicht zu wenig, um valide Ergebnisse zu produzieren.

In diesem Zusammenhang hat sich bzgl. der Zeitungsauswahl (s. Tab 5) gezeigt, dass die beiden Boulevardblätter – *Bild* und *USA Today* – kaum Nutzen gebracht haben (da sie eher regional ausgerichtet und von geringerem Umfang sind). Statt ihnen sollten das nächste Mal weitere auflagenstarke Qualitätszeitungen einbezogen werden, um die Themen-Ausbeute möglichst zu erhöhen.

Bzgl. der Suchanfragen hat sich gezeigt, dass diese z.T. zu allgemein gewählt waren (etwa bei den Themen 4 – Hillary Clinton und Spendengelder: ‚hillary clinton geld/money‘ und 11 – Entwicklungshilfe- und Migrantenproblematik Afrikas: ‚afrika migranten‘/‚africa migrants‘). Dies führte zu allzu diversen bzw. irrelevanten Suchergebnissen. Es sollte bei der Anfragenformulierung mit Bedacht vorgegangen werden, da die Suchergebnisse natürlich auf den Anfragen beruhen. Auch durch Ergebnisse, die nicht aufgrund der Anfrage, sondern aufgrund der Suchmaschinen irrelevant waren, variierte die Anzahl der auswertbaren Suchergebnisse bei den Einzelauswertungen z.T. erheblich (z.B. waren bei Thema 14 statt 40 (4 Suchmaschinen x 10 Ergebnisse) nur 5 relevant) – zu diverse und irrelevante Ergebnisse führten zum Verwerfen kompletter Einzelauswertungen und teilw. erheblichen Verzerrungen in (trotzdem) durchgeführten (s. Kap. 5.2.2). Das schmälert die Validität der Untersuchungsergebnisse zusätzlich. Davon abgesehen erfolgte die Definition ‚irrelevanter‘ Ergebnisse (s. Kap. 5.2.2) natürlich in gewissem Maße subjektiv – ein anderer Bearbeiter hätte möglicherweise anders über Irrelevanz entschieden.

Weitere Einschränkungen ergeben sich aufgrund der Probleme beim Scraping, durch die etwa organische Ergebnisse, die vorher als Universal-Search-News-Treffer erfasst worden waren, ausgelassen wurden, oder Universal-Search-News-Treffer nicht oder unvollständig erfasst wurden (s. Kap. 5.3). Zwar konnten diese Fehler bei der Auswertung oft, aber eben nicht immer, korrigiert werden.

Natürlich wurden die inhaltlichen Aspekte (und ihre Definitionen) der Themen subjektiv bestimmt – ein anderer Bearbeiter hätte sie möglicherweise anders gewählt und/oder definiert. Bei der Bestimmung der Aspekte gab es ein paar knifflige Fälle; z.B. lag bei Thema 3 – Erdbeben in Nepal – bei den deutschen Suchergebnissen z.T. viel Gewicht auf der Hilfe aus Deutschland, während es bei den amerikanischen Ergebnissen kein vergleichbares Gewicht auf der amerikanischen Hilfe gab. Es

wurde daher ein Aspekt für die deutsche und keiner für die amerikanische Hilfe definiert (s. beiliegende CD). Es leuchtet allerdings ein, dass das keine gute Lösung war, denn so fehlte den amerikanischen Suchmaschinen natürlich dieser Aspekt, was ja aber nicht bedeutet, dass amerikanischen Suchmaschinennutzern ein (für sie) wichtiger inhaltlicher Aspekt vorenthalten würde. Rückblickend wäre es in diesem Fall z.B. besser gewesen, einen variablen Aspekt ‚Deutsche/amerikanische Hilfe‘ zu definieren, der bei der Auswertung nur bei Vorhandensein zu berücksichtigen gewesen wäre, und diese Problematik damit zu umgehen.

Die inhaltlichen Aspekte sind das Herzstück der Untersuchung, sie stellen das Gerüst dar und fordern auch den größten kognitiven Aufwand – daher sollte bei ihrer Bestimmung und Definition mit Umsicht vorgegangen, und versucht werden, bei den Einzelauswertungen nicht das große Ganze aus dem Blick zu verlieren und Entscheidungen, wie bspw. die eben erläuterte eines variablen Aspekts, so zu treffen, dass sie in Einklang mit der gesamten Untersuchung stehen (eine Sichtung des gesamten Materials vorab in Hinsicht auf evtl. auftretende Probleme ist dabei sicher hilfreich).

Auch die Bewertungen der Suchergebnisse wurden natürlich subjektiv vorgenommen – ein anderer Bearbeiter hätte sie möglicherweise anders gewählt.

Die gesamte Auswertung ist also sehr subjektiv, mögliche Eigenarten des Bearbeiters konnten nicht durch einen anderen oder mehrere andere ausgeglichen werden und schlugen voll zu Buche – Interrater-Reliabilität ist nicht vorhanden. Das nächste Mal sollten mindestens zwei Bearbeiter zusammenarbeiten, um die Validität der Untersuchungsergebnisse zu erhöhen.

Aufgrund all dieser Punkte haben die Ergebnisse dieser empirischen Untersuchung im Grunde keine Aussagekraft (sie können höchstens als eine mögliche Tendenz gesehen werden).

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

Es hat sich gezeigt, dass sich für eine empirische Untersuchung wie die vorgenommene am besten singuläre Themen, die hervorstechen und möglichst noch nicht da waren, eignen (wie hier Thema 3 – Erdbeben in Nepal – oder 9 – Rücktritt Südkoreas Ministerpräsident Lee).

Freilich ist man dabei auch vom Weltgeschehen abhängig: So wären, hätte die Untersuchung früher stattgefunden, etwa der *Germanwings*-A320-Absturz in Südfrankreich im März, oder, hätte sie später stattgefunden, der Terroranschlag in Tunesien oder die Legalisierung der gleichgeschlechtlichen Ehe in den USA im Juni, oder das Griechenland-Referendum im Juli 2015 geeignete Themen gewesen. (Optimal wäre es natürlich, wenn man bzgl. des Zeitpunkts der Untersuchung einen Spielraum hätte, um möglichst viele geeignete Themen abzudecken.)

Dementsprechend hat sich auch gezeigt, dass Google und Bing erhebliche Probleme damit haben, bei längerfristigen, mit der Zeit ‚verwaschenen‘ Themen aktuelle und damit relevante Ergebnisse

auszugeben (hier scheint es sich also um einen klassischen Konflikt der Rankingfaktoren-Bereiche Popularität und Aktualität zu handeln (s. Kap. 3.1)). Dies wurde etwa bei den Themen 1 – Einsparmaßnahmen der Deutschen Bank –, 8 – USA-Japan-Verteidigungsabkommen – und 14 – erneute Kämpfe in Mariupol – deutlich (s. Kap. 5.3).

Sozusagen nebenbei traten während der Untersuchung zwei grobe Mängel bzgl. der Ergebnispräsentation zutage: Universal-Search-News-Treffer wurden oft nochmals als organische Ergebnisse ausgegeben, und unter verschiedenen URLs und Titeln tauchte wiederholt der gleiche Inhalt auf (s. Kap. 5.2.2) – was wieder nach eigenen Untersuchungen ruft.

Eine mögliche Erklärung für die Dubletten ergibt sich aus der Einbindung der Universal Search, wie sie in Kap. 3.2 beschrieben wurde: Möglicherweise wurden diese Ergebnisse sowohl von den allgemeinen als auch von den Universal-Search-News-Suchmaschinen gefunden, und da kein gemeinsames Ranking durchgeführt, sondern nur festgelegt wird, wo die relevanten Universal-Search-Ergebnisse auf den SERPs platziert werden, findet womöglich auch kein Dublettenabgleich statt und so kommt es zu doppelten Ausgaben. Das wäre jedoch ein relativ simpler Fehler, und es erscheint seltsam, dass so etwas nach ca. sechs Jahren Universal-Search-Praxis (immer noch (?)) auftritt.

Es wäre interessant, die empirische Untersuchung dieser Arbeit unter Berücksichtigung der verbessernden Anmerkungen in diesem Kapitel zu wiederholen und so zu valideren Ergebnissen zu kommen, denn das Thema ist nach wie vor sehr relevant.

Ebenfalls interessant wäre es, einen Ansatz zur automatisierten Inhaltserschließung von Suchergebnissen zu entwickeln. Dazu gibt es diverse *Text-mining*-, *Text-analysis*- und weitere Tools, oft aus der Wirtschaft. Bekannt sind z.B. *MAXQDA* ([www.maxqda.de/](http://www.maxqda.de/)) oder *Aylien* ([aylien.com/](http://aylien.com/)). Mit Aylien lassen sich bspw. Konzepte und Entitäten aus Texten extrahieren. Aus dem akademischen Forschungskontext klingt bspw. der *DigiDoc Metaeditor* vielversprechend, der HTML-Dokumente durch das Identifizieren und Vorschlagen von Keywords aus einem Thesaurus anhand enthaltener Metainformationen (die ja auch für textspezifische Rankingfaktoren (und damit auch für Suchmaschinenoptimierung) wichtig sind (vgl. Kap. 3.1)) halbautomatisch indexieren soll (vgl. Vázquez et al 2015).

Allerdings ist die automatisierte Inhaltserschließung sehr komplex und technisch, und es hapert immer noch an relativ grundlegenden Dingen, wie z.B. der Zuordnung von Pronomen zu ihren Nomen, sofern diese nicht im selben Satz stehen. Ein weiteres Problem ist, dass diese Tools meist aus dem englischsprachigen Raum stammen und oft kein Deutsch verarbeiten können. Trotzdem wäre die Entwicklung eines automatisierten Textanalyse-Ansatzes eine dankbare Aufgabe, da der manuelle Aufwand doch sehr hoch ist.

