

Digitale Methoden in der Kommunikationswissenschaft

Maireder, Axel (Ed.); Ausserhofer, Julian (Ed.); Schumann, Christina (Ed.);
Taddicken, Monika (Ed.)

Erstveröffentlichung / Primary Publication

Sammelwerk / collection

Empfohlene Zitierung / Suggested Citation:

Maireder, A., Ausserhofer, J., Schumann, C., & Taddicken, M. (Hrsg.). (2015). *Digitale Methoden in der Kommunikationswissenschaft* (Digital Communication Research, 2). Berlin. <https://doi.org/10.17174/dcr.v2.0>

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>

Was bedeuten Big-Data-Untersuchungen für die Entwicklung von Theorien und für forschungsethische Aspekte? Wie können öffentliche Spuren digitaler Kommunikation eingefangen, analysiert und interpretiert werden? Wie lassen sich Metriken von Social-Media-Plattformen in empirisch fundierte Forschung überführen? Welche Strategien gibt es, um in algorithmische Blackboxes wie Suchmaschinen und News Feeds zu schauen?

Dieser Band beschäftigt sich mit diesen und vielen ähnlichen Fragen, die bei der kommunikationswissenschaftlichen Forschungsarbeit im digitalen Zeitalter auftreten. Das Buch versammelt sowohl theoretische und ethische Auseinandersetzungen wie auch Aufsätze, die empirische Forschung zu digitaler Kommunikation dokumentieren. Immer im Mittelpunkt: Jene Praktiken, die sich an das Medium anpassen, die seine Objekte, Akteure und Infrastrukturen erforschen – also das, was wir „digitale Methoden“ der Kommunikationswissenschaft nennen.

Maireder, Ausserhofer, Schumann & Taddicken

Digitale Methoden

Band 2

Axel Maireder
Julian Ausserhofer
Christina Schumann
Monika Taddicken

Digitale Methoden in der Kommunikations- wissenschaft



dx.doi.org/10.17174/dcr.v2.0
digitalcommunicationresearch.de

ISSN 2198-7610
ISBN 978-3-945681-02-2



Digital
Communication
Research.de

Digital Communication Research

Herausgegeben von Martin Emmer, Christina Schumann, Monika Taddicken,
Martin Welker & Jens Wolling

Band 2

*Axel Maireder, Julian Ausserhofer,
Christina Schumann & Monika Taddicken*

Digitale Methoden in der Kommunikationswissenschaft

Redaktion der Reihe *Digital Communication Research*
Christian Strippel, M.A.
Freie Universität Berlin
Institut für Publizistik- und Kommunikationswissenschaft
Garystraße 55
D-14195 Berlin
info@digitalcommunicationresearch.de

Bibliografische Information der Deutschen Nationalbibliothek
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Angaben sind im Internet über <http://dnb.d-nb.de> abrufbar.

ISSN 2198-7610
ISBN 978-3-945681-02-2

Die persistente Langzeitarchivierung dieses Buches erfolgt mit Hilfe des Social Science Open Access Repository sowie der Registrierungsagentur für Sozial- und Wirtschaftsdaten da|ra.

DOI 10.17174/dcr.v2.0

Dieses Werk erscheint Open Access und ist lizenziert unter Creative Commons Attribution 4.0 (CC-BY 4.0): <http://creativecommons.org/licenses/by/4.0/>

Berlin, 2015
digitalcommunicationresearch.de

Inhaltsverzeichnis

<i>Christina Schumann, Julian Ausserhofer, Axel Maireder & Monika Taddicken</i>	
Digitale Methoden in der Kommunikationswissenschaft: Ansätze zur Analyse öffentlicher Kommunikation im Internet	9-19
I THEORIE UND ETHIK	
<i>Merja Mahrt</i>	
Mit Big Data gegen das „Ende der Theorie“?	23-37
<i>Nele Heise</i>	
Big Data – small problems? Ethische Dimensionen der Forschung mit Online-Kommunikationsspuren	39-58
<i>Christian Pentzold</i>	
Forschungsethische Prämissen und Problemfelder teilnehmenden Beobachtens auf Online-Plattformen	61-85
<i>Julian Ausserhofer</i>	
„Der Methode liegt im Code“: Routinen und digitale Methoden im Datenjournalismus	87-111

II STRUKTURANALYSEN

Axel Maireder & Stephan Schlögl

Twitter-Öffentlichkeiten: Identifikation und Interpretation
der Strukturen von Follower-Netzwerken 115-139

Axel Bruns & Theresa Sauter

Anatomie eines Trending Topics: Methoden zur
Visualisierung von Retweet-Ketten 141-161

Jürgen Grimm & Christiane Grill

Politische Online-Pluriversen. Aggregatanalyse der Twitter-
Kommunikation zur Salzburger Landtagswahl 2013 163-195

Felix Sattelberger & Wolfgang Seufert

Lassen sich mit Social-Media-Analysen die Absatzprognosen
für Vertrauensgüter verbessern? Wechselwirkungen bei der
Kommunikation über Filme vor dem Kinostart 197-221

III EXPERIMENTE

Christin Hildebrandt, Christina Schumann & Jens Wolling

Licht in die Black Box?! Experimente zum Einfluss von Social
Signals auf das organische Ranking von Google 225-249

Pascal Jürgens, Birgit Stark & Melanie Magin

Messung von Personalisierung in computer-
vermittelter Kommunikation 251-270

IV TEXT- UND INHALTSANALYSEN

Martin Emmer & Christian Strippel

Stichprobenziehung für Online-Inhaltsanalysen: Suchmaschinen
und Filter Bubbles

275-300

Katrin Jungnickel & Axel Maireder

Der Multi-Step-Flow in Social Media: Ein Mehrmethodenansatz
zur Analyse des Facebook-Newsfeed

303-327

Michael Eble & Daniel Stein

Utilisation of Audio Mining Technologies for Researching Public
Communication on Multimedia Platforms

329-345

Empfohlene Zitierung: Schumann, C., Ausserhofer, J., Maireder, A., & Taddicken, M. (2015). Digitale Methoden in der Kommunikationswissenschaft: Ansätze zur Analyse öffentlicher Kommunikation im Internet. In A. Maireder, J. Ausserhofer, C. Schumann, & M. Taddicken (Hrsg.), *Digitale Methoden in der Kommunikationswissenschaft* (S. 9-19). doi: 10.17174/dcr.v2.1

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

*Christina Schumann, Julian Ausserhofer,
Axel Maireder & Monika Taddicken*

Digitale Methoden in der Kommunikationswissenschaft

Ansätze zur Analyse öffentlicher Kommunikation im Internet

1 Zum Begriff „digitale Methoden“

Tagtäglich werden unzählige Mitteilungen auf Sozialen Netzwerkseiten und Online-Foren abgesetzt, Blogbeiträge und Webseiten veröffentlicht, Fotos, Videoclips und Dokumente über mobile und stationäre Geräte hochgeladen, bewertet, verlinkt und geteilt. Verstanden als Artefakte vermittelter Kommunikation liegt in ihrer Analyse sowie in der Analyse der Akteure, Objekte, Infrastrukturen und ihrer Relationen zueinander ein immenses Potenzial für reichhaltige Erkenntnisse zur Struktur und Kultur unserer Gesellschaft. Für die Sozialwissenschaften ergeben sich daraus zahlreiche methodologische, forschungspraktische und ethische Herausforderungen. Dabei handelt es sich vielfach um neue, zuvor noch nicht diskutierte Problematiken.

Die Kommunikationswissenschaft erscheint für eine Auseinandersetzung mit den neuen methodischen Problemen prädestiniert. Zum einen beschäftigt sie sich seit Jahrzehnten mit medienvermittelter, zuletzt auch computervermittelter Kommunikation; zum anderen manifestieren sich hier Indikatoren für jene sozialen Prozesse, die für den Kernbereich kommunikationswissen-

schaftlicher Theoriebildung maßgeblich sind: die Entwicklung von Öffentlichkeit und öffentlicher Meinung. Tatsächlich befassen sich inzwischen zahlreiche Forschungsprojekte mit der Analyse der Akteure, Artefakte und Infrastrukturen sozialer Kommunikation im Internet. Dabei werden sowohl datenintensive quantitative Methoden angewandt als auch qualitative, zum Beispiel medienethnografische Zugänge oder semantische Analysen. In den vergangenen Jahren taucht in der Literatur in diesem Zusammenhang immer öfter der Begriff der *digitalen Methoden* auf.

„Digitale Gesellschaft“, „digitale Moderne“, „digitale Demenz“, „digitale Revolution“, „digitaler Journalismus“, „digitale Demokratie“. Nun also auch noch „digitale Methoden“? In den vergangenen Jahren wurde häufig der Versuch unternommen, durch die Verwendung des Adjektivs „digital“ zusätzliche Aufmerksamkeit für ein Thema herzustellen: Ein in sich breit angelegter Begriff, der mit Wandel assoziiert ist (oder damit assoziiert werden soll), wird mit dem Zusatz „digital“ versehen. Im Optimalfall beginnt der Begriff noch mit einem „D“, so dass eine sprachlich attraktive Alliteration entsteht. Dabei hatte der Zusatz „digital“ dann seine Berechtigung, wenn damit tatsächlich ein Umstand charakterisiert wurde, der sich in seiner Form vom „analogen Zeitalter“ (ein weiterer aufgeladener Ausdruck) unterschied. Oft aber war das Adjektiv überflüssig, denn „digital“ erklärte einzig und allein, dass es um „irgend etwas mit Computern“ ging – und nicht mehr. Heute haben wir in unserem Verständnis „digital“ die digitalen Offline-Medien hinter uns gelassen und verbinden den Ausdruck nur mehr mit Phänomenen aus dem Internet beziehungsweise dem World Wide Web.

„Digital“ ist also ein Modewort, ähnlich wie es „cyber“ oder „virtuell“ ab Ende der 1990er Jahre war. Heute klingen Worte wie „Cyberspace“, „Cyberwar“, „Cyber Monday“ oder „virtuelle Realität“ altbacken, auch wenn sie noch in Verwendung sind. Digital macht da nach wie vor einen dynamischeren Eindruck. Wie steht es also um das „digital“ als Addendum zu „Methoden“? Handelt es sich dabei auch um nicht mehr als um ein „Buzzword“, das wenig klärt, aber modern klingt? Gibt es überhaupt so etwas wie „analoge Methoden“ in der Erforschung computervermittelter Kommunikation?

Der Begriff „digitale Methoden“ ist eng mit Richard Rogers (2013) und seiner Digital Methods Initiative an der Universität Amsterdam verknüpft. Rogers stellt digitale Methoden dem Begriff der „virtuellen Methoden“ (Hine, 2005) gegenüber. Der methodologische Unterschied liegt dabei im Detail: Während For-

schungen, die virtuelle Methoden einsetzen, darauf abzielen, die Dynamiken und Kulturen von Online-Umgebungen zu ergründen, erfassen digitale Methoden gesamtgesellschaftliche Phänomene durch die Analyse computervermittelter Kommunikation. Auf der einen Seite geht es also stärker darum, das Virtuelle als Raum zu verorten und diesen durch Methoden wie Online-Diskursanalyse oder -Ethnografie auszuloten. Auf der anderen Seite sollen Dynamiken des Webs untersucht werden, um bessere Einblicke in gesamtgesellschaftliche Diskurse entwickeln zu können.

2 Rahmenbedingungen digitaler Methodenentwicklung

Digitale Methoden haben in den vergangenen Jahren einen großen Aufschwung erlebt. Diese Popularisierung hängt nicht zuletzt mit dem starken Plattformfokus des Ansatzes zusammen. Die Konzentration auf Plattformen kommt auch uns Erforschenden dieser Dynamiken gelegen, weil diese Infrastrukturen und ihre Metriken leichter zu operationalisieren sind als die Akteure, die an diesen Kommunikationsprozessen beteiligt sind. Im Unterschied zu analoger Massenkommunikation ist in der vernetzten Plattformkommunikation oft nicht klar, wer mit wem kommuniziert: Individuen, Kollektive, Institutionen, Profis, Laien, Algorithmen und viele andere Akteure sind voneinander abhängig und interagieren miteinander. Die Diversität ist groß, empirisch fundierte Rollen- und Urheberzuschreibungen werden schwierig.

Die Erforschung dieser Dynamiken inmitten solcher verschwimmenden Grenzen ist mit traditionellen kommunikationswissenschaftlichen Methoden in vielen Fällen nur mehr schwer bis gar nicht zu bewerkstelligen. Deshalb müssen Methoden weiterentwickelt und an den Forschungsgegenstand angepasst werden. Dabei lassen sich die Rahmenbedingungen, an denen sich die Entwicklung digitaler Methoden orientieren, unserer Ansicht nach in folgende fünf Dimensionen zusammenfassen:

1. Methodenanpassungen aufgrund der Datenmenge

Mediale Pluralität und Diversität haben dazu geführt, dass die Anzahl potenziell relevanter Daten ins vermeintlich Unmessbare steigen kann: Sei es die Menge an

Mitteilungen, die auf Twitter, Facebook oder in Kommentarbereichen zu einem Thema abgesetzt wird und die sich sekundlich verändert, sei es die Menge an unterschiedlichen Akteuren, die zu einem Thema Stellung beziehen, oder sei es die Vielzahl an unterschiedlichen Plattformen, die zur öffentlichen Kommunikation genutzt werden – Big Data unterscheiden sich in vielerlei Hinsicht von „herkömmlichen“ Daten (boyd & Crawford, 2012; Burrows & Savage, 2014). In der empirischen Forschung führt dies in erster Linie zu Problemen bei a) der Stichprobenziehung, b) der Speicherung und Verarbeitung und c) der Auswertung und Aufbereitung von Datenmengen dieser Größe: Müssen überhaupt noch Stichproben aus diesem Material gezogen werden und wenn ja, welche Verfahren sind hier anzuwenden? Manuell jeden zehnten Tweet aus einer sekundlich steigenden Menge an Twitter-Nachrichten zu einem Thema abzuspeichern ist zeitraubend und unökonomisch. Wie können diese riesigen Mengen an Daten überhaupt gesichert und verarbeitet werden? Denn die „herkömmliche“ Speicherung auf dem Rechner des Forschers stößt hier schnell an Grenzen. Und wie können und sollen Daten dieser Menge ausgewertet und deren Ergebnisse visualisiert werden? An welchen Stellen benötigen wir Automatisierung? Und lassen sich die bekannten, statistischen Verfahren überhaupt auf solche Datenmengen übertragen (Stichwort „Signifikanztests“)?

2. Methodenanpassungen aufgrund der Unstrukturiertheit von Kommunikationsartefakten

Unmittelbar verknüpft mit dem Mengenproblem ist die Frage, wie die Vielfalt der Artefakte, die durch soziale Kommunikationsprozesse im Netz entsteht, einzuordnen ist und wie sie analysierbar gemacht werden kann. Das methodische Hauptproblem liegt dabei in der Unstrukturiertheit der Daten. Während beispielsweise „herkömmliche“ Nachrichten leicht anhand ihrer Merkmale beispielsweise in Meldung, Kommentar oder Glosse eingeordnet werden können, werfen Artefakte sozialer Kommunikationsprozesse neue Fragen auf: Wer ist der Kommunikator eines von einer Privatperson über Facebook geteilten Links zu einer Tagesthemennachricht? Die Privatperson oder „die Tagesthemennachricht“? Und welche Auswirkungen hat es, wenn die Privatperson den Inhalt des Beitrags kommentiert? Ist der Beitrag dann Meldung oder Kommentar? Problematisch ist das Fehlen dieser Systematisierungsansätze insbesondere deshalb, weil die Analyse großer Datenmengen

manuell nicht mehr zu leisten ist. Doch gerade Computerprogramme benötigen verlässliche Kriterien, durch die die Daten automatisiert aufbereitet und systematisiert werden können. Die empirische Kommunikationswissenschaft wird sich daher zunehmend solchen Systematisierungsansätzen widmen müssen.

3. Methodenanpassungen aufgrund von Individualisierung

Auch das zunehmend individualisierte Nutzungsverhalten im Netz stellt Herausforderungen an die empirische Kommunikationswissenschaft. Dieses wirkt sich insbesondere auf die nutzerseitige Analyse von Rezeptionsprozessen aus. Bei einer Untersuchung zur politischen Informiertheit von Bildzeitungs- und SZ-Leserinnen und -Lesern können Forschende mit hoher Wahrscheinlichkeit auf den Medien-Input schließen, den diese beiden Gruppen bekommen haben und muss „nur“ die Unwägbarkeit der Selektion und gegebenenfalls einige situative Faktoren wie zeitliche Restriktionen in Kauf nehmen. Im Gegensatz dazu kann der Input, den Rezipientinnen und Rezipienten auf einer Webseite oder auch in den Suchergebnissen von Suchmaschinen erhalten, erheblich von Person zu Person variieren. Insbesondere bei den großen Anbietern sorgen Personalisierungsstrategien dafür, dass jeder Nutzerin und jedem Nutzer ein individuelles Potpourri an Informationen zusammengestellt wird – und somit auch den Forschenden, deren Analyseobjekte sich von denjenigen unterscheiden können, die den Rezipientinnen und Rezipienten in ihrer seiner Stichprobe präsentiert wurden.

4. Methodenanpassungen aufgrund der Nicht-Zugänglichkeit von Daten

Ein Resultat dieser Individualisierung ist, dass viele Daten den Forschenden nicht oder nur schwer zugänglich sind. Wie erhält man beispielsweise Einsicht in die hochgradig individualisierten Newsfeeds der Facebook-User? Wie kann man wissen, welche Inhalte einem User auf einer Website angezeigt wurden, die mit Personalisierungsstrategien arbeitet? Sicherlich verfügen die Anbieter und Plattform-Betreiber über diese Daten, doch selbst wenn diese für die Wissenschaft zugänglich gemacht werden – inwieweit kann man diesen Daten vertrauen? Und ist die Nutzung solcher Daten in methodisch-ethischer Hinsicht überhaupt angemessen? Was würde es für die Reputation und die Akzeptanz der Wissenschaft bedeuten, solche Daten zu verwenden?

5. Methodenanpassungen aufgrund der Nicht-Beobachtbarkeit sozio-technischer Determinanten

In der digitalen Kommunikation greifen neue „Akteure“ in soziale Kommunikationsprozesse ein: Algorithmen wirken sich zunehmend auf den Informationsfluss im Netz aus, zum Beispiel indem der Algorithmus einer Suchmaschine darüber entscheidet, welche Informationen einem suchenden Nutzer präsentiert werden oder wenn Facebooks News-Feed-Ranking-Algorithmus festlegt, welche Neuigkeiten aus dem individuellen Netzwerk ein User in seiner Timeline angezeigt bekommt. Für die empirische Kommunikationswissenschaft ist die Beforschung dieser „neuen Akteure“ zentral und herausfordernd zugleich. Das methodische Hauptproblem liegt darin, dass die Funktionsweise dieser Algorithmen nicht offen liegt und sie sich damit weder inhaltsanalytisch noch durch Beobachtung erforschen lassen. Auch sind mit ihnen enorme wirtschaftliche Interessen verbunden, weswegen eine Befragung ihrer Programmierer bestenfalls von geringer Validität wäre. Damit stehen die empirischen Wissenschaften vor der Frage, wie sich die Funktionsweise dieser Algorithmen analysieren lässt und ob dies mit dem bekannten Methodenkanon überhaupt zu realisieren ist.

3 Interdisziplinarität und ethische Standards

Die Kommunikationswissenschaft steht allerdings nicht alleine vor diesen Herausforderungen. Auch Disziplinen wie die Informatik, die Informationswissenschaft, die Linguistik, die Medienwissenschaft oder Data Science sind in ihren Forschungen mit den oben beschriebenen Problemen konfrontiert. In den Geisteswissenschaften hat sich mit den Digital Humanities in den vergangenen Jahren eine fächerübergreifende Interessensgemeinschaft gebildet, die sich über die Anwendung technologieintensiver Methoden definiert und austauscht. Diese Parallelentwicklungen bedeuten für die Kommunikationswissenschaft eine große Chance zur Verbesserung der eigenen Forschungsstrategien und -methoden.

Bei all dem Fokus auf technologiezentrierte Methoden sollten wir nicht aus den Augen verlieren, dass sich auch andere Domänen abseits der Wissenschaft für das Erfassen öffentlicher Online-Kommunikationsprozesse interessieren: Plattformbetreiber nutzen Ansätze, die digitalen Methoden ähneln, um ihre Pro-

dukte zu verändern und etwa Effekte von Werbung zu perfektionieren. Auch in diversen praxisnäheren Feldern wie zum Beispiel dem Datenjournalismus und der Marktforschung kommen digitale Methoden zum Einsatz. Und nicht zuletzt werden diese Methoden skaliert und hinter verschlossenen Türen von Geheimdiensten zur Überwachung großer Bevölkerungsteile eingesetzt, wie wir seit den Enthüllungen durch Edward Snowden wissen. Es liegt an uns Forschenden, die Ethik- und Privatssphäre-Standards hoch zu setzen und gegen das „anything goes“, das vielerorts praktiziert wird, Position zu beziehen. Etwa bedeutet die Öffentlichkeit von Kommunikationsinhalten nicht automatisch, dass sie akquiriert und analysiert werden dürfen.

4 Dieser Band: Theoretische, ethische und forschungspraktische Perspektiven auf digitale Methoden

Dieser Band versammelt 13 Aufsätze, die sich der Anwendung digitaler Methoden und den damit verknüpften Herausforderungen und Problemfeldern widmen. Es finden sich sowohl theoretische und ethische Auseinandersetzungen mit der Thematik wie auch Aufsätze, die empirische Forschungsarbeiten dokumentieren. Bei letzteren gehen die Autorinnen und Autoren ausführlicher als üblich auf ihre Methoden ein. Sie beschreiben ihre Auseinandersetzungen mit digitalen Objekten und ihre Schwierigkeiten, in diverse Blackboxes zu blicken. Selbstreflexiv erklären sie, wo sie an Grenzen stoßen und wofür sich digitale Methoden gut eignen.

Abschnitt 1: Theoretische und ethische Perspektiven auf digitale Methoden

Merja Mahrt diskutiert in ihrem Beitrag die Frage, in welchem Verhältnis Theorie und Empirie angesichts von „Big Data“ zukünftig stehen werden. Dabei fokussiert sie sowohl die Perspektive der Skeptiker, die befürchten, dass Big-Data-Analysen zu einer Theorielosigkeit der Forschung führen könnten, als auch die der Verfechter von Big-Data-Studien. Als Synopsis erarbeitet sie Vorschläge für eine fundierte Einbindung von Theorie in die empirische Big-Data-Forschung und erläutert ihre Ausführungen an Beispielen des Agenda-Setting sowie der Sozialkapital- und Diffusionsforschung.

Mit dem Thema Big Data befasst sich auch der Beitrag von *Nele Heise*. Allerdings beleuchtet sie die ethischen Implikationen und geht insbesondere der Frage nach, welchen Status die Nutzenden in dieser neuen Forschungsrichtung haben. Mit Big Data werden die forschungsethischen Probleme nicht gerade kleiner.

Ethische Aspekte stehen ebenfalls bei *Christian Pentzold* im Fokus. Am Beispiel der teilnehmenden Beobachtung zeigt er auf, mit welchen forschungsethischen Prämissen und Problemfeldern die qualitative Forschung zu Kommunikationsprozessen auf Online-Plattformen verknüpft ist. Besonderes Augenmerk liegt dabei auf den Spannungsfeldern, die sich zwischen der (bedingten) Öffentlichkeit der Daten auf solchen Plattformen und der Zusicherung von Anonymität in der empirischen Forschung, sowie der auf manchen Plattformen zugesicherten Anonymität der Autoren und dem Einholen der informierter Zustimmung der Beforschten ergeben können.

Julian Ausserhofer befasst sich in seinem Beitrag mit dem Datenjournalismus. Er zeigt auf, dass Digital Methods in diesem zunehmend bedeutender werdenden Zweig des Journalismus fester Bestandteil der alltäglichen Arbeitsroutinen sind. Im Mittelpunkt seines Beitrags steht die Frage, ob und was von den Erfahrungswerten der Praktikerinnen und Praktikern in die kommunikationswissenschaftliche Methodenentwicklung übertragen werden kann. In elf semistrukturierten Interviews mit DatenjournalistInnen untersucht er die im Datenjournalismus eingesetzten digitalen Methoden und leitet aus ihnen Implikationen für die Weiterentwicklung des sozialwissenschaftlichen Methodenkanons ab.

Abschnitt 2: Strukturanalysen mittels digitaler Methoden

Hier analysieren *Axel Maireder* und *Stefan Schlögl* zunächst die Struktur der Follower-Verknüpfungen von TwitternutzerInnen. Verdichtungen innerhalb von Twitter verstehen die Autoren hier als (Teil-)öffentlichkeiten, deren Analyse Einblicke in die Partizipationsstrukturen in öffentlichen Kommunikationsprozessen ermöglicht. Anhand zweier Beispiele aus dem Feld der politischen Kommunikation legen sie die Vorgehensweise bei der Datensammlung, -aufbereitung und -analyse dar und diskutieren sie vor dem Hintergrund des methodologischen Konzepts.

Um Twitter geht es auch in der Studie von *Axel Bruns* und *Theresa Sauter*. Sie zeigen auf, wie bestehende Netzwerkanalysemethoden ergänzt werden können, um Erfassung, Analyse und Visualisierung von Retweet-Ketten zu ermöglichen.

Am Beispiel eines Videos zu einer Wutrede der australischen Premierministerin Julia Gilliard zeichnen sie die Diffusion des Videos auf Twitter nach. Hierzu sammeln sie automatisiert Daten der jeweiligen Twitter-Nutzerinnen und -Nutzer, die an der Weiterleitung des Videos beteiligt waren und verdeutlichen, wie sich unter Zuhilfenahme von automatisierten Methoden auch der zeitliche Verlauf der Dissemination nachzeichnen lässt.

Jürgen Grimm und *Christiane Grill* stellen Aspekte einer Methode vor, die sie Twitter-Aggregatanalyse nennen. Das Ziel dieses Ansatzes ist es unter anderem, zusammenhängende Konversationen auf Twitter besser erfassen zu können. Durch den kombinierten Einsatz automatisierter und manueller Sampling-Methoden versuchen sie, zur Standardisierung in der Twitterforschung beizutragen.

Am Beispiel des deutschen Kinomarktes stellen *Felix Sattelberger* und *Wolfgang Seufert* die Frage, ob sich mittels Social-Media-Analysen die Absatzprognosen für Vertrauensgüter verbessern lassen. Zur Beantwortung erarbeiten sie ein methodisches Design, in dem sie automatisiert Screenshots von Social-Media-Kommunikation sowie klassischer Pressekritiken (=institutionalisierter Kommunikation) über demnächst anlaufende Filme erheben. Die Dynamik dieser beiden Kommunikationsformen modellieren sie mithilfe latenter Wachstumskurven. Die so gewonnenen Daten integrieren sie in einem nächsten Schritt in ein Pfadmodell zur Beurteilung der Wechselwirkungen beider Kommunikationsprozesse. Ihre Ergebnisse zeigen unter anderem, dass zwischen der Social-Media-Kommunikation und der institutionalisierten Kommunikation vielfältige Wechselbeziehungen bestehen.

Abschnitt 3: Digitale Experimente

Christin Hildebrandt, *Christina Schumann* und *Jens Wolling* gehen der Frage nach, wie sich mithilfe sozialwissenschaftlicher Methoden Erkenntnisse über die Funktionsweise von Suchmaschinenalgorithmen generieren lassen. Am Beispiel der Social Signals – das sind Links, die durch Empfehlungen in sozialen Netzwerken entstehen – entwerfen sie ein digitalisiertes Experiment, mit dem sie untersuchen, ob und wie diese Social Signals von der Suchmaschine Google als Rankingkriterium verwendet werden. Ihre Studie zeigt, dass sich mit digitalisierten, sozialwissenschaftlichen Methoden durchaus Erkenntnisse über die Funktionsweise von Algorithmen gewinnen lassen.

Im Beitrag von *Pascal Jürgens*, *Birgit Stark* und *Melanie Magin* stehen Personalisierungsalgorithmen im Mittelpunkt. In einem ersten Schritt demonstrieren sie, wieso Personalisierungsstrategien von Online-Anbietern für die kommunikationswissenschaftliche Forschung ein Problem darstellen. Danach stellen sie die Vorgehensweise eines automatisierten Online-Experiments am Beispiel der Google-Suche vor, das Erkenntnisse über Funktionsweise und Inhalte personalisierter Angebote liefert.

Abschnitt 4: Probleme und Herausforderungen digitaler Text- und Inhaltsanalysen

Martin Emmer und *Christian Strippel* befassen sich mit der Stichprobenziehung für Online-Inhaltsanalysen. Sie zeigen, dass hier in der kommunikationswissenschaftlichen Forschungspraxis häufig Suchmaschinen zum Einsatz kommen. Sowohl anhand einer theoretischen Argumentation als auch auf Basis einer empirisch-kooordinierten Google-Suche von 16 Accounts verdeutlichen sie, welche Probleme bezüglich der Konstruktion von Öffentlichkeit sowie der Personalisierungsstrategien der Suchmaschinenbetreiber mit der bisherig gängigen Vorgehensweise verknüpft sind und zeigen Lösungsansätze für diese Probleme auf.

Katrin Jungnickel und *Axel Maireder* widmen sich in ihrer Studie der nutzerseitigen Analyse der Facebook-Timeline. Dem bereits thematisierten Problem der Nicht-Zugänglichkeit dieser Inhalte begegnen sie in einem Methodendesign, das Befragung und Inhaltsanalyse kombiniert: Insgesamt 557 Facebook-Nutzer haben ihnen dazu bis zu fünf externe Links aus ihrer Facebook-Timeline, die sie zuletzt erhalten hatten, in einen Fragebogen kopiert und weitere Fragen, zum Beispiel zur Beziehungsqualität zum jeweiligen Sender beantwortet. Die Links wurden dann inhaltsanalytisch analysiert. Neben der Ergebnisdarstellung fokussiert das Autorenteam auf die methodischen Besonderheiten der Studie, zum Beispiel zur Schwierigkeit der inhaltsanalytischen Link-Klassifikation.

Um Audio-Mining-Technologies geht es im Beitrag von *Michael Eble* und *Daniel Stein*. Angesichts der kontinuierlich steigenden Verbreitung und Bedeutung audio-visueller Online-Inhalte zum Beispiel in publizistischen Medien wie auch auf Social Media Plattformen verdeutlichen die Autoren, dass die manuelle Datensammlung und Aufbereitung zunehmend an Grenzen stößt. Sie erläutern daher, wie dieser Arbeitsschritt anhand automatisierter Verfahren der Sprachanalyse (Audio-Mining) unterstützt und vereinfacht werden kann.

Abschließende Hinweise und Danksagung

Die Beiträge, die dieser Band versammelt, wurden auf der Jahrestagung der Fachgruppe Computervermittelte Kommunikation der Deutschen Gesellschaft für Publizistik- und Kommunikationswissenschaft (DGPK) präsentiert. Sie fand vom 7. bis 9. November 2013 an der Universität Wien zum Thema „Digital Methods“ statt. Die ursprünglichen Abstracts sind unter <http://www.univie.ac.at/digitalmethods/> abrufbar. Die Einhaltung geschlechtergerechter Formulierungsstandards wurde den Autorinnen und Autoren selbst überlassen. Kathrin Braun war für den organisatorischen Support dieses Sammelbands verantwortlich. Mark Dang-Anh übernahm die Begutachtung eines Beitrags. Joseph Robinson lektorierte die englischen Übersetzungen der Zusammenfassungen. Vielen Dank an alle Beitragenden!

Dr. Christina Schumann ist wissenschaftliche Mitarbeiterin am Fachgebiet „Empirische Medienforschung und Politische Kommunikation“ der Technischen Universität Ilmenau

Mag. (FH) Julian Ausserhofer ist wissenschaftlicher Mitarbeiter am Institut für Journalismus und Public Relations der FH Joanneum Graz und Doktorand am Institut für Publizistik- und Kommunikationswissenschaft der Universität Wien

Dr. Axel Maireder ist wissenschaftlicher Leiter des Social Media Intelligence Center der GfK und war bis Ende 2014 Mitarbeiter am Institut für Publizistik- und Kommunikationswissenschaft der Universität Wien

Prof. Dr. Monika Taddicken ist Professorin für Kommunikations- und Medienwissenschaften an der Technischen Universität Braunschweig

Quellenverzeichnis

- boyd, d., & Crawford, K. (2012). Critical Questions for Big Data: Provocations for a Cultural, Technological, and Scholarly Phenomenon. *Information, Communication & Society*, 15(5), 662-679. doi: 10.1080/1369118X.2012.678878
- Burrows, R., & Savage, M. (2014). After the Crisis? Big Data and the Methodological Challenges of Empirical Sociology. *Big Data & Society*, 1(1), 2053951714540280. doi: 10.1177/2053951714540280
- Hine, C. (2005). *Virtual Methods*. Oxford: Berg.
- Rogers, R. (2013). *Digital Methods*. Cambridge MIT Press.

I THEORIE UND ETHIK

Empfohlene Zitierung: Mahrt, M. (2015). Mit Big Data gegen das „Ende der Theorie“? In A. Maireder, J. Ausserhofer, C. Schumann, & M. Taddicken (Hrsg.), *Digitale Methoden in der Kommunikationswissenschaft* (S. 23-37). doi: 10.17174/dcr.v2.2

Zusammenfassung: Die Analyse großer Datensätze aus Online-Plattformen („Big Data“) wirft eine Reihe von methodologischen Fragen auf, die insbesondere das Verhältnis von Theorie und Empirie betreffen. Während an vielen Stellen Begeisterung für neue, datenorientierte Forschungsmöglichkeiten zu hören ist, befürchten andere ein „Ende der Theorie“. Der Beitrag zeigt verschiedene Aspekte dieses Problemereichs auf und macht Vorschläge, wie Big-Data-Studien theoretisch gehaltvoll gestaltet werden können. Beispiele aus der Agenda-Setting-, Sozialkapital- und Diffusionsforschung werden diskutiert. Interdisziplinäre Zusammenarbeit und systematische Validierung von Big-Data-Verfahren erscheinen dabei als zentrale Desiderate. Welchen Stellenwert Big-Data-Ansätze in der kommunikationswissenschaftlichen Forschung und Lehre in Zukunft haben werden, hängt mutmaßlich auch von ihrer Fähigkeit zu theoretischer Innovation ab.

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

Merja Mahrt

Mit Big Data gegen das „Ende der Theorie“?

In den letzten Jahren hat sich eine Reihe kommunikationswissenschaftlicher Zeitschriften und Tagungen mit „Big Data“ befasst. Die Verfügbarkeit großer Mengen an Daten aus Online-Plattformen zur Informationssuche, Rezeption medialer Inhalte und interpersonalen Kommunikation hat neue Arten der Analyse hervorgebracht, die allerdings im Fach nicht unumstritten sind. Der vorliegende Beitrag beleuchtet vor allem die Diskussion um den theoretischen Gehalt von solchen oft eher datengetriebenen Untersuchungen und versucht aufzuzeigen, wie Big-Data-Verfahren auch einen theoretischen Mehrwert erbringen können. In einem ersten Schritt werden Hoffnungen zusammengefasst, die sich mit Big Data verbinden. Das zweite Kapitel beschreibt drei Problembereiche, die sich aus typischen Vorgehensweisen in Big-Data-Untersuchungen ergeben. Zu diesen werden daraufhin Lösungen vorgeschlagen und abschließend einige Annahmen zur Zukunft von Big-Data-Forschung in der Kommunikationswissenschaft aufgestellt.

1 Euphorie für Big Data

Mit dem Schlagwort Big Data verbinden sich verschiedene Formen der Analyse „großer“ Datensätze. Im ursprünglichen, informatischen Sinn handelt es sich dabei um Datenmengen, die zu groß sind, um sie mit üblicher Hard- und Software zu verarbeiten (Manovich, 2012). Die absolute Größe von Daten, die als

Big Data bezeichnet werden, wächst dadurch im Laufe der Zeit. Für die Sozialwissenschaften verbindet sich mit dem Begriff in der Regel ein bestimmter Typ von Datensatz, der nach informatischen Maßstäben nicht unbedingt groß sein muss, aber dennoch deutlich umfangreicher sein kann als etwa übliche Befragungs- oder inhaltsanalytische Datensätze.

Vor allem geht es in der sozial- und speziell kommunikationswissenschaftlichen Diskussion um den Begriff aber um die neuartige Herkunft und Struktur dieser Daten. Während klassischerweise Daten mittels eines im Forschungsprozess entwickelten Messinstruments (Interviewleitfaden, Fragebogen, Codebuch, Beobachtungsschema...) erfasst werden, nutzen viele Big-Data-Studien Daten, die nicht speziell zum Zweck der Analyse entstanden sind. Sie fallen etwa als Spuren der Nutzung digitaler Technologien und Medien an. Häufig unterscheidet sich daher ihre Fall- und/oder Variablenstruktur von typischen sozialwissenschaftlichen Datensätzen. Es können für einen Nutzungsvorgang zum Beispiel sehr viele Datenpunkte in kurzem Abstand erfasst werden, die einen Zeitverlauf sehr engmaschig dokumentieren. Wenn es sich um Kommunikationsdaten handelt, ergeben sich häufig außerdem Kontakt- oder Netzwerkstrukturen. Beide Arten von Strukturen lassen sich mit klassischen Erhebungsverfahren nur sehr aufwändig in ähnlicher Detailgenauigkeit oder Menge erfassen. Auf der anderen Seite sind in Big-Data-Studien häufig die Informationen nicht verfügbar, die man mit üblichen Verfahren als unabhängige oder Kontrollvariablen über die Personen erheben würde, deren Verhalten untersucht wird. Groß sind Big-Data-Datensätze in diesen Fällen dann nur in einer Dimension der Datenmatrix: Es können zwar sehr viele Fälle, aber oft nur wenige Variablen, also Informationen pro Fall, untersucht werden.

Während etwa Telefongesellschaften oder Einzelhandelsunternehmen durch den Einsatz elektronischer Verfahren schon seit längerem Kommunikations- bzw. Konsumverhalten detailliert erfassen und analysieren können (Savage & Burrows, 2007), hat die sozialwissenschaftliche Diskussion um ähnliche Ansätze etwas später eingesetzt. Sie kam insbesondere mit der steigenden Nutzung von Online-Plattformen auf, die Spuren von Nutzerverhalten entweder öffentlich machen oder über die diese Spuren durch Schnittstellen auslesbar werden. Von der angewandten Markt- und Meinungsforschung werden diese Möglichkeiten nach wie vor als große Verheißung angepriesen, die zum Beispiel ein besseres, schnelleres oder umfassenderes Verständnis von Kundenwünschen und -verhalten erlaubten (z. B. Kearon & Harrison, 2011; Russom, 2011).

Dagegen ist die Reaktion der Sozialwissenschaften vielschichtiger: Auf der einen Seite lassen sich auch hier Euphorie für neue Forschungsmöglichkeiten und Datenstrukturen verzeichnen (Lazer et al., 2009; Miller, 2011). Auf Tagungen und in jüngeren Publikationen finden sich in den letzten Jahren so immer wieder Untersuchungen, die mit großen Datensätzen aus Online-Plattformen arbeiten, etwa Twitter-Meldungen, Google-Suchanfragen oder Anzahl von „Freunden“ auf Facebook. Auf der anderen Seite hat sich aber auch eine kritische Diskussion um Big-Data-Forschung dieser Art entwickelt. Diese beleuchtet etwa methodische Unterschiede zwischen traditioneller Sozialforschung und Big-Data-Ansätzen und Implikationen, die sich hieraus für die Aussagekraft der gewonnenen Befunde ergeben (boyd & Crawford, 2012; Couper, 2013; Mahrt & Scharkow, 2013; Manovich, 2012; Tinati, Halford, Carr, & Pope, 2014). Auch rücken neue ethische Probleme in den Vordergrund, die etwa die fehlende Einwilligung der Nutzer, deren Verhalten untersucht wird, oder den Schutz privater Daten betreffen (siehe Heise in diesem Band; Markham & Buchanan, 2012; Zimmer, 2010). Neben solchen eher forschungspraktischen Aspekten werden in Bezug auf Big Data aber vor allem theoretische Fragen diskutiert. In verschiedenen Variationen ist hierbei von einem „Ende der Theorie“ die Rede.

2 Big Data und das „Ende der Theorie“

Als Theorien werden in den Sozialwissenschaften streng genommen in sich widerspruchsfreie Systeme von Aussagen verstanden, aus denen sich Hypothesen ableiten lassen (Schnell, Hill, & Esser, 2005). Auf dieses Verständnis von Theorie bezieht sich auch die Diskussion um deren (vermeintliches) Ende durch Big Data. In einer positiven Sichtweise machen die neuen großen Datensätze die Beschäftigung mit Theorien überflüssig oder stellen zumindest ihre traditionelle, hypothesengenerierende und forschungsleitende Rolle in Frage. Anderson (2008) etwa beschreibt, wie mit ausreichend großen Mengen an Daten auch Korrelationen für Wissenszuwachs sorgen können, anstatt aus Theorien Annahmen abzuleiten und anschließend entsprechende Kausalmodelle zu prüfen. Durch die Fülle an Daten und heute verfügbare Rechnerkapazität, diese zu analysieren, werde der übliche Prozess der Aufstellung und Überprüfung von Hypothesen durch letztlich aussagekräftigere datengetriebene Befunde ersetzt. Während theoretisch abgelei-

tete Modelle immer vereinfachen und dadurch auch verfälschen, könnten nach Anderson schier unüberschaubare Mengen an Daten die Notwendigkeit solcher Modelle (die er als wenig mehr als Hilfskonstruktionen darstellt) überflüssig machen. Echte – im Sinne von: nicht durch ein Modell um ihre natürliche Komplexität reduzierte – und in unbeschränkter Menge verfügbare Daten sieht er daher gegenüber traditionell erhobenen im Vorteil.

Ein konkretes Beispiel soll diese Vorgehensweise illustrieren: In seinem Beitrag zur Abschlussdiskussion der ICA-Konferenz in Phoenix 2012 hat Bailenson (2012) eine entsprechende Studie aus der Kommunikationswissenschaft vorgestellt. Um das Verhalten von Lehrenden im Klassenraum und die Reaktionen von Schülerinnen und Schülern darauf zu untersuchen, hat er Sensorleisten aus Computerspielen eingesetzt, die Körperhaltungen und Bewegungen im Raum aufzeichnen können. Die resultierenden Daten sind in Zeitverlauf und Detailreichtum wesentlich umfangreicher als bei klassischen Beobachtungsverfahren. Dadurch konnte der Forscher Zusammenhänge zwischen Lehrer- und Schülerverhalten entdecken, die anderweitig nicht erfassbar gewesen wären. Es ergaben sich Korrelationen zwischen minimalen Unterschieden in Haltung und Bewegung der Lehrenden sowie der Reaktionen der Schülerinnen und Schüler. Bailenson suchte also tatsächlich in riesigen Datenmengen nach Mustern und Zusammenhängen, die in einer kleineren Untersuchung entweder nicht erfasst oder im Rauschen von Messfehlern und statistischer Irrtumswahrscheinlichkeit übersehen worden wären. Provokativ spitzte er zu, dass Theorie im Forschungsprozess damit überflüssig sei, denn auf Basis des vorhandenen Wissens hätte er niemals zu Hypothesen über die sich als relevant erweisenden Verhaltensaspekte und ohne den Big-Data-Ansatz auch nicht zu vergleichbaren Befunden gelangen können.

In seiner Keynote zur „Digital Methods“-Tagung hat Jürgen Pfeffer (2013) ein solches Vorgehen als typisch für die Arbeit mit Big Data beschrieben: Das Prinzip des klassischen Forschungsprozesses (von ihm in folgender Reihenfolge skizziert: Problem, Fragestellung/Hypothesen, Theorie, Methode, Daten, Analyse, Ergebnispräsentation) sei quasi auf den Kopf gestellt, weil erstens mit der Methode begonnen werde, woraus sich zweitens Daten ergäben, die drittens einer Analyse unterzogen würden, deren Ergebnisse dann viertens präsentiert würden. Und erst daran schließe sich fünftens die Darstellung eines Problems an, auf das die Untersuchung eine Antwort gegeben habe. Wohlgermerkt geht es in Pfeffers Worten hier um ein gesellschaftliches oder anderes *Problem*. Ein

Bezug zu *Theorie* im eigentlichen Sinn wird in solchen Forschungsarbeiten nicht notwendigerweise hergestellt.

Es mag genügend gesellschaftliche Probleme geben, über die man mit einem solchen Vorgehen Befunde liefern kann. Und nicht jede Untersuchung muss zur Theorieentwicklung beitragen wollen. Ein Vorgehen wie von Pfeffer zusammengefasst schließt aber theoretische Fortschritte tatsächlich aus, weil es ohne jeglichen Bezug zu Theorie auskommt. Im Gegensatz zum induktiven Vorgehen im herkömmlichen sozialwissenschaftlichen Verständnis wird an die Auswertung der Daten eben keine Theoriediskussion angeschlossen. Dieses erste „Ende der Theorie“ wird wie oben beschrieben von einigen begrüßt und kann durchaus zu neuen Erkenntnissen führen. Im eigentlichen Sinn wird Theorie aber nicht weiterentwickelt, weil man sich bei dieser Variante des Forschungsprozesses nicht mit übergeordneten widerspruchsfreien Systemen von Aussagen beschäftigt und keine Gesetzmäßigkeiten oder Regelmäßigkeiten jenseits des untersuchten Problems ableiten möchte.

Daneben existiert noch eine zweite Variante des „Endes der Theorie“, die sich zuweilen in Big-Data-Studien zeigt. Dies ist dann der Fall, wenn es zwar dem Wortlaut nach in einer Untersuchung Bezüge zu bestehenden Theorien gibt, diese aber durch die methodische Umsetzung nicht wirklich eingelöst werden können. Insbesondere drei Bereiche erscheinen hierbei problematisch.

(1) Aussagekraft von Big-Data-Stichproben

Big-Data-Analysen mit online abrufbaren Daten sind häufig auf ad-hoc-Stichproben angewiesen, deren Repräsentativität und Vergleichbarkeit mit anderen Stichproben unklar ist. Zum einen gibt es in der Regel keine Möglichkeit, Verteilungen der Stichprobe mit bekannten Parametern der Grundgesamtheit zu vergleichen. Zum anderen sind die genauen Bedingungen der Datenerhebung häufig nicht reproduzierbar und können daher nur schwer überprüft werden (Gaffney & Puschmann, 2014). Dies schränkt die Aussagekraft der daraus abgeleiteten Befunde ein, etwa in einer Agenda-Setting-Studie mit Twitter-Daten: Um die Erfassung einer Publikumsagenda haben sich in der Agenda-Setting-Forschung eine umfassende Diskussion und darauf aufbauend verschiedene Standardbefragungsinstrumente entwickelt (Dearing & Rogers, 1996). Ein Big-Data-Ansatz kann dagegen darin bestehen, Twitter-Meldungen für die Messung der Wichtigkeit von Themen

in der Bevölkerung zu verwenden (z. B. Verdegem & D'heer, 2013). Je nach Art des Zugriffs auf die Daten gibt es aber mehrere Stellen, an denen die Plattform Twitter möglicherweise bereits eine Auswahl aus allen Meldungen trifft, die das Aufgreifkriterium erfüllen. Dies ist in der Regel für Forschende nicht transparent (Gaffney & Puschmann, 2014).

Ein weiteres Stichprobenproblem ergibt sich aus der Frage, wofür die Twitter-Meldungen stehen, wen oder was sie repräsentieren sollen. Welche Begriffe oder Personen auf Twitter im Zeitverlauf wie häufig erwähnt werden, kann in der Tat Aufschluss über deren sich verändernde Wichtigkeit geben. Aber dies betrifft nur Twitter-Nutzerinnen und -Nutzer, die keinen repräsentativen Ausschnitt der Bevölkerung bilden (Busemann & Gscheidle, 2012). Eine Liste von Twitter-Themen ist damit nicht ohne Weiteres mit einer über repräsentative Befragung erhobenen Publikumsagenda vergleichbar. Selbst wenn Begriffe aus etablierten Theorien verwendet werden (wie Publikumsagenda), sind die Ergebnisse von Big-Data-Studien damit nicht unbedingt an die traditioneller Untersuchungen anschlussfähig. Die Stichproben sind zwar häufig groß, müssen aber nicht zwingend auch aussagekräftig sein. Dadurch wird der theoretische Bezug nur ein scheinbarer, und es wäre möglicherweise besser, die Daten schlicht in ihrem ursprünglichen Sinn zu beschreiben (bspw. als Themenkarriere in Twitter), anstatt sie mit theoretischer Bedeutung zu überfrachten, die sie nicht einlösen können (weil Twitter-Trends eben nicht als Indikator für breitere Themeninteressen der Bevölkerung stehen können; siehe auch Mitchell & Hitlin, 2013).

(2) Messprobleme in Big-Data-Untersuchungen

Ein zweites Problem kann entstehen, wenn Maßzahlen, die aus Online-Plattformen einfach zu generieren sind, eins zu eins auf komplexe theoretische Konstrukte übertragen werden. Beispielsweise kann man in sozialen Netzwerkseiten (etwa Facebook) in der Regel die Anzahl der Kontakte einer Nutzerin oder eines Nutzers einsehen. Dies als alleinigen Indikator für das Sozialkapital einer Person anzunehmen (wie z. B. bei Brooks, Welser, Hogan, & Titsworth, 2011), wird diesem Begriff allerdings wohl kaum gerecht, weder theoretisch (Adler & Kwon, 2002), noch in Anbetracht der Befunde zu Sozialkapital und Facebook-Beziehungen (Burke, Kraut, & Marlow, 2011; Yoder & Stutzmann, 2011). Eine Studie, die Kontakte in sozialen Netzwerkseiten mit Sozialkapital gleichsetzt, lässt sich damit nur

schwer mit dem Forschungsstand zu Sozialkapital in Beziehung setzen. Ähnliche Probleme haben alle Big-Data-Studien, die auf Daten beschränkt sind, die von Online-Plattformen vorstrukturiert worden sind. Die Art und Weise, wie Eigenschaften oder Verhalten von einer Plattform erfasst werden, lässt sich nur bedingt auf das Theorieinventar der Sozialwissenschaften übertragen und ist nicht mit der üblichen Vorgehensweise sozialwissenschaftlicher Operationalisierung von Konstrukten vergleichbar. Die Vorstrukturierung in Kategorien durch eine Plattform ist aber Voraussetzung dafür, dass große Mengen an Daten ausgelesen werden können. Wenn die Bedeutung dieser somit vorgegebenen Kategorien aber unklar ist, bleibt auch hier offen, was diese Daten jenseits des rein deskriptiven Werts eigentlich bedeuten – insbesondere für die Theorie, in deren Kontext sie erhoben werden sollen (boyd & Crawford, 2012; Mahrt & Scharnow, 2013).

(3) Fehlende theoretische Bezüge in interdisziplinären Forschungsfeldern

Ein dritter Problembereich ist nicht auf Big-Data-Forschung beschränkt, sondern entsteht dort häufig durch den besonders stark ausgeprägten interdisziplinären Charakter des Forschungsfeldes Online-Kommunikation. Stellvertretend für viele andere Bereiche kann man das am Beispiel der Diffusionsforschung illustrieren. Eine viel zitierte Studie hat untersucht, welche Rolle die Nähe zwischen Facebook-Nutzerinnen und -Nutzern beim Weiterleiten von Informationen innerhalb der Plattform spielt (Bakshy, Rosenn, Marlow, & Adamic, 2012). Mit 253 Millionen Probanden ist hierzu ein wahrlich großes Experiment durchgeführt worden. Der Titel der Studie, „The role of social networks in information diffusion“, klingt aus kommunikationswissenschaftlicher Perspektive zunächst vertraut; schaut man sich die in der Studie zitierte Literatur an, so fällt dagegen auf, dass hier ganz andere Quellen zitiert werden, als man erwarten würde. Mit Greenberg und Granovetter kommen zwar am Rande auch für dieses Fach einschlägige Forscher vor, der Fokus der theoretischen Auseinandersetzung liegt bei Bakshy et al. aber anderswo. Dies ist ihr gutes Recht und nicht der Kern des Problems. Es geht hier nicht um den Vorwurf, Bakshy oder andere hätten aus der Perspektive der Kommunikationswissenschaft zentrale Studien nicht gelesen. Ebenso kann kein Fach den einzigen (oder einzig richtigen) Zugang zu einem Forschungsgegenstand für sich beanspruchen. Es fällt aber auf, dass mit der (datenorientierten) Online-Forschung wiederum ein Bereich entsteht, in dem kommunikationswissenschaft-

liche Forschung mehr oder weniger unverbunden neben Arbeiten aus anderen Fächern steht, und das selbst in Fällen wie der Diffusionsforschung, in denen ein Gegenstand in zwei Fächern ähnlich untersucht wird.

Wie oben angedeutet, ist dieses Phänomen nicht neu und nicht auf Big-Data-Forschung beschränkt. In der Medizin sind ab den 1980er Jahren unabhängig von der damals bereits existierenden kommunikationswissenschaftlichen Forschung Skalen zur Fernsehnutzung sowie ihren Motiven entwickelt worden (Bickham, 2009). Umgekehrt wird ein Großteil der gut 2.000 in der medizinischen Datenbank *PubMed* verzeichneten Studien zu Fernsehnutzung in unserem Fach eher selten rezipiert. Psychologische Darstellungen zu Einstellungen, kognitiver Dissonanz und selektiver Zuwendung kommen ohne Verweise auf kommunikationswissenschaftliche Forschung aus, während in der Psychologie geläufige Dimensionen von Einstellungen in der Kommunikationswissenschaft nicht berücksichtigt werden (Mahrt, 2014). In solchen Fällen geht Potenzial verloren, Theorie (und auch Empirie) in den betreffenden Fächern weiterzuentwickeln, weil nicht berücksichtigt wird, was in nahen oder auch entfernteren Nachbardisziplinen erarbeitet wird. Dies bedeutet zwar für keines der betreffenden Fächer das Ende der (eigenen) Theorie, aber es werden doch Möglichkeiten zur interdisziplinären theoretischen Weiterentwicklung verschenkt. Dies muss sich in der Big-Data-Forschung nicht im gleichen Maße wiederholen wie in den beiden genannten älteren Forschungsbereichen, und es wäre wünschenswert, dass es nicht so kommt. Die Verfügbarkeit von disziplinübergreifenden Suchmaschinen und Katalogen macht es heute einfacher, auf fachfremde, aber inhaltsverwandte Forschung aufmerksam zu werden. Unter anderem auf diesem Weg könnte auch Big-Data-Forschung zu Online-Kommunikation integrierte Theoriebildung vorantreiben und disziplinäre Parallelentwicklungen vermeiden, wozu im folgenden Abschnitt weitere Anregungen für die drei dargestellten Problembereiche gegeben werden sollen.

3 Theoretische Fortschritte durch Big-Data-Forschung

Für die drei im vorigen Abschnitt skizzierten Problem- und Forschungsfelder sind zum Teil bereits Lösungsansätze erarbeitet worden. In anderen Fällen könnten Arbeiten jenseits der Big-Data-Forschung im hier diskutierten Sinne Anregungen für theoretisch gehaltvollere Forschung sein.

(1) Vergleichbarkeit von Big-Data-Stichproben

Wo Befunde aus Big-Data-Untersuchungen bis dato nicht mit dem Forschungsstand verglichen werden können, bietet es sich natürlich an, Studien zur Vergleichbarkeit durchzuführen. Ob sich Themenkarrieren in Social-Media-Plattformen mit üblichen Verfahren der Erfassung von Agenden vergleichen lassen, ist ja letztlich eine empirische Frage. Groshek und Groshek (2013) beispielsweise vergleichen die Beliebtheit von Nachrichtenthemen auf Twitter und Facebook mit Nachrichten des TV-Senders CNN und der New York Times. Für Facebook ergeben sich deutliche und signifikante Rangkorrelationen, für Twitter dagegen nicht. Auch zwischen der Häufigkeit von Google-Suchanfragen zu einem politischen Thema und der Wichtigkeit dieses Themas in repräsentativen Befragungen desselben Zeitraums gibt es einen Zusammenhang, dieser fällt aber eher schwach aus (Scharnow & Vogelgesang, 2011). Weitere Studien dieser Art werden es hoffentlich in Zukunft ermöglichen, die Aussagekraft von Big-Data-Stichproben und damit auch ihrer Ergebnisse besser einzuschätzen. Es wäre sogar denkbar zu prüfen, ob sich zum Beispiel für bestimmte Substichproben einer Befragung (etwa alle Twitter-Nutzerinnen und -Nutzer) ähnliche Befunde zeigen wie in den Daten, die aus der betreffenden Plattform in großer Zahl ausgelesen wurden. Die bislang vorliegenden vergleichenden Studien verdeutlichen aber auch, dass Nutzungsdaten aus Onlineplattformen eben nicht ohne weiteres für etwas anderes stehen können.

(2) Vergleichbarkeit von Big-Data-Maßzahlen

Die im vorigen Abschnitt vorgestellten Arbeiten erlauben es nicht nur, die Übertragbarkeit von Stichprobenergebnissen zu untersuchen. Auch die von einer Plattform bereitgestellten Kategorien (beliebte Begriffe einer Social-Media-Plattform, Suchanfrage bei Google) können dabei mit üblichen inhaltsanalytischen bzw. Befragungskategorien verglichen werden. Darüber hinaus sind auch systematische Validierungen von Messverfahren möglich und werden für die Big-Data-Forschung eingefordert (boyd & Crawford, 2012; Lazer, Kennedy, King, & Vespignani, 2014). In Abschnitt 2 wurde beispielsweise bezweifelt, dass die Anzahl von Kontakten auf Facebook ein valider Indikator für Sozialkapital sei. Auch dies ließe sich empirisch prüfen. Vorbild dafür könnte zum Beispiel Williams' (2006) Studie sein, in der verschiedene Maße für Online- und Offline-Sozialkapital

entwickelt und mit anderen, etablierten Indikatoren verglichen wurden. Ohne solche Validierungen bleibt letztlich unklar, was die Kategorien aus Online-Plattformen eigentlich erfassen.

(3) *Theoriegeleitete Big-Data-Studien*

Neben vergleichenden Studien ist es auch denkbar, die Vorteile von Big Data für die Weiterentwicklung von Theorien fruchtbar zu machen. Vorbilder dafür können Simulationsstudien sein. Watts und Dodds (2007) beispielsweise prüfen mit großen, aber simulierten Datensätzen theoretische Annahmen und Modelle der Diffusionsforschung. Sie können zeigen, wie unwahrscheinlich ein reiner Zwei-Stufen-Fluss in sozialen Netzwerken ist, bei dem Informationen aus Massenmedien zunächst nur Meinungsführer erreichen, die diese dann an ihre Kontakte (Follower) weitergeben. Wenn man wie Bakshy et al. (2012) privilegierten Zugang zu Facebook-Nutzerdaten hat, wäre es denkbar, in der immensen Stichprobe von 253 Millionen Nutzerkonten ebenfalls nach solchen Mustern zu suchen – in diesem Fall sogar unter natürlichen, alltäglichen Bedingungen. Ob Facebook-Netzwerke in Bezug auf Diffusion mit anderen sozialen Netzwerken vergleichbar sind, lässt sich damit möglicherweise auch einschätzen. Das anders gelagerte Forschungsinteresse von Bakshy et al. ist natürlich auch ohne diese Erweiterung legitim – aber die verschiedenen Fächer, die sich mit ähnlichen Big-Data-Datenquellen oder sogar ähnlichen theoretischen Konstrukten beschäftigen, könnten sich sicherlich stärker befruchten, als es zur Zeit der Fall zu sein scheint – und zwar gegenseitig.

4 Ausblick: Vorschläge für den zukünftigen Umgang mit Big Data

Zusammenfassend erscheinen vergleichende Studien als zentrales Desiderat für die Big-Data-Forschung (nicht nur) in der Kommunikationswissenschaft, wenn sie theoretischen Mehrwert haben und sich von der reinen Deskription lösen soll. Je mehr Studien publiziert werden, umso eher wird es hoffentlich möglich, auf deskriptiven und häufig auch explorativen Untersuchungen aufzubauen, um durch Big-Data-Studien zu Theoriebildung und -weiterentwicklung beizutragen. Daneben kann auch die Zusammenarbeit mit Kolleginnen und Kollegen aus anderen Fächern den theoretischen Mehrwert von Big-Data-Studien

erhöhen. Ein naheliegender Weg besteht darin, den Kontakt zur Informatik oder Informationswissenschaft zu suchen und damit die eigenen methodischen Kompetenzen und auch technischen Ressourcen zu ergänzen. Dies ermöglicht eventuell komplexere Erhebungs- und Auswertungsverfahren, die ebenfalls über Deskription und Einzelfallstudien hinausgehen. Gerade bei der Arbeit mit ungewohnten Datenmengen und -strukturen, was auf Big Data in der Kommunikationswissenschaft in der Regel zutrifft, können fehlende Kenntnisse und Ressourcen ein zentrales Hemmnis für die Aussagekraft einer Untersuchung sein (boyd & Crawford, 2012; Manovich, 2012).

Dabei kann es aber nicht darum gehen, wie Pfeffer (2013) es in seiner Keynote zuspitzte, sich „einen Informatiker im Nebenzimmer zu halten“. Denn auch der theoretische Austausch mit anderen Disziplinen kann die eigene Arbeit befruchten, sei es durch wirkliche Kollaboration (statt Betrachtung fachfremder Kolleginnen oder Mitarbeiter als reine Dienstleister) oder auch schlicht durch Lektüre der Literatur anderer Fächer. Selbst wenn das Forschungsinteresse fachfremder Untersuchungen von Online-Plattformen vom eigenen Theoriebestand weit entfernt ist, können deren Befunde dennoch Anregungen für die Entwicklung von Theorie (und Methode) einer eigenen Studie liefern. Das Phänomen der Diffusion von Online-Inhalten wird, wenn auch mit anderen Begriffen, beispielsweise auch in der Informatik (z. B. Brodersen, Scellato, & Wattenhofer, 2012; Rodrigues, Benevenuto, Cha, Gummadi, & Almeida, 2011) und der Marketing-Forschung untersucht (z. B. Berger & Milkman, 2012; Nelson-Field, Riebe, & Newstead, 2013). Die Beschäftigung mit Big Data aus Online-Plattformen zeigt damit wiederum auf, dass das interdisziplinäre Fach Kommunikationswissenschaft auch jenseits der Sozialwissenschaften Nachbarn hat, deren Arbeit es lohnt, zur Kenntnis zu nehmen.

Letztlich kann auch die fachinterne methodische Ausbildung dazu beitragen, den theoretischen Ertrag von Big-Data-Untersuchungen zu erhöhen (boyd & Crawford, 2012; Manovich, 2012). Eine Reflexion über Big-Data-Verfahren im Lichte üblicher sozialwissenschaftlicher Gütekriterien erscheint dabei insbesondere dringlich (Mahrt & Scharrow, 2013). In den jüngsten Empfehlungen zur Methodenausbildung der DGPK (Matthes et al., 2011) wird Big Data nicht erwähnt. Und sicherlich wird es in absehbarer Zeit nicht die klassischen sozialwissenschaftlichen Verfahren der Datenerhebung und -analyse ersetzen. Es wäre aber denkbar, an Standorten mit forschungsstarken Studiengängen auch Big-Data-Verfahren in die Methodenlehre aufzunehmen. Im Bereich fortgeschrittener quantitativer Analy-

severfahren gibt es zur Zeit keine festen Standards in der Lehre, und neben oder anstelle von Mehrebenenanalyse oder Strukturgleichungsmodellen – wie von Matthes et al. (2011) vorgeschlagen – könnten auch datenorientierte Verfahren stehen, etwa in Masterstudiengängen mit umfangreichem Methodenanteil. Darüber hinaus könnte die Ausbildung in speziellen Analyseverfahren für große Datenmengen auch zu höherem Anspruch bei der Auswertung von Big-Data-Studien führen. Wie oben dargestellt, sehen Big-Data-Enthusiasten wie Anderson (2008) und Bailenson (2012) Korrelationsanalysen mit sehr großen Datensätzen als vielversprechenden Weg zu Erkenntnis, der Kausalanalysen (zumindest in bestimmten Fällen) ersetzen kann. Viele kommunikationswissenschaftliche Big-Data-Studien führen aber keine Korrelationsanalysen durch, sondern präsentieren streng genommen rein deskriptive und zudem häufig univariate Auswertungen. Mit komplexeren Analysen, die Zusammenhänge zwischen verschiedenen Variablen berücksichtigen, könnte sich auch der theoretische Ertrag einer Studie erhöhen.

Big-Data-Forschung ist nicht per se gut oder schlecht, theoretisch gehaltvoll oder theoriearm. Es liegt an der konkreten Ausgestaltung jeder Untersuchung, ob sie z. B. Problembeschreibung oder Theorienentwicklung leisten will – und kann. Die Banalisierung von theoretischen Konstrukten wird hoffentlich nicht das bestimmende Muster von Big Data in der Kommunikationswissenschaft sein. Welche Rolle der Ansatz langfristig in diesem Fach spielen wird, wird aber voraussichtlich auch und gerade von den theoretischen Impulsen abhängen, die durch Verfahren mit großen Datensätzen geleistet werden.

Dr. Merja Mahrt ist wissenschaftliche Mitarbeiterin am Institut für Sozialwissenschaften an der Heinrich-Heine-Universität Düsseldorf

Quellenverzeichnis

- Adler, P. S., & Kwon, S.-W. (2002). Social capital: Prospects for a new concept. *Academy of Management Review*, 27(1), 17-40. doi: 10.5465/amr.2002.5922314
- Anderson, C. (23. Juni 2008). The end of theory: The data deluge makes the scientific method obsolete. *Wired*. Abgerufen von http://archive.wired.com/science/discoveries/magazine/16-07/pb_theory

- Bailenson, J. N. (2012). Beitrag zur Closing Plenary der ICA-Jahrestagung: "The Internet is the end of communication theory as we know it". 62. Jahrestagung der International Communication Association, Phoenix, AZ. Abgerufen von <http://www.icaahdq.org/conf/2012/closing.asp>
- Bakshy, E., Rosenn, I., Marlow, C., & Adamic, L. (2012). *The role of social networks in information diffusion*. Proceedings of the 21st international conference on World Wide Web, Lyon, Frankreich. doi: 10.1145/2187836.2187907
- Berger, J., & Milkman, K. L. (2012). What makes online content viral? *Journal of Marketing Research*, 49(2), 192-205. doi: 10.1509/jmr.10.0353
- Bickham, D. S. (2009). Media effects research as applied science: Social scientists surviving (and prospering) in a newly medicalized field. 59. Jahrestagung der International Communication Association, Chicago, IL. Abgerufen von http://citation.allacademic.com/meta/p297804_index.html
- boyd, d., & Crawford, K. (2012). Critical questions for Big Data. Provocations for a cultural, technological, and scholarly phenomenon. *Information, Communication & Society*, 15(5), 662-679. doi: 10.1080/1369118x.2012.678878
- Brodersen, A., Scellato, S., & Wattenhofer, M. (2012). *YouTube around the world: Geographic popularity of videos*. Proceedings of the 21st international conference on World Wide Web, Lyon, Frankreich. doi: 10.1145/2187836.2187870
- Brooks, B., Welser, H. T., Hogan, B., & Titsworth, S. (2011). Socioeconomic status updates. Family SES and emergent social capital in college student Facebook networks. *Information, Communication & Society*, 14(4), 529-549. doi: 10.1080/1369118x.2011.562221
- Burke, M., Kraut, R., & Marlow, C. (2011). *Social capital on Facebook: Differentiating uses and users*. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Vancouver, BC, Kanada. doi: 10.1145/1978942.1979023
- Busemann, K., & Gscheidle, C. (2012). Web 2.0: Habitualisierung der Social Communities. *Media Perspektiven*, (7-8), 380-390.
- Couper, M. P. (2013). Is the sky falling? New technology, changing media, and the future of surveys. *Survey Research Methods*, 7(3), 145-156.
- Dearing, J. W., & Rogers, E. (1996). *Agenda-setting*. Thousand Oaks, CA: Sage.
- Gaffney, D., & Puschmann, C. (2014). Data collection on Twitter. In K. Weller, A. Bruns, J. Burgess, M. Mahrt, & C. Puschmann (Hrsg.), *Twitter and society* (S. 55-67). New York, NY: Lang.

- Groshek, J., & Groshek, M. C. (2013). Agenda trending: Reciprocity and the predictive capacity of social networking sites in intermedia agenda setting across topics over time. *Media and Communication*, 1(1), 15-27. doi: 10.12924/mac2013.01010015
- Kearon, J., & Harrison, P. (2011). Research robots. A dramatic new way to conduct research & generate insights. Abgerufen von http://www.brainjuicer.com/xtra/BrainJuicer_DigiViduals_Research_Robots_Paper.pdf
- Lazer, D., Kennedy, R., King, G., & Vespignani, A. (2014). The parable of Google Flu: Traps in big data analysis. *Science*, 343(6176), 1203-1205. doi: 10.1126/science.1248506
- Lazer, D., Pentland, A., Adamic, L., Aral, S., Barabási, A.-L., Brewer, D., ... Van Alstyne, M. (2009). Computational social science. *Science*, 323(5915), 721-723. doi: 10.1126/science.1167742
- Mahrt, M. (2014). Einstellung. In C. Wünsch, H. Schramm, V. Gehrau & H. Bilandzic (Hrsg.), *Handbuch Medienrezeption* (S. 113-127). Baden-Baden: Nomos.
- Mahrt, M., & Scharkow, M. (2013). The value of big data in digital media research. *Journal of Broadcasting & Electronic Media*, 57(1), 20-33. doi: 10.1080/08838151.2012.761700
- Manovich, L. (2012). Trending: The promises and the challenges of big social data. In M. K. Gold (Hrsg.), *Debates in the digital humanites* (S. 460-475). Minneapolis: University of Minnesota Press.
- Markham, A. N., & Buchanan, E. (2012). Ethical decision-making and Internet research: Version 2.0. Recommendations from the AoIR Ethics Working Committee. Abgerufen von <http://www.aoir.org/reports/ethics2.pdf>
- Matthes, J., Kuhlmann, C., Gehrau, V., Jandura, O., Möhring, W., Vogelgesang, J., & Wünsch, C. (2011). Zur Methodenausbildung in kommunikationswissenschaftlichen Bachelor- und Masterstudiengängen. *Publizistik*, 56(4), 461-481. doi: 10.1007/s11616-011-0133-6
- Miller, G. (2011). Social scientists wade into the tweet stream. *Science*, 333(6051), 1814-1815. doi: 10.1126/science.333.6051.1814
- Mitchell, A., & Hitlin, P. (2013). Twitter reaction to events often at odds with overall public opinion. *Pew Research Center*. Abgerufen von <http://www.pewresearch.org/2013/03/04/twitter-reaction-to-events-often-at-odds-with-overall-public-opinion>
- Nelson-Field, K., Riebe, E., & Newstead, K. (2013). The emotions that drive viral video. *Australasian Marketing Journal*, 21(4), 205-211. doi: 10.1016/j.ausmj.2013.07.003

- Pfeffer, J. (2013). *Big data, big research? Opportunities and constraints for computer supported social science*. Keynote zur „Digital methods“-Tagung der DGPUK-Fachgruppe Computervermittelte Kommunikation, Wien. Abgerufen von <http://www.pfeffer.at/slides/DigitalMethods-BigData.pdf>
- Rodrigues, T., Benevenuto, F., Cha, M., Gummadi, K., & Almeida, V. (2011). *On word-of-mouth based discovery of the web*. Proceedings of the 2011 ACM SIGCOMM conference on Internet measurement conference, Berlin. doi: 10.1145/2068816.2068852
- Russom, P. (2011). Big data analytics. Abgerufen von http://www.cloudtalk.it/wp-content/uploads/2012/03/1_17959_TDWIBigDataAnalytics.pdf
- Savage, M., & Burrows, R. (2007). The coming crisis of empirical sociology. *Sociology*, 41(5), 885-899. doi: 10.1177/0038038507080443
- Scharkow, M., & Vogelgesang, J. (2011). Measuring the public agenda using search engine queries. *International Journal of Public Opinion Research*, 23(1), 104-113. doi: 10.1093/ijpor/edq048
- Schnell, R., Hill, P. B., & Esser, E. (2005). *Methoden der empirischen Sozialforschung* (7. Aufl.). München: Oldenbourg.
- Tinati, R., Halford, S., Carr, L., & Pope, C. (2014). Big data: Methodological challenges and approaches for sociological analysis. *Sociology*, 48(4), 663-681. doi: 10.1177/0038038513511561
- Verdegem, P., & D'heer, E. (2013). *The role of Twitter in political agenda-setting during elections: Findings from the 2012 elections in Belgium*. Vortrag auf der Konferenz „Twitter and Microblogging: Political, Professional and Personal Practices“, Lancaster, UK. Abgerufen von <http://hdl.handle.net/1854/LU-3200218>
- Watts, Duncan J., & Dodds, Peter S. (2007). Influentials, networks, and public opinion formation. *Journal of Consumer Research*, 34(4), 441-458. doi: 10.1086/518527
- Williams, D. (2006). On and off the 'Net: Scales for social capital in an online era. *Journal of Computer-Mediated Communication*, 11(2), 593-628. doi: 10.1111/j.1083-6101.2006.00029.x
- Yoder, C., & Stutzmann, F. (2011). *Identifying social capital in the Facebook interface*. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Vancouver, BC, Kanada. doi: 10.1145/1978942.1979025
- Zimmer, M. (2010). "But the data is already public": On the ethics of research in Facebook. *Ethics and Information Technology*, 12(4), 313-325. doi: 10.1007/s10676-010-9227-5

Empfohlene Zitierung: Heise, N. (2015). Big Data – small problems? Ethische Dimensionen der Forschung mit Online-Kommunikationsspuren. In A. Maireder, J. Ausserhofer, C. Schumann, & M. Taddicken (Hrsg.), *Digitale Methoden in der Kommunikationswissenschaft* (S. 39-58). doi: 10.17174/dcr.v2.3

Zusammenfassung: Kaum ein Thema hat in den vergangenen Jahren derart kontroverse Debatten in akademischen Kreisen hervorgerufen wie Big Data. Die automatisierte Verarbeitung riesiger Datenmengen für sozialwissenschaftliche Zwecke wirft nicht nur neue methodologische und theoretische Fragen auf – auch die ethischen Implikationen dieses datengetriebenen Forschungsansatzes geraten zusehends in den Blick. Der Beitrag diskutiert einige Besonderheiten von Big-Data-Analysen, insbesondere in Social Media, und welche Herausforderungen sich für die Umsetzung forschungsethischer Standards ergeben. Im Fokus steht auch die Frage, welchen Status die Nutzerinnen – die Menschen hinter den Daten – bei dieser Art der Forschung haben. Zudem sollen einige generelle Implikationen von Big-Data-Verfahren für Praxis und grundlegende Prinzipien sozialwissenschaftlicher Forschung beleuchtet werden. Ziel des Beitrages ist es, Forscherinnen für forschungsethische Fragen zu sensibilisieren sowie eine Verständigung der scientific community zum Umgang mit Big Data anzuregen.

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

Nele Heise

Big Data – small problems?

Ethische Dimensionen der Forschung mit Online-Kommunikationsspuren

Im März 2014 publizierte das US-Journal PNAS eine Studie, in der die Timeline von fast 690.000 Nutzerinnen¹ der sozialen Netzwerkplattform Facebook experimentell beeinflusst wurde, um die Theorie der emotionalen Ansteckung zu untersuchen (Kramer et al., 2014). Nicht nur die immens hohe Zahl an verarbeiteten Daten und die Manipulation der Nutzungsoberfläche erregte die Gemüter: Vor allem das Zustandekommen der Studie ohne das explizite Einverständnis der Untersuchungsteilnehmerinnen führte international zu heftig geführten Diskussionen um die Rolle von Unternehmen wie Facebook, wissenschaftliche Standards und ethische Aspekte bei der Nutzung von Big Data (Deterding, 2014).

Die kontroversen Debatten um das „Facebook-Experiment“ verdeutlichen, dass es im Hinblick auf Big-Data-Analysen in der (sozial-)wissenschaftlichen Forschung noch großen Verständigungsbedarf gibt. Zugleich wirft der Fall zahlreiche Fragen nach den ethischen Implikationen dieses neuen, datenintensiven Forschungsansatzes auf. Vor diesem Hintergrund diskutiert der vorliegende Beitrag zunächst einige

1 Zur leichteren Lesbarkeit wird auf die Aufführung des männlichen Begriffsäquivalents verzichtet.

Besonderheiten von Big-Data-Forschung und welche Herausforderungen sich daraus für die Umsetzung forschungsethischer Standards ergeben. Im Fokus steht dabei auch die Frage, welchen Status die beforschten Nutzerinnen, also die Menschen hinter den Daten, bei dieser Art der Forschung haben. Darüber hinaus sollen auch generelle Implikationen von Big-Data-Verfahren für die Praxis und grundlegende Werte sozialwissenschaftlicher Forschung beleuchtet werden.

1 **Big Data und ethisch relevante Besonderheiten online-basierter Forschung**

1.1 *Was Big-Data-Forschung ist*

Um den Begriff Big Data gibt es viele Diskussionen: Ist er lediglich ein Marketing-Buzzword, ein generalisierter und daher unpräziser Begriff, oder verweist er auf ein gänzlich neues Forschungsparadigma im Sinne einer datengetriebenen Wissenschaft? In der einschlägigen sozialwissenschaftlichen Literatur ist man aufgrund dieser Unbestimmtheit um alternative Begriffe bemüht, wie etwa „traceable information“ oder „massified research“ (Neuhaus & Webmoor, 2012). Im Folgenden meinen wir damit ganz pragmatisch die Erfassung, Speicherung und Verwertung von „massive quantities of information produced by and about people, things, and their interactions“ (boyd & Crawford, 2012, S. 662).

Vereinfacht ausgedrückt geht es bei Big Data im kommunikationswissenschaftlichen Kontext also um große Datensätze von Online-Kommunikations Spuren und deren Verwendung für datenbasierte Forschung. Ein wesentliches Merkmal „datenintensiver“ Ansätze ist die weitgehende Automatisierung des wissenschaftlichen Prozesses „von der Datenerfassung über die Verarbeitung bis hin zur Modellbildung“, so dass die Wissenschaftlerinnen „erst ziemlich spät einen Einblick in ihre Daten“ (Pietsch, 2013, S. 58-59) erhalten. Erhoben werden die Daten meist automatisiert, beispielsweise mittels der APIs² von sozialen Netzwerkwartplattformen wie Facebook oder Twitter. Auch die Verarbeitung, Auswertung

2 Die Abkürzung API steht für „application programming interface“, also Programmierschnittstellen eines Softwaresystems, die eine Anbindung an das System ermöglichen, zum Beispiel für Zugriffe auf Datenbanken.

und Darstellung der meist sehr großen Datensätze erfolgt bei Big-Data-Verfahren weitgehend automatisiert über verschiedene Software-Tools, mit deren Hilfe kausale Zusammenhänge, Korrelationen und Abhängigkeiten zwischen einer großen Zahl an Variablen und Indikatoren ermittelt werden können. Während so genannte „predictive analytics“ in Bereichen wie der Informatik, Medizin, Wirtschaft oder Politik bereits vor Jahren Einzug hielten, sind Big-Data-Analysen in den Sozialwissenschaften ein noch eher junges Phänomen. In der Sozialforschung werden sie unter anderem zur Beobachtung der Dynamiken öffentlicher Meinungsbildung in Echtzeit genutzt, zum Beispiel anhand politischer Diskussionen auf Twitter (Kaczmirek et al., 2014; Zeller, 2014).

Der Stellenwert solcher Analysen ist in den Sozialwissenschaften angesichts zahlreicher methodologischer und theoretischer Herausforderungen jedoch nicht unumstritten (siehe Mahrt in diesem Band).³ Zum einen stören sich Kritikerinnen an der Annahme, dass durch große Datensätze allein Ergebnisse mit höherem Anspruch an Wahrheit, Objektivität oder Genauigkeit erzielt werden (boyd & Crawford, 2012). Zum anderen widerstrebt einigen die mit Big-Data-Analysen einhergehende „Mathematisierung gesellschaftlicher Zusammenhänge“ (Pietsch, 2013, S. 62), weil sie hierin die Gefahr einer positivistisch-reduktionistischen Sozialforschung sehen (Uprichard, 2013).

Seltener indes werden die forschungsethischen Implikationen von Big-Data-Forschung dezidiert in den Blick genommen. Zunächst wird daher im Folgenden diskutiert, inwieweit Besonderheiten onlinebasierter Kommunikation(sdaten) sowie Rahmenbedingungen onlinebasierter Forschung für Big-Data-Analysen forschungsethisch relevant sind, denn: „the information generated by users of social media platforms and services cannot be considered equivalent to conventional types of offline information collected by social researchers“ (Neuhaus & Webmoor, 2012, S. 46). Die Übertragung gängiger Standards gestaltet sich bei Big-Data-Analysen vor allem deshalb schwierig, weil Status und Identität der Teilnehmerinnen und der „Daten“ sowie die Bedingungen, unter denen Forscherinnen mit ihnen interagieren, weniger eindeutig sind als bei offline durchgeführter Forschung (Neuhaus & Webmoor, 2012). Insbesondere die (räumlich-zeitliche)

3 Nähere Ausführungen dazu in Mahrt und Scharkow (2013), Tinati, Halford, Carr und Pope (2014) oder Zeller (2014).

De-Kontextualisierung und Entgrenzung privater und öffentlicher Kommunikation sowie die Anonymität und kommunikative Distanz zwischen Forscherin und Forschungssubjekten können die Prozesse ethischer Entscheidungsfindung bei Big-Data-Analysen erschweren.

1.2 *De-Kontextualisierung und (kommunikative) Distanz als Ursachen ethischer Dilemmata*

„Taken out of context, Big Data loses its meaning“ (boyd & Crawford, 2012, S. 670). Mit dieser Aussage bringen die Autorinnen ein zentrales Merkmal von Big-Data-Analysen auf den Punkt: Daten oder vielmehr Spuren sozialer Kommunikation werden bei Big-Data-Analysen aus dem Kontext ihrer Entstehung gegriffen. Dies ist forschungsethisch insofern relevant, als der Entstehungskontext und damit verknüpfte Annahmen über die Privatheit/Öffentlichkeit der Kommunikation die Erwartungen der Nutzerinnen, wer die von ihnen mitgeteilten Informationen wahrnehmen soll, beeinflussen. Insbesondere die Kommunikation in Social Media ist jedoch von einer gewissen Entgrenzung oder Vermischung privater und öffentlicher Kontexte geprägt (Zeller, 2014). Da längst nicht alle Nutzerinnen den Zugang zu ihren Profilen mithilfe von Privacy-Einstellungen regulieren, können Forscherinnen nicht per se davon ausgehen, dass die Vorstellungen der Nutzerinnen mit der tatsächlichen Reichweite der Kommunikation übereinstimmen (Heise & Schmidt, 2014). Die Erfassung und Verwertung von explizit/implizit produzierten Inhalten kann daher mitunter die „kontextuelle Integrität“ (Nissenbaum, 2004) verletzen, das heißt einen Bruch der Handlungs- und Erwartungszusammenhänge hinsichtlich der Frage darstellen, was mit diesen Informationen geschehen soll – und was nicht:

„Data are most often not produced with the intention that it be used as raw research material. Additionally, the original data are, in the process of mining and visualizing, taken out of its context of production and assembled as a new instance“ (Neuhaus & Webmoor, 2012, S. 58).

In diesem Zusammenhang weist boyd (2010) darauf hin, dass es sich bei der Aggregation, Zusammenführung und Weiterverarbeitung gesammelter Daten letztlich um eine Art Zweckentfremdung handelt, die potenziell eine Verletzung

der Privatheit bedeuten kann (siehe Abschnitt 2). Auch das „frictionless sharing“ – das Teilen von Inhalten über verschiedene Plattformen hinweg – stellt eine Herausforderung dar, da die Daten teils automatisch von einem Kontext in einen anderen überführt werden, beispielsweise von einem öffentlichen Twitter- auf ein nicht-öffentliches Facebook-Profil.

In solchen Fällen ist fraglich, auf welchen Kontext oder welche Online-Umgebung sich die Erwartungen der Nutzerinnen beziehen. Dies wirkt sich auch auf die Berechtigung der Forscherinnen aus, deren Daten in ihre Analysen einzubeziehen, „because people have different values or expectations in different online settings“ (Eynon, Schroeder & Fry, 2009, S. 188). Hierfür gilt es also Beurteilungskriterien zu entwickeln, die sich etwa an der Zugänglichkeit und Sensibilität der Informationen bemessen können (Heise & Schmidt, 2014); vor allem, da vielen Nutzerinnen die unterschwellige Erhebung ihrer Daten mitunter gar nicht bewusst ist: „Many are not aware of the multiplicity of agents and algorithms currently gathering and storing their data (...) [they] are not necessarily aware of all the multiple uses, profits, and other gains that come from information they have posted“ (boyd & Crawford, 2012, S. 673).

Eine weitere forschungsethische relevante Herausforderung onlinebasierter Forschung ist der im Vergleich zu Offline-Forschung geringere Grad an sozialer Präsenz und wechselseitiger Sichtbarkeit (Heise & Schmidt, 2014). So kann die Anonymität der Beziehung zwischen Forscherin und Beforschter bei Big-Data-Analysen insbesondere im Hinblick auf Verantwortlichkeit problematisch sein: “In terms of accountability, the issue of (researcher) anonymity now becomes a problem rather than being a solution for minimizing risk to the individual” (Neuhaus & Webmoor, 2012, S. 57).

Hinzu kommt die große kommunikative Distanz bei Big-Data-Verfahren. Den gängigen reaktiven sozialwissenschaftlichen Forschungsmethoden ist ein unterschiedlicher Grad an Austauschprozessen und Interaktionen inhärent, über die kommunikative Beziehungen zwischen Forscherin und Forschungsobjekten hergestellt werden (Russel & Kelly, 2002; Ziegau, 2009). Bei gewissen onlinebasierten Forschungsverfahren erfolgt diese Kommunikation, im Rahmen derer die Teilnehmerinnen in der Regel bei der Kontaktaufnahme oder zu Forschungsbeginn über den Kontext einer Erhebung (Ziele, involvierte Akteurinnen oder Ähnliches) informiert werden, technisch vermittelt und (a-)synchron. Dadurch haben die Teilnehmerinnen ein gewisses Maß an Selbstbestimmung und Kontrolle

darüber, inwieweit sie ihr Einverständnis zur Erhebung, Auswertung und Veröffentlichung bestimmter Daten geben (Heise & Schmidt, 2014).

Bei Big-Data-Analysen ist es nun möglich (oder angesichts hoher Fallzahlen sogar notwendig), auf einen kommunikativen Austausch mit den beforschten Individuen zu verzichten: „a researcher can conveniently collect data from his/her computer without the knowledge of the data producer or the ‘participant’“ (Neuhaus & Webmoor, 2012, S. 57). Das heißt, bei automatisierten Erhebungen findet eine Interaktion zwischen beiden Parteien kaum oder nicht mehr statt. Dies unterminiert zum einen ein basales kommunikationsethisches Prinzip: Reziprozität, also die Wechselseitigkeit der Kommunikation, die als eine Maßgabe der Interaktion zwischen Forscherin und Nutzerin im Online-Kontext verstanden werden kann (Wolff, 2007; Heise, 2013). Eine solche Wechselseitigkeit ist unter den Bedingungen von Big-Data-Forschung nicht mehr gegeben oder aufgrund der großen Anzahl involvierter Nutzerinnen gar unmöglich, was vor allem Folgen für das Einholen einer informierten Einwilligung (siehe Abschnitt 2) hat. Zum anderen hat sich der Austausch verlagert und findet nun vorrangig in Form eines Datenaustauschs zwischen Forscherin und Plattformbetreiberin statt. Durch derartig automatisierte Datenzugänge werden aber nicht nur die Nutzerinnen – als Produzentinnen der Daten – um- oder vielmehr übergangen. Ethisch problematisch ist es auch, dass die Nutzerinnen und ihr Handeln aufgrund der Nicht-Interaktion als Individuen „verschwinden“ und nur als maschinenlesbarer Text wahrnehmbar sind.

1.3 *Big Data: Bits und Bytes oder Spuren kommunikativen Handelns?*

Der Status und die Identität der Teilnehmerinnen sind bei Big-Data-Analysen weniger eindeutig als bei offline durchgeführter Forschung (Neuhaus & Webmoor, 2012). Dabei stellt die kommunikative Distanz der Forscherin zu den Nutzerinnen insofern eine forschungsethische Herausforderung dar, als sie zu einem gewissen Wahrnehmungsparadox führt: Die Forscherinnen haben es bei Big-Data-Analysen oftmals nur mit Datensätzen oder (Text-)„Objekten“ zu tun, wodurch die Nutzerinnen für sie vor allem als „einrollende Datensätze“ und nicht mehr als Einzelpersonen wahrnehmbar sind. In dieser Distanz zum betroffenen „actual human“ bestehe, so Zimmer (2012), einerseits das Risiko, dass Forscherinnen ihre Arbeit

nicht oder nur bedingt als „‘human subject‘ research“ begreifen, was die Verletzung oder Nicht-Einhaltung ethischer Standards im Umgang mit Forschungsteilnehmerinnen befördern könne.

Andererseits ist fraglich, inwieweit dabei das kommunikationsethische Prinzip der Personalität als eine Norm onlinebasierter Forschung Berücksichtigung findet (Heise, 2013). Bezugsnehmend auf das Grundrecht der Menschenwürde postuliert dieses Prinzip die Anerkennung des Personseins und das Verbot einer Verdinglichung, Objektivierung oder Instrumentalisierung anderer Nutzerinnen (Wolff, 2007). In dieser Hinsicht hält McFarland (2012) die Prozesse der automatisierten Datenerhebung und -verarbeitung im Rahmen von Big-Data-Analysen für ethisch bedenklich, „because it dehumanizes those being judged (...) and thus threatens those profiled objects as collections of facts, rather than as persons“. Auch die Masse von Daten könne dazu führen, dass Risiken für einzelne involvierte Nutzerinnen falsch eingeschätzt werden:

„the individual disappears in the mass, but the issue and potential risk remain on the scale of the individual. (...) [it is] important for both researchers and institutions to accept the fact that this kind of large-scale data mining still involves human subjects“ (Neuhaus & Webmoor, 2012, S. 58-59).

Dies verweist auf ein zentrales Problem im Zusammenhang mit Big-Data-Analysen: Und zwar die Frage, ob es sich bei den gewonnenen Daten um Spuren sozialer Kommunikation, also Handlungsakte, oder Artefakte/Texte handelt – eine Unterscheidung, die bei der Beurteilung ethischer Fragen hochgradig relevant ist. Während Big-Data-Analysen als Erhebungsmethode zwar nicht-reaktiv und Verfahren wie der Inhaltsanalyse durchaus ähnlich sind, untersuchen sie häufig nicht publizistische Inhalte, sondern Nutzer*handeln* („moments of people“), für das bestimmte Standards wie das Einholen der informierten Einwilligung gelten (Heise & Schmidt, 2014).

2 Forschungsethische Standards – wie umgehen mit den Menschen hinter den Daten?

Wir gehen hier von der Prämisse aus, dass Big-Data-Analysen in Abhängigkeit zum Inhalt nicht mit Inhaltsanalysen gleichzusetzen sind, da es sich ins-

besondere bei Daten aus Social-Media-Diensten immer auch um Spuren sozialer Kommunikation handelt. Daher müssen auch ethische Standards berücksichtigt werden, die den Umgang mit Forschungsteilnehmerinnen regulieren und sich insbesondere auf deren Persönlichkeits- und Datenschutzrechte beziehen. Dazu gehören neben der *Freiwilligkeit* die *informierte Einwilligung* („informed consent“), die Wahrung der *Anonymität* und der *Schutz der Privatsphäre* sowie die *Nicht-Schädigung* der Teilnehmerinnen (Diener & Crandall, 1978; Hopf, 2009; Strohm Kitchener & Kitchener, 2009).

Ein zentraler ethischer Standard besagt, dass die Einwilligung zur Teilnahme an wissenschaftlicher Forschung freiwillig von Personen erteilt wird, die dafür kompetent sind und in der Regel im Vorfeld angemessen informiert wurden. Die informierte Einwilligung zählt zu den „primary technical and procedural steps for minimizing risks to human subjects in non-medical research contexts“ (Neuhaus & Webmoor, 2012, S. 45). Sie liegt einmal in der Freiheit des Individuums begründet, selbst darüber zu entscheiden, unter welchen Umständen es anderen Personen Einstellungen, Überzeugungen oder Meinungen mitteilt. Und zum anderen im Recht auf informationelle Selbstbestimmung, also der Kontrolle einer Person über die von ihr selbst mitgeteilten, über sie gesammelte bzw. sie betreffende, von anderen mitgeteilte Daten.⁴

Bei Big-Data-Analysen sind herkömmliche Consent-Verfahren angesichts der Größe der Datensätze und der Anzahl beteiligter Individuen oft nur schwer oder gar nicht umsetzbar. Adäquate Verfahren sind bislang nicht entwickelt. Ungeklärt ist zudem, wie etwa mit einzelnen Postings auf sozialen Netzwerkplattformen umgegangen werden soll, die nach der Datenerhebung von den Nutzerinnen wieder gelöscht wurden. Oder wie man beispielsweise mit Postings verfahren soll, in denen sich Personen nicht selbst äußern, sondern von anderen Nutzerinnen erwähnt werden. Derartige Inhalte können bei einigen Fragestellungen, etwa zu politischen Einstellungen, als durchaus sensibel eingestuft werden (Kaczmarek et al., 2014).

Da es zudem fraglich ist, ob die Kommunikation auf bestimmten Diensten per se als öffentlich einzustufen ist, sofern sie öffentlich zugänglich ist, können wir

4 Dabei handelt es sich laut Schmidt (2012) nicht nur um ein normatives Konzept, sondern auch eine Kompetenz im praktischen Handlungsvollzug.

nicht automatisch von einer impliziten Zustimmung der Nutzerinnen zur Sammlung und Auswertung ihrer Daten ausgehen, denn: „In light of current research (...), it is questionable whether users can effectively distinguish private from public messages and behaviour” (Mahrt & Scharrow, 2013, S. 26).

Ebenfalls nicht geklärt ist, ob die Zustimmung zu den AGBs einzelner Dienste als impliziter informed consent gewertet werden kann bzw. eine explizite Einwilligung „überlagert“ oder gar ersetzt. Die empörten Reaktionen auf das „Facebook-Experiment“ zeigen, dass diese Lesart aus Nutzerinnensicht problematisch ist und die Forscherinnen keinesfalls von der Umsetzung des informed consent entbinden sollte: „Individuals may (...) have agreed to the service conditions of a third party. However, this cancels neither the responsibilities of the researcher nor the institutions that undertake such research” (Neuhaus & Webmoor, 2012, S. 59).

Die bereits in Abschnitt 1.2 angesprochene Frage nach der Öffentlichkeit/Privateit bestimmter Daten oder Kommunikationsformen ist auch in Bezug auf die Standards der Anonymisierung und Nicht-Schädigung relevant (boyd & Crawford, 2012). Der forschungsethische Standard der Nicht-Schädigung postuliert die Vermeidung möglicher Risiken oder Gefährdungen. Das heißt die an der Forschung beteiligten Personen/-gruppen dürfen keinen Nachteilen oder Gefahren ausgesetzt werden, die das Maß des im Alltag Üblichen überschreiten, und müssen über etwaige Risiken aufgeklärt werden (Hopf, 2009). Zur Vermeidung möglicher Schäden, wie der Verletzung der Privatsphäre der Teilnehmerinnen durch die Weitergabe und mangelhafte Sicherung personenbezogener Daten oder unzureichend anonymisierte Publikationen, sind die Daten zu anonymisieren, insbesondere im Hinblick auf die Sensibilität bestimmter Informationen.⁵

Bei Big-Data-Analysen ist die Frage umstritten, was konkret unter persönlichen/sensiblen Informationen zu verstehen ist und welche Schädigungen für Individuen potenziell auftreten können. Einige Autorinnen argumentieren, dass

5 Der Datenschutz zählt zum rechtlich bindenden Kontext empirischer Forschung. So treffen das Bundesdatenschutz- oder das Telemediengesetz Regelungen zur Erhebung, Verarbeitung und Nutzung sowie sicheren Verwahrung von Daten. Demnach dürfen personenbezogene Daten besonderer Art, wie etwa ethnische Herkunft oder religiöse Überzeugungen, nur dann erhoben und verwendet werden, wenn sie zur Durchführung der Forschung unerlässlich sind.

über die Analyse oder Zusammenführung von Datensätzen völlig neue Formen persönlich identifizierender Informationen entstehen:

„By (...) analysing metadata, such as a set of predictive and aggregated findings, or by combining previously discrete data sets, Big Data approaches are not only able to manufacture novel PII [personally identifiable information; Anm. d. Verf.], but often do so outside the purview of current privacy protections.” (Crawford & Schultz, 2013, S. 94)

Individuen sollten daher das Recht haben, zu erfahren, wie persönliche Informationen konkret genutzt werden. Dies scheint umso drängender angesichts der Tatsache, dass Nutzerinnen dazu tendieren, über mehrere Dienste hinweg ähnliche „Identifier“ zu verwenden. Über die (Re-)Kombination oder das „cross screening“ verschiedener Datensätze lassen sich so sekundäre Informationen generieren, über die Einzelpersonen identifizierbar sind: „This is especially the case if traces are available over a longer period of time or in greater quantity so that a more comprehensive picture about an individual may be pieced together“ (Neuhaus & Webmoor, 2012, S. 58). Diesbezüglich argumentieren Crawford und Schultz, dass das Sammeln von Informationen über das Verhalten Einzelner eine neuartige Form der Verletzung von Privatsphäre („predictive privacy harms“) darstellen und durchaus persönliche Schädigungen zur Folge haben könne: „Personal harms emerge from the inappropriate inclusion and predictive analysis of an individual’s personal data without their knowledge or express consent“ (Crawford & Schultz, 2013, S. 94).⁶ Ein weiteres Problem bei der Umsetzung des Anonymisierungsstandards ist die potenzielle De-Anonymisierung von Datensätzen nach der Veröffentlichung (siehe Abschnitt 4.1).⁷

- 6 Ein Beispiel für unberechtigte Privatsphäreneingriffe ist der Fall Target: Das Unternehmen hatte mittels der Analyse verschiedener (Meta-)Daten schwangere Nutzerinnen identifiziert und deren Kontaktdaten ohne Zustimmung an Marketingfirmen weitergegeben. Einige Betroffene hatten ihre Schwangerschaft noch nicht öffentlich gemacht, wodurch es möglicherweise auch zu persönlichen Schädigungen kam (Crawford & Schultz, 2013).
- 7 Zimmer (2010) zeigte am Beispiel des US-amerikanischen Forschungsprojekts T3, dass sich selbst größere Datensätze de-anonymisieren lassen: Nachdem die auf Facebook gesammelten Daten eines ganzen Studierendenjahrgangs online gestellt wurden, konnte die betreffende Universität anhand des öffentlich einsehbaren Codebuchs der Studie identifiziert werden, womit die Anonymisierung des Datensatzes hinfällig war.

Insgesamt zeigt sich, dass forschungsethische Standards auch bei Big-Data-Analysen relevant sind, insbesondere im Kontext der Kommunikation auf sozialen Netzwerkplattformen. Und auch wenn deren konkrete Umsetzung noch ausgehandelt werden muss, scheint klar: „Concerns over consent, privacy and anonymity do not disappear simply because subjects participate in online social networks; rather, they become even more important“ (Zimmer, 2010, S. 324).

3 Was Big Data für Praxis und Grundwerte sozialwissenschaftlicher Forschung bedeutet

Neben forschungsethischen Standards, die den Umgang mit Akteurinnen *außerhalb* der Wissenschaft, insbesondere den Untersuchungsteilnehmerinnen regeln, gehören auch normative Leitlinien zur Wertbasis wissenschaftlicher Tätigkeit, die das Handeln *innerhalb* der Wissenschaft steuern. Darunter fallen forschungstechnische Mindeststandards und Grundsätze wissenschaftlichen Arbeitens sowie höhere Prinzipien, die Forscherinnen bei der Beurteilung ihrer Arbeit und der ethischen Entscheidungsfindung im Hinblick auf grundlegende Werte, Ziele und Verantwortung von Forschung helfen sollen (Diener & Crandall, 1978; Strohm Kitchener & Kitchener, 2009; Fenner, 2010). Im Folgenden soll diskutiert werden, inwieweit Big-Data-Verfahren eine Herausforderung für die Praxis und grundlegende Werte sozialwissenschaftlicher Forschung darstellen.

3.1 Big Data als Herausforderung für sozialwissenschaftliche Praxis

In der konkreten Anwendungspraxis berührt Big-Data-Forschung einige Grundsätze guten wissenschaftlichen Arbeitens, wie sie etwa die Deutsche Forschungsgemeinschaft (DFG, 2013) formuliert, beispielsweise den Grundsatz der Nachvollziehbarkeit aller Forschungsschritte. Im Vergleich zu anderen Verfahren sind bei Big-Data-Analysen die weitgehend automatisierten Arbeitsschritte und algorithmische Modellierung, bei denen die Ergebnisse unter Berücksichtigung einer großen Anzahl von Parametern und Daten erlangt werden, zum Teil nur schwer und auch nicht für alle nachvollziehbar (Pietsch, 2013): „the advent of Big Data makes it easier for those with programming skills and knowledge of

social media APIs to ask social questions” (Tinati, Halford, Carr & Pope, 2014, S. 5). Derzeit verfügen nur wenige Forscherinnen über das notwendige Wissen oder die Ressourcen, um etwa die Qualität von Big-Data-Analysen beurteilen zu können. Dies kann ein Problem für innerwissenschaftliche Kontrollprozesse, wie peer feedback und die Beurteilung der Qualität des Vorgehens darstellen: „instead of (...) [an] information silo, we potentially have instead the information scientist silo“ (Neuhaus & Webmoor, 2012, S. 59).

Auch bei der Veröffentlichung von Ergebnissen sind der Offenlegung algorithmischer Datenerhebung und -auswertung aufgrund ihrer Komplexität und Datenmenge Grenzen gesetzt. Zwar werden, ähnlich wie bei Codebüchern zu Inhaltsanalysen, häufig Hashtag- oder Schlagwortlisten publiziert – die Abbildung des zugrunde liegenden Codes oder Algorithmus ist aber weitaus problematischer bis unmöglich. Ein aus ethischer Perspektive schwerwiegendes Problem bei der Veröffentlichung ganzer Datensätze ist zudem die potenzielle De-Anonymisierung, weshalb laut Mahrt und Scharrow (2013) nur noch wenige Anbieter zur Kooperation und Weitergabe anonymer Datensätze bereit seien. Neben datenschutzrechtlichen Problemen besteht beim (internationalen) Austausch und der Bereitstellung von Forschungsdaten ein nachträgliches Risiko darin, dass sich Content-Providerinnen oder einzelne Nutzerinnen gegen die Veröffentlichung ihrer Daten aussprechen: „The post-hoc withdrawal (...) makes replications of the findings impossible and therefore violates a core principle of empirical research” (Mahrt & Scharrow, 2013, S. 26; siehe auch Zeller, 2014).

Aus diesen und anderen Gründen sollten Big-Data-Forscherinnen als Teil guter wissenschaftlicher Praxis verstärkt Maßnahmen zur Gewährleistung von Transparenz, Sicherheit und „Accountability“ (Verantwortlichkeit/Zurechenbarkeit) ergreifen (Wen, 2012). Einerseits trägt dies zur Qualitätssicherung und Güte wissenschaftlichen Arbeitens bei. Andererseits können Maßnahmen, die Zurechenbarkeit und einen verantwortungsvollen Umgang mit den Daten (Stichwort: Datenschutz) anzeigen, das Vertrauen in die Forschung steigern. Bei der Veröffentlichung von Datensätzen und Grafiken könnte dies über Erklärungen zur Autorinnenschaft bei der Datenerhebung und -verarbeitung erfolgen: „This would both break the spell of anonymity (...) surrounding these data sets and assign a claimed responsibility“ (Neuhaus & Webmoor, 2012, S. 59).

3.2 Big Data und ethische Grundprinzipien

Neben der Umsetzung der Grundsätze wissenschaftlichen Arbeitens stellt sich auch die Frage, inwieweit Big-Data-Analysen ethische Grundprinzipien wissenschaftlicher Forschung berühren. Dazu gehören laut Strohm Kitchener und Kitchener (2009) die vier Dimensionen *Non-Maleficence* (Nicht-Schädigung/Schadensvermeidung), *Beneficence* (Fürsorge), *Justice* (Gleichheit/Gerechtigkeit) und *Fidelity* (Vertrauen), aber auch Grundwerte wie Forschungsfreiheit. Für uns sind hier zunächst das Gerechtigkeitsprinzip (das Ideal einer Gleichbehandlung aller) und das Prinzip der *Beneficence* interessant, welches besagt, dass wissenschaftliche Forschung nicht nur um das Wohlergehen aller bemüht sein soll, sondern auch die Autonomie und Selbstbestimmung sowie die Werte und Entscheidungen der Teilnehmerinnen anerkennen und respektieren soll.

Im Hinblick auf diese Prinzipien stellt sich bei Big-Data-Analysen zum einen die Frage, wer in den Datensätzen repräsentiert ist und welche Menschen davon ausgeschlossen sind, etwa weil sie bestimmte soziale Netzwerkplattformen nicht nutzen. Dabei geht es nicht nur um demographische Diskrepanzen, sondern auch um Unterschiede in der Nutzung einzelner Angebote, die die Generalisierbarkeit der Ergebnisse erheblich einschränken. So erlaubt ein Dienst wie Twitter unterschiedlichste Beziehungsgeflechte und eine breite Vielfalt an Kommunikationsmodi, von denen die Nutzerinnen in verschiedener Weise Gebrauch machen. Die Aussagekraft von Analysen einzelner Hashtags oder @-replies ist daher oft fraglich:

„users take advantage of these communicative facilities in different ways. Some may enjoy the increased ability to speak, whilst others may appreciate merely the wider array of content available to consume. (...) analysing only tweet data may not offer a fair overview of Twitter usership per se, let alone a national population more generally“ (Cows, 2014).

Angesichts dieser dateninhärenten „biases“ ist fraglich, inwieweit Big-Data-Forschung überhaupt den Grundprinzipien einer Gleichbehandlung und der Förderung des Wohlergehens aller genügen kann. Hier sollte auch in den Sozialwissenschaften eine Debatte zu den gesellschaftlichen Folgen von Big-Data-Analysen stattfinden, etwa zum Einfluss so genannter „predictive analytics“ (Wen, 2012)

auf politische Entscheidungen.⁸ Zudem sollten Sozialwissenschaftlerinnen ein Bewusstsein dafür entwickeln, dass die Interpretation von Big-Data-Analysen, wie bei jeder anderen Forschung auch, gewisse Nachteile für gesellschaftliche Gruppen nach sich ziehen kann, zum Beispiel durch die Erzeugung oder Verstärkung von Stereotypen (Hopf, 2009). Letztlich stellt das algorithmenbasierte Erfassen, Auswerten und Darstellen von Daten immer auch eine Klassifikation sozialen Handelns dar, die durchaus (schwer abschätzbare) gesellschaftliche Folgen haben kann: „How we shape the social through our counting and classifying are highly political and ethical issues“ (Uprichard, 2013).

Neben dieser gesellschaftlichen Ebene spielen ethische Grundprinzipien bei Big-Data-Analysen auch auf individueller Ebene der Nutzerinnen eine Rolle, insbesondere im Hinblick auf die Wahrung und Anerkennung ihrer Autonomie. Dies betrifft nicht nur ihr Recht auf informationelle Selbstbestimmung (siehe Abschnitt 2), sondern auch die Machtverhältnisse zwischen Forscherinnen und Nutzerinnen, die bei Big-Data-Forschung hinsichtlich der Kontrolle und des Zugangs zu den Daten häufig verschoben sind: “researchers have the tools and access, while social media users as a whole do not“ (boyd & Crawford 2012, S. 673).

Die Frage nach Handlungsmacht betrifft aber nicht nur die Beziehung zwischen Forscherin und Beforschten, da bei Big-Data-Analysen eine dritte Partei hinzukommt: die „Besitzerinnen“ der Daten. Bei sozialwissenschaftlichen Studien handelt es sich hier meist um Dienste wie Twitter, Facebook oder Google, also kommerzielle Unternehmen, die als Intermediäre die zur Forschung genutzten Informationen in ihren Datenbanken bereithalten und (partiell) bereitstellen. Diese Mittler-Position der Intermediäre trägt nicht nur zur kommunikativen Distanz zwischen Forscherinnen und Nutzerinnen bei (siehe Abschnitt 1), sondern macht sie auch zu einer neuartigen Anspruchsgruppe ethischer Entscheidungsfindung, vor allem bezüglich der informierten Einwilligung (siehe Abschnitt 2).

Fraglich ist, inwieweit dadurch neue Abhängigkeiten entstehen und ein zentrales wissenschaftliches Grundprinzip, nämlich die innere und äußere Freiheit der Forscherinnen in der Definition ihres Untersuchungsgegenstandes, der Wahl

8 Dass Big-Data-Analysen durchaus realweltliche Folgen haben, zeigte sich 2013 in den USA: Dort erfolgte die Zuteilung von Notversorgungsgütern nach dem Hurrikan Sandy unter anderem auf Basis einer Analyse der Twitter-Kommunikation aus betroffenen Gebieten (Crawford, 2013).

ihres Forschungsweges sowie das Recht auf Veröffentlichung der Ergebnisse beeinträchtigt wird. Dies betrifft zum einen die Datenerhebung: Viele Plattformbetreiberinnen schließen die automatisierte Erfassung von Nutzerinneninformationen/ -interaktionen in ihren AGBs aus oder knüpfen sie im Sinne ihres „Hausrechtes“ an bestimmte Bedingungen. Facebook etwa fordert von Forscherinnen explizit, die betroffenen Nutzerinnen über Art und Zweck der gesammelten Informationen in Kenntnis zu setzen und ihre Zustimmung einzuholen (Heise & Schmidt, 2014). Mitunter sind Wissenschaftlerinnen daher auf Kooperationen mit Plattformbetreiberinnen angewiesen, die aber eher selten sind und, wie das „Facebook-Experiment“ zeigt, aus forschungsethischer Perspektive nicht immer glücklich verlaufen. Da die Unternehmen in der Regel auf datenschutzrechtliche Absicherung und den Schutz ihrer Geschäftsmodelle – zu denen auch Datenbanken gehören – bedacht sind, können solche Kooperationen zudem mit Einschränkungen für die Forscherinnen verbunden sein, beispielsweise bezüglich der Publikation von Daten (siehe Abschnitt 4.1).

Auch über die Art und Form der verfügbaren Daten haben die Forscherinnen häufig keine unmittelbare Entscheidungsgewalt, was die Bandbreite an möglichen Fragestellungen bzw. ihre Forschung generell einschränken kann: „the tools we use can limit the range of questions that might be imagined, simply because they do not fit the affordances of the tool. (...) This then introduces serious limitations in terms of the scope of research that can be done“ (Vis, 2013). Entscheidungen über die Gestaltung von Nutzungsoberflächen oder bestimmte Nutzungsoptionen werden von den Intermediären gemäß eigener Zielvorstellungen (wie Profit- oder Reichweitenmaximierung) implementiert und sind daher nicht wertneutral. Mit Blick auf das Prinzip der Forschungsfreiheit gilt es bei Big-Data-Analysen daher die Rolle der Plattformbetreiberinnen zu berücksichtigen und zugleich kritisch zu reflektieren, inwieweit sie den Möglichkeitshorizont der Forscherinnen (mit-)bestimmen.

4 Ausblick

Im öffentlichen Diskurs um Big Data wurden zuletzt immer wieder eine gesellschaftliche Diskussion sowie politische Entscheidungen darüber eingefordert, für welche Bereiche derartige Analysen eingesetzt werden dürfen – und für

welche nicht. An diesen Diskussionen sollten auch wir Sozialwissenschaftlerinnen uns beteiligen: „We need to think about what it means to measure the social world and how our models of causality are constructed (...), how these counts and measurements are used and for whom? These answers are not trivial and social scientists need to be part of those conversations” (Uprichard, 2013). Die kontroversen Debatten um das „Facebook-Experiment“ machen deutlich, dass auch eine Reflexion der ethischen Implikationen und ein Verständigungsprozess über unseren Umgang mit Big-Data-Forschung stattfinden muss – gerade jetzt, da die Forschung mit „traceable information“ zögerlich, aber unumkehrbar Einzug in unsere Forschungsfelder erhält.

Dabei gilt es nicht, den Wert sozialwissenschaftlicher Big-Data-Forschung grundsätzlich anzuzweifeln, sondern „to examine exactly what is known through such totalizing inquiries (...), to query the discursive horizon they construct, as well as what vanishes beyond that horizon” (Baym, 2013). Zugleich gilt es auch den Nutzen von Big-Data-Analysen und ihren Beitrag zum Gemeinschaftswohl klar herausstellen, um den Eindruck zu vermeiden, Daten würden schlichtweg aufgrund ihrer Verfügbarkeit oder rein um des Datensammeln-Willens erhoben. Nicht zuletzt führen uns die Enthüllungen von Abhör- und Big-Data-Zugriffstaktiken verschiedener Regierungsinstitutionen und die dadurch erneut angestoßene Datenschutzdebatte einmal mehr vor Augen: Es ist wichtig, datenintensive Forschung klar vom auch nur ansatzweisen Verdacht der verdeckten Beobachtung oder „Überwachung“ abzugrenzen, um das gesellschaftliche Vertrauen in Wissenschaft nicht aufs Spiel zu setzen. Denn selbst wenn die kommerzielle Verwendung von Nutzungsdaten durch die Plattformbetreiberinnen fragwürdig erscheint, so beruht sie doch auf einem „exchange or transaction of mutual benefit between service providers and customers. However, for academics (...) who deploy such data, the ethical implications must be carefully evaluated“ (Neuhaus & Webmoor, 2012, S. 57).

Die Absicht des vorliegenden Beitrags war es, einige ethische Dimensionen und Herausforderungen von Big-Data-Forschung aufzuzeigen, um die Sensibilität für forschungsethische Fragen zu schärfen und die Annahme zu entkräften, diese Art der Forschung sei „disconnected from ethics“ (Neuhaus & Webmoor, 2012, S. 60). Da es sich insbesondere bei Social-Media-Analysen um die Untersuchung menschlich-sozialer Ausdrucksformen handelt, müssen auch in Big-Data-Studien sozialwissenschaftliche ethische Grundprinzipien greifen (Zeller, 2014).

Angesichts vieler offener Fragen und Handlungsunsicherheiten braucht es eine Verständigung innerhalb der Scientific Community zum Umgang mit ethischen Dilemmata: „Big Data research has to solve the problem of guaranteeing privacy and ethical standards while also being replicable and open to scholarly debate” (Mahrt & Scharnow, 2013, S. 26).

Die Ergebnisse einer solchen Aushandlung sollten bestenfalls in der Formulierung von Richtlinien münden, die zur Orientierung im Prozess ethischer Entscheidungsfindung, der wissenschaftlichen Selbstkontrolle sowie der Legitimation gegenüber Akteurinnen außerhalb der Wissenschaft dienen können. Aufgrund der Besonderheiten onlinebasierter Kommunikation stoßen Forscherinnen bei der Übertragung ethischer Standards bislang häufig an Grenzen und gängige formalisierte Verfahrensweisen erweisen sich oft als ineffektiv oder impraktikabel. Es braucht daher praxisorientierte Handlungsleitlinien, die flexibel für die jeweiligen Forschungsvorhaben adaptierbar sind. Neuhaus und Webmoor (2012) etwa schlagen eine „agile ethics for massified research and visualization“ vor, bei der es weniger um standardisierte Ethik-Protokolle, sondern um situativ anzupassende „working solutions“ und eine Sensibilität für gute Praxis geht. Letztlich tragen ein verantwortungsvoller Umgang mit den Daten sowie Transparenz und Zurechenbarkeit nicht nur zur Sicherung der Qualität und Güte wissenschaftlichen Arbeitens bei, sondern steigern auch das Vertrauen in Big-Data-Forschung.

Die im Beitragstitel aufgeworfene Frage „Big Data – small problems?“ muss angesichts der vielschichtigen forschungsethischen Herausforderungen verneint werden. Wieder einmal stehen Wissenschaftlerinnen vor der Aufgabe, technische Machbarkeit und ethische Vertretbarkeit ihres Handelns abzuwägen. Das Aufkommen einer „data-driven science“ sollte auch als Erinnerung an die besondere Verantwortung sozialwissenschaftlicher Forschung verstanden werden. Denn, wie Diener und Crandall bereits vor über dreißig Jahren anmahnten, „the social sciences can (...) have a tremendous impact on society, even to the point of revolutionizing our conceptions of human nature, society, and culture. (...) [We] must accept responsibility for its social consequences” (Diener & Crandall, 1978, S. 195).

Nele Heise, M.A. ist wissenschaftliche Mitarbeiterin an der Universität Hamburg und Doktorandin an der Graduate School Media and Communication Hamburg

Quellenverzeichnis

- Baym, N. (2013). Data not seen: The uses and shortcomings of social media metrics. *First Monday*, 18(10). doi: 10.5210/2Ffm.v18i10.4873
- boyd, d. (April 2010). Privacy and Publicity in the Context of Big Data. Vortrag auf der WWW. Raleigh: North Carolina. Abgerufen von <http://www.danah.org/papers/talks/2010/WWW2010.html>
- boyd, d., & Crawford, K. (2012). Critical questions for Big Data. *Information, Communication & Society*, 15(5), 662-679. doi: 10.1080/1369118X.2012.678878
- Crawford, K. (1. April 2013). The Hidden Biases in Big Data. *Harvard Business Review*. Abgerufen von http://blogs.hbr.org/cs/2013/04/the_hidden_biases_in_big_data.html
- Crawford, K., & Schultz, J. (2013). Big Data and Due Process: Toward a Framework to Redress Predictive Privacy Harms. *Boston College Law Review*, 55(1), 93-128. Abgerufen von <http://lawdigitalcommons.bc.edu/bclr/vol55/iss1/4/>
- Cowls, J. (28. Januar 2014). "Twitter says..." – Can big social data tell us about public opinion? Abgerufen von <http://joshcowls.com/2014/01/28/twitter-says-can-big-social-data-tell-us-about-public-opinion/>
- Deterding, S. (9. Juli 2014). The Facebook Loophole. Abgerufen von <http://codingconduct.tumblr.com/post/91239582975/the-facebook-loophole>
- DFG (2013). *Vorschläge zur Sicherung guter wissenschaftlicher Praxis*. Weinheim: Wiley-VCH. Abgerufen von http://www.dfg.de/download/pdf/dfg_im_profil/reden_stellungnahmen/download/empfehlung_wiss_praxis_1310.pdf
- Diener, E., & Crandall, R. (1978). *Ethics in Social and Behavioral Research*. Chicago: University of Chicago Press.
- Eynon, R., Schroeder, R., & Fry, J. (2009). New Techniques in Online Research. Challenges for Research Ethics. *21st Century Society*, 4(2), 187-199. doi: 10.1080/17450140903000308
- Fenner, D. (2010). *Einführung in die Angewandte Ethik*. Tübingen: Francke.
- Heise, N. (2013). 'Doing it for real' – Authentizität als kommunikationsethische Voraussetzung onlinebasierter Forschung. In M. Emmer, A. Filipovic, J.-H. Schmidt, & I. Stapf (Hrsg.), *Echtheit, Wahrheit, Ehrlichkeit. Authentizität in der computervermittelten Kommunikation* (S. 88-109). Weinheim: Juventa.

- Heise, N., & Schmidt, J. (2014). Ethik der Onlineforschung. In M. Welker, M. Taddicken, J.-H. Schmidt, & N. Jakob (Hrsg.), *Handbuch Online-Forschung* (S. 532-553). Köln: Herbert von Halem.
- Hopf, C. (2009). Forschungsethik und qualitative Forschung. In U. Flick, E. Kardoff, & I. Steinke (Hrsg.), *Qualitative Sozialforschung* (S. 589-600). Reinbek: Rowohlt.
- Kaczmirek, L., Mayr, P., Vatrapu, R., Bleier, A., Blumenberg, M., Gummer, T., et al. (2014). Social Media Monitoring of the Campaigns for the 2013 German Bundestag Elections on Facebook and Twitter. *Gesis Working Papers* (31). Abgerufen von http://www.gesis.org/fileadmin/upload/forschung/publikationen/gesis_reihen/gesis_arbeitsberichte/WorkingPapers_2014-31.pdf
- Kramer, A. D. I., Guillory, J. E., & Hancock, J. T. (2014). Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings of the National Academy of Sciences*, *111*(24), 8788-8790. doi: 10.1073/pnas.1320040111
- Mahrt, M. & Scharnow, M. (2013). The Value of Big Data in Digital Media Research. *Journal of Broadcasting & Electronic Media*, *57*(1), 20-33. doi: 10.1080/08838151.2012.761700
- McFarland, M. (2012). Ethical Implications of Data Aggregation. Abgerufen unter <http://www.scu.edu/ethics/practicing/focusareas/technology/internet/privacy/data-aggregation.html>
- Neuhaus, F., & Webmoor, T. (2012). Agile ethics for massified research and visualization. *Information, Communication & Society*, *15*(1), 43-65. doi: 10.1080/1369118X.2011.616519
- Nissenbaum, H. F. (2004). Privacy as Contextual Integrity. *Washington Law Review*, *79*(1), 119-158.
- Pietsch, W. (2013). Big Data – über Chancen und Risiken einer neuen Wissenschaft des Komplexen. *Universitas*, *806*.
- Russel, G., & Kelly, N. (2002). Research as Interacting Dialogic Processes: Implications for Reflexivity. *Forum: Qualitative Social Research*, *3*(3). urn: urn:nbn:de:0114-fqs0203181
- Schmidt, J. (2012). Persönliche Öffentlichkeiten und informationelle Selbstbestimmung im Social Web. In J. Schmidt & T. Weichert (Hrsg.), *Datenschutz* (S. 215-225). Bonn: Bundeszentrale für politische Bildung.
- Strohm Kitchner, K., & Kitchner, R. F. (2009). Social Research Ethics. Historical and Philosophical Issues. In D. Mertens, & P. Ginsberg (Hrsg.), *The Handbook of Social Research Ethics* (S. 5-22). Thousand Oaks: Sage.

- Tinati, R., Halford, S., Carr, L., & Pope, C. (2014). Big Data: Methodological Challenges and Approaches for Sociological Analysis. *Sociology*, 48(4), 663-681. doi: 10.1177/0038038513511561
- Uprichard, E. (1. Oktober 2013). Big Data, Little Questions? *Discover Society*. Abgerufen von <http://www.discoverociety.org/focus-big-data-little-questions/>
- Vis, F. (2013). A critical reflection on Big Data: Considering APIs, researchers and tools as data makers. *First Monday*, 18(10). doi: 10.5210/2Ffm.v18i10.4878
- Wen, H. (21. Juni 2012). The Ethics of Big Data. *Forbes*. Abgerufen von <http://www.forbes.com/sites/oreillymedia/2012/06/21/the-ethics-of-big-data>
- Wolff, O. (2007). *Kommunikationsethik des Internets: eine anthropologisch-theologische Grundlegung*. Hamburg: Dr. Kovač.
- Zeller, F. (2014). Online-Forschung und Big Data. In M. Welker, M. Taddicken, J.-H. Schmidt, & N. Jakob (Hrsg.), *Handbuch Online-Forschung* (S. 453-480). Köln: Herbert von Halem.
- Ziegau, S. (2009). *Die Abhängigkeit der Sozialwissenschaften von ihren Medien. Grundlagen einer kommunikativen Sozialforschung*. Bielefeld: transcript.
- Zimmer, M. (2010). 'But the data is already public': on the ethics of research on Facebook. *Ethics and Information Technology*, 12(4), 313-325. doi: 10.1007/s10676-010-9227-5
- Zimmer, M. (26. Mai 2012). New media, new ethics. How social media-based research demands new attention to research ethics. Vortrag auf der 62. Jahrestagung der International Communication Association (ICA). Phoenix: Arizona. Abgerufen von <http://de.slideshare.net/michaelzimmer/new-media-new-ethics-ica-2012>

Empfohlene Zitierung: Pentzold, C. (2015). Forschungsethische Prämissen und Problemfelder teilnehmenden Beobachtens auf Online-Plattformen. In A. Maireder, J. Ausserhofer, C. Schumann, & M. Taddicken (Hrsg.), *Digitale Methoden in der Kommunikationswissenschaft* (S. 61-85). doi: 10.17174/dcr.v2.4

Zusammenfassung: Ziel dieses Beitrags ist es, forschungsethische Prämissen und Problemfelder des qualitativen kommunikationswissenschaftlichen Untersuchens von Interaktion und Kommunikation auf online-medialen Plattformen zu bestimmen. Der Umgang mit zwei Problemfeldern wird für das teilnehmende Beobachten in der deutsch- und englischsprachigen Wikipedia demonstriert. Es wird im ersten Problemfeld betrachtet, welche Konsequenzen die (bedingte) Öffentlichkeit von selbst-, fremd- oder systemgenerierten Informationen für das Gewährleisten von Anonymität hat. Im zweiten Problemfeld wird diskutiert, welche Konsequenzen die Anonymität auf Plattformen für das Absichern von informierter Zustimmung mit sich bringt.

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

Christian Pentzold

Forschungsethische Prämissen und Problemfelder teilnehmenden Beobachtens auf Online-Plattformen

1 Einleitung: Kodifizieren und Diversifizieren als doppelte forschungsethische Dynamik

Sozialwissenschaftliche Untersuchungen sozialen Lebens sollen forschungsethisch plausibel und integer vorbereitet sein, ablaufen und dargestellt werden. Die dazu gesetzten datenschutzrechtlichen Vorgaben und normativen ethischen Prämissen gelten für Analysen online-medial (mit-)konstituierter Lebenswelten ebenso wie sie Anforderung an jede methodisch-empirische Auseinandersetzung mit sozialen Zusammenhängen im weitesten Sinn sind (Hopf, 2005; Gola & Schomerus, 2010).

Die besondere Dringlichkeit, nach der forschungsethischen Dimension der Untersuchung online-medialer Kommunikation und Interaktion zu fragen, ergibt sich aus zwei Entwicklungen: Zum einen aus der stärkeren fachlichen Auseinandersetzung mit forschungsethischen Problemen und entsprechenden disziplinären Kodifizierungsversuchen, wie sie beispielsweise in den soziologischen und kommunikationswissenschaftlichen deutschen und englischsprachigen Fachgesellschaften in den letzten zwei Jahrzehnten stattgefunden haben (ASA, 1999; DG-PuK, 1999; DGS, 2014; ICA, o.J.). Im Kern zielen diese Anstrengungen darauf, den

Konsens ethisch-integren und professionellen Handelns in allen Stationen wissenschaftlicher Praxis festzuhalten, um Studien forschungsethisch zu rechtfertigen bzw. begründet zu kritisieren. Im Fokus steht dabei vor allem das Wahren der Rechte der Untersuchten. Hinzu kommen weitere forschungsethische Entscheidungsbereiche, insbesondere bezüglich des Zuschreibens wissenschaftlicher Leistungen wie korrektes Zitieren oder das Nennen von Forschungsbeteiligten, des Sicherns von Autorenrechten, des unparteiischen Begutachtens und eines nicht-diskriminierenden wissenschaftlichen Betriebs.

Zum anderen machen es die fortschreitende Entwicklung medialer Kommunikationsformen und medienbezogener Praktiken einerseits und die Innovationen digitaler Methoden für Datenerhebung und Datenauswertung andererseits fraglich, inwiefern die darauf gerichteten kommunikationswissenschaftlichen Untersuchungen forschungsethische Musterlösungen übernehmen können oder ob hier neue forschungsethische Herausforderungen vorliegen. Auf das Diversifizieren des empirischen Forschungsfeldes und der analytischen Prozedere reagierten folglich damit befasste Komitees mit gegenstandsbezogen adaptierten forschungsethischen Richtlinien (DGOF, 2007; Ess & AoIR Ethics Working Committee, 2002; Eynon, Fry & Schroeder, 2008; ICC/ESOMAR, 2011; McKee & Porter, 2008; Markham, Buchanan & AoIR Ethics Working Committee, 2012; Heider & Massanari, 2012). Diese gehen zum Beispiel auf das Nutzen persönlichen, online publizierten Materials, das Rekrutieren von Nutzern für Interviewstudien oder das Sichern von elektronischen Kopien des Forschungsmaterials ein. In ihrem Versuch, die sich ergebenden Anforderungen forschungsethisch zu reflektieren und in Handreichungen abzubilden, stehen diese Vorhaben jedoch vor dem Problem, dass sie umso schneller veralten und unpraktikabel werden, je konkreter sie auf die Spezifika einzelner medialer oder kommunikativer Zusammenhänge eingehen.

Gemäß dieser doppelten Dynamik hat jüngst insbesondere die Option massiven Sammelns und Auswertens von Daten („Big Data“) forschungsethische Reflexionen angestoßen (boyd & Crawford, 2012; Mahrt & Scharrow, 2013; siehe auch Mahrt sowie Heise in diesem Band). Jedoch stellt sich die Notwendigkeit forschungsethischer Überlegungen ebenso dringlich für qualitativ orientierte kommunikationswissenschaftliche Vorhaben, denn hier ist der Rapport zum studierten Feld und seinen Akteuren gerade persönlich und individuell zu gestalten, statt dass Daten reaktiv und ohne direkte Interaktion aggregiert werden können (Roth, 2004; von Unger, Narimani, & M’Bayo, 2014). Zudem kann der Forschungsbericht selten auf

persönliche Stimmen und individualisierte Beispiele verzichten, um seine Analyse valide darzustellen, was forschungsethisch für den Online-Bereich besonders problematisch erscheint, da hier zitierte Dokumente trotz Anonymisierung ohne Weiteres per Suchmaschine in ihrem originalen Kontext aufgefunden werden können. Vor diesem Hintergrund ist das Ziel des Beitrags erstens, forschungsethische Problemfelder einer qualitativen Erforschung der Kommunikation und Interaktion auf online-medialen Plattformen zu bestimmen. Zweitens wird der praktische Umgang mit den forschungsethischen Axiomen am Beispiel einer teilnehmenden Beobachtung in der Online-Enzyklopädie Wikipedia erklärt. Das forschungsethisch relevante analytische Spektrum digitaler Methoden, das von der hier vorgestellten, mehrmethodisch angelegten Studie bedient wurde, umfasst das Erheben von Daten per Online-Kommunikation, zum Beispiel durch Führen von Interviews per Skype, das Studium von Internetnutzung, das Untersuchen von technologischen Infrastrukturen und die Analyse online-medialer Textkorpora.

Wikipedia als kooperatives Projekt steht exemplarisch für eine neue Generation digital vernetzter Medien, die sich als Plattformen präsentieren. Plattformen können hier als technologische Architekturen verstanden werden, die durch ihre Hardware und Software bestimmte Nutzungsweisen unterstützen und ihnen eine Arena bieten (Gillespie, 2010, S. 350). Plattformen greifen also einzelne, oft alltägliche Aktivitäten auf, setzen sie in programmierte Anwendungen um, kombinieren sie mit anderen Funktionalitäten und popularisieren sie unter ihrem Namen. Auf diese Weise werden plattformbasiert Nachrichten ausgetauscht, Fotos gesammelt, geordnet und gezeigt, Notizen gekritzelt oder Videos geteilt. Seit Ende der 1990er Jahre dominieren auf diese Weise Plattformen wie Blogger, Wikipedia, Myspace, Facebook, Flickr, YouTube, Skype oder Twitter die digitalisierte Netzwerkkommunikation (van Dijck, 2013, S. 4-5).

2 Basis: Forschungsethische Prämissen und forschungspraktische Konsequenzen

Um die forschungsethischen Problemfelder auf online-medialen Plattformen zu bestimmen, werden in einem ersten Schritt forschungsethische Axiome und daraus folgende Konsequenzen diskutiert. Die forschungsethisch gebotene Abwägung zwischen normativen Forderungen an die Forschung und den Einsich-

ten, die diese Forschung verspricht, kann grundsätzlich mittels zweier Strategien begründet werden. Sie ist im Prinzip entweder, wie Ess (2006) erklärt, deontologisch, also mit Verweis auf den epistemologischen Eigenwert der gewonnenen Erkenntnis, oder utilitaristisch und so mit dem weiterführenden Nutzen der erbrachten Ergebnisse in anderen Handlungsbereichen, etwa der Wirtschaft oder Pädagogik, zu rechtfertigen.

Generell kann Forschungsethik dazu aus drei einander bedingenden Perspektiven betrachtet werden. Erstens stellen sich forschungsethische Ansprüche als normative Prinzipien dar und umfassen so allgemeine Menschenrechte auf Basis des Prinzips der Menschenwürde wie das Recht auf informationelle Selbstbestimmung sowie das Recht auf Sicherheit und Schutzwürdigkeit der Person (Albers, 2005; Ess, 2009; Hudson & Bruckman, 2004). Sie bilden in ihrer rechtsphilosophisch grundlegenden Dimension das Fundament forschungsethischer Vorkehrungen und müssen daher in konkrete forschungspraktische Entscheidungen erst übersetzt werden. Dazu wurden forschungsethische Anforderungen zweitens in regulatorische Satzungen und Einrichtungen gebracht, wie sie mit gesetzlichen Persönlichkeitsrechten, Urheberrechten und Datenschutzrechten (Basset & O’Riordan, 2002; Schmidt & Weichert, 2012; Walther, 2002; Waskul & Douglass, 1996) sowie in Form von Ethikkommissionen oder Institutional Review Boards (IRBs) vorliegen (Buchanan & Ess, 2009; Eynon et al., 2008; Kraut, Olson, Banaji, Bruckman, Cohen, & Couper, 2004). Als solche sind sie institutionalisierte und fallweise auszulegende Standards, anhand derer Forschungsvorhaben geprüft werden können. Schließlich werden die Prinzipien dafür drittens in methodische Gütekriterien übersetzt, um so die ethische Integrität des wissenschaftlichen Prozesses und seiner Ergebnisse abzusichern (Hopf, 2005). Hinsichtlich der Wahrung der Rechte der Untersuchten sind dies im Wesentlichen fünf Prämissen, um (1) freiwillige Teilnahme zu gewähren, (2) Anonymität abzusichern, (3) Vertraulichkeit zuzusagen, (4) informierte Zustimmung einzuholen sowie (5) Schädigungen und ungünstige Folgen oder Risiken der in die Untersuchung einbezogenen Teilnehmer zu vermeiden.

In der Konsequenz steht jedes kommunikationswissenschaftliche empirische Forschungsvorhaben vor der prinzipiellen Aufgabe, die axiomatischen Forderungen in ihrer Allgemeingültigkeit und Verbindlichkeit im Blick auf den jeweiligen Fall und die gebrauchten Methoden adäquat umzulegen. Ein Weg, universell geltende forschungsethische Prämissen in spezifische forschungspraktische Lösungen zu übersetzen, ist der Gebrauch gestaffelter Heuristiken. Diese helfen

dabei, im Forschungsprozess vom Konzipieren über das Durchführen, Auswerten und Darstellen jeweils forschungsethisch begründete theoretische, methodologische und methodische Entscheidungen treffen zu können. In diesem Sinn raten die Empfehlungen der *Association of Internet Researchers* (Markham, Buchanan & AoIR Ethics Working Committee, 2012, S. 7), dass „ethical decision making is best approached through the application of practical judgement attentive to the specific context“. Als adaptives Vorgehen, um forschungsethische Prämissen und forschungspraktische Konsequenzen schrittweise in Einklang zu bringen und so die Plausibilität und Legitimität des analytischen Prozedere und seiner Ergebnisse zu steigern, entwickeln McKee und Porter (2008, S. 23-28) ein kasuistisches Modell forschungsethischen Entscheidens. Es umfasst folgende vier Aspekte: (1) Das Identifizieren relevanter ethischer Normen, (2) das Bestimmen der jeweiligen Umstände des untersuchten Bereichs, für den die ethischen Prämissen angewandt werden sollen, (3) der Vergleich zwischen verschiedenen möglicherweise in einem Forschungsvorhaben einbezogenen empirischen Feldern und Daten und ihren fallweisen forschungsethischen Anforderungen und Antworten, (4) die Konsultation mit Bezugsgruppen, also den beforschten Personen und der engeren und weiteren Forschergemeinschaft. Diese sind einmal vertreten durch wissenschaftliche Kooperationen, wissenschaftliche Foren (Tagungen, Workshops, Lehre), disziplinäre Ethikausschüsse und Präzedenzfälle sowie weiterhin durch akademische und gesellschaftliche Interessengruppen wie Geldgeber, Wissenschaftsverwaltung und -politik, andere Wissenschaftssektoren, allgemeine wissenschaftsethische Kontrollgremien und die publizistische Öffentlichkeit.

3 Fall: Teilnehmende Beobachtung in der Online-Enzyklopädie Wikipedia

Im Fokus der Studie, deren forschungsethische Kasuistik hier diskutiert wird, stand die Frage nach den Praktiken und den institutionellen Bedingungen online-medialen, gelingenden und produktiven, also das Generieren informationeller Güter bezweckenden Zusammenarbeitens. Diese wurde exemplarisch und exponiert am Beispiel der freien Gemeingüterproduktion studiert, genauer gesagt am Fall der Online-Enzyklopädie Wikipedia (Benkler, 2006; Pentzold, 2013). Die Wikipedia ist einerseits ein extremer Fall für die Untersuchung freiwilliger, plattformbasierter

Kooperation, da sie, gemessen an ihrer Bestandsdauer, der Zahl der Artikel, der Zahl editierender und lesender Nutzer, der Güte der bereitgestellten Informationsgüter sowie hinsichtlich ihrer Bekanntheit, extrem erfolgreich ist.¹ Andererseits ist die Wikipedia ein extremer Fall, weil die technologischen Anlagen und das Designprinzip des Wikis einen hohen Grad aufeinander bezogenen und ineinander verschränkten Interagierens möglich machen und Wikipedia somit stark interdependentes Kooperieren statt unzusammenhängende Einzelaktionen erlaubt. Rekonstruiert wurden auf der einen Seite die Teilnehmungsweisen, die es dem stark aktiven und längerfristig engagierten Kern der Wikipedianer erlauben, an dem Projekt konstruktiv teilzunehmen. Auf der anderen Seite wurden die institutionellen Einrichtungen an Regeln, Normen und programmierter Software (Code) ermittelt, die von den Nutzern gebraucht wurden, um ihr Interagieren und Kommunizieren zu organisieren (Pentzold, 2013).

Die Untersuchung wurde als ethnografische teilnehmende Beobachtung in einem Zeitraum von 33 Monaten vom Januar 2009 bis September 2011 durchgeführt. Das untersuchte Feld waren die deutsch- und englischsprachigen Plattformen des Projekts. Methodisch umfasste die ethnografische teilnehmende Beobachtung entsprechend ihrer Anlage als Multi-Methoden-Studie das Mitmachen an und das Beobachten von Handlungsmustern, das Sammeln von Dokumenten wie etwa von Wiki-Seiten und anderen Web-Dokumenten, dann der Besuch von Wikipedianertreffen bei Stammtischen, Workshops und Tagungen sowie das Führen von Gesprächen und problemzentrierten Interviews (siehe zum Vorgehen allgemein

1 Im Frühsommer 2014 war Wikipedia in knapp 285 Sprachversionen mit insgesamt etwa 31 Millionen Artikeln verfügbar, womit sie eines der größten bislang erschienenen Allgemeinlexika ist. Die umfangreichste Version auf Englisch zählte mehr als 4,5 Millionen angelegte Artikel und insgesamt knapp 33 Millionen Wiki-Seiten. Rund 21 Millionen Nutzer hatten sich registriert (WikipediaSpecial:Statistics, 2014). Die deutschsprachige als die, gemessen an der Artikelzahl, zweitgrößte Version versammelte zu diesem Zeitpunkt knapp 1,7 Millionen enzyklopädische Artikel und mehr als 4,7 Millionen Wiki-Seiten. Bis Mai 2014 hatten sich hier mehr als 1,8 Millionen Nutzer angemeldet und es wurden insgesamt über 136 Millionen Edits getätigt (WikipediaSpecial:Statistik, 2014). Nur ein Bruchteil der registrierten Nutzer brachte sich jedoch mit eigenen Beiträgen ein. So waren im Februar 2014 nur knapp 30.000 Benutzer in der englischsprachigen Ausgabe mit mehr als fünf Edits aktiv und etwa 3.000 kamen auf über 100 Beiträge in diesem Monat. Im deutschsprachigen Pendant waren es im gleichen Zeitraum rund 6.200, die sich mehr als fünfmal beteiligten, und etwa 940, die mehr als 100 Edits leisteten (Wikipedia Statistik, 2014).

Hammersley & Atkinson, 2007; und konkret Boellstorff, Nardi, Pearce, & Taylor, 2012, S. 113-128; Kozinets, 2010, S. 95-103).²

Die Datenerhebung und -auswertung richteten sich nach der Forschungslogik der Grounded Theory (Glaser & Strauss, 1967/2005). Dadurch war es möglich, die forschungsethischen Anforderungen mit den analytischen Phasen forschungsbegleitend abzugleichen, statt sie a priori festzulegen. Als Verfahrensrahmen bot sie eine analytische Systematik, die eine Schrittfolge zum systematischen Erschließen des Feldes, Erheben empirischen Materials und interpretativen Rekonstruierens bereitstellte und das Offenlegen und Plausibilisieren der Arbeitsschritte bei Datenauswahl und Datenauswertung einforderte (Krotz, 2005, S. 167-179).

Angelehnt an das Gerüst kasuistischen Klärens forschungsethischer Fragen lag das Hauptaugenmerk darauf, in vier einander bedingenden Schritten erstens die fallweisen Umstände des jeweiligen Feldausschnitts, der genutzten Methode und des hier gewonnenen Materials festzustellen und zweitens die relevanten forschungsethischen Prämissen zu identifizieren. Um die Praxis forschungsethischen Prüfens und Entscheidens zu erklären, wird im folgenden Abschnitt der Schwerpunkt besonders auf diese beiden Stationen gelegt. In der Studie selbst schloss sich daran drittens das Recherchieren von Musterlösungen bzw. *best practices* an, wie

- 2 Die Übersicht der geleisteten Editierung zeigt an, dass das für die Teilnahme angelegte Nutzerprofil *Christianpentzold* in der deutschsprachigen Wikipedia für 780 Edits an 124 Wiki-Seiten in verschiedenen Namensräumen verantwortlich ist. Die Aktionsepisoden, Beobachtungen, Gespräche, Erfahrungen und davon ausgehende methodische und konzeptuelle Überlegungen wurden über diese systemgenerierte Dokumentation hinaus schriftlich (Protokolle, Feldnotizen, Forschungstagebücher) sowie visuell (Screenshots, Bilddateien, Fotos) und audiovisuell (Bildschirmmitschnitte) erfasst. Dabei kamen mehr als 600 maschine- und handgeschriebene Seiten an Aufzeichnungen, rund 130 Fotos, Screenshots und Bilddateien sowie 116 Videos mit einer Gesamtdauer von etwas mehr als 72 Stunden zusammen. Zudem wurden 12 problemzentrierte Interviews mit deutschsprachigen Wikipedianern geführt und transkribiert (ca. 386.000 Zeichen, ca. 8.000 Zeilen). Während des teilnehmenden Beobachtens wurde außerdem ein Korpus von 1.605 sehr unterschiedlich umfangreichen Dokumenten aus den Namensräumen der deutsch- und englischsprachigen Wikipedia und aus weiteren Online- und Offline-Quellen zusammengetragen. Zu den Aktionsepisoden und Interviews kamen Beobachtungen bei zwei Dutzend Zusammenkünften von Wikipedianern in Deutschland, Dänemark, den USA und Polen sowie rund 100 informelle Gespräche (auf Deutsch oder Englisch) mit Wikipedia-Autoren. Umgesetzt wurde die Auswertungsarbeit mit Unterstützung der Datenanalysesoftware Atlas.ti.

sie in gegenstandsbezogenen Kodices, ethnografischen Studien digital vernetzter Umgebungen sowie in Anleitungen und forschungsethischen Hinweisen der Wikipedianer zu finden waren (Materialgrundlage: Ayers, Matthews, & Yates, 2008; Coleman, 2013; van Dijk, 2010; Fiebig, 2005; Stöcklin, 2010; Wikipedia:Ethically researching Wikipedia, 2012). Viertens wurde das Vorgehen in Kolloquien und mit den Interaktionspartnern im Feld selbst diskutiert. Dazu gehörten die internationalen Treffen von Wikipedianern, Trägerorganisation und Wikimedia-Vereinen (Wikimania), die Tagungen der Wikiforschung und -entwicklung (WikiSym) sowie Werkstattberichte in sozialwissenschaftlichen Interpretationsgruppen und Doktorandenforen (Pentzold, 2013).

4 Anwendung: Forschungsethische Problemfelder

Ausgangspunkt des teilnehmenden Beobachtens auf der Plattform waren *Episoden* als situative, raum-zeitlich begrenzte Verrichtungen. Die beobachteten und getätigten Aktionen in Wikipedia wurden in diesem Sinn in einzelne Episoden als prozesshafte, sequenzierbare Handlungs- und Interaktionsweisen gegliedert (Goffman, 1983; Schmidt, 2009; S. 48).

Wesentlich für die forschungsethisch zu reflektierenden Spezifika der teilnehmenden Beobachtung an Wikipedia war die gewählte Untersuchungsstrategie, ein möglichst breites Spektrum an verschieden öffentlichen *Stationen* und *Stationenbereichen* des Beschäftigens mit Wikipedia zu erfassen.³ Stationen wurden mit Giddens (1997) definiert als „Orte, an denen sich die Routineaktivitäten verschiedener Individuen überschneiden“ (S. 171). In Bezug auf die Plattform Wikipedia waren solche Stationen zum Beispiel die Diskussionsseite eines Nutzers, die Seite zum Melden von Vandalismus, das Portal zum Reviewen von Artikeln oder die Seite zum Kandidieren als Administrator. Das variierte Merkmal war die Öffentlichkeit bzw. Privatheit dieser Stationen und es wurde davon ausgegangen, dass ein stationenweise variabler

3 Darüber hinaus orientierte sich die Felderschließung und -erhebung an *Aktivitäten*. Ziel dieser Strategie war es, an möglichst vielen ähnlichen und unterschiedlichen Tätigkeiten teilzuhaben bzw. diese zu beobachten, um so den Aktionsraum zu erfassen. Zudem folgte das Vorgehen *kritischen Ereignissen* geleitet von der Frage, wie mit ihnen als krisenhafte Störungen im Kontrast zum routinemäßigen Projektalltag umgegangen wurde (Pentzold, 2013).

‚sense of audience‘ die Bereiche auf der Plattform sowie die plattformunabhängige Interaktion und Kommunikation online und offline differenziere.

Mit Nissenbaums (2011, S. 2) Konzept von Privatheit als „contextual integrity“ unterschieden sich die Stationen also entsprechend der jeweils von den Wikipedianern zugeschriebenen Kommunikationsnormen und der bei einer Station erwartbar aktiven Akteure und relevanten Informationen. Im Fall der Plattform Wikipedia waren demgemäß zwar im Wiki alle Seiten ohne Zugangsschutz gleich zugänglich und auffindbar, doch hatte sich im Projekt ein Verständnis von verschiedenen öffentlichen Bereichen mit differenzierten Aufmerksamkeitshorizonten und Leser- bzw. Teilnehmerkreisen konventionalisiert. „Nicht alle Teilnehmer bewegen sich in allen ‚Räumen der Wikipedia‘ gleichzeitig“ (Stegbauer, 2009, S. 65) und zumindest jeder angemeldete Nutzer konnte sich aus einem Angebot an Tools und Spezialseiten bedienen, um Wiki-Seiten für sich besonders öffentlich zu machen. Neben dem potenziellen Publikum, für das die Seiten prinzipiell erreichbar wären, war das variierende empirische Publikum von Wiki-Seiten an der Zahl der Beobachter, also der Autoren, bei denen eine Seite in ihrer Beobachtungsliste auftauchte, und an der Zahl der Seitenaufrufe messbar.⁴ Von diesen potenziellen und empirischen Publika waren wiederum das intendierte Publikum als die vorgestellte Leser- und Autorenschaft einer Seite bzw. eines dort gemachten Edits sowie das mit bestimmten Botschaften adressierte Publikum zu unterscheiden.

Öffentlichkeit wurde also analytisch und im Feld selbst als abgestufte *publicness* von Stationen auf der Plattform und der plattformunabhängigen Kommunikation und Interaktion verstanden. Deren potenzielle Wahrnehmungschancen waren im Prinzip zwar gleich hoch, ihre empirische Rezeptionswahrscheinlich-

4 In seiner Auswertung von diachronen Daten von 1.501 angemeldeten Nutzern in der deutschsprachigen Wikipedia (im September 2006) zeigt Stegbauer (2009: 83), dass von diesen nur 46 Prozent überhaupt editierend aktiv wurden, 38 Prozent einen Artikel bearbeiteten, zwölf Prozent an einer Artikeldiskussion teilnahmen, 15 Prozent eine Benutzerseite editierten, zehn Prozent auch die Diskussionsseite einer Benutzerseite editierten, acht Prozent in eine Seite im Wikipedia-Namensraum eingriffen und ein Prozent auch im WikipediaDiskussion-Namensraum aktiv wurde. Zugleich zeigt er, dass die Namensräume insgesamt nicht nur sehr verschieden editierend erschlossen wurden, sondern dass diese ungleichen Aktivitätsraten mit einem zeitlich gestreckten Übergang von einem in einen anderen Namensraum einhergingen, sodass der WikipediaDiskussion-Namensraum auch als letzter nach Anmelden editierend erreicht wurde.

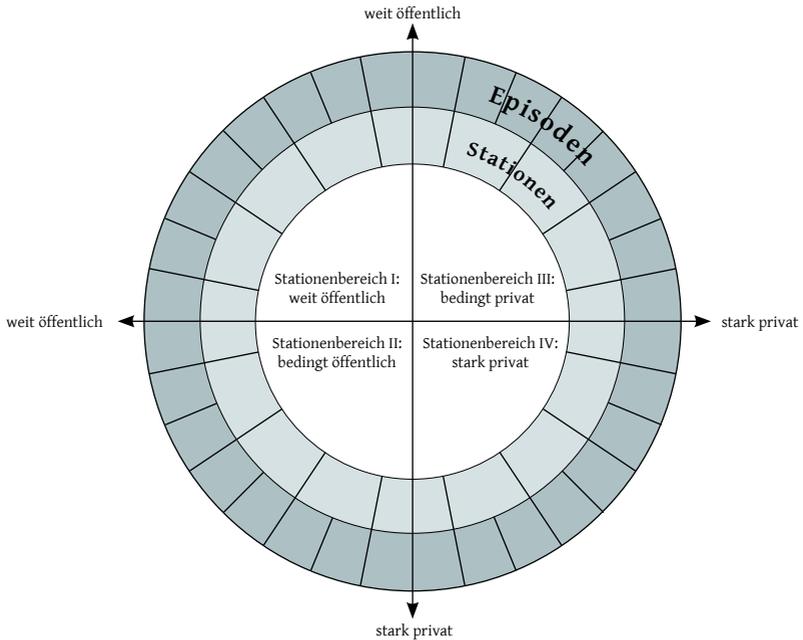
keit sowie die bei einer Station intendierten und adressierten Publika variierten indessen (Baym & boyd, 2012; Schmidt, 2012).⁵

Es war konsequenterweise forschungsethisch zu berücksichtigen, dass die Nutzer ihre Beiträge an verschieden große und verschieden zusammengesetzte Publika adressierten. Dabei konnte in der Analyse nicht ohne Weiteres davon ausgegangen werden, dass die Wikipedianer hierbei auch eine (fachwissenschaftliche) Öffentlichkeit im Blick hatten (Hudson & Bruckman, 2004). Folglich kombinierte die kontrastive Erhebung mehr oder minder öffentlicher Stationen Zahlenwerte zu Besuchern, Editoren und Edits mit im Projekt ausgebildeten und an neue Nutzer weitergegebenen Überzeugungen, welche Stationen stark oder schwach in der Aufmerksamkeit der Wikipedia-internen und der Wikipedia-externen Öffentlichkeit ständen. Heuristisch konnte über das quantitative Bestimmen unterschiedlich großer empirischer Wikipedia-Öffentlichkeiten hinaus das Gruppieren der plattformbasierten Stationen in softwareseitig und funktional getrennte Namensräume und der damit zusammenhängenden Einteilung verschieden öffentlicher, weil unterschiedlich intensiv editierter und von Nutzern aufgerufenen, Stationen eingesetzt werden (Pentzold, 2007, S. 36-44).

Folglich wurden vier Stationenbereiche über die Trennung von Online- und Offline-Arenen hinweg entsprechend ihrer vergleichsweise großen bzw. kleinen empirischen sowie breiten oder engen intendierten und adressierten Publika getrennt in (1) den *weit öffentlichen Stationenbereich* der Vorträge und Plenen auf internationalen Wikipedia-Tagungen, der Wikipedia-Artikel und der Wiki-Seiten, über die das Organisieren des Projekts ablief, (2) den *bedingt öffentlichen Stationenbereich* der an verschiedenen Orten stattfindenden Autoren-Workshops und der plattformbasierten Diskussionen zu Wikipedia-Artikeln, (3) den *bedingt privaten Stationenbereich* der lokalen Stammtische und Exkursionen, der auf der Plattform zu findenden Nutzerseiten und Nutzerdiskussionsseiten sowie anderweitig betriebener Webpräsenzen (z. B. Blogs, Twitter-Profile, Webseiten, andere Wiki-Projekte) und (4) den

5 Die empirische Öffentlichkeit einzelner Wiki-Seiten und Botschaften war zudem potenziell skalierbar und konnte relativ zu verschiedenen Faktoren auch in kurzer Zeit stark zunehmen, etwa wenn sie in (massen-)medialen Publikationen auftauchten. Spoeris (2007) Untersuchung der Top 100 an aufgerufenen Seiten zeigte, dass zwischen September 2006 und Januar 2007 25 Prozent der monatlich meist aufgerufenen Seiten auch nur in diesem Monat darunter zählten und dass nur 39 Prozent in allen Monaten prominent waren.

Abbildung 1: Episoden, Stationen und Stationenbereiche der teilnehmenden Beobachtung



Quelle: Eigene Darstellung

stark privaten Stationenbereich der Interviews und persönlichen Gespräche, online wie offline (siehe Abbildung 1).

Im Blick auf diese vier Stationenbereiche und die sich durch das Differenzierungskriterium ‚Öffentlichkeit‘ ergebenden forschungsethischen Umstände wurden zwei forschungsethische Problemfelder spezifiziert. Im *ersten Problemfeld Öffentlichkeit vs. Anonymität* wurde demnach pro Stationenbereich betrachtet, welche Konsequenzen die (bedingte) Öffentlichkeit von selbst-, fremd- oder systemgenerierten Informationen für das Gewährleisten von Anonymität hatte. Im *zweiten Problemfeld Anonymität vs. informierte Zustimmung* wurde für jeden Stationenbereich erfasst, welche Schwierigkeiten wiederum die Anonymität bzw. Pseudonymität des Interagierens online für das Absichern von informierter Zustimmung mit sich brachten.

Die jeweiligen Aspekte der beiden Problemfelder wurde dabei in Anlehnung an den Vorschlag von McKee und Porter (2008, S. 97) pro Stationenbereich in schematischen Rastern zusammengefasst, die zum einen die fallbezogenen Entscheidungen in den vier Stationenbereichen unterstützten und die zum anderen die gefundenen forschungsethischen Lösungen dokumentierten und so doppelt zur Validierung des Vorgehens beitrugen (siehe Abbildungen 2 und 3). Bei aller versuchten Schematisierung und Veranschaulichung des Vorgehens ist darauf hinzuweisen, dass diese Entscheidungen pro Problemfeld und Stationenbereich heuristisch getroffen werden mussten. Das bedeutet, dass forschungsethische Entscheidungen, wie dargelegt wurde, nicht notwendigerweise in dichotome Ja/Nein-Optionen zerfallen müssen. Vielmehr eröffnen sie ein Entscheidungskontinuum und verlangen das Abwägen der verschiedenen Ansprüche und das Finden von Kompromissen zwischen akzeptablen und inakzeptablen Entscheidungen.

In der Studie wurden auf diesem Weg forschungsethische Entscheidungen – angepasst an das jeweilige Problemfeld – in viererlei Hinsicht getroffen:

1. Darf eine Situation beobachtet werden bzw. darf ein Dokument gesampelt werden? Hier lag das Spektrum an Optionen zwischen: ohne Zustimmung beobachtbar/zu sampeln, bedingt ohne Zustimmung beobachtbar/zu sampeln, nur mit Zustimmung beobachtbar/zu sampeln und unbeobachtbar/nicht zu sampeln.
2. Darf Material zitiert werden? Hier lag das Spektrum an Optionen zwischen: ohne Bedenken zitierbar, mit leichten Bedenken zitierbar, mit starken Bedenken zitierbar und nicht zitierbar.
3. Kann informierte Zustimmung eingeholt werden? Hier lag das Spektrum an Optionen zwischen: völlig gewährleistet, gewährleistet, bedingt gewährleistet und unmöglich gewährleistet.
4. Muss informierte Zustimmung eingeholt werden? Hier lag das Spektrum an Optionen zwischen: notwendigerweise einzuholender informierter Zustimmung, sehr ratsam einzuholender informierter Zustimmung, ratsam einzuholender informierter Zustimmung und nicht notwendigerweise einzuholender informierter Zustimmung.

Im Folgenden wird dieser Entscheidungsprozess für beide Problemfelder ausgeführt und exemplarisch für zwei Stationenbereiche veranschaulicht.

4.1 *Forschungsethisches Problemfeld 1: Öffentlichkeit vs. Anonymität*

Die erste forschungsethische Schwierigkeit stellte sich bezüglich der Öffentlichkeit des Beteiligens an Wikipedia einerseits und der Wahrung der Anonymität der an Wikipedia beteiligten Autoren andererseits (Sveningsson-Elm, 2009). Forschungspraktisch war die *Öffentlichkeit des Beteiligens an Wikipedia und der von den Nutzern und von der Plattform erzeugten Informationen*, also das Fehlen von technologisch implementierten Zugangs- und Zugriffsbarrieren, von Vorteil. Wikipedia war also zum einen ein öffentliches und selbstdokumentierendes Projekt und daher der überwiegende Teil der Wiki-Seiten mit enzyklopädischen Inhalten, Aktivitätsaufzeichnungen, Diskussionen und Selbstauskünften ohne Hindernisse auch unangemeldeten Besuchern zugänglich. Zum anderen ging ein Teil der hochaktiven Nutzer offen mit ihrer Teilnahme und Autorenschaft um. Manche Autoren gaben auf ihren Benutzerseiten an, bei welchen Artikeln sie mitgewirkt hatten, Nutzernamen wurden aus dem Projekt herausgetragen und tauchten in der journalistischen Berichterstattung und in online-medialen Publikationen auf, Wikipedianer gaben Interviews und ließen sich fotografieren, sie traten mit ihrem bürgerlichen Namen und Tätigkeitsbeschreibungen auf Podien oder Konferenzen auf und gebrauchten ihre Nutzerseite, um mehr oder weniger ausführliche Informationen über sich preis zu geben. Die meistens Beiträge lagen archiviert vor und waren in der internen Wiki-Suche als auch in externen Suchmaschinen referenziert. Sie waren auf diesem Weg ähnlich einfach bzw. schwer zugänglich und auf Autoren rückführbar. Selbst wenn daher die Pseudonyme oder IP-Adressen der Autoren in den Ergebnisdarstellungen nochmals maskiert werden würden, könnten wörtliche Zitate über Suchfunktionen ohne Weiteres auf den jeweiligen angemeldeten Autor bzw. die in der Datenbank abgespeicherte IP-Adresse rückgeführt, mit anderen Daten kombiniert und eventuell zu anderen als den mit der Studie verfolgten Zwecken ausgewertet werden (Tilley & Woodthorpe, 2011). Darüber hinaus war das öffentliche Dokumentieren, welcher Account welchen Textedit getätigt hatte, gemäß der auf der Plattform gültigen Auslegung der Creative-Commons-Attribution-ShareAlike 3.0 (CC-BY-SA) eine lizenzrechtliche

Verpflichtung. Entsprechend wurden die Editoren in den Nutzungsbedingungen der Wikimedia Foundation (WMF) als US-amerikanischer Trägerorganisation der Plattform darauf hingewiesen, dass das lizenzkonforme Nennen ihres Nutzernamens bzw. ihrer IP-Adresse „1. durch einen Hyperlink (wenn möglich) oder eine URL auf die Seite, zu der du beigetragen hast, 2. durch einen Hyperlink (wenn möglich) oder eine URL auf eine alternative, stabile, frei zugängliche, lizenzkonforme Online-Kopie, welche die Autoren in einer zum ursprünglichen Text gleichwertigen Weise nennt, oder 3. durch eine Liste sämtlicher Autoren“ erfolgen müsste (WMF:Terms of Use, 2012).⁶

Forschungsethisch war indessen zu reflektieren, dass den Autoren der Wikipedia versichert wurde, sie hätten ein *Recht auf Anonymität*, sodass ihre Arbeit und Identität als Wikipedianer mit entsprechend frei gewähltem Pseudonym und ihr ziviles Leben unter bürgerlichem Namen abseits der Wikipedia im Prinzip zu trennen wären. Neben der Option, per IP-Adresse oder Pseudonym tätig zu werden, gewährleistete die Plattform Verfahren, um Klarnamen zu löschen und Pseudonyme von bürgerlichen Namen zu trennen. So konnten Administratoren einzelne kritische Versionen mit persönlich identifizierenden Inhalten entfernen lassen, die Benutzergruppe der Oversighter hatte Zugang zu depublizierten Informationen und konnte Edits entsprechend verbergen, mit einer IP registrierte Eingriffe konnten nachträglich Benutzernamen zugeordnet werden, für das unerlaubte ‚Outing‘ verantwortlich gemachte Nutzer konnten gesperrt werden und Unterlassungs- bzw. Schadenersatzklagen wurden angedroht. Mit ihrer 2008 vom *Board of Trustees* erlassenen *Data Retention Policy* erklärte die WMF zudem, nur „the least amount of personally identifiable information consistent with maintenance of its services, with its privacy policy, or as required by state or federal legal provisions under United States of America law“ (WMF:Data Retention Policy, 2008) von den

6 „All Projects of the Wikimedia Foundation are collaboratively developed by its users using the MediaWiki software. Anyone with Internet access (and not otherwise restricted from doing so) may edit the publicly editable pages of these sites with or without logging in as a registered user. By doing this, editors create a published document, and a public record of every word added, subtracted, or changed. This is a public act, and editors are identified publicly as the author of such changes. All contributions made to a Project, and all publicly available information about those contributions, are irrevocably licensed and may be freely copied, quoted, reused and adapted by third parties with few restrictions.“ WMF:Privacy Policy (2011).

Nutzern erheben zu wollen. Das Gewährleisten informationeller Selbstbestimmung der Wikipedia-Autoren umfasste somit auf der Plattform das Aufklären der Nutzer über ihre datenschutzrelevanten öffentlichen Aktionen und das Zusichern von Anonymität unter der Erwartung, die Nutzer machten kompetent von ihren Persönlichkeitsrechten Gebrauch (Schmidt, 2012). Der erlangte bzw. beibehaltene Grad an Anonymität stand so gesehen in der Verfügung der einzelnen Nutzer.⁷

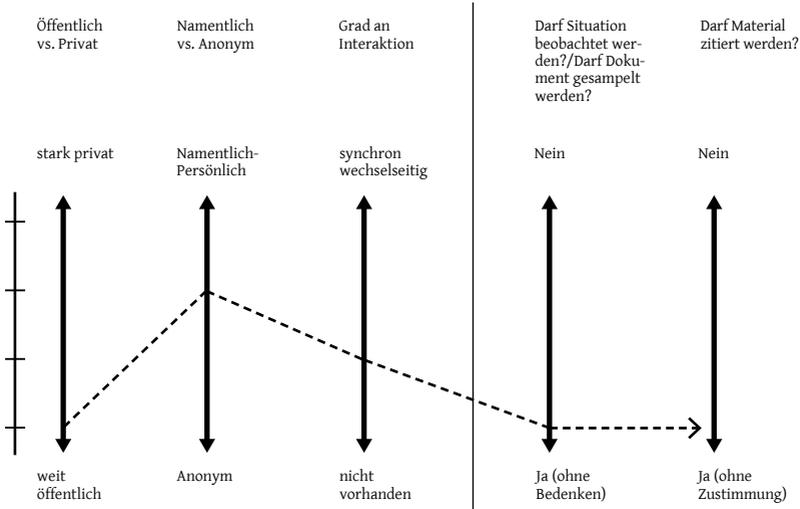
Im Blick auf die beiden in diesem Problemfeld zu treffenden forschungsethischen Entscheidungen, um Anonymität wenn möglich zu wahren, – Darf eine Situation beobachtet werden bzw. darf ein Dokument gesampelt werden?; Darf das Material zitiert werden? – waren hier pro Stationenbereich die Öffentlichkeitsdimensionen zwischen weit öffentlich, bedingt öffentlich, bedingt privat und stark privat sowie die namentliche bzw. anonyme Kennung der Akteure zwischen namentlich-persönlich, pseudonym-identifizierbar, pseudonym-nicht identifizierbar und anonym heuristisch zu bestimmen. Darüber hinaus war als zusätzlicher Faktor der Grad an Interaktion von Feld und Beobachter zwischen synchron wechselseitig, asynchron wechselseitig, einseitig und nicht vorhanden einzustufen. Unberücksichtigt blieben in diesem Entscheidungsraaster und im folgenden Entscheidungsraaster somit andere mögliche Faktoren wie die potenzielle Verletzbarkeit der einbezogenen Subjekte oder die Sensibilität der jeweiligen Episode, da diese nicht für Stationenbereiche generell bestimmt werden konnten.

Das Entscheidungsraaster und die Entscheidungsheuristik, exemplarisch für den weit öffentlichen Stationenbereich, gestaltete sich folglich wie in Abbildung 2 (siehe nächste Seite) dargestellt. Ein Beispiel für diesen Stationenbereich wäre die Episode einer Wikipedia-internen Reviewdiskussion, bei der ein Autor einen für gewöhnlich hauptsächlich von ihm editierten Artikel auf der Station einer eigens dafür eingerichteten Wiki-Seite zur Diskussion stellt, um so Feedback von anderen Autoren einzuholen und den Artikel zu verbessern.

Zusammenfassend wurden folgende forschungsethische Entscheidungen getroffen: Es wurden wörtliche Belegstellen in der Ergebnisdarstellung ohne Zustimmung nur aus den beiden Stationenbereichen zitiert, die von den Nutzern als stark bis bedingt öffentlich eingestuft wurden. Konkret wurde also darauf verzichtet, ohne Zu-

7 Die WMF Privacy Policy machte Einschränkungen dieser Zusicherung bzgl. der Nutzung von E-Mails, Mailinglisten und dem Open Ticket Request System (OTRS), mit dem Anfragen an das Projekt organisiert wurden.

Abbildung 2: Heuristisches forschungsethisches Entscheidungsrastrer:
Öffentlichkeit vs. Anonymität



Quelle: Eigene Darstellung

stimmung Zitate aus den Benutzer- und Benutzerdiskussion-Namensräumen, den Stammtischen und anderen plattformenunabhängigen Webpräsenzen anzuführen, wenn diese erkennbar von den jeweiligen Benutzern nicht für größere (Wikipedia-) Publika längerfristig öffentlich gehalten wurden. Das von der Lizenz geforderte Nennen der jeweiligen Autoren wurde dadurch abgegolten, dass im Forschungsbericht ein Link auf die jeweilige zitierte Wiki-Seite angegeben wurde.

4.2 Forschungsethisches Problemfeld 2: Anonymität vs. informierte Zustimmung

Die zweite Problemfeld stellte sich hinsichtlich der forschungsethischen Forderung nach *Anonymität der Autoren* und dem ebenso geforderten *Einholen informierter Zustimmung* von anonym (als IP-Adresse) bzw. pseudonym (mit Nutzernamen) agierenden Autoren, die dokumentierte Beiträge auf der Plattform

leisteten und unzuverlässig über ihre Nutzerseite erreichbar waren. Aktive Nutzer mit dokumentierten Beiträgen und gestalteter Nutzerseite konnten die Anfragen des Forschenden ignorieren, sie konnten aber auch ihre Mitarbeit beendet haben und deshalb entgingen ihnen die Aktivitäten auf ihrer Nutzerseite. Unangemeldete Nutzer wiederum hatten nur dann überhaupt Zugriff auf eine für sie angelegte Nutzerseite, wenn sie mit einer statischen IP-Adresse editierten, andernfalls verhinderten dynamische IP-Adressen den dauerhaften Zugang zu einer festen Nutzerseite.

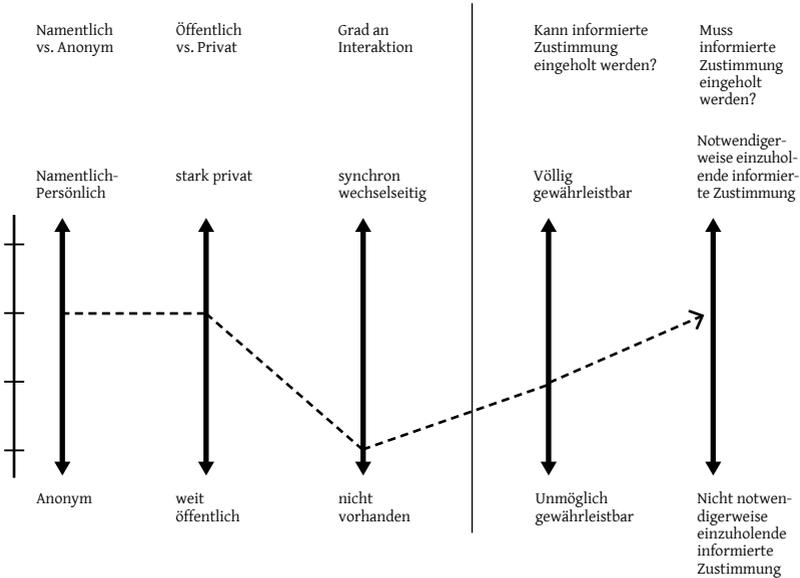
Grundsätzlich problematisch war dabei, dass informierte Zustimmung forschungsethisch plausibel als Aushandlung stattfinden sollte, in der das Vorgehen und die Erwartungen der Forschenden mit den Ansprüchen der in die Untersuchung einbezogenen Personen im Fortgang der Studie abzugleichen wären. So wurde erst im Laufe der Analyse festgelegt, welche Wikipedia-Seiten, Webdokumente, Beobachtungen und Gespräche in die Datensammlung aufgenommen wurden, welche davon wiederum gemäß Grounded Theory ausgewertet wurden und ob sie eventuell in ihrer originalen Form oder paraphrasiert in den Forschungsbericht eingehen sollten. Auf diese Weise war der Prozesshaftigkeit des methodischen Erkenntnisvorgangs Rechnung zu tragen und zugleich die Freiwilligkeit der Teilnahme und das angemessene Informieren der Entscheidung der studierten Akteure, fortwährend teilzunehmen, zu berücksichtigen (Hopf, 2005, S. 592). In der Praxis war dies vor allem im Sammeln von Dokumenten und dem retrospektiven Beobachten dokumentierter Aktionen nur bedingt möglich. In beiden Fällen konnten die betroffenen anonymen oder pseudonymen Editoren nur nachträglich mittels ihrer Nutzerseite angefragt und um Zustimmung gebeten werden, weil die Autoren selten eine E-Mail-Adresse angegeben hatten. Diese Kontaktaufnahme per Nutzerseite war nicht privat, sondern in ihrer Form öffentlich und stellte so Öffentlichkeit für die erfolgte bzw. eventuell zu erfolgende Beobachtung und Analyse her und schuf bereits eine Verbindung zwischen der Studie und den Nutzern, noch bevor sie überhaupt die Chance hatten, sich begründet gegen eine Teilnahme zu entscheiden. Wenn die Nutzer dynamischer IP-Adressen keinen dauerhaften Zugriff auf ihre Nutzer- und ihre Nutzerdiskussionsseite hatten, Autoren unter statischen IP-Adressen ihre Diskussionsseite nicht verfolgten oder einmal aktive Accounts verwaist waren, dann blieben die Anfragen zudem unbeantwortet.

Im Blick auf die zwei in diesem Problemfeld zu treffenden forschungsethischen Entscheidