

Stichprobenziehung für Online-Inhaltsanalysen: Suchmaschinen und Filter Bubbles

Emmer, Martin; Strippel, Christian

Erstveröffentlichung / Primary Publication

Sammelwerksbeitrag / collection article

Empfohlene Zitierung / Suggested Citation:

Emmer, Martin ; Strippel, Christian: Stichprobenziehung für Online-Inhaltsanalysen: Suchmaschinen und Filter Bubbles. In: Maireder, Axel (Ed.) ; Ausserhofer, Julian (Ed.) ; Schumann, Christina (Ed.) ; Taddicken, Monika (Ed.): *Digitale Methoden in der Kommunikationswissenschaft*. Berlin, 2015 (Digital Communication Research 2). - ISBN 978-3-945681-02-2, pp. 275-300. URN: <http://dx.doi.org/10.17174/dcr.v2.12>

Nutzungsbedingungen:

Dieser Text wird unter einer CC BY Lizenz (Namensnennung) zur Verfügung gestellt. Nähere Auskünfte zu den CC-Lizenzen finden Sie hier:
<https://creativecommons.org/licenses/by/4.0/deed.de>

Terms of use:

This document is made available under a CC BY Licence (Attribution). For more information see:
<https://creativecommons.org/licenses/by/4.0>

Empfohlene Zitierung: Emmer, M., & Strippel, C. (2015). Stichprobenziehung für Online-Inhaltsanalysen: Suchmaschinen und Filter Bubbles. In A. Maireder, J. Ausserhofer, C. Schumann, & M. Taddicken (Hrsg.), *Digitale Methoden in der Kommunikationswissenschaft* (S. 275-300). doi: 10.17174/dcr.v2.12

Zusammenfassung: Stichproben für Inhaltsanalysen im Internet sind aufgrund der spezifischen Eigenschaften von Online-Inhalten nur schwer zu realisieren. Häufig greift die Forschung auf Suchmaschinen zurück, um mit ihrer Hilfe öffentliche Kommunikation im Internet zu erfassen. Dabei ergeben sich jedoch verschiedene Probleme, die einerseits mit einem noch stark massenmedial geprägten Medienbegriff und andererseits mit den Algorithmen der eingesetzten Suchmaschinen zu tun haben. In einer Studie anhand verschiedener Suchanfragen bei Google wurde untersucht, welche Folgen ihr Einsatz für Online-Stichprobenziehungen haben kann. Die Ergebnisse zeigen, dass Suchmaschinen erstens eine Öffentlichkeit konstruieren, die weit über die herkömmliche massenmediale Öffentlichkeit hinausgeht und zweitens zu erheblichen Unterschieden je nach suchender Person führt. Als Schlussfolgerung sollten in Zukunft verschiedene Nutzungspraktiken berücksichtigt und Analysen auf alle auffindbaren anstatt nur auf journalistische Angebote bezogen werden.

Lizenz: Creative Commons Attribution 4.0 (CC-BY 4.0)

Martin Emmer & Christian Strippel

Stichprobenziehung für Online-Inhaltsanalysen: Suchmaschinen und Filter Bubbles

1 Einleitung

Die Qualität der Empirie in der quantitativen Sozialforschung wird in ganz erheblichem Maße von dem Auswahlverfahren beeinflusst, mit dem die zugrunde liegenden Daten erhoben werden. Kann von den Befunden in einer Stichprobe nicht mit Sicherheit auf die jeweils angestrebte Grundgesamtheit geschlossen werden, so ist die Aussagekraft einer Studie – egal wie gut begründet das theoretische Modell und wie elaboriert die Auswertungsmethoden sein mögen – gering. Deshalb haben sich in der empirischen Sozialforschung Standardmethoden entwickelt, die sicherstellen sollen, dass solche Schlussfolgerungen möglich sind (Gehrau, Fretwurst, Krause, & Daschmann, 2005).

Zufallsstichproben bieten dabei grundsätzlich die größte Sicherheit hinsichtlich der Repräsentativität einer analysierten Stichprobe für die definierte Grundgesamtheit. Allerdings sind reine, einfache Zufallsstichproben nur selten eine Option, denn meist fehlen dafür die Auswahlgrundlagen: Streng genommen ist eine vollständige Liste der Elemente einer Grundgesamtheit erforderlich, damit alle diese Elemente die gleiche Auswahlchance haben und die Voraussetzungen für eine möglichst strukturtreue Abbildung der Grundgesamtheit gegeben sind. Deshalb wird häufig auf mehrstufige Verfahren zurückgegriffen, bei denen sich

schrittweise einer solchen Auswahl angenähert wird – bei Bevölkerungsstichproben etwa zuerst über Regionen, dann über Haushalte und dann über die Haushaltsmitglieder (Arbeitsgemeinschaft ADM-Stichproben & Bureau Wendt, 1994; Häder & Gabler, 1998).

Bei Inhaltsanalysen wird in der Regel nur teilweise auf Zufallsauswahlen gesetzt: Da auf bestimmten Auswahlstufen nur wenige Fälle zur Verfügung stehen – etwa bei der Art des Mediums (Fernsehen, Radio, Tageszeitung etc.) oder den vorhandenen Medienangeboten (überregionale Abonnementzeitungen, Kaufzeitungen) – und eine Zufallsauswahl dabei mit einem hohen Verzerrungsrisiko verbunden ist, wird hier häufig zu bewussten, theoretisch begründeten Auswahlen gegriffen, die dann erst auf weiteren Auswahlstufen mit Zufallsverfahren kombiniert werden (Kops, 1977). Ein Beispiel dafür ist die „Künstliche Woche“, mit deren Hilfe einzelne Ausgaben und zu analysierende Beiträge eines Mediums selektiert werden (Rössler, 2010, S. 53).

Durch die Digitalisierung der Kommunikationswelt ergeben sich im Bereich der Stichprobenziehung nun sowohl neue Möglichkeiten als auch vielfältige Herausforderungen: Während die Ziehung von Personenstichproben im Internet für online-gestützte Befragungen ein bisher ungelöstes Problem darstellt (Hauptmanns, 1999; Hauptmanns & Lande, 2001; Welker, Werner, & Scholz, 2005, S. 31-62; Starsetzki, 2007), erlaubt die digitalisierte Archivierung journalistischer Medienangebote die systematische Berücksichtigung größerer Datengrundlagen und auch stärker zufallsbasierte Auswahlen. So können etwa über Stichwortsuchen aus umfassenden Mediendatenbanken Auswahlgrundlagen erstellt werden, auf die dann eine einstufige Zufallsauswahl von Artikeln angewendet wird.

Allerdings erweitert das Internet nicht nur die Recherche- und Analysemöglichkeiten in Archiven klassischer journalistischer Medien, sondern es stellt auch eine erhebliche Erweiterung von Öffentlichkeit dar: Insbesondere über Social Media beteiligen sich an der öffentlichen bzw. öffentlich zugänglichen Kommunikation nun neue Akteure, die bisher kaum sichtbar waren und einem inhaltsanalytischen Zugang deshalb verschlossen blieben (Schmidt, 2011; Volpers, 2013). Folglich müssen sie auch bei Inhaltsanalysen öffentlicher Kommunikation zumindest theoretisch einbezogen werden, was jedoch ein gewichtiges Folgeproblem aufwirft: Denn während klassische, journalistische Medien in überschaubarer und auch bekannter Form und Zahl vorliegen, ist die Menge alternativer,

online-gestützter Angebote weder in ihrer Zahl, noch in ihrem (digitalen) Erscheinungsort, noch bezüglich ihrer Relevanz für die Rezipienten öffentlicher Kommunikationsangebote wirklich bekannt. Damit fehlt jede Grundlage für eine systematische Auswahl.

Bereits früh wurde deshalb in der sozialwissenschaftlichen Forschung auf Suchmaschinen zurückgegriffen, mit der Begründung, dass sie für den Zugang zu Online-Inhalten zentral seien und Inhalte, die mit ihrer Hilfe nicht aufgefunden werden, wenig öffentliche Relevanz hätten (Wolling & Kuhlmann, 2003, S. 151). Dieses Vorgehen kann heute jedoch insofern kritisiert werden, als dass diese Suchmaschinen zunehmend nutzerbasierte Filter-Algorithmen einsetzen, um in der Auswahl der angezeigten Inhalte den vermeintlichen Interessen der jeweiligen Nutzerinnen und Nutzer möglichst nahe zu kommen (siehe Jürgens, Stark, & Magin in diesem Band). Eine reliable Stichprobenziehung wird auf diese Weise erheblich erschwert.

In diesem Beitrag wird nun auf Basis einer empirischen Untersuchung gezeigt, welche konkreten Folgen diese Filter in beispielhaften Internetnutzungssituationen haben und was dies für die Ziehung von Inhaltsstichproben im Internet bedeutet. So wird danach gefragt, wie sich die aktuellen Filteralgorithmen auf die präsentierten Angebote an Online-Quellen zu einem bestimmten Thema auswirken. Anschließend wird diskutiert, welche Schlussfolgerungen sich daraus ziehen lassen, um trotz dieser Probleme sinnvolle Auswahlen im Internet treffen zu können.

2 Stichproben für Inhaltsanalysen im Internet

Bei der Ziehung von Inhaltsstichproben im Internet besteht das grundsätzliche Problem darin, dass sich in vielen Fällen die Grundgesamtheit höchstens theoretisch, nicht aber praktisch als Auswahlgesamtheit beschreiben lässt. Das hängt im Wesentlichen mit dem bereits diskutierten Umstand zusammen, dass die „Öffentlichkeit“ der Online-Kommunikation nicht mehr an traditionelle massenmediale Publikationsformen geknüpft ist: Insbesondere nutzergenerierte Inhalte auf privaten Webseiten, in sozialen Netzwerken, Blogs, Foren oder Wikis können durchaus als relevante Beiträge zur öffentlich zugänglichen Online-Kommunikation aufgefasst werden. Durch exponentielles Wachstum und stetige Veränderung

verleihen sie dieser Öffentlichkeit jedoch eine große Dynamik, sodass sie nur noch schwer valide abgebildet werden kann (Weare & Lin, 2000; Rössler & Wirth, 2001; Seibold, 2002).

Suchmaschinenbasierte Auswahl

In der gängigen Forschung wird deshalb die angebotsseitige Stichprobenziehung häufig durch eine nutzerseitige ersetzt oder ergänzt, indem zum Beispiel auf Online-Suchmaschinen zurückgegriffen und das Suchverhalten der Nutzerinnen und Nutzer simuliert wird (Galliker & Männel, 1999; McMillan 2000; Erlhofer, 2007; Gerhards & Schäfer, 2007; Schweiger & Weihermüller, 2008). Doch auch Suchmaschinen stehen vor dem Problem der exponentiell wachsenden und sich dynamisch verändernden Online-Inhalte. Sie begegnen diesem Problem, indem sie das Internet mit Hilfe algorithmischer Suchsoftware („Bot“, „Robot“, „Spider“) durchkämmen, die wie in einem Schneeballverfahren den Verlinkungen bereits gefundener Websites folgt („Crawling“). Dann indizieren sie die so gesammelten Inhalte und stellen diesen Index schließlich ihren Nutzerinnen und Nutzern für eine Stichwortsuche zur Verfügung (Wolling, 2002, S. 17-18; Griesbaum, 2003, S. 5-7; Hartmann, Näf, & Schäuble, 2000, S. 64-71; Hildebrandt, Schumann, & Wolling in diesem Band). Die Such-, Indizierungs- und Darstellungsalgorithmen, die bei diesem Prozess zum Einsatz kommen, folgen spezifischen Selektions- und Gewichtungskriterien, die die Suchergebnisse für die Nutzer nach Relevanz vorsortieren, damit jedoch auch nur einen kleinen, spezifischen und durch das abschließende Ranking auch stark verzerrten Ausschnitt des Internets zeigen (Cothey, 2004; Gulli & Signorini, 2005; Hartmann et al., 2000, S. 34-36, 72-74; Ke, Deng, Ng, & Lee, 2005; Machill, Neuberger, & Schindler, 2002, S. 7-22, 29-40).

Es handelt sich bei Online-Suchmaschinen also keineswegs um neutrale Informationsvermittler, sondern vielmehr um algorithmische „Selektiermaschinen“ und „Gatekeeper“ (Wolling, 2002; 2005; Griesbaum, 2003; Loosen, 1999; 2001; Röhle, 2009). Gleichwohl sind sie für eine Online-Stichprobenziehung insofern geeignet, als mit ihrer Hilfe ermittelt werden kann, auf welche Online-Inhalte die Nutzerinnen und Nutzer stoßen, wenn sie im Internet nach bestimmten Schlagworten suchen (Welker & Wünsch, 2010, S. 510; Wolling & Kuhlmann, 2003, S. 150). Entsprechend bemühte sich die Forschung in den vergangenen Jahren verstärkt darum, Suchma-

schinen für eine solche nutzerzentrierte Stichprobenziehung nutzbar zu machen (Anagnostopoulos, Broder, & Carmel, 2006; Bar-Yossef & Gurevich, 2008).

Google-Suche und Personalisierung

Die Simulation des Nutzerverhaltens ist für eine Stichprobenziehung allerdings nur so lange tragfähig, wie die Suchergebnisse auch eine hohe Reliabilität zwischen den suchenden Nutzerinnen und Nutzern aufweisen. Insbesondere die Google-Suche, die aufgrund ihrer nahezu konkurrenzlosen Dominanz auf dem Suchmaschinenmarkt (Webhits, 2014; Netmarketshare, 2013) für die meisten suchmaschinenbasierten Stichprobenziehungen herangezogen werden dürfte, bringt ein spezifisches Problem mit sich, das die obige Argumentation für eine nutzerseitige Stichprobenziehung mittels prominenter Suchmaschinen in ihrem Kern aushebelt: Denn da Google – wie andere Suchmaschinen auch – seine Suchergebnisse zunehmend personalisiert, verliert dieses Verfahren der nutzerseitigen Stichprobenziehung an Reliabilität.

Den ersten Schritt hin zur Personalisierung seiner Suchergebnisse unternahm Google bereits im Juni 2005. Im Google-Blog wurde damals angekündigt: „With the launch of Personalized Search, you can use that search history you’ve been building to get better results. You probably won’t notice much difference at first, but as your search history grows, your personalized results will gradually improve.“ (Google, 2005) Personalisiert wurden zunächst also nur die Suchergebnisse angemeldeter Nutzer, indem die Einträge der individuellen „Web History“ des Nutzers bei jeder neuen Suche mit berücksichtigt wurden. Diese Web History enthält dabei alle bisherigen Google-Suchanfragen des Nutzers, die daraufhin angezeigten Ergebnisse (inklusive „privater Ergebnisse“ von Google+, Gmail und Google Kalender), die Interaktion mit eingeblendeten Anzeigen (einschließlich den auf diesen Seiten ggf. erfolgten Transaktionen), die IP-Adresse des Nutzers, die Spracheinstellung seines Browsers sowie alle seine Aktivitäten auf Google Maps (Google, 2015).

Diese Personalisierung weitete Google dann im Dezember 2009 mit Hilfe der Browser-Cookies auch auf alle nicht-angemeldeten Nutzerinnen und Nutzer aus. Unter dem Titel „Personalized Search for everyone“ schrieb Google: „What we’re doing today is expanding Personalized Search so that we can provide it to signed-out users as well. This addition enables us to customize search results for you based upon 180 days of search activity linked to an anonymous cookie in your

browser“ (Google, 2009b). Google bot seinen Nutzern allerdings die Möglichkeit, über den Link „View Customizations“ die Personalisierung der Ergebnisse nachzuvollziehen und sie ggf. zu deaktivieren.

Die vorerst letzte Ausweitung der Personalisierungsfunktion bei Google erfolgte schließlich unter dem Stichwort „Social Search“, worunter allgemein die Verknüpfung von Online-Suche und Sozialen Netzwerken verstanden wird (Mislove, Gummadi, & Druschel, 2006). Seit Oktober 2009 integriert Google schrittweise öffentlich zugängliche Inhalte (Blogeinträge, Bilder, Links etc.) von Personen, mit denen man in diversen Netzwerken (Facebook, Twitter, YouTube, Flickr etc.) vernetzt ist (Google, 2009a; 2010; 2011) oder die ähnliche Interessen haben (Google, 2012), in seinen Darstellungsalgorithmus.

Auf diese Weise werden nicht mehr nur angebotsbezogene Relevanzkriterien berücksichtigt (wie z.B. die Anzahl der Verlinkungen auf eine Seite), sondern auch personenbezogene Daten wie demografische Charakteristika oder die jeweiligen Nutzungsgewohnheiten. Der für den so genannten PageRank bei Google verantwortliche Algorithmus umfasst mittlerweile zahlreiche Parameter, deren genaue Anzahl, Kombination und Gewichtung als wichtigstes Kapital von Google nicht bekannt sind (Pasquinelli, 2009). Es muss jedoch davon ausgegangen werden, dass sich vor diesem Hintergrund die Suchergebnisse verschiedener Nutzerinnen und Nutzer zunehmend und in erheblichem Umfang unterscheiden.

Selective Exposure und Filter Bubble

Nun ist selbst die Nutzung klassischer Medienangebote nicht gänzlich frei von individueller Selektion auf Basis persönlicher Präferenzen: Der Selective-Exposure-Ansatz geht davon aus, dass Rezipienten sich grundsätzlich Inhalten auf Basis ihrer Voreinstellungen zuwenden und bevorzugt solche Inhalte nutzen, die mit ihren Einstellungen und Vorstellungen konsonant sind (Zillmann & Bryant, 1985). So kann die Nutzung und Wahrnehmung von Inhalten ein und derselben Zeitung oder Fernsehsendung je nach individuellen Präferenzen sehr unterschiedlich ausfallen. Unter den neuen digitalen Kommunikationsbedingungen muss man nun jedoch annehmen, dass diese präferenzbasierten Abweichungen durch personalisierte Algorithmen noch verstärkt werden (Bozdog, 2013; Stalder & Mayer, 2009). Die Folge könnten so genannte „Filter Bubbles“ sein, personalisierte Sphären, in denen jeder Nutzerin und jedem Nutzer ein an die individu-

ellen Präferenzen angepasstes Angebot konstruiert wird. Eli Pariser (2011), der diesen Begriff der „Filterblase“ eingeführt hat, beschreibt das damit verbundene Problem wie folgt:

„The new generation of Internet filters looks at the things you seem to like – the actual things you’ve done, or the things people like you like – and tries to extrapolate. They are prediction engines, constantly creating and refining a theory of who we are and what you’ll do and want next. Together, these engines create a unique universe of information for each of us [...] which fundamentally alters the way we encounter ideas and information.“ (S. 9)

Verstärkt werden können solche Effekte noch dadurch, dass nicht nur für jeden Suchenden selbst eine individuelle „Suchergebnis-Realität“ konstruiert wird, sondern dass Suchmaschinen auch von anderen Akteuren – wie etwa Journalistinnen und Journalisten – als Recherche-Instrument genutzt werden (Springer & Wolling, 2008). Über diesen Umweg können die jeweiligen Verzerrungen auch Eingang in die Angebote anderer Medien finden.

3 Forschungsfragen und methodisches Vorgehen

Für die Untersuchung, welchen konkreten Einfluss die verschiedenen Algorithmen von Google auf die Zusammenstellung der präsentierten Angebote haben und was dies für die Ziehung von Online-Inhaltsstichproben bedeutet, waren die zwei folgenden Forschungsfragen leitend:

FF 1: Welche medialen Angebotsformen werden von der personalisierten Google-Suche grundsätzlich zugänglich gemacht?

FF 2: Wie stark unterscheiden sich die dargebotenen Suchergebnisse unterschiedlicher Nutzerinnen und Nutzer bei gleichlautenden Suchbegriffen?

Hintergrund der ersten Frage ist die eingangs erwähnte Annahme, dass im Internet das Spektrum der Stimmen, die sich zu gesellschaftlich relevanten Themen öffentlich zu Wort melden, breiter sei als in der Offline-Welt: Während dort die öffentliche Debatte durch die Inhalte professionell-journalistischer Mas-

senmedien repräsentiert wird, ermöglicht das Internet auch anderen Akteuren, öffentlich sichtbar zu kommunizieren. Dies können erstens neue *journalistisch-mediale Akteure* sein, die nicht mit herkömmlichen Medienorganisationen verbunden sind. Zweitens eröffnet das Internet aber auch anderen gesellschaftlichen Akteuren wie *Unternehmen, Behörden, politische Parteien oder NGOs* die Möglichkeit, an journalistischen Gatekeepern vorbei eigene Inhalte in öffentliche Debatten einzubringen und mit einzelnen Nutzerinnen und Nutzern direkt zu kommunizieren. Drittens existieren mit Foren und Blogs auch spezifische Angebotsformen, die es selbst *schwach organisierten Gruppen und Einzelpersonen* ermöglichen, sichtbare Beiträge zu öffentlichen Debatten zu leisten. Eine Sonderform dieser individuellen Kommunikationsbeiträge sind viertens die *Nutzerkommentare*, die zwar von Einzelpersonen geschrieben werden, jedoch nicht unabhängig von den jeweiligen zum Beispiel journalistischen Online-Angeboten sind. In der vorliegenden Untersuchung wurden alle Fundstellen in diese vier Kategorien unterschieden, um zu prüfen, welche Angebotsformen in welchem Verhältnis und mit welchen nutzerabhängigen Unterschieden Teil der Suchergebnisdarstellung von Google sind.

Die zweite Frage bezieht sich auf die Annahme, dass Google seine Ergebnislisten personalisiert und es so zu individuellen Filterblasen kommt, die die Öffentlichkeit im Internet in viele persönliche Online-Öffentlichkeiten fragmentiert (Schmidt, 2011, S. 107-116). Dabei ist nicht nur von Interesse, wie stark die Suchergebnisse der einzelnen Nutzerinnen und Nutzer voneinander abweichen, sondern auch, welche Angebote bzw. Angebotsformen von Google häufiger selektiert und präsentiert werden als andere.

Datenerhebung

Um diese Fragen zu klären, wurde am Abend des 22. Oktober 2013 eine koordinierte Google-Suche durchgeführt: 16 Studierende des Master-Studiengangs „Medien und politische Kommunikation“ an der Freien Universität Berlin starteten auf ihren privaten Rechnern außerhalb der Universität je eine Google-Suchanfrage zu den beiden Suchstrings „Salafismus Deutschland“ und „Tunesien“. Die Begriffskombination „Salafismus Deutschland“ hat dabei angesichts der medialen Berichterstattung in den Jahren 2012 und 2013 eine klar aktuelle und politische Konnotation. Zur Kontrolle wurde – der Strategie und Argumentation von Pari-

ser (2011) folgend – zudem der eher ambivalente Begriff „Tunesien“ gewählt: Im Kontext des „Arabischen Frühlings“ ist auch er stark mit politischer Berichterstattung verknüpft; da Tunesien als beliebtes Urlaubsziel aber auch im Reise- und Geografiekontext vorkommt, kann hier ein stärkerer Einfluss individueller Präferenzen und damit eine größere Varianz erwartet werden.

Da sich die meisten Nutzerinnen und Nutzer bei ihrer Online-Suche üblicherweise auf die Fundstellen der ersten Ergebnisseite beschränken (Nielsen & Loranger, 2006), wurden auch nur die Links dieser ersten Seite protokolliert und analysiert. Vor dem Hintergrund, dass Google sein Ranking zwar personalisiert, dieses Ranking sich aber für alle Suchanfragen aus dem gleichen Pool indexierter Quellen speist, kann davon ausgegangen werden, dass die individuellen Unterschiede in den Suchergebnissen geringer werden (und ggf. vollständig verschwinden), würde man auch die nachfolgenden Ergebnisseiten vollständig berücksichtigen. Vor der Auswertung der Links wurden die Daten noch bereinigt, indem unterschiedliche URLs, die inhaltlich auf die gleichen Angebote verwiesen, zusammengefasst wurden (etwa die Angebote verschiedener Verfassungsschutzbehörden, die auf die gleiche Broschüre verlinkten).

Limitationen

Die technischen Bedingungen (Browser, Cookies, Account-Verknüpfungen), unter denen mit Google gesucht wurde, wurden nicht experimentell variiert, sondern nur als Kontrollvariablen protokolliert. Da gewisse Informationen über IP-Nummern, Computer- und Browser-Identifikationsmerkmale bei der Google-Suche immer übertragen und von der Suchmaschine auch ausgewertet werden, lässt sich kein völlig von Personalisierungselementen unbeeinflusstes Ergebnis konstruieren, selbst wenn alle Cookies, Browserverläufe und Verknüpfungen mit Google- und sonstigen Accounts gelöscht würden. So richtete sich das Interesse entsprechend der Forschungsfragen stärker auf möglichst „realistische“ Suchverläufe, die – wenn auch die recht homogene Studierendens Stichprobe nicht repräsentativ für die deutschen Internetnutzerinnen und -nutzer sein kann – einen Eindruck von einem Personalisierungseffekt der Google-Suche in einem alltäglichen Nutzungskontext geben.

Erhoben wurde die Art des genutzten Rechners (Desktop, Laptop, Tablet etc.), das Betriebssystem (Windows, Mac OS) und ob die Teilnehmerinnen und Teilneh-

mer während der Suche in einem Google-Account (durch den eine stärkere Personalisierung ermöglicht würde) eingeloggt waren oder nicht. 13 der 16 Studierenden verwendeten für ihre Suche einen Windows-Rechner, nur drei arbeiteten mit einem Mac-Betriebssystem. Die Browser-Vielfalt war hingegen etwas größer: Zwei Mac-Nutzer suchten mit Hilfe des Safari-Browsers, viermal kam Google-Chrome zum Einsatz und die restlichen zehn Personen verwendeten Firefox. Mit sieben Teilnehmerinnen und Teilnehmern war etwa die Hälfte während der Suche mit einem Google-Konto eingeloggt.

Eine zusätzliche Unschärfe entsteht dadurch, dass der Zeitpunkt der Suchanfragen nicht sekundengenau koordiniert werden konnte. Vorgabe an die Probanden war es, die Suche abends gegen 19 Uhr zu Hause von einem privaten Rechner aus durchzuführen, die Ergebnisse in einer Excel-Datei zu dokumentieren und diese Datei dann bis zum nächsten Morgen an den Versuchsleiter zu übermitteln. Entsprechend kann die Möglichkeit, dass sich zwischen den einzelnen Suchzeitpunkten der Versuchspersonen die Grundgesamtheit an auffindbaren Quellen verändert hat, nicht ausgeschlossen werden. Eine stärker koordinierte Suche – eventuell sogar mit einer viel größeren Zahl an Teilnehmern – könnte jedoch zu Reaktionen der Google-Algorithmen und damit zu verzerrten Ergebnissen führen. So könnten die parallel durchgeführten Suchen zum Beispiel als technische Robot-Anfragen oder gar als Angriff auf den Server interpretiert werden.

4 Ergebnisse

Einen ersten aggregierten Überblick über die Ergebnisse der zwei durchgeführten Suchen liefert Tabelle 1. Deutlich wird, dass die Zahl der je Studienteilnehmer angezeigten Fundstellen relativ stark variiert: Es wurden immer mindestens zehn Quellen angezeigt, im Einzelfall umfasste die Ergebnisliste allerdings bis zu 16 Links. Zusammengenommen wurden den 16 Teilnehmerinnen und Teilnehmern auf der jeweils ersten Ergebnisliste einmal 164 und einmal 171 Fundstellen angezeigt, was im Durchschnitt elf bis zwölf Links ergibt. Diese Links verwiesen dabei aber nur auf insgesamt 26 bzw. 34 Quellen. Dies deutet einerseits darauf hin, dass es einen gewissen Überlappungsbereich der Suchergebnisse gibt und somit nicht jede und jeder Suchende vollständig andere Ergebnisse erhält. Auf der anderen Seite übersteigt die Zahl der eindeutig unterscheidbaren

Quellen doch deutlich die Zahl der verfügbaren Plätze auf der ersten Ergebnisseite, so dass es offenbar auch klare individuelle Unterschiede geben muss.

Interessant ist insbesondere, dass die Gesamtzahl der indizierten Fundstellen, die Google zu jeder Suche ausgibt, ebenfalls variiert: Für die Begriffskombination „Salafismus Deutschland“ schwankte die Angabe zwischen 97.000 und 135.000, für den Begriff „Tunesien“ sogar zwischen 9 und 37 Millionen. Dies widerspricht der oben formulierten Annahme, dass sich die Personalisierung vor allem auf ein individualisiertes Ranking aus der grundsätzlich gleichen Menge an zu einem Begriff indizierten Quellen beschränkt. Stattdessen liegt es nahe, dass es selbst bei einer Analyse der vollständigen Suchergebnisse über die erste Seite hinaus zu relevanten Differenzen kommen würde.

Tabelle 1: Zusammengefasste Ergebnisse der Google-Suchen der 16 Probanden

	„Salafismus Deutschland“	„Tunesien“
Summe aller Fundstellen (Anzahl der Quellen)	164 (26)	171 (34)
Durchschnittlich angezeigte Fundstellen (Range)	12,1 (10-16)	11,0 (10-15)
Durchschnittliche Gesamtzahl der Fundstellen laut Google (Range)	ca. 118.000 (97.000-135.000)	ca. 25 Mio. (9-37 Mio.)
Übereinstimmende Quellen nach Holsti (Range)	0,63 (0,36-0,91)	0,64 (0,20-1,0)

4.1 Zugänglich gemachte Angebote

Zur Beantwortung der ersten Forschungsfrage, welche medialen Angebotsformen Google grundsätzlich zugänglich macht, wurden die Fundstellen entsprechend der oben definierten Kategorien nach verschiedenen Quellenarten unterschieden. Da sich in den Ergebnissen allerdings nicht alle oben definierten theoretisch relevanten Angebotsformen fanden, unterscheiden wir im Folgenden nur noch die

folgenden drei Kategorien: Erstens *professionell-journalistische Medienangebote*, die im Rahmen von Medienorganisationen entstanden sind. Hier handelt es sich in der Regel um Links auf konkrete Artikel oder um Verweise auf Themenkategorien oder spezielle Archive. Davon unterschieden werden zweitens *nutzergenerierte Angebote* wie allgemeine oder thematische Online-Foren oder unabhängige Blogs, in denen eher schwach organisierte Gruppen und Einzelpersonen Beiträge veröffentlichen (können). Die dritte Kategorie umfasst schließlich alle *Angebote anderer (organisierter) Akteure*, wobei es sich sowohl um staatliche Einrichtungen als auch um NGOs oder Wirtschaftsunternehmen handeln kann. Im Einzelfall fiel die Abgrenzung allerdings nicht ganz leicht: So fand sich auf den Seiten der Deutschen Welle zum Beispiel ein Filmbeitrag, der dort jedoch nicht unter den journalistischen Beiträgen gelistet, sondern Teil eines interaktiven Deutsch-Sprachkurses war, weshalb er auch in der Auswertung nicht als journalistische Quelle gewertet wurde. Ebenso wurde Wikipedia, obwohl es sich dabei um eine Plattform mit nutzergenerierten Inhalten handelt, in die dritte und nicht in die zweite Kategorie eingeordnet, da es sich dabei letztlich nicht um Beiträge von Einzelpersonen und schwach organisierten Gruppen zu einer öffentlichen Debatte handelt, sondern eher um ein Amalgam aus stark moderierten und komprimierten Einzelbeiträgen.

Inwiefern tauchen diese drei Angebotsformen nun also in der Google-Suche auf? Zunächst ist bemerkenswert, dass bei beiden Suchen immerhin knapp die Hälfte (46 bzw. 44%) der gefundenen Quellen als *klassisch journalistische Medienangebote* identifiziert werden konnten (siehe Tabelle 2), sie aber nur 35 bzw. 27 Prozent der allen Teilnehmerinnen und Teilnehmern angezeigten Fundstellen ausmachen. Ihre relative Bedeutung ist über mehrere Suchen hinweg also eher gering. Im Durchschnitt wurde eine journalistische Quelle von nur vier bis fünf Studienteilnehmern gefunden, also gerade mal etwa einem Viertel der Suchenden. *Foren und Blogs* sind hingegen in beiden Fällen nur einmal anzutreffen, wurden aber immerhin von 12 bzw. 14 Studienteilnehmern und damit von einer Mehrheit der Suchenden gefunden. Am stärksten waren entsprechend die *Angebote anderer Einzelakteure* zu finden. Obwohl sie ebenfalls nur rund die Hälfte der Quellen stellen, machen sie bis zu zwei Drittel (57 bzw. 65%) der insgesamt angezeigten Fundstellen aus. Eine Quelle aus dieser Kategorie wurde im Durchschnitt von sechs bis sieben Studienteilnehmern – also von immerhin rund 40 Prozent der Suchenden – gefunden.

Diese Befunde sind theoretisch wie methodisch brisant, denn sie bedeuten, dass die individuell wahrgenommene „Öffentlichkeit“ für diejenigen, die ihre

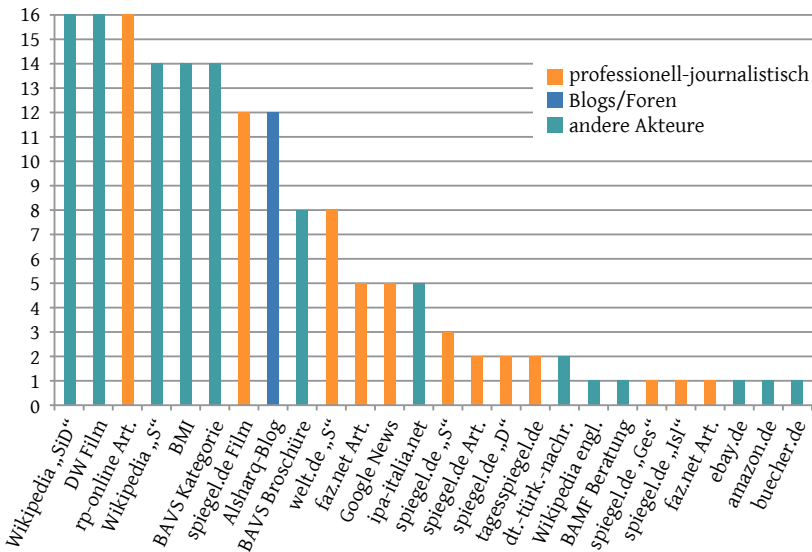
Informationen in erster Linie aus dem Internet beziehen, vermutlich deutlich weniger als früher von klassischen Medienangeboten definiert wird. Öffentlichkeit ist demzufolge heute weniger massenmedial-journalistisch geprägt, sondern besteht aus einem vielfältigeren Mix aus Quellen verschiedenster Herkunft und Qualität. Im Fall der Salafismus-Suche handelt es sich dabei vor allem um staatliche Behörden wie das Bundesinnenministerium (BMI), das Bundesamt für Verfassungsschutz (BAVS) und das Bundesamt für Migration und Flüchtlinge (BAMF); sie alleine machten immerhin 37 (22,5%) der insgesamt 164 Fundstellen aus (siehe Abbildung 1). Durch diese Dominanz staatlicher Akteure wird das Thema in einen ganz bestimmten Kontext gestellt, sodass bestimmte Interpretationsmuster eines Problems näher liegen als andere. In einer begleitenden Inhaltsanalyse der online-publizierten Debattenbeiträge zeigte sich zum Beispiel, dass die Argumentationsmuster in journalistischen Medien einem auch durch diese Google-Fundstellenstruktur nahegelegten Framing von Salafismus als Problem staatlicher Sicherheitsbehörden folgte (Emmer, 2013). In alternativen öffentlichen Arenen im Internet wie Foren, Blogs oder auch Kommentarbereichen anderer Medien und Webseiten, die stärker durch Nutzerbeiträge dominiert sind, spielte hingegen eine eher zivile Auseinandersetzung mit dem Problem des Salafismus eine

Tabelle 2: Summe der Fundstellen (und Anzahl der Quellen) nach Angebotsform

	„Salafismus Deutschland“		„Tunesien“	
	Summe aller Fund- stellen	Anzahl der Quellen	Summe aller Fund- stellen	Anzahl der Quellen
Professionell-journalisti- sche Medienangebote	58 (35%)	12 (46%)	46 (27%)	15 (44%)
Alternative Formen (Foren, Blogs, etc.)	12 (7%)	1 (4%)	14 (8%)	1 (3%)
Angebote anderer Akteure	94 (57%)	13 (50%)	111 (65%)	18 (53%)
Gesamt	164 (100%)	26 (100%)	171 (100%)	34 (100%)

größere Rolle. Google bildet dies hier jedoch kaum ab. Auch dieser Befund spricht dafür, dass absolute Relevanzindikatoren wie Abrufzahlen, Links auf vielgenutzte Angebote et cetera einen erheblichen Einfluss auf die dargestellten Suchergebnisse haben.

Abbildung 1: Absolute Häufigkeiten der Suchergebnisse „Salafismus Deutschland“

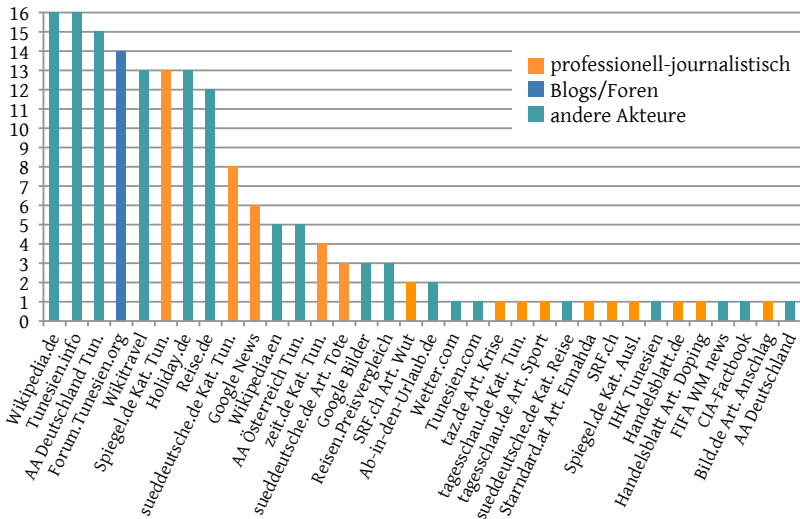


Bei den Angaben in Anführungszeichen handelt es sich nicht um Links auf konkrete Texte sondern auf thematische Kategorien innerhalb des jeweiligen Angebots. So haben zahlreiche Medien eigene Archive, die mit „Salafismus“ verschlagwortete Artikel versammeln. Die vollständigen Links können dem digitalen Anhang zu diesem Beitrag entnommen werden (siehe Emmer & Strippel, 2015).

Auch bei der Suche nach dem Begriff „Tunesien“ (siehe Abbildung 2) finden sich häufig staatliche bzw. andere politische Akteure: Hier sind es die Außenministerien Deutschlands und Österreichs sowie eine deutsch-tunesische Handelskammer. Vor allem dominieren aber ökonomische Akteure aus dem Tourismusbereich. Insgesamt zeigt sich bei diesem etwas ambivalenteren Thema eine stärkere Heteroge-

nität der Ergebnisse: Während journalistische Fundstellen die aktuelle politische Situation nach dem politischen Umbruch des Jahres 2011 thematisieren, drehen sich die Links auf die sonstigen Quellen von Reiseveranstaltern und Reiseportalen wie *Tunesien.info* bis zu *Wetter.de* eher um das Urlaubsziel Tunesien.

Abb. 2: Absolute Häufigkeiten der Suchergebnisse „Tunesien“



Bei den Angaben in Anführungszeichen handelt es sich nicht um Links auf konkrete Texte, sondern um Links auf thematische Kategorien innerhalb des jeweiligen Angebots. So haben zahlreiche Medien eigene Archive, die mit „Reise“, „Ausland“ o.ä. vorschlagwortete Artikel versammeln. Als professionell-journalistisch sind hier nur Quellen kategorisiert, die einen allgemeinen Nachrichten- oder Politikinhalt hatten; Reisejournalismus wurde unter die anderen Quellen subsummiert, weil hier die Abgrenzung zu Reiseportalen zum Teil nicht mehr eindeutig war. Die vollständigen Links können dem digitalen Anhang zu diesem Beitrag entnommen werden (siehe Emmer & Strippel, 2015).

Methodisch lässt sich schlussfolgern, dass Analysen, die „öffentliche Debatten“ über bestimmte Themen in den Fokus nehmen – in welchem theoretischen Kontext auch immer – zu kurz greifen, wenn sie sich ausschließlich auf die Analyse klassischer Medienstichproben verlassen. Denn andere Quellenarten werden un-

seren Ergebnissen zufolge dabei systematisch unterschätzt: So ist zum Beispiel das kollaborative Wissensmanagementsystem „Wikipedia“ die einzige Quelle, die in beiden Suchen tatsächlich allen Suchenden angezeigt wurde. Sie dominiert auch die ersten drei Plätze des Google-Rankings: Im Falle der Suche zu „Salafismus in Deutschland“ waren 35 dieser insgesamt 48 Plätze im Datensatz von Wikipedia-Artikeln belegt. Nur in einem Fall verwies das erste Ergebnis auf die Beratungsseite des Bundesamtes für Migration und Flüchtlinge (BAMF). Erst auf Rangplatz drei folgte dann häufiger eine Verfassungsschutzseite (11 Mal) sowie ein Link auf ein journalistisches Angebot (welt.de). Dieser Befund stützt bisherige Befragungsstudien, die ebenfalls zu dem Ergebnis kamen, dass Wikipedia eine relevante Informationsquelle für Bürgerinnen und Bürger ist – und zwar durchaus auch im Kontext des auf aktuelle politische Zusammenhänge bezogenen Informationsverhaltens (Emmer, Vowe, & Wolling, 2011, S. 136-137). Die andere vermutlich unterschätzte Quellenart sind die Angebote politischer und gesellschaftlicher Akteure, die sich direkt mit ihren Angeboten an die Nutzerinnen und Nutzer wenden. Diese Möglichkeit für solche Organisationen, die klassischen journalistischen Gatekeeper zu umgehen, scheint tatsächlich eine gewisse Wirkung zu entfalten.

4.2 *Unterschiede in den Suchergebnissen*

Die zweite Forschungsfrage bezieht sich schließlich auf die Frage, inwieweit Google bzw. Suchmaschinen-Ergebnisse prinzipiell geeignet sind, die öffentliche Kommunikation im Internet reliabel abzubilden. Die Vermutung ist hier, dass die Personalisierung der Ergebnisauswahl und -darstellung zu einem stark individualisierten Ranking führt, so dass Suchmaschinen alleine nicht mehr als geeignete Auswahlgrundlagen dienen können, weil sie in unkontrollierbarer Weise die persönlichen Eigenschaften der jeweiligen Forscherinnen und Forscher widerspiegeln.

Tatsächlich zeigt sich bei beiden Suchen, dass nur wenige Quellen pro Suchbegriff wirklich ohne Unterschied allen Studienteilnehmerinnen und -teilnehmern angezeigt werden (siehe Abbildungen 1 und 2): Wikipedia scheint sich hier zu einem Standard entwickelt zu haben, der unabhängig von individuellen Präferenzen grundsätzlich immer ausgegeben wird, und zwar meistens auch als erstplatziertes Ergebnis. Etwa ein Drittel aller Quellen erreicht noch die Mehrheit

der Versuchspersonen, danach beginnen die Häufigkeiten deutlich abzufallen bis hin zum sprichwörtlichen „Long Tail“ von Einzelangeboten, die keine besonders große Reichweite haben, aber in der Summe doch einen relevanten Teil der öffentlich zugänglichen Online-Kommunikation ausmachen.

Über alle Suchenden hinweg war in Paarvergleichen im Durchschnitt etwa die Hälfte der jeweiligen Quellen gleich, in nur einem von 240 Einzelpaarungen wurden vollkommen identische Suchergebnisse ermittelt, und auch hier nicht in der identischen Reihenfolge. Die geringste Übereinstimmung lag bei nur 20 Prozent bzw. zwei von zehn Quellen. Bei der Tunesien-Suche war die Vielfalt unterschiedlicher Quellen mit 34 insgesamt etwas größer als bei der Suche nach Salafismus mit 26, was angesichts der größeren Ambivalenz des reinen Länderbegriffs „Tunesien“ nicht verwundert. Die Reliabilität der Suchen lag mit einem Koeffizienten um .64 nach Holsti unter den Werten, die man für eine einfache Nominalskala im Rahmen einer quantitativen Inhaltsanalyse akzeptieren würde. Für die Stichprobenziehung ist damit ein nicht unerhebliches Risiko verbunden, mit einer einzelnen, am Computer des Forschers bzw. der Forscherin gezogenen Stichprobe die für eine breitere Öffentlichkeit relevanten Quellen zu verfehlen.

5 Schlussfolgerungen und Lösungsansätze

In einem Test mit jeweils 16 Google-Suchanfragen zu zwei unterschiedlich eng definierten Themen hat sich gezeigt, dass den suchenden Nutzerinnen und Nutzern erstens ein wesentlich durch nicht-journalistische Quellen geprägtes Informationsangebot geliefert wird und dass zweitens das Suchergebnis in erheblichem Maße individuell variiert. Auch wenn die Befunde aufgrund der Stichprobe keinen Repräsentativitätsanspruch erheben können, so war das Untersuchungsdesign insgesamt doch so angelegt, dass durchaus angenommen werden kann, dass sich mit anderen Teilnehmern ähnliche Ergebnisse zeigen würden.

Vor diesem Hintergrund stellt sich die Frage, wie in der empirischen, inhaltsbezogenen Kommunikationsforschung mit den daraus resultierenden Problemen für die Stichprobenziehung mittels Suchmaschinen umgegangen werden soll. Mögliche technische Lösungen, etwa das Eliminieren aller Filter- und Personalisierungsquellen, sind wenig realistisch: Erstens kann auch durch ein solches Vorgehen nicht sichergestellt werden, dass tatsächlich alle im Netz existierenden Quellen

gefunden werden, da alleine im Crawling- und Indexierungsverfahren der Suchmaschinen bereits eine Auswahl getroffen wird. Zweitens sind die eingesetzten Algorithmen als Geschäftsgeheimnis der Suchmaschinenbetreiber nicht ausreichend transparent, um sie sicher zu kontrollieren. Und drittens sind in die Funktionsweise des Internets gewisse Identifikationsmechanismen eingebaut (Ortsinformationen über die IP-Nummer, Details zu PC- und Softwarenutzung), die sich grundsätzlich nicht verbergen, sondern bestenfalls zum Beispiel über Anonymisierungsdienste verändern oder verschleiern lassen. Ein viertes, sehr grundsätzliches Argument gegen solche technischen Lösungen ist schließlich die Tatsache, dass solche Suchen wenig realitätsnah wären und die Gefahr bestünde, zwar technisch reliable, aber dafür kaum valide Ergebnislisten zu produzieren.

Orientierung am Nutzerverhalten

Statt zu versuchen, eine „unverfälschte“ Suchvariante zu konstruieren, sollte sich also an den üblichen Suchstrategien der Nutzer orientiert und die entsprechenden Filter und Personalisierungen als deren Bestandteil akzeptiert und im Auswahlverfahren berücksichtigt werden (für eine automatisierte Simulation von Nutzerverhalten siehe den Beitrag von Jürgens et al. in diesem Band). Eine konsequente Orientierung am Nutzerverhalten bedeutet jedoch, dass sich Angebots- bzw. Inhaltsstichproben nicht mehr unabhängig vom Nutzerverhalten ziehen lassen: Sowohl die Nutzerinnen und Nutzer als auch die auf sie ausgerichteten und auf sie reagierenden Suchmaschinen-Algorithmen definieren ständig neu, was öffentlich zugängliche Inhalte sind. Die ideale Variante einer suchmaschinenbasierten Stichprobenziehung wäre demnach, eine repräsentative Stichprobe von Nutzerinnen und Nutzern nach Inhalten suchen zu lassen, und anschließend die von ihnen mit Hilfe der Suchmaschine gefundenen Angebote zu analysieren (Wolling & Kuhlmann, 2003, S. 150-153). Die Überlappungen und Häufigkeiten verschiedener Quellen könnten dann zugleich als Gewichte für die analysierten Angebote dienen und die Daten um eine Dimension bereichern, die bei klassischen Inhaltsanalysen herkömmlicher Medien nicht oder nur sehr begrenzt zur Verfügung steht.

Dazu wäre jedoch eine detailliertere Erforschung des Online-Nutzungsverhaltens erforderlich, damit die Relevanz der verschiedenen Angebote und Angebotsformen besser eingeschätzt werden kann und so zum Beispiel eine theo-

retisch fundiertere Auswahl möglich wird. Trotz einer beinahe zwanzigjährigen Forschungstradition ist bis heute nur sehr grob bekannt, wie die alltägliche Mediennutzung im Online-Kontext abläuft. Während angebotsbezogene Forschung nur Ausschnitte des kommerziellen Online-Angebots analysiert und über den „Long Tail“ kleiner und nutzergenerierte Angebote nichts aussagen kann, bleibt die nutzerorientierte Forschung, die im Wesentlichen auf Befragungsmethoden basiert (z.B. ARD/ZDF-Arbeitsgruppe Marketing, 1997; Eimeren & Frees, 2013), zumeist auf einer allgemeinen, kategorialen Ebene, ohne konkrete Aussagen über die Quantität der Nutzung von einzelnen Angeboten, Beiträgen und Informationen machen zu können. Hier könnten neuere Forschungsmethoden mit Hilfe von User-Tracking und so genannten „Big Data“-Analysen in Zukunft jedoch helfen, eine Bewertungsgrundlage für die nutzerseitige Relevanz verschiedener Angebotsformen zu liefern (Strippel & Emmer, 2015).

Berücksichtigung verschiedener Nutzungspraktiken

In diesem Zusammenhang kann vermutet werden, dass eine exklusiv suchmaschinenbasierte Auswahl die Bedeutung alternativer Informationswege unterschätzt: Die gezielte Suche ist schließlich nur eine von zahlreichen Strategien, mit deren Hilfe sich über das aktuelle Geschehen informiert wird. Es gibt zahlreiche andere Informationswege, die ohne Suchmaschinen auskommen, etwa die regelmäßige Nutzung von Portalseiten mit Informationsanteilen (Emmer, Vowe, & Wolling, 2011, S. 134) oder auch die mehr oder weniger habitualisierte Nutzung der immer gleichen Informationsangebote (Emmer, Wolling, & Vowe, 2012), ganz zu schweigen von eher beiläufiger Information durch Freunde über Soziale Online-Netzwerke wie Facebook (Busemann, 2013, S. 394). Hierüber werden sicher in größerem Umfang Informationen wahrgenommen, ohne dass sie vorher den Nutzerinnen und Nutzern auf der ersten Ergebnisseite von Google präsentiert wurden.

Allerdings sind diese Quellen vermutlich auch nicht gänzlich unabhängig von den Google-Relevanzkriterien: Die intensivere Nutzung bestimmter Quellen, zum Beispiel die Verlinkung in verschiedenen anderen Online-Medien wie Twitter, Facebook oder in Blogs, wird sehr wahrscheinlich die Relevanz eines solchen Beitrags stärken und zumindest die Wahrscheinlichkeit, dass andere Nutzer sie eventuell doch über eine Suchmaschine finden, wieder erhöhen (siehe Hildebrandt et al. in diesem Band). So kann der Einsatz von Suchmaschinen sehr

wohl dazu genutzt werden, relevante Foren und Teilöffentlichkeiten im Internet zu identifizieren. Diese können dann gezielt – und über interne Such- und Archiv-Funktionen in der Regel auch vollständig – nach jeweils relevanten Inhalten durchsucht werden. Individuelle Filter, die für einzelne Suchende bestimmte Angebote ausblenden, können somit zumindest teilweise umgangen werden. Im Rahmen eines solchen Vorgehens ist es darüber hinaus sinnvoll, auch andere Suchmaschinen als Google zu nutzen (Meier, Wunsch, Pentzold, & Welker 2010, S. 113). Diese haben zwar keine großen Marktanteile, können aber – wie zum Beispiel Yahoo! für die eigenen Foren- und Gruppenangebote oder Technorati für Blogs – einzelne Bereiche des Internets zum Teil besser erfassen. Die Wahrscheinlichkeit, Stichproben durch ein solches Vorgehen noch stärker zu verzerren (Meier et al., 2010), ist angesichts der Dominanz von Google und dessen vielfältigen Filtern nicht besonders hoch. Vielmehr kann eine solche Kombination von Quellen als eine Form von Triangulation dienen und dabei helfen, „Forscher-Filter-Bubbles“ zu vermeiden. Hier können Anleihen bei einer theoriegeleiteten Auswahl, etwa dem „Theoretical Sampling“ der Grounded Theory, hilfreich sein (Meier & Pentzold, 2010).

Ein solches mehrstufiges Vorgehen, das nach einem Screening der relevanten Online-Angebote in einem zweiten Schritt innerhalb der für relevant erachteten Angebote konkrete Texte auswählt, hat den zusätzlichen Vorteil, dass darüber bestimmte Angebotsformen ermittelt und in die Analyse integriert werden können, die über Suchmaschinen nur schwer systematisch auffindbar oder archivierbar sind. Dies gilt etwa für Nutzerkommentare, die sich inhaltlich in der Regel auf einen bestimmten Ausgangstext beziehen und – je nach Forschungsfrage – ohne diesen Zusammenhang wenig Aussagekraft haben. Solche Kommentare können je nach Angebot verschiedene technische Formen haben und deswegen unterschiedlich gut durch Suchmaschinen gefunden werden, zum Beispiel als eigene Webpage und damit in technischer Hinsicht eigener Artikel oder direkt integriert in den entsprechenden Ausgangsartikel.

Teil einer solchen mehrstufigen Auswahl können dann in der letzten Auswahlstufe klassische Zufallsauswahlen sein. Hat man ein relevantes Angebot identifiziert und die relevanten Angebotsformen (Beiträge, Artikel, Kommentare etc.) erfasst, hat man es häufig mit sehr umfangreichen Materialsammlungen zu tun: Viele Angebote im Netz haben immer weiter wachsende Archive online, darüber hinaus ermöglichen sie Nutzern die Kommentierung von Beiträgen, die zum

Korpus hinzukommen. Geht man von einer nutzerbasierten Stichprobenziehung aus, sind letztlich auch solche Kommentare relevanter Gegenstand einer Inhaltsanalyse zu öffentlichen Debatten über bestimmte Themen. Da sich das Verhältnis von Originalbeiträgen zu Nutzerkommentaren in verschiedenen Angeboten durchaus im Bereich von 1:40 bewegen kann und im Einzelfall Beiträge mehr als 1.000 Kommentare erhalten (Emmer, 2013, S. 10), sind hier weitergehende Stichprobenverfahren nötig, um die Materialmenge auf ein handhabbares Maß zu reduzieren. Auch hier bieten sich in der Regel Zufallsverfahren an.

Letztlich stellt die Stichprobenziehung für Inhaltsanalysen im Internet noch immer eine große Herausforderung dar, die eine komplexe Kombination aus theoretischen und zufallsgesteuerten Auswahlritten erfordert. Ebenso ist der Zeit- und Mitteleinsatz höher als bei der Analyse herkömmlicher massenmedialer Kommunikationsinhalte. Dabei wird es angesichts der immer weiter wachsenden und sich ausdifferenzierenden digitalen Kommunikationswelt eine Illusion bleiben, einmal eine Lösung für alle speziellen Stichprobenprobleme zu finden. Inhaltsanalyse-Projekte im Netz erfordern daher eine intensive Auseinandersetzung mit stichprobentheoretischen Problemen. Allerdings lassen sich die Schwierigkeiten, zerlegt man sie in einzelne Probleme, schrittweise durchaus soweit eingrenzen, dass sich am Ende wissenschaftlich tragfähige Verfahren für viele Forschungsfragen finden.

Prof. Dr. Martin Emmer ist Professor für Mediennutzung am Institut für Publizistik- und Kommunikationswissenschaft an der Freien Universität Berlin

Christian Strippl, M.A. ist wissenschaftlicher Mitarbeiter am Institut für Publizistik- und Kommunikationswissenschaft an der Freien Universität Berlin

Quellenverzeichnis

- Anagnostopoulos, A., Broder, A. Z., & Carmel, D. (2006). Sampling Search-Engine Results. *World Wide Web*, (9), 397-429. doi: 10.1007/s11280-006-0222-z
- Arbeitsgemeinschaft ADM-Stichproben & Bureau Wendt (1994). Das ADM Stichproben-System: Stand 1993. In S. Gabler, J. H. P. Hoffmeyer-Zlotnik, & D. Krebs (Hrsg.), *Gewichtung in der Umfragepraxis* (S. 188-202). Opladen: Westdeutscher Verlag.

- ARD/ZDF-Arbeitsgruppe Marketing (1997). *Was Sie über Rundfunk wissen sollten. Materialien zum Verständnis eines Mediums*. Berlin: Vistas.
- Bar-Yossef, Z., & Gurevich, M. (2008). Random Sampling from a Search Engine's Index. *Journal of the ACM*, 55(5), Artikel 24.
- Bozdag, E. (2013). Bias in algorithmic filtering and personalization. *Ethics and Information Technology*, 15(3), 209-227. doi: 10.1007/s10676-013-9321-6
- Busemann, K. (2013). Wer nutzt was im Social Web? Ergebnisse der ARD/ZDF-Onlinestudie 2013. *Media Perspektiven*, (7-8), 391-399.
- Cothey, V. (2004). Web-Crawling Reliability. *Journal of the American Society for Information Science and Technology*, 55(14), 1228-1238.
- Eimeren, B. v., & Frees, B. (2013). Ergebnisse der ARD/ZDF-Online-Studie 2013: Rasanter Anstieg des Internetkonsums – Onliner fast drei Stunden täglich im Netz. *Media Perspektiven*, (7-8), 358-372.
- Emmer, M. (2013). Abschlussbericht BaSiD-Modul Medienmonitoring (Memo-S), Fallstudie 5: Web-Debatten: Die Auseinandersetzung im Netz um die Gefährlichkeit von Salafisten. Unveröffentlichter Bericht. FU Berlin: Berlin.
- Emmer, M., & Strippel, C. (2015). Liste der Fundstellen der Google-Suche. Anhang zum Beitrag „Stichprobenziehung für Online-Inhaltsanalysen: Suchmaschinen und Filter Bubbles“. In A. Maireder, J. Ausserhofer, C. Schumann, & M. Taddicken (Hrsg.), *Digitale Methoden in der Kommunikationswissenschaft* (S. A1-A4). doi: 10.17174/dcr.v2.12a
- Emmer, M., Vowe, G., & Wolling, J. (2011). *Bürger online. Die Entwicklung der politischen Online-Kommunikation in Deutschland*. Konstanz: UVK.
- Emmer, M., Wolling, J., & Vowe, G. (2012). Changing Political Communication in Germany. Findings from a Longitudinal Study on the Influence of the Internet on Political Information, Discussion and Participation of Citizens. *Communications: The European Journal of Communication Research*, 37(3), 233-252.
- Erlhofer, S. (2007). *Informationssuche im World Wide Web. Taktiken und Strategien bei der Nutzung von Suchmaschinen*. Berlin: WVB.
- Galliker, M., & Männel, O. (1999). Suchmaschinen als Datenerhebungsinstrument von WWW-Inhaltsanalysen. In B. Batinić, A. Werner, L. Gräf, & W. Bandilla (Hrsg.), *Online Research. Methoden, Anwendungen und Ergebnisse* (S. 245-261). Göttingen, Bern, Toronto, Seattle: Hogrefe.
- Gehrau, V., Fretwurst, B., Krause, B., & Daschmann, G. (2005). *Auswahlverfahren in der Kommunikationswissenschaft*. Köln: Herbert von Halem.

- Gerhards, J., & Schäfer, M. S. (2007). Demokratische Internet-Öffentlichkeit? Ein Vergleich der öffentlichen Kommunikation im Internet und in den Printmedien am Beispiel der Humangenomforschung. *Publizistik*, 52(2), 210-228. doi: 10.1007/s11616-007-0084-0
- Google (2005). Search gets personal. Abgerufen von <http://googleblog.blogspot.de/2005/06/search-gets-personal.html>
- Google (2009a). Introducing Google Social Search: I finally found my friend's New York blog! Abgerufen von <http://googleblog.blogspot.de/2009/10/introducing-google-social-search-i.html>
- Google (2009b). Personalized Search for everyone. Abgerufen von <http://googleblog.blogspot.de/2009/12/personalized-search-for-everyone.html>
- Google (2010). Search is getting more social. Abgerufen von <http://googleblog.blogspot.de/2010/01/search-is-getting-more-social.html>
- Google (2011). An update to Google Social Search. Abgerufen von <http://googleblog.blogspot.de/2011/02/update-to-google-social-search.html>
- Google (2012). Search, plus Your World. Abgerufen von <http://googleblog.blogspot.de/2012/01/search-plus-your-world.html>
- Google (2015). Web- und App-Aktivitäten in Bezug auf Google. Abgerufen von <https://support.google.com/accounts/answer/54068>
- Griesbaum, J. (2003). Unbeschränkter Zugang zu Wissen? Leistungsfähigkeit und Grenzen von Suchdiensten im Web. Zwischen informationeller Absicherung und manipulierter Information. Abgerufen von <http://fiz1.fh-potsdam.de/volltext/konstanz/04092a.pdf>
- Gulli, A., & Signorini, A. (2005). The Indexable Web is More than 11.5 Billion Pages. *Proceedings of 14th International World Wide Conference*, 902-903.
- Häder, S., & Gabler, S. (1998). Ein neues Stichprobendesign für telefonische Umfragen in Deutschland. In S. Gabler, S. Häder, & J. H. P. Hoffmeyer-Zlotnik (Hrsg.), *Telefonstichproben in Deutschland* (S. 69-88). Opladen, Wiesbaden: Westdeutscher Verlag.
- Hartmann, W., Näf, M., & Schäuble, P. (2000). *Informationsbeschaffung im Internet. Grundlegende Konzepte verstehen und umsetzen*. Zürich: Orell Füssli.
- Hauptmanns, P. (1999). Grenzen und Chancen von quantitativen Befragungen mit Hilfe des Internet. In B. Batinic, A. Werner, L. Gräf, & W. Bandilla (Hrsg.), *Online Research. Methoden, Anwendungen und Ergebnisse* (S. 21-38). Göttingen: Hogrefe.

- Hauptmanns, P., & Lander, B. (2001). Zur Problematik von Internet-Stichproben. In A. Theobald, M. Dreyer, & T. Starsetzki (Hrsg.), *Online-Marktforschung. Theoretische Grundlagen und praktische Erfahrungen* (S. 27-40). Wiesbaden: Gabler.
- Ke, Y., Deng, L., Ng, W., & Lee, D.-L. (2005). Web Dynamics and their Ramifications for the Development of Web Search Engines. *Computer Networks*, 50(10), 1430-1447.
- Kops, M. (1977). *Auswahlverfahren in der Inhaltsanalyse*. Meisenheim am Glan: Verlag Anton Hain.
- Loosen, W. (1999). Suchmaschinen. „Informations- und Wissensverwalter“ im World Wide Web. *Medien Journal*, (3), 42-48.
- Loosen, W. (2001). Das Ende der Verheißungen? Technisch und ökonomisch determinierte Beschränkungen des „freien“ Informationszugangs im World Wide Web. In U. Maier-Rabler, & M. Latzer (Hrsg.), *Kommunikationskulturen zwischen Kontinuität und Wandel. Universelle Netzwerke für die Zivilgesellschaft* (S. 71-93). Konstanz: UVK.
- Machill, M., Neuberger, C., & Schindler, F. (2002). *Transparenz im Netz. Funktionen und Defizite von Internet-Suchmaschinen*. Gütersloh: Verlag Bertelsmann Stiftung.
- McMillan, S. J. (2000). The Microscope and the Moving Target: The Challenge of Applying Content Analysis to the World Wide Web. *Journalism & Mass Communication Quarterly*, 77(1), 80-98. doi: 10.1177/107769900007700107
- Meier, S., & Pentzold, C. (2010). Theoretical Sampling als Auswahlstrategie für Online-Inhaltsanalysen. In M. Welker & C. Wünsch (Hrsg.), *Die Online-Inhaltsanalyse. Forschungsobjekt Internet* (S. 124-143). Köln: Herbert von Halem.
- Meier, S., Wünsch, C., Pentzold, C., & Welker, M. (2010). Auswahlverfahren für Online-Inhalte. In M. Welker & C. Wünsch (Hrsg.), *Die Online-Inhaltsanalyse. Forschungsobjekt Internet* (S. 103-123). Köln: Herbert von Halem.
- Mislove, A., Gummadi, K. P., & Druschel, P. (2006). Exploiting Social Networks for Internet Search. Abgerufen von <http://iris.csail.mit.edu/irisbib/papers/exploiting-social-hotnets06/paper.pdf>
- Netmarketshare (2013). Desktop Search Engine Market Share. Abgerufen von <http://www.netmarketshare.com/search-engine-market-share.aspx?qprid=4&qpcustomd=0>
- Nielsen, J., & Loranger, H. (2006). *Web Usability*. München u.a.: Addison-Wesley.
- Pariser, E. (2011). *The Filter Bubble: What the Internet is Hiding from You*. New York: Penguin Press.

- Pasquinelli, M. (2009). Googles PageRank. Diagramm des kognitiven Kapitalismus und Rentier des gemeinsamen Wissens. In K. Becker & F. Stalder (Hrsg.), *Deep Search. Politik des Suchens jenseits von Google* (S. 171-181). Innsbruck: Studienverlag.
- Röhle, T. (2009). Die Demontage der Gatekeeper. Relationale Perspektiven zur Macht der Suchmaschinen. In K. Becker & F. Stalder (Hrsg.), *Deep Search. Politik des Suchens jenseits von Google* (S. 133-149). Innsbruck: Studienverlag.
- Rössler, P. (2010). *Inhaltsanalyse* (2. Aufl.). Konstanz: UVK.
- Rössler, P., & Wirth, W. (2001). Inhaltsanalysen im World Wide Web. In W. Wirth & E. Lauf (Hrsg.), *Inhaltsanalyse: Perspektiven, Probleme, Potentiale* (S. 280-302). Köln: Herbert von Halem.
- Schmidt, J.-H. (2011). *Das neue Netz: Merkmale, Praktiken und Folgen des Web 2.0* (2., überarbeitete Auflage). Konstanz: UVK.
- Schweiger, W., & Weihermüller, M. (2008). Öffentliche Meinung als Online-Diskurs – ein neuer empirischer Zugang. *Publizistik*, 53(4), 535-559.
- Seibold, B. (2002). Die flüchtigen Web-Informationen einfangen: Lösungsansätze für die Online-Inhaltsanalyse bei dynamischen Inhalten im Internet. *Publizistik*, 47(1), 45-56.
- Springer, N., & Wolling, J. (2008). Recherchegoogeln. Wie Zeitungsjournalisten das Internet für ihre Arbeit nutzen. In T. Quandt & W. Schweiger (Hrsg.), *Journalismus online: Partizipation oder Profession?* (S. 45-59). Wiesbaden: VS Verlag.
- Stalder, F., & Mayer, C. (2009). Der zweite Index. Suchmaschinen, Personalisierung und Überwachung. In K. Becker & F. Stalder (Hrsg.), *Deep Search. Politik des Suchens jenseits von Google* (S. 112-131). Innsbruck: Studienverlag.
- Starsetzki, T. (2007). Rekrutierung von Befragungsteilnehmern in der Online-Marktforschung. In M. Welker & O. Wenzel (Hrsg.), *Online-Forschung 2007. Grundlagen und Fallstudien* (S. 77-84). Köln: Herbert von Halem.
- Strippel, C., & Emmer, M. (2015). Proxy-Logfile-Analyse: Möglichkeiten und Grenzen der automatisierten Messung individueller Online-Nutzung. In O. Hahn, R. Hohlfeld, & T. Knieper (Hrsg.), *Digitale Öffentlichkeit(en)* (S. 85-103). Konstanz: UVK.
- Volpers, H. (2013). Online-Inhaltsanalyse. In K. Scherfer & H. Volpers (Hrsg.), *Methoden der Webwissenschaft* (S. 15-42). Münster: LIT Verlag.
- Weare, C., & Lin, W.-Y. (2000). Content Analysis of the World Wide Web. Opportunities and Challenges. *Social Science Computer Review*, 18(3), 272-292. doi: 10.1177/089443930001800304

- Webhits (2014). Web Barometer. Abgerufen von <http://www.webhits.de/webhits/browser.htm>
- Welker, M., Werner, A., & Scholz, J. (2005). *Online-Research. Markt- und Sozialforschung mit dem Internet*. Heidelberg: Dpunkt.
- Welker, M., & Wünsch, C. (2010). Methoden der Online-Forschung. In W. Schweiger & K. Beck (Hrsg.), *Handbuch Online-Kommunikation* (S. 487-517). Wiesbaden: VS Verlag.
- Wolling, J. (2002). Suchmaschinen – Gatekeeper im Internet. *Medienwissenschaft Schweiz*(2), 15-23.
- Wolling, J. (2005). Suchmaschinen? – Selektiermaschinen! In H. Krömer & P. Klimsa (Hrsg.), *Handbuch Medienproduktion. Produktion von Film, Fernsehen, Hörfunk, Print, Internet, Mobilfunk und Musik* (S. 529-537). Wiesbaden: VS Verlag.
- Wolling, J., & Kuhlmann, C. (2003). Das Internet als Gegenstand und Instrument der empirischen Kommunikationsforschung. In M. Löffelholz & T. Quandt (Hrsg.), *Die neue Kommunikationswissenschaft. Theorien, Themen und Berufsfelder im Internet-Zeitalter. Eine Einführung* (S. 131-161). Wiesbaden: Westdeutscher Verlag.
- Zillmann, D., & Bryant, J. (1985). *Selective Exposure to Communication*. Hillsdale: Lawrence Erlbaum.