Interessant wäre in diesem Zusammenhang auch ein Vergleich zwischen einem manuellen und einem automatisierten Verfahren.

Mit der vorliegenden Untersuchung wurde jedenfalls gerade einmal an der Oberfläche des Themas gekratzt – es gibt im Bereich inhaltlicher Suchmaschinenverzerrungen noch viel zu erforschen.

## **6. Fazit**

In diesem Abschnitt werden die wichtigsten Punkte der Arbeit insbes. der empirischen Untersuchung noch einmal zusammengefasst.

Im ersten Teilbereich der vorliegenden Arbeit, dem theoretischen Hintergrund (Kap. 2), wurde sich Suchmaschinen als Kommunikationsmedien theoretisch über Kommunikation und Medien sowie Massenkommunikation und Massenmedien genähert. Dabei wurden sie auf Grundlage des Internets als technische Infrastruktur als Medien zweiter Ordnung (vgl. Joerges & Braun 1994, S. 19; Kubicek et al. 1997, S. 32ff, 218ff) und weiter als Metamedien verortet (vgl. Merten 2007, S. 321; Beck 2012, S. 172).

Anschließend wurde das Mediensystems Deutschlands betrachtet, das besonders von der Neuordnung durch die Alliierten nach dem Zweiten Weltkrieg und dem Beitritt der neuen Bundesländer nach der Wiedervereinigung 1989 geprägt ist (vgl. Beck 2012, S. 333ff), und das der USA, das einen marktwirtschaftlich-kapitalistischen Prototyp unter den Systemen der westlichen Welt darstellt (vgl. Kleinsteuber 2009, S. 1211). In Hinsicht auf eine rechtsvergleichende Perspektive bzgl. Suchmaschinen überformen die Rechtssetzungsakte der EU vielfach die nationale Rechtsordnung, während die USA bereits eine längere Historie der Auseinandersetzung mit Suchmaschinen-sachverhalten aufweisen (vgl. Paal 2012, S. 62-65).

Ein ausführlicheres Zwischenfazit zum ersten Teilbereich findet sich am Ende von Kap. 2.5.

Im zweiten Teilbereich der Arbeit, Suchmaschinen und dem Stand der Forschung (Kap. 3 und 4), wurde sich Suchmaschinen auf technischerer Ebene genähert. Dabei wurde ihre enorme gesellschaftliche Bedeutung (vgl. Lewandowski 2015, S. 2) aufgezeigt und Kritik an ihnen (z.B. als Gatekeeper) beleuchtet. Es wurde die Funktionsweise von Suchmaschinen anhand ihrer modellhaften Komponenten Crawler, Indexer und Searcher (vgl. Risvik & Michelsen 2002, S. 290) erläutert und kurz auf Suchmaschinenoptimierung und Personalisierung eingegangen, bevor Spezialsuchmaschinen (z.B. Google News) mit ihrer häufigen Einbindung in der Universal Search beleuchtet wurden.

Anschließend wurde sich dem Kernthema Suchmaschinenverzerrungen gewidmet. Dabei wurde gezeigt, dass es keine unverzerrten Suchergebnisse geben kann (vgl. Weber 2011, S. 277; Lewan-

dowski 2015, S. 272), es wurden die Ursachen – Implementierung, Anbieterverhalten, Nutzerverhalten und Kombinationen der drei (vgl. Weber 2011) – erläutert und Gegenstrategien aufgezeigt.

Schließlich wurden kurz neun Arbeiten aus dem Zeitraum 2002 bis 2013 vorgestellt, die sich auf vielfältige Weise mit der Erforschung von Suchmaschinenverzerrungen beschäftigen.

Ein ausführlicheres Zwischenfazit zum zweiten Teilbereich findet sich am Ende von Kap. 4.2.

Im dritten Teilbereich der vorliegenden Arbeit (Kap. 5) schließlich ging es um die empirische Untersuchung. Dabei wurden zunächst die Forschungsfragen und -hypothesen zur Vollständigkeit, Abdeckung und Gewichtung inhaltlicher Aspekte in Suchergebnissen dargelegt.

Anschließend wurde der theoretische Hintergrund der angewandten Methodik erläutert – dieser bestand aus den Methoden Inhaltsanalyse zur Analyse und Ziehung von Rückschlüssen auf fixierte Kommunikation (vgl. Früh 2011, S. 133, Mayring 2010, S. 13), und Retrievaltests zur Messung der Retrievaleffektivität von Information-Retrieval-Systemen, die für Suchmaschinen adaptiert wurde (vgl. Lewandowski 2015, S. 215-221).

Dann wurde die praktische Methodik erläutert, bei der an einem bestimmten Tag aus den drei auflagenstärksten deutschen und amerikanischen Tageszeitungen 14 Themenüberschneidungen identifiziert wurden, zu denen Anfragen formuliert und noch am selben Tag jeweils die Top-10 textlichen Suchergebnisse (organische und Universal-Search-News) von Google und Bing jeweils unter .de- und .com-Domain gescrapet wurden.

Anschließend wurden zu jedem der neun auswertbaren Themen die inhaltlichen Aussagen der Suchergebnisse zu Aspekten gruppiert, die dann wiederum in den Suchergebnissen gesucht und anhand einer 5er-Likert-Skala hinsichtlich ihrer Ausführlichkeit bewertet wurden. Diese Daten wurden schließlich nach Suchmaschinen und Regionen in Hinsicht auf die Forschungsfragen und -hypothesen ausgewertet.

Das Ergebnis ist, dass Google und Bing in Deutschland und den USA keine vollständigen Bilder aktueller internationaler Themen vermitteln, wobei es Unterschiede hinsichtlich der wichtigen Aspekte gibt; dass sie auf den ersten Trefferpositionen nicht die drei wichtigsten inhaltlichen Aspekte abdecken; und dass es bei der Gewichtung der wichtigen inhaltlichen Aspekte keine signifikanten Unterschiede gibt.

Allerdings erfährt die Untersuchung erhebliche Einschränkungen, etwa durch zu allgemein gewählte Suchanfragen (die zu allzu diversen Ergebnissen führten), die geringe Anzahl der (auswertbaren) Themen, Scrapingfehler und die Subjektivität der gesamten Auswertung (es gab nur einen Bearbeiter). Daher haben die Ergebnisse praktisch keine Aussagekraft. Die Untersuchung kann allerdings als Ausgangspunkt für weitere Forschung zu inhaltlichen Suchmaschinenverzerrungen dienen.

## Literaturverzeichnis

### **ADOMAVICIUS & TUZHILIN 2005**

ADOMAVICIUS, GEDIMINAS; Tuzhilin, Alexander: *Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions*. In: *IEEE Transactions on Knowledge and Data Engineering* 17 (2005) 6 –S. 734-749 – Online verfügbar unter:

<http://web.stanford.edu/class/ee378b/papers/adomavicius-recsys.pdf> – Abruf: 2015-05-24

### **AGOF 2015**

AGOF E.V. (Hrsg.): *Internet Facts 2015-01*. Online verfügbar unter:

[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) – Abruf: 2015-04-13

### **BAILEY ET AL. 2007**

BAILEY, Peter; THOMAS, Paul; HAWKING, David: *Does brandname influence perceived search result quality? Yahoo!, Google and WebKumara*. In: *Proceedings of the 12th Australasian Document Computing Symposium*. Melbourne, Australia, December 10, 2007 – Online verfügbar unter:

<http://goanna.cs.rmit.edu.au/~aht/adcs2007/papers/8.pdf> – Abruf: 2015-06-22

### **BAR-ILAN ET AL. 2009**

BAR-ILAN, Judit; KEENOY, Kevin; LEVENE, Mark; YAARI, Eti: *Presentation bias is significant in determining user preference for search results—A user study*. In: *Journal of the American Society for Information Science and Technology* 60 (2009) 1, S. 135-149

### **BECK 2006**

BECK, Klaus: *Computervermittelte Kommunikation im Internet*. München : Oldenbourg, 2006

### **BECK 2012**

BECK, Klaus: *Das Mediensystem Deutschlands : Strukturen, Märkte, Regulierung*. Wiesbaden : Springer VS, 2012

### **BECK 2013**

BECK, Klaus: *Kommunikationswissenschaft*. 3. Aufl. Stuttgart : UVK, 2013

### **BECK 2014**

BECK, Klaus: *Soziologie der Online-Kommunikation*. Wiesbaden : Springer VS, 2014

### **BEILER 2005**

BEILER, Markus: *Selektionsverhalten in den Ergebnislisten von Suchmaschinen : Modellentwicklung und empirische Überprüfung am Beispiel von Google*. In: MACHILL, Marcel; SCHNEIDER, Norbert (Hrsg.): *Suchmaschinen: Neue Herausforderungen für die Medienpolitik*. Berlin : Vistas, 2005 – S. 165-190 –

Online verfügbar unter: [http://eprints.rclis.org/16085/1/Journalistische\\_Aktualitaet.pdf](http://eprints.rclis.org/16085/1/Journalistische_Aktualitaet.pdf) – Abruf: 2015-03-19

### **BEILER 2008**

BEILER, Markus: *Nachrichtensuchmaschinen als journalistisches Rechercheinstrument: Inhaltsanalyse zur Quellen- und Beitragsvielfalt*. In: MACHILL, Marcel; BEILER, Markus; ZENKER, Martin: *Journalistische Recherche im Internet : Bestandsaufnahme journalistischer Arbeitsweisen in Zeitungen, Hörfunk, Fernsehen und Online*. Berlin : Vistas, 2008 – S. 165-189

**BEILER 2013**

BEILER, Markus: *Nachrichtensuche im Internet : Inhaltsanalyse zur journalistischen Qualität von Nachrichtensuchmaschinen*. Konstanz, München : UVK, 2013

**BERELSON 1952**

BERELSON, Bernard: *Content analysis in communications research*. Glencoe, Ill. : Free Press, 1952

**BERGHAUS 1994**

BERGHAUS, Margot: *Multimedia-Zukunft : Herausforderungen für die Medien- und Kommunikationswissenschaft*. In: *Rundfunk und Fernsehen* 42 (1994) 3, S. 404-412

**BERTELSMANN 2014**

BERTELSMANN (Hrsg.): *Bertelsmann übernimmt Gruner + Jahr vollständig*. Stand: 2014-10-06  
<http://www.bertelsmann.de/news-und-media/nachrichten/bertelsmann-uebernimmt-gruner-jahr-vollstaendig.jsp> – Abruf: 2015-07-18

**BRODER 2002**

BRODER, Andrei: *A taxonomy of web search*. In: *SIGIR Forum* 36 (2002) 2, S. 3-10 – Online verfügbar unter: <http://www.cis.upenn.edu/~nenkova/Courses/cis430/p3-broder.pdf> – Abruf: 2015-07-18

**BRODER ET AL. 2000**

BRODER, Andrej; KUMAR, Ravi; MAGHOUL, Farzin; RAGHAVAN, Prabhakar; RAJAGOPALAN, Sridhar; STATA, Raymie; TOMKINS, Andrew; WIENER, Janet: *Graph structure in the web*. In: *Computer Networks* 33 (2000) 1-6, S. 309-320. Online verfügbar unter: <http://www9.org/w9cdrom/160/160.html> – Abruf: 2015-02-22

**BROWN ET AL. 2008**

BROWN, Lauren; DUBOIS, Alain; SHEPARD, Donald: *Inefficiency and Bias of Search Engines in Retrieving References and Containing Scientific Names of Fossil Amphibians*. In: *Bulletin of Science, Technology & Society* 28 (2008) 4, S. 279-288 – Online verfügbar unter: [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) – Abruf: 2015-06-06

**BÜNDNIS 90 / DIE GRÜNEN 2005**

BÜNDNIS 90 / DIE GRÜNEN (Hrsg.): *Suchmaschinen: Das Tor zum Netz : Hintergrund und Information*. – Informationsbroschüre März 2005 – Online verfügbar unter: [www.alle.de/assets/includes/sendtext.cfm?aus=11&key=728](http://www.alle.de/assets/includes/sendtext.cfm?aus=11&key=728) – Abruf: 2015-04-20

**BURGHARDT 1972**

BURGHARDT, Anton: *Einführung in die allgemeine Soziologie*. München : Vahlen, 1972

**BURKART 2002**

BURKART, Roland: *Kommunikationswissenschaft : Grundlagen und Problemfelder : Umriss einer interdisziplinären Sozialwissenschaft*. 4. überarb. und akt. Aufl. Wien Köln : Böhlau, 2002

**BVDW 2013**

BUNDESVERBAND DIGITALE WIRTSCHAFT (BVDW) (Hrsg.): *OVK Online-Report 2013/02 : Zahlen und Trends im Überblick*. Düsseldorf, 2013 – Online verfügbar unter: <http://www.bvdw.org/mybvdw/media/download/ovk-report-2013-02.pdf?file=2950> – Abruf: 2014-04-21

**CERI ET AL. 2013**

CERI, Stefano; BOZZON, Alessandro; BRAMBILLA, Marco; DELLA VALLE, Emanuelle; FRATERNALI, Piero; QUARTERONI, Silvia: *Web Information Retrieval*. Berlin u.a. : Springer, 2013

**CHO & ROY 2004**

CHO, Junghoo; ROY, Sourahis: *Impact of Search Engines on Page Popularity*. In: *Proceedings of the 13th international conference on World Wide Web*. New York : ACM, 2004 – S. 20-29 – Online verfügbar unter: <http://oak.cs.ucla.edu/~cho/papers/cho-bias.pdf> – Abruf: 2015-06-06

**COMPUTERWOCHE 2015**

COMPUTERWOCHE (Hrsg.): *Yahoo sucht einstweilen weiter mit Bing*. Stand: 2015-04-16  
<http://www.computerwoche.de/a/yahoo-sucht-einstweilen-weiter-mit-bing,3097454> – Abruf: 2015-07-10

**CREECH 2013**

CREECH, Kenneth: *Electronic Media Law and Regulation*. 6th Ed. Waltham : Focal Press, 2013

**DESTATIS 2015A**

DESTATIS (Hrsg.): *Deutschland*. – Stand: 2015  
<https://www.destatis.de/DE/ZahlenFakten/LaenderRegionen/Internationales/Land/Europa/Deutschland.html> – Abruf: 2015-02-16

**DESTATIS 2015B**

DESTATIS (Hrsg.): *Vereinigte Staaten von Amerika*. – Stand: 2015  
<https://www.destatis.de/DE/ZahlenFakten/LaenderRegionen/Internationales/Land/Amerika/VereinigeteStaatenvonAmerika.html> – Abruf: 2015-02-16

**DOBAL & WERNER 1997**

DOBAL, Raoul; WERNER, Andreas: *Das World Wide Web aus funktionalistischer Sicht*. In: LUDS, Peter; WERNER, Andreas (Hrsg.): *Multimedia-Kommunikation : Theorien, Trends und Praxis*. Opladen : Westdeutscher Verlag, 1997 – S. 105-122

**DREIER 2009**

DREIER, Hardy: *Das Mediensystem Deutschlands*. In: HANS BREDOW INSTITUT (Hrsg.): *Internationales Handbuch Medien*. [ehem. *Internationales Handbuch für Hörfunk und Fernsehen*] 28. Aufl. Hamburg : Hans Bredow Verlag, 2009 – S. 257-272

**EDELMAN 2010**

EDELMAN, Benjamin: *Hard-Coding Bias in Google "Algorithmic" Search Results*. Stand: 2010-11-15  
<http://www.benedelman.org/hardcoding/> - Abruf: 2015-05-04

**EDELMAN & LOCKWOOD 2011**

EDELMAN, Benjamin; LOCKWOOD, Benjamin: *Measuring Bias in "Organic" Web Search*. Stand: 2011-01-19  
<http://www.benedelman.org/searchbias/> - Abruf: 2015-05-04

**ELEKTRONIK KOMPENDIUM 2015**

ELEKTRONIK KOMPENDIUM (Hrsg.): *ISO/OSI-7-Schichtenmodell*. – Stand: 2015 <http://www.elektronik-kompodium.de/sites/kom/0301201.htm> – Abruf: 2015-04-13

**ELLIANCE 2007**

ELLIANCE INC. (Hrsg.): *Google Universal Search*. – Stand: 2007  
<http://www.elliance.com/aha/infographics/google-universal-search.aspx> – Abruf: 2015-03-15

**ERLHOFER 2013**

ERLHOFER, Sebastian: *Suchmaschinen-Optimierung : Das umfassende Handbuch*. 6. akt. und erw. Aufl. Bonn : Galileo Press, 2013

**EU 2010**

EUROPÄISCHE UNION (EU) (Hrsg.): *Kartellrecht: Die Kommission untersucht mögliche Kartellrechtsverstöße durch Google*. Stand: 2010-11-30 [http://europa.eu/rapid/press-release\\_IP-10-1624\\_de.pdf](http://europa.eu/rapid/press-release_IP-10-1624_de.pdf) – Abruf: 2015-05-04

**EU 2015**

EUROPÄISCHE UNION (EU) (Hrsg.): *Kommission übermittelt Google Mitteilung der Beschwerdepunkte zu seinem Preisvergleichsdienst und leitet außerdem förmliche Untersuchung zu Android ein*. Stand: 2015-04-15 [http://europa.eu/rapid/press-release\\_IP-15-4780\\_de.pdf](http://europa.eu/rapid/press-release_IP-15-4780_de.pdf) – Abruf: 2015-05-04

**FAULSTICH 2004A**

FAULSTICH, Werner: *Medienwissenschaft*. Paderborn : Fink, 2004

**FAULSTICH 2004B**

FAULSTICH, Werner: *Medium*. In: FAULSTICH, Werner (Hrsg.) 2004: *Grundwissen Medien*. 5. Aufl. Paderborn : Fink, 2004 - S. 13-102

**FCC 2011**

FEDERAL COMMUNICATIONS COMMISSION (FCC) (Hrsg.): *Second Report and Order*. – Stand: 2011-07-15 [https://apps.fcc.gov/edocs\\_public/attachmatch/FCC-11-110A1.pdf](https://apps.fcc.gov/edocs_public/attachmatch/FCC-11-110A1.pdf) – Abruf: 2015-02-16

**FRÜH 2011**

FRÜH, Werner: *Inhaltsanalyse*. 7. Aufl. München : Ölschläger, 2011

**FSM 2015**

FREIWILLIGE SELBSTKONTROLLE MULTIMEDIA-DIENSTANBIETER E.V. (FSM) (Hrsg.): *Selbstverpflichtung der Wirtschaft : Suchmaschinen*. Stand: 2015 <https://www.fsm.de/selbstverpflichtungen/suchmaschinen> – Abruf: 2015-05-04

**GOOGLE 2009**

GOOGLE INC. (Hrsg.): *Personalized Search for everyone*. Stand: 2009-12-04 <http://googleblog.blogspot.de/2009/12/personalized-search-for-everyone.html> – Abruf: 2015-05-18

**GOOGLE 2015A**

GOOGLE INC. (Hrsg.): *Algorithms*. Stand: 2015 <http://www.google.com/intl/en/insidesearch/howsearchworks/algorithms.html> – Abruf: 2015-03-09

**GOOGLE 2015B**

GOOGLE INC. (Hrsg.): *Alles über Google News*. Stand: 2015 [http://news.google.com/intl/de\\_de/about\\_google\\_news.html](http://news.google.com/intl/de_de/about_google_news.html) – Abruf: 2015-03-14

**GOOGLE 2015C**

GOOGLE INC. (Hrsg.): *Richtlinien für Webmaster*. Stand: 2015 <https://support.google.com/webmasters/answer/35769?hl=de> – Abruf: 2015-05-05

**GORDON & 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), S. 141-180 – Online verfügbar unter: [http://www.jasonmorrison.net/iakm/cited/Gordon\\_Pathak.pdf](http://www.jasonmorrison.net/iakm/cited/Gordon_Pathak.pdf) - Abruf: 2015-06-22

**GRANKA 2010**

GRANKA, Laura: *The Politics of Search: A Decade Retrospective*. In: *The Information Society* 26 (2010), S. 364-374 – Online verfügbar unter: <http://www.australianscience.com.au/research/google/36914.pdf> – Abruf: 2015-05-17

**GRAUMANN 1972**

GRAUMANN, Carl: *Interaktion und Kommunikation*. In: GRAUMANN, Carl: *Handbuch der Psychologie : 7. Band 7: Sozialpsychologie : 2. Halbband*. Göttingen : Verlag für Psychologie, 1972 – S. 1109-1262

**GRIESBAUM ET AL. 2009**

GRIESBAUM, Joachim; BEKAVAC, Bernard; RITTBERGER, Marc: *Typologie der Suchdienste im Internet*. In: LEWANDOWSKI, Dirk (Hrsg.): *Handbuch Internet-Suchmaschinen : Nutzerorientierung in Wissenschaft und Praxis*. Heidelberg : AKA, 2009 – S. 18-52 – Online verfügbar unter: [http://eprints.rclis.org/12711/1/Typologie\\_der\\_Suchdienste\\_im\\_Internet.pdf](http://eprints.rclis.org/12711/1/Typologie_der_Suchdienste_im_Internet.pdf) – Abruf: 2015-03-19

**GÜNTHER 1963**

GÜNTHER, Gotthard: *Das Bewußtsein der Maschinen*. Krefeld : Agis, 1963

**HARMAN & VOORHEES 2006**

HARMAN, Donna; VOORHEES, Ellen: *TREC: An overview*. In: *Annual Review of Information Science and Technology* 40 (2006), S. 113-155

**HAWKING ET AL. 2001**

HAWKING, David; CRASWELL, Nick; BAILEY, Peter; GRIFFITHS, Kathy: *Measuring Search Engine Quality*. In: *Information Retrieval* 4 (2001), S. 33-59 – Online verfügbar unter: [http://es.csiro.au/pubs/hawking\\_ir01.pdf](http://es.csiro.au/pubs/hawking_ir01.pdf) - Abruf: 2015-06-22

**HINMAN 2005**

HINMAN, Lawrence: *Esse est indicato in Google: Ethical and Political Issues in Search Engines*. In: *International Review of Information Ethics* 6 (2005) 3, S. 19-25 – Online verfügbar unter: [http://www.i-r-i-e.net/inhalt/003/003\\_hinman.pdf](http://www.i-r-i-e.net/inhalt/003/003_hinman.pdf) – Abruf: 2015-07-18

**HÖFLICH 2003**

HÖFLICH, Joachim: *Mensch, Computer und Kommunikation : Theoretische Verortungen und empirische Befunde*. Frankfurt a.M. u.a. : Lang, 2003

**HÖLIG & HASEBRINK 2014**

HÖLIG, Sascha; HASEBRINK, Uwe: *Nachrichtennutzung im Wandel: Neue Plattformen, Endgeräte und Zugänge . International vergleichende Befunde auf Basis des Reuter Institute Digital News Survey 2014*. In: *Media Perspektiven* 11 (2014), S. 530-538 – Online verfügbar unter: [http://www.media-perspektiven.de/fileadmin/user\\_upload/media-perspektiven/pdf/2014/11-2014\\_Hoelig\\_Hasebrink.pdf](http://www.media-perspektiven.de/fileadmin/user_upload/media-perspektiven/pdf/2014/11-2014_Hoelig_Hasebrink.pdf) – Abruf: 2015-07-26

**HORSTER 1999**

HORSTER, Detlef: *Jürgen Habermas zur Einführung*. Hamburg : Junius, 1999

**IEONG ET AL. 2012**

IEONG, Samuel; MISHRA, Nina; SADIKOV, Eldar; ZHANG, Li: *Domain Bias in Web Search*. In: *Proceedings of the fifth ACM international conference on Web search and data mining* New York : ACM, 2012 - S. 413-422 – Online verfügbar unter: <http://research.microsoft.com/pubs/155941/domainbias.pdf> – Abruf: 2015-06-14

**IFM 2014**

INSTITUT FÜR MEDIEN- UND KOMMUNIKATIONSPOLITIK (IFM) (Hrsg.): *Ranking - Die 50 größten Medienkonzerne 2014*. – Stand: 2014-06-12 <http://www.mediadb.eu/rankings/intl-medienkonzerne-2014.html> – Abruf: 2015-02-14

**INTRONA & NISSENBAUM 2000**

INTRONA, Lucas, NISSENBAUM, Helen: *Shaping the Web: Why the Politics of Search Engines Matters*. In: *The Information Society* 16 (2000), S. 169-185 – Online verfügbar unter: <http://www.nyu.edu/projects/nissenbaum/papers/ShapingTheWeb.pdf> – Abruf: 2015-05-18

**JANSEN ET AL. 2007A**

JANSEN, Jim; ZHANG, Mimi; ZHANG, Ying: *Brand awareness and the evaluation of search results*. In: *Proceedings of the 16th international conference on World Wide Web, Banff, Alberta, Canada*. New York : ACM, 2007 - S. 1139-1149 – Online verfügbar unter: <http://www2007.org/posters/poster969.pdf> – Abruf: 2015-07-18

**JANSEN ET AL. 2007B**

JANSEN, Jim; ZHANG, Mimi; ZHANG, Ying: *The effect of brand awareness on the evaluation of search engine results*. In: *Conference on Human Factors in Computing Systems – Proceedings* New York : ACM, 2007 - S. 2471-2476

**JANSSEN & LAATZ 2013**

JANSSEN, Jürgen; LAATZ, Wilfried: *Statistische Datenanalyse mit SPSS : Eine anwendungsorientierte Einführung in das Basissystem und das Modul Exakte Tests*. 8. Aufl. Berlin Heidelberg : Springer Gabler, 2013

**JARREN & BONFADELLI 2001**

JARREN, Ottfried; BONFADELLI, Heinz (Hrsg.): *Einführung in die Publizistikwissenschaft*. Bern u.a. : Haupt, 2001

**JOACHIMS ET AL. 2007**

JOACHIMS, Thorsten; GRANKA, Laura; PAN, Bing; HEMBROOKE, Helene; RADLINSKI, Filip; GY, Geri: *Evaluating the Accuracy of Implicit Feedback from Clicks and Query Reformulations in Web Search*. In: *ACM Transactions on Information Systems* 25 (2007) 2, Article No. 7 – Online verfügbar unter: [http://www.cs.cornell.edu/people/tj/publications/joachims\\_etal\\_07a.pdf](http://www.cs.cornell.edu/people/tj/publications/joachims_etal_07a.pdf) – Abruf: 2015-06-13

**JOERGES & BRAUN 1994**

JOERGES, Bernward; BRAUN, Ingo: *Große technische Systeme – erzählt, gedeutet, modelliert*. In: BRAUN, Ingo; JOERGES, Bernward (Hrsg.): *Technik ohne Grenzen*. Frankfurt a.M. : Suhrkamp, 1994 – S. 7-49

**JÜRGENS ET AL. 2014**

JÜRGENS, Pascal; STARK, Birgit; MAGIN, Melanie: *Gefangen in der Filter Bubble? : Search Engine Bias und Personalisierungsprozesse bei Suchmaschinen*. In: STARK, Birgit; DÖRR, Dieter; AUFENANGER, Stefan (Hrsg.): *Die Googleisierung der Informationssuche : Suchmaschinen zwischen Nutzung und Regulierung*. Berlin u.a. : de Gruyter, 2014 – S. 98-135

**KELLER 2012**

KELLER, Daniela: *Parametrisch oder nichtparametrisch? Das ist hier die Frage*. Stand: 2012-09-06  
<http://www.statistik-und-beratung.de/2012/09/parametrisch-oder-nichtparametrisch-das-ist-hier-die-frage/> - Abruf: 2015-08-13

**KLEINSTEUBER 2007A**

KLEINSTEUBER, Hans: *Massenmedien und öffentliche Meinung*. In: JÄGER, Wolfgang; HAAS, Christoph; WELZ, Wolfgang (Hrsg.): *Das Regierungssystem der USA : Lehr- und Handbuch*. 3. Aufl. München : Oldenbourg, 2007 – S. 377-394

**KLEINSTEUBER 2007B**

KLEINSTEUBER, Hans: *Nordamerika*. In: THOMAS, Barbara (Hrsg.): *Mediensysteme im internationalen Vergleich*. 2. überarb. Aufl. Konstanz : UTB, 2013 [Beitrag von 2007] – S. 257-270

**KLEINSTEUBER 2009**

KLEINSTEUBER, Hans: *Das Mediensystem der USA*. In: HANS BREDOW INSTITUT (Hrsg.): *Internationales Handbuch Medien*. [ehem. *Internationales Handbuch für Hörfunk und Fernsehen*] 28. Aufl. Hamburg : Hans Bredow Institut, 2009 – S. 1211-1226

**KLEINSTEUBER & KUTZ 2008**

KLEINSTEUBER, Hans; KUTZ, Magnus-Sebastian: *Vorreiter auf neuem Kurs? : Die Zukunft der elektronischen Medien in den USA*. Berlin : Friedrich-Ebert-Stiftung, 2008

**KNEER & NASSEHI 1993**

KNEER, Georg; NASSEHI, Armin: *Luhmanns Theorie sozialer Systeme : Eine Einführung*. München : Fink, 1993

**KRIEGER 1998**

KRIEGER, David: *Einführung in die allgemeine Systemtheorie*. 2. unv. Aufl. München : Fink, 1998

**KRIPPENDORFF 1994**

KRIPPENDORFF, Klaus: *Der verschwundene Bote*. In: MERTEN, Klaus; SCHMIDT, Siegfried; WEISCHENBERG, Siegfried (Hrsg.): *Die Wirklichkeit der Medien : Eine Einführung in die Kommunikationswissenschaft*. Opladen : Westdeutscher Verlag, 1994 - S. 79-113

**KUBICEK ET AL. 1997**

KUBICEK, Herbert; SCHMID, Ulrich; WAGNER, Heiderose: *Bürgerinformation durch „neue“ Medien : Analysen und Fallstudien zur Etablierung elektronischer Informationssysteme im Alltag*. Opladen : Westdeutscher Verlag, 1997

**LANG & BEKAVAC 2004**

LANG, Norbert; BEKAVAC, Bernard: *World Wide Web*. In: FAULSTICH; Werner (Hrsg.): *Grundwissen Medien*. 5. Aufl. Berlin : Fink, 2004 - S. 433-453

**LAURIN 1997**

LAURIN, Stefan: *Das Internet ist kein Massenmedium : Niklas Luhmann über Medien, Journalisten und Wahrheit*. In: *Unicum* (1997) 2, S. 20 – Online verfügbar unter: <http://www.ruhrbarone.de/niklas-luhmann-%E2%80%9Edas-internet-ist-kein-massenmedium%E2%80%9C/1109> – Abruf: 2015-04-13

**LAZARFELD ET AL. 1944**

LAZARFELD, Paul; BERELSON, Bernard; GAUDET, Hazel: *The People's Choice. How the Voter Makes Up his Mind in a Presidential Campaign*. New York : Duell, Sloan & Pearce, 1944

**LEWANDOWSKI 2005**

LEWANDOWSKI, Dirk: *Web Information Retrieval : Technologien zur Informationssuche im Internet*. Frankfurt a.M. : DGI, 2005 – Online verfügbar unter: [http://www.bui.haw-hamburg.de/fileadmin/user\\_upload/lewandowski/doc/Web\\_Information\\_Retrieval\\_Buch.pdf](http://www.bui.haw-hamburg.de/fileadmin/user_upload/lewandowski/doc/Web_Information_Retrieval_Buch.pdf) – Abruf: 2015-04-21

**LEWANDOWSKI 2008**

LEWANDOWSKI, Dirk: *Spezielsuchmaschinen*. In: LEWANDOWSKI, Dirk (Hrsg.): *Handbuch Internet-Suchmaschinen : Nutzerorientierung in Wissenschaft und Praxis*. Heidelberg : AKA, 2008 - S. 53–69 – Online verfügbar unter: [http://searchstudies.org/tl\\_files/Publikationen\\_PDFs/2008/spezielsuchmaschinen.pdf](http://searchstudies.org/tl_files/Publikationen_PDFs/2008/spezielsuchmaschinen.pdf) – Abruf: 2015-07-18

**LEWANDOWSKI 2011**

LEWANDOWSKI, Dirk: *Evaluierung von Suchmaschinen*. In: LEWANDOWSKI, Dirk (Hrsg.): *Handbuch Internet-Suchmaschinen 2 : Neue Entwicklungen in der Web-Suche*. Heidelberg : AKA, 2011 – S. 203–228 – Online verfügbar unter: [http://www.searchstudies.org/tl\\_files/Publikationen\\_PDFs/2011/Evaluierung%20von%20Suchmaschinen%20282011%29.pdf](http://www.searchstudies.org/tl_files/Publikationen_PDFs/2011/Evaluierung%20von%20Suchmaschinen%20282011%29.pdf) – Abruf: 2015-06-22

**LEWANDOWSKI 2013**

LEWANDOWSKI, Dirk: *Suchmaschinen*. In: KUHLEN, Rainer; SEMAR, Wolfgang; STRAUCH, Dietmar (Hrsg.): *Grundlagen der praktischen Information und Dokumentation*. 6. Ausg. Berlin : de Gruyter - S. 495–508 – Online verfügbar unter: [http://searchstudies.org/tl\\_files/Publikationen\\_PDFs/2013/GdPluD\\_D1.pdf](http://searchstudies.org/tl_files/Publikationen_PDFs/2013/GdPluD_D1.pdf) – Abruf: 2015-07-18

**LEWANDOWSKI 2014**

LEWANDOWSKI, Dirk: *Die Macht der Suchmaschinen und ihr Einfluss auf unsere Entscheidungen*. In: *Information – Wissenschaft & Praxis* 65 (2014) 4-5, S. 231-238

**LEWANDOWSKI 2015**

LEWANDOWSKI, Dirk: *Suchmaschinen verstehen*. Berlin Heidelberg : Springer Vieweg, 2015

**LEWANDOWSKI & SÜNKLER 2012**

LEWANDOWSKI, Dirk; SÜNKLER, Sebastian: *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, 2012 – S. 237-249 – Online verfügbar unter: [http://searchstudies.org/tl\\_files/Publikationen\\_PDFs/2012/RAT\\_DGI\\_Lewandowski\\_Suenkler\\_preprint.pdf](http://searchstudies.org/tl_files/Publikationen_PDFs/2012/RAT_DGI_Lewandowski_Suenkler_preprint.pdf) [Preprint] – Abruf: 2015-07-10

**LEWIN 1947**

LEWIN, Kurt: *Channels of Group Life, Social Planning and Action Research*. In: *Human Relations* 1 (1947), S. 143-153 – Online verfügbar unter: <http://www.sagepub.com/mertler2study/articles/Lewin.pdf> – Abruf: 2015-04-26

**LfM 2005**

LANDESANSTALT FÜR MEDIEN NORDRHEIN-WESTFALEN (LFM) (Hrsg.): *Suchmaschinen als „Gatekeeper“ in der öffentlichen Kommunikation*. [Pressemitteilung] – Stand: 2005-06-09 <https://www.lfm-nrw.de/aktuell/pressemitteilungen/pressemitteilungen-detail/article/suchmaschinen-als-gatekeeper-in-der-oeffentlichen-kommunikation.html> – Abruf: 2015-04-20

**LUHMANN 1970**

LUHMANN, Niklas: *Soziologische Aufklärung : 1. Aufsätze zur Theorie sozialer Systeme*. Opladen : Westdeutscher Verlag, 1970

**LUHMANN 1971**

LUHMANN, Niklas: *Sinn als Grundbegriff der Soziologie*. In: HABERMAS, Jürgen; LUHMANN, Niklas (Hrsg.): *Theorie der Gesellschaft oder Sozialtechnologie – Was leistet die Systemforschung?*. Frankfurt a.M. : Suhrkamp, 1971 - S. 25-100

**LUHMANN 1990**

LUHMANN, Niklas: *Die Wissenschaft der Gesellschaft*. Frankfurt a.M. : Suhrkamp, 1990

**LUHMANN 1996**

LUHMANN, Niklas: *Die Realität der Massenmedien*. Opladen : Westdeutscher Verlag, 1996

**LUHMANN 2002**

LUHMANN, Niklas: *Soziale Systeme : Grundriß einer allgemeinen Theorie*. Lizenzausgabe 2002 für die Wissenschaftliche Buchgesellschaft. Frankfurt a.M. : Suhrkamp, 2002

**LUHMANN 2004**

LUHMANN, Niklas: *Die Realität der Massenmedien*. 3. Aufl. Opladen : Westdeutscher Verlag, 2004

**MACHILL & BEILER 2008**

MACHILL, Marcel; BEILER, Markus: *Suchmaschinen als Vertrauensgüter : Internet-Gatekeeper für die Informationsgesellschaft?*. In: KLUMPP, Dieter; KUBICEK, Herbert; ROßNAGEL, Alexander; SCHULZ, Wolfgang (Hrsg.): *Informationelles Vertrauen für die Informationsgesellschaft*. Berlin u.a. : Springer, 2008 – S. 159-172

**MACHILL ET AL. 2003**

MACHILL, Marcel; NEUBERGER, Christoph; SCHWEIGER, Wolfgang; WIRTH, Werner: *Wegweiser im Netz : Qualität und Nutzung von Suchmaschinen*. In: MACHILL, Marcel; WELP, Carsten (Hrsg.): *Wegweiser im Netz : Qualität und Nutzung von Suchmaschinen*. Gütersloh : Bertelsmann Stiftung, 2003 - S. 13-490 – Online verfügbar unter:  
[http://www.academia.edu/2580970/Machill\\_Marcel\\_Neuberger\\_Christoph\\_Schweiger\\_Wolfgang\\_Wirth\\_Werner\\_2003\\_Wegweiser\\_im\\_Netz\\_Qualit%C3%A4t\\_und\\_Nutzung\\_von\\_Suchmaschinen](http://www.academia.edu/2580970/Machill_Marcel_Neuberger_Christoph_Schweiger_Wolfgang_Wirth_Werner_2003_Wegweiser_im_Netz_Qualit%C3%A4t_und_Nutzung_von_Suchmaschinen). In: [Machill Marcel Welp Carsten Hrsg. Wegweiser im Netz. Qualität und Nutzung von Suchmaschinen. Gütersloh Bertelsmann Stiftung S. 17-490](http://www.academia.edu/2580970/Machill_Marcel_Neuberger_Christoph_Schweiger_Wolfgang_Wirth_Werner_2003_Wegweiser_im_Netz_Qualit%C3%A4t_und_Nutzung_von_Suchmaschinen) – Abruf: 2015-07-18

**MACHILL ET AL. 2005**

MACHILL, Marcel; LEWANDOWSKI, Dirk; KARZAUNINKAT, Stefan: *Journalistische Aktualität im Internet : Ein Experiment mit den „News-Suchfunktionen“ von Suchmaschinen*. In: MACHILL, Marcel; SCHNEIDER, Norbert (Hrsg.): *Suchmaschinen : Neue Herausforderungen für die Medienpolitik*. Berlin : Vistas, 2005 – S. 105-164 – Online verfügbar unter:  
[http://eprints.rclis.org/16085/1/Journalistische\\_Aktualitaet.pdf](http://eprints.rclis.org/16085/1/Journalistische_Aktualitaet.pdf) – Abruf: 2015-03-19

**MALETZKE 1963**

MALETZKE, Gerhard: *Psychologie der Massenkommunikation : Theorie und Systematik*. Hamburg : Hans Bredow-Institut, 1963

**MATURANA & VARELA 1982**

MATURANA, Humberto; VARELA, Francisco: *Autopoietische Systeme: eine Bestimmung der lebendigen Organisation*. In: MATURANA, Humberto (Hrsg.): *Erkennen. Die Organisation und Verkörperung von Wirklichkeit*. Braunschweig/Wiesbaden : Springer Vieweg, 1982 - S. 170-235

**MAYRING 2010**

MAYRING, Philipp: *Qualitative Inhaltsanalyse : Grundlagen und Techniken*. 11. akt. und überarb. Aufl. Weinheim : Beltz, 2010

**MCLUHAN 1997**

MCLUHAN, Marshall: *Understanding Media : The Extensions of Man*. Fifth printing. New York : Signet Books, 1997

**MEAD 1934**

MEAD, George: *Mind, Self, and Society*. Chicago : University of Chicago Press, 1934 – Online verfügbar unter: <http://morephilosophystuff.pbworks.com/f/Mead-George-Herbert-Mind-Self-And-Society.pdf> – Abruf: 2015-07-18

**MERTEN 1977**

MERTEN, Klaus: *Kommunikation : Eine Begriffs- und Prozeßanalyse*. Opladen : Westdeutscher Verlag, 1977

**MERTEN 2006**

MERTEN, Klaus: *Kommunikationstheorie*. In: TSVASMAN, Leon (Hrsg.): *Das große Lexikon Medien und Kommunikation : Kompendium interdisziplinärer Konzepte*. Würzburg : Ergon, 2006 – S. 194-206

**MERTEN 2007**

MERTEN, Klaus: *Einführung in die Kommunikationswissenschaft : Band 1: Grundlagen der Kommunikationswissenschaft* 3. Aufl. Münster u.a. : Lit, 2007

**MIZZARO 1997**

MIZZARO, Stefano: *Relevance : The whole history*. In: *Journal of the American Society for Information Science*, 48 (1997) 9, S. 810-832 . Online verfügbar unter: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.50.8058&rep=rep1&type=pdf> – Abruf: 2015-07-18

**MORRIS & OGAN 1996**

MORRIS, Merrill; OGAN, Christine: *The Internet as Mass Medium*. In: *Journal of Communication* 46 (1996) 1, S. 39-50

**MOWSHOWITZ & KAWAGUCHI 2002A**

MOWSHOWITZ, Abbe; KAWAGUCHI, Akira: *Assessing bias in search engines*. In: *Information Processing and Management* 38 (2002) 1, S. 141-156 – Online verfügbar unter: [http://www.researchgate.net/profile/Akira\\_Kawaguchi2/publication/220229065\\_Assessing\\_bias\\_in\\_search\\_engines/links/0deec524cab3264220000000.pdf](http://www.researchgate.net/profile/Akira_Kawaguchi2/publication/220229065_Assessing_bias_in_search_engines/links/0deec524cab3264220000000.pdf) – Abruf: 2015-05-18

**MOWSHOWITZ & KAWAGUCHI 2002B**

MOWSHOWITZ, Abbe; KAWAGUCHI, Akira: *Bias on the web*. In: *Communications of the ACM* 45 (2002) 9, S. 56-60 – Online verfügbar unter: [https://homepages.fhv.at/se/ss2003/im01/acm\\_comm\\_0209\\_p56-mowshowitz.pdf](https://homepages.fhv.at/se/ss2003/im01/acm_comm_0209_p56-mowshowitz.pdf) – Abruf: 2015-05-18

**MUSGRAVE 1957**

MUSGRAVE, Richard: *A Multiple Theory of Budget Determination*. In: *Finanzarchiv* 17 (1956/1957) 3, S. 333-343 – online verfügbar unter: [http://www.mohr.de/fileadmin/user\\_upload/Zeitschriften/musgrave\\_budget\\_determination.pdf](http://www.mohr.de/fileadmin/user_upload/Zeitschriften/musgrave_budget_determination.pdf) – Abruf: 2015-01-09

**NAPOLI 2011**

NAPOLI, Philip: *Audience Evolution : New Technologies and the Transformation of Media Audiences*. New York : Columbia University Press, 2011

**NEUBERGER 2001**

NEUBERGER, Christoph: *Journalismus im Internet : Theoriekontext und empirische Exploration*. Habilitationsschrift, Sprach- und Literaturwissenschaftliche Fakultät der Katholischen Universität Eichstatt, 2001

**NEUBERGER 2005**

NEUBERGER, Christoph: *Funktionen, Probleme und Regulierung von Suchmaschinen im Internet*. In: *International Review of Information Ethics* 6 (2005) 3, S. 3-17 – Online verfügbar unter: [http://www.i-r-i-e.net/inhalt/003/003\\_neuberger.pdf](http://www.i-r-i-e.net/inhalt/003/003_neuberger.pdf) – Abruf: 2015-04-20

**NEUBERGER & TONNEMACHER 1999**

NEUBERGER, Christoph; TONNEMACHER, Jan (Hrsg.): *Online – die Zukunft der Zeitung? : Das Engagement deutscher Tageszeitungen im Internet*. Opladen : Westdeutscher Verlag, 1999

**NEUMAHR 2000**

NEUMAHR, Andreas: *Leistungsbewertung von Kommunikationsnetzen*. [Diplomarbeit], 2000 – Online verfügbar unter: <http://neumahr.de/study/diplomarbeit.pdf> – Abruf: 2015-04-13

**PAAL 2012**

PAAL, Boris: *Suchmaschinen, Marktmacht und Meinungsbildung*. Baden-Baden : Nomos, 2012 – Online verfügbar unter: <http://www.i-comp.org/wp-content/uploads/2013/07/PAAL-Suchmaschinen-Marktmacht-und-Meinungsbildung-2012-final.pdf> – Abruf: 2015-07-19

**PARISER 2012**

PARISER, Eli: *Filter Bubble : Wie wir im Internet entmündigt werden*. München : Hanser, 2012

**Pross 1972**

Pross, Harry: *Medienforschung : Film, Funk, Presse, Fernsehen*. Darmstadt : Habel, 1972

**PÜRER 2003**

PÜRER, Heinz: *Publizistik- und Kommunikationswissenschaft : Ein Lehrbuch*. Konstanz : UVK, 2003

**PURCELL ET AL. 2012**

PURCELL, Kristen; BRENNER, Joanna; RAINIE, Lee: *Search Engine Use 2012*. Online verfügbar unter: [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) – Abruf: 2015-04-13

**QUIRMBACH 2009**

QUIRMBACH, Sonja: *Universal Search : Kontextuelle Einbindung von Ergebnissen unterschiedlicher Quellen und Auswirkungen auf das User Interface*. In: LEWANDOWSKI, Dirk (Hrsg.): *Handbuch Internet-Suchmaschinen*. Heidelberg : AKA, 2009 – S. 220-248. – Online verfügbar unter: [http://eprints.rclis.org/12741/1/universal\\_search.pdf](http://eprints.rclis.org/12741/1/universal_search.pdf) – Abruf: 2015-03-15

**RADLINSKI & JOACHIMS 2006**

RADLINSKI, Filip; JOACHIMS, Thorsten: *Active exploration for learning rankings from clickthrough data*. In: *ACM International Conference On Knowledge Discovery and Data Mining (KDD), 2007* – Online verfügbar unter: [http://www.cs.cornell.edu/people/tj/publications/radlinski\\_joachims\\_07a.pdf](http://www.cs.cornell.edu/people/tj/publications/radlinski_joachims_07a.pdf) – Abruf: 2015-06-13

**RAU 2013**

RAU, Harald: *Einladung zur Kommunikationswissenschaft*. Baden-Baden : Nomos, 2013

**RIEMER & BRÜGGEMANN 2009**

RIEMER, Kai; BRÜGGEMANN, Fabian: *Personalisierung der Internetsuche – Lösungstechniken und Marktüberblick*. In: LEWANDOWSKI, Dirk: *Handbuch Internet-Suchmaschinen : Nutzerorientierung in Wissenschaft und Praxis*. Heidelberg : AKA, 2009 – S. 148-171 – Online verfügbar unter: [http://www.researchgate.net/profile/Kai\\_Riemer/publication/28809880\\_Personalisierung\\_der\\_Internetsuche\\_Lsungstechniken\\_und\\_Marktberblick/links/00b7d52a9402b67006000000.pdf](http://www.researchgate.net/profile/Kai_Riemer/publication/28809880_Personalisierung_der_Internetsuche_Lsungstechniken_und_Marktberblick/links/00b7d52a9402b67006000000.pdf) – Abruf: 2015-05-18

**RISVIK & MICHELSEN 2002**

RISVIK, Knut; MICHELSEN, Rolf: *Search engines and Web dynamics*. In: *Computer Networks* 39 (2002) 3, S. 289-302 – Online verfügbar unter: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.16.8435&rep=rep1&type=pdf> – Abruf: 2015-07-18

**RÖHLE 2010**

RÖHLE, Theo: *Der Google-Komplex : Über Macht im Zeitalter des Internets*. Bielefeld : transcript, 2010 – Online verfügbar unter: [http://www.netzmedium.de/uploads/Roehle\\_Google-Komplex\\_2010.pdf](http://www.netzmedium.de/uploads/Roehle_Google-Komplex_2010.pdf) – Abruf: 2015-04-21

**RÖSSLER 1998**

RÖSSLER, Patrick: *Wirkungsmodelle: die digitale Herausforderung : Überlegungen zu einer Inventur bestehender Erklärungsansätze der Medienwirkungsforschung*. In: RÖSSLER, Patrick (Hrsg.): *Online-Kommunikation : Beiträge zur Nutzung und Wirkung*. Opladen : Westdeutscher Verlag, 1998 – S. 17-46

**RONNEBERGER 1971**

RONNEBERGER, Franz: *Sozialisation durch Massenkommunikation*. In: RONNEBERGER, Franz (Hrsg.): *Sozialisation durch Massenkommunikation*. Stuttgart : Enke, 1971 – S. 32-101

**RONNEBERGER 1974**

RONNEBERGER, Franz: *Die politischen Funktionen der Massenkommunikation*. In: LANGENBUCHER, Wolfgang (Hrsg.): *Zur Theorie der politischen Kommunikation*. München : Piper, 1974 – S. 193-205

**RUNKEHL ET AL. 1998**

RUNKEHL, Jens; SCHLOBINSKI, Peter; SIEVER, Torsten: *Sprache und Kommunikation im Internet : Überblick und Analysen*. Opladen/Wiesbaden : Westdeutscher Verlag, 1998 – Online verfügbar unter: <http://www.mediensprache.net/archiv/pubs/3-531-13267-9.pdf> – Abruf: 2015-04-11

**RUßMANN 2007**

RUßMANN, Uta: *Agenda Setting und Internet : Themensetzung im Spannungsfeld von Onlinemedien und sozialen Netzwerken*. München : Fischer, 2007

**SARACEVIC 2007A**

SARACEVIC, Tefko: *Relevance : A review of the literature and a framework for thinking on the notion in information science : Part II : Nature and manifestations of relevance*. In: *Journal of the American Society for Information Science and Technology* 58 (2007) 13, S. 1915-1933

**SARACEVIC 2007B**

SARACEVIC, Tefko: *Relevance : A review of the literature and a framework for thinking on the notion in information science : Part III : Behavior and effects of relevance*. In: *Journal of the American Society for Information Science and Technology* 58 (2007) 13, S. 2126-2144

**SAXER 1974**

SAXER, Ulrich: *Funktionen der Massenmedien in der modernen Gesellschaft*. In: KURZROCK, Ruprecht (Hrsg.): *Medienforschung*. Berlin : Colloquium, 1974 – S. 22-33

**SAXER 1980**

SAXER, Ulrich: *Grenzen der Publizistikwissenschaft : Wissenschaftswissenschaftliche Reflexionen zur Zeitungs-/Publizistik-/Kommunikationswissenschaft seit 1945*. In: *Publizistik* 4 (1980), S. 525-543

**SAXER 1998**

SAXER, Ulrich: *Mediengesellschaft : Verständnisse und Mißverständnisse*. In: SARCINELLI, Ulrich (Hrsg.): *Politikvermittlung und Demokratie in der Mediengesellschaft : Beiträge zur politischen Kommunikationskultur*. Opladen/Wiesbaden : Westdeutscher Verlag, 1998 - S. 52-73

**SCHAER ET AL. 2010**

SCHAER, Philipp; MAYR, Philipp; MUTSCHKE, Peter: *Implications of Inter-Rater Agreement on a Student Information Retrieval Evaluation*. LWA 2010, Kassel – Online verfügbar unter: <http://arxiv.org/ftp/arxiv/papers/1010/1010.1824.pdf> – Abruf: 2015-06-22

**SCHMITT-WALTER 2004**

SCHMITT-WALTER, Nikolaus: *Online-Medien als funktionale Alternative? Über die Konkurrenz zwischen den Mediengattungen*. München : Fischer, 2004

**SCHULZ 2007**

SCHULZ, Winfried: *Politische Kommunikation : Theoretische Ansätze und Ergebnisse empirischer Forschung zur Rolle der Massenmedien in der Politik*. Opladen : Westdeutscher Verlag, 2007

**SCHULZ 2009**

SCHULZ, Winfried: *Kommunikationsprozess*. In: NOELLE-NEUMANN, Elisabeth; WILKE, Jürgen; SCHULZ, Winfried (Hrsg.): *Fischer Lexikon Publizistik*. Akt., vollst. überarb. und erg. Aufl. Frankfurt a.M. : Fischer, 2009 - S. 169-200

**SHERMAN & PRICE 2001**

SHERMAN, Chris; PRICE, Gary: *The invisible web : Uncovering information sources search engines can't see*. Medford, NJ : Information Today, 2001 – Online verfügbar unter: [https://www.ideals.illinois.edu/bitstream/handle/2142/8528/librarytrendsv52i2h\\_opt.pdf](https://www.ideals.illinois.edu/bitstream/handle/2142/8528/librarytrendsv52i2h_opt.pdf) – Abruf: 2015-03-19

**SPIEGEL ONLINE 2013**

SPIEGEL ONLINE (Hrsg.): *News Corp: Rupert Murdoch verkauft 33 Lokalzeitungen*. – Stand: 2013-09-04 <http://www.spiegel.de/kultur/gesellschaft/rupert-murdoch-verlagsgruppe-news-corp-verkauft-33-lokalzeitungen-a-920373.html> – Abruf: 2015-02-13

**STARK 2014**

STARK, Birgit: „Don't be evil“. In: STARK, Birgit; DÖRR, Dieter; AUFENANGER, Stefan (Hrsg.): *Die Googleisierung der Informationssuche : Suchmaschinen zwischen Nutzung und Regulierung*. Berlin u.a. : de Gruyter, 2014 – S. 1-19

**STARK ET AL. 2014**

STARK, Birgit; MAGIN, Melanie; JÜRGENS, Pascal: *Navigieren im Netz : Befunde einer qualitativen und quantitativen Nutzerbefragung*. In: STARK, Birgit; DÖRR, Dieter; AUFENANGER, Stefan (Hrsg.): *Die Googleisierung der Informationssuche : Suchmaschinen zwischen Nutzung und Regulierung*. Berlin : de Gruyter, 2014 – S. 20-74

**STARKULLA 1963**

STARKULLA, Heinz: *Publizistik und Kommunikation*. In: *Publizistik* 8 (1963) 5, S. 562-571

**STATISTA 2015A**

STATISTA (Hrsg.): *Höhe der Werbeumsätze von Google von 2001 bis 2014 (in Milliarden US-Dollar)*. – Stand: 2015 <http://de.statista.com/statistik/daten/studie/75188/umfrage/werbeumsatz-von-google-seit-2001/> - Abruf: 2015-02-16

**STATISTA 2015B**

STATISTA (Hrsg.): *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/> - Abruf: 2015-02-13

**STATISTA 2015C**

STATISTA (Hrsg.): *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/> - Abruf: 2015-07-10

**STATISTA 2015D**

STATISTA (Hrsg.): *Suchmaschinenverteilung in Deutschland im Jahr 2015 (sowie Vorjahresvergleich)*. Stand: 2015 <http://de.statista.com/statistik/daten/studie/167841/umfrage/marktanteile-ausgewaehlter-suchmaschinen-in-deutschland/> - Abruf: 2015-07-10

**STATISTA 2015E**

STATISTA (Hrsg.): *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/> - Abruf: 2015-07-10

**STOCK 2007**

STOCK, Wolfgang: *Information Retrieval : Informationen suchen und finden*. München : Oldenbourg, 2007

**SULLIVAN 2003**

SULLIVAN, Danny: *Searching with Invisible Tabs*. – Stand: 2003-12-01 <http://searchenginewatch.com/article/2064036/Searching-With-Invisible-Tabs> - Abruf: 2015-03-14

**TAGESSPIEGEL 2013**

TAGESSPIEGEL (Hrsg.): *"Reader's Digest" ist pleite*. – Stand: 2013-02-19 <http://www.tagesspiegel.de/medien/verlag-readers-digest-ist-pleite/7802934.html> – Abruf: 2015-02-13

**TAGUE-SUTCLIFFE 1992**

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

**THOMAS 2013**

THOMAS, Barbara: *Westeuropa*. In: THOMAS, Barbara (Hrsg.): *Mediensysteme im internationalen Vergleich*. 2. überarb. Aufl. Konstanz : UTB, 2013 - S. 220-238

**TNS INFRATEST 2013**

TNS INFRATEST (Hrsg.): *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) – Abruf: 2015-07-26

**TURNSTALL 2008**

TURNSTALL, Jeremy: *The media were American : U.S. mass media in decline*. Oxford : Oxford University Press, 2008

**UYAR 2009A**

UYAR, Ahmet: *Google stemming mechanisms*. In: *Journal of Information Science* 35 (2009) 5, S. 499-514

**UYAR 2009B**

UYAR, Ahmet: *Investigation of the accuracy of search engine hit counts*. In: *Journal of information Science* 35 (2009) 4, S. 469-480

**UZH 2010A**

UNIVERSITÄT ZÜRICH (UZH) (Hrsg.): *Datenanalyse*. Stand: 2010-12-02

<http://www.methodenberatung.uzh.ch/datenanalyse.html> – Abruf: 2015-07-23

**UZH 2010B**

UNIVERSITÄT ZÜRICH (UZH) (Hrsg.): *Der Kruskal-Wallis-Test*. Stand: 2010-12-02

<http://www.methodenberatung.uzh.ch/datenanalyse/unterschiede/zentral/kruskal.html> – Abruf: 2015-07-25

**UZH 2010c**

UNIVERSITÄT ZÜRICH (UZH) (Hrsg.): *Mann-Whitney-U-Test*. Stand: 2010-12-02

<http://www.methodenberatung.uzh.ch/datenanalyse/unterschiede/zentral/mann.html> – Abruf: 2015-07-25

**UZH 2014A**

UNIVERSITÄT ZÜRICH (UZH) (Hrsg.): *Unterschiede*. Stand: 2014-08-11

<http://www.methodenberatung.uzh.ch/datenanalyse/unterschiede.html> – Abruf: 2015-07-23

**UZH 2014B**

UNIVERSITÄT ZÜRICH (UZH) (Hrsg.): *Zentrale Tendenz*. Stand: 2014-08-11

<http://www.methodenberatung.uzh.ch/datenanalyse/unterschiede/zentral.html> – Abruf: 2015-07-23

**VÁLLEZ ET AL. 2015**

VÁLLEZ, Mari; PEDRAZA-JIMÉNEZ, Rafael; CODINA, Lluís; VLANCO, Saúl; ROVIRA, Cristòfol: *A semi-automatic indexing system based on embedded information in HTML documents*. In: *Library Hi Tech* 33 (2015) 2, S. 195-210 – Online verfügbar unter: <http://eprints.rclis.org/25382/1/Preprint-Semi-automaticIndexingSystemBasedEmbeddedInformationHTMLDocuments.pdf> – Abruf: 2015-07-27

**VAN EIMEREN & FREES 2014**

VAN EIMEREN, Birgit; FREES, Beate: *Ergebnisse der ARD/ZDF-Online-Studie : 79 Prozent der Deutschen online – Zuwachs bei mobiler Internetnutzung und Bewegtbild*. In: *Media Perspektiven* 45 (2014) 7-8, S. 378-396. Online verfügbar unter: [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) – Abruf: 2015-04-13

**VAN EIMEREN ET AL. 2003**

VAN EIMEREN; Birgit; GERHARD, Heinz; FREES, Beate: *Internetverbreitung in Deutschland: Unerwartet hoher Zuwachs : ARD/ZDF-Online-Studie 2003*. In: *Media Perspektiven* 8 (2003), S. 338-358 – Online verfügbar unter: [http://www.media-perspektiven.de/fileadmin/user\\_upload/media-perspektiven/pdf/2003/08-2003\\_Eimeren\\_verbessert\\_neu.pdf](http://www.media-perspektiven.de/fileadmin/user_upload/media-perspektiven/pdf/2003/08-2003_Eimeren_verbessert_neu.pdf) – Abruf: 2015-07-26

**VAUGHAN & THELWALL 2004**

VAUGHAN, Liwen; THELWALL, Mike: *Search Engine Coverage Bias: Evidence and Possible Causes*. In: *Information Processing & Management* 40 (2004) 4, S. 693-707 – Online verfügbar unter: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.65.5130&rep=rep1&type=pdf> – Abruf: 2015-06-06

**WEBER 2011**

WEBER, Karsten: *Search Engine Bias*. In: LEWANDOWSKI, Dirk (Hrsg.): *Handbuch Internet-Suchmaschinen 2 : Neue Entwicklungen in der Web-Suche*. Berlin : AKA, 2011 – S. 265-285 – Online verfügbar unter: [http://www.researchgate.net/publication/256473910\\_Search\\_Engine\\_Bias](http://www.researchgate.net/publication/256473910_Search_Engine_Bias) – Abruf: 2015-05-04

**WEBER & DRÜEKE 2010**

WEBER, Karsten; DRÜEKE, Ricarda: *Konvergente Medien, Fragmentierung der Öffentlichkeit und medienethische Anmerkungen*. In: *Zeitschrift für Kommunikationsökologie und Medienethik* 12 (2010) 1, S. 68-72

**WEBSTER & KSIAZEK 2012**

WEBSTER, James; KSIAZEK, Thomas: *The Dynamics of Audience Fragmentation: Public Attention in an Age of Digital Media*. In: *Journal of Communication* 62 (2012) 1 – S. 39-56 – Online verfügbar unter: <http://www.looooker.com/wp-content/uploads/2013/05/j.1460-2466.2011.01616.x2.pdf> – Abruf: 2015-05-24

**WEINREICH 1998**

WEINREICH, Frank: *Nutzen- und Belohnungsstrukturen computergestützter Kommunikationsformen : Zur Anwendung des Uses und Gratifications Approach in einem neuen Forschungsfeld*. In: *Publizistik* 43 (1998) 1 – S. 130-142

**WEISCHENBERG 2004**

WEISCHENBERG, Siegfried: *Journalistik. Medienkommunikation: Theorie und Praxis : Band 1: Mediensysteme – Medienethik – Medieninstitutionen*. 3. Aufl. Wiesbaden : VS, 2004

**WHITE 1950**

WHITE, David: *The „Gate Keeper“: A Case Study in the Selection of News*. In: *Journalism Quarterly* 27 (1950), S. 383-390 – Online verfügbar unter: <http://www.aejmc.org/home/wp-content/uploads/2012/09/Journalism-Quarterly-1950-White-383-90.pdf> – Abruf: 2015-04-26

**WHITE 2013**

WHITE, Ryan: *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, 2013 – S. 3-12 – Online verfügbar unter: <http://research.microsoft.com/en-us/um/people/ryenw/papers/WhiteSIGIR2013.pdf> – Abruf: 2015-06-13

**WHITE & HORVITZ 2015**

WHITE, Ryan; HORVITZ, Eric: *Belief Dynamics and Biases in Web Search*. In: *ACM Transactions on Information Systems (ACM TOIS)* [In Press] – Online verfügbar unter: <http://research.microsoft.com/en-us/um/people/ryenw/papers/WhiteTOIS2015.pdf> – Abruf: 2015-06-14

**WILDENMANN & KALTEFLEITER 1965**

WILDENMANN, Rudolf; KALTEFLEITER, Werner: *Funktionen der Massenmedien*. Frankfurt a.M. : Athenäum, 1965

**WILKE 2009**

WILKE, Jürgen: *Multimedia/Online-Medien*. In: NOELLE-NEUMANN, Elisabeth; WILKE, Jürgen; SCHULZ, Winfried (Hrsg.): *Fischer Lexikon Publizistik*. Akt., vollst. überarb. und erg. Aufl. Frankfurt a.M. : Fischer, 2009 - S. 329-358

**WINKLER 1997**

WINKLER, Hartmut: *Suchmaschinen : Metamedien im Internet?* In: BECKER, Barbara; PAETAU, Michael (Hrsg.): *Virtualisierung des Sozialen : Die Informationsgesellschaft zwischen Fragmentierung und Globalisierung*. Frankfurt a.M. u.a. : Campus, 1997 - S. 217-236 – Online verfügbar unter: [http://homepages.uni-paderborn.de/winkler/suchm\\_d.html](http://homepages.uni-paderborn.de/winkler/suchm_d.html) – Abruf: 2015-04-21

**WIRTH & SCHWEIGER 1999**

WIRTH, Werner; SCHWEIGER, Wolfgang: *Selektion neu betrachtet: Auswahlentscheidungen im Internet*. In: WIRTH, Werner; SCHWEIGER, Wolfgang (Hrsg.): *Selektion im Internet : Empirische Analysen zu einem Schlüsselkonzept*. Opladen : Westdeutscher Verlag, 1999 – S. 43-70

**WIRTH ET AL. 2007**

WIRTH, Werner; BÖCKING, Tabea; KARNOWSKI, Veronika; VON PAPE, Thilo: *Heuristic and Systematic Use of Search Engines*. In: *Journal of Computer Mediated Communication* 12 (2007) 3, S. 778-800 – Online verfügbar unter: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.453.4665&rep=rep1&type=pdf> – Abruf: 2015-05-08

**YUE ET AL. 2010**

YUE, Yisong; PATEL, Rajan; ROEHRIG, Hein: *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 : Association for Computing Machinery, 2010 – S. 1011-1018 – Online verfügbar unter: [http://www.yisongyue.com/publications/www2010\\_bias.pdf](http://www.yisongyue.com/publications/www2010_bias.pdf) – Abruf: 2015-06-13

## **Anhang**

### **A Beigabe: Inhalt der CD**

Masterarbeit als PDF-Version

Inhaltliche Aspekte

Auswertungen

Scraping\Suchergebnisse

Scraping\Suchergebnisse lokal [nach Suchmaschinen – s. „Readme“-Datei]

Scraping\Suchergebnisseiten

## **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, 24. August 2015

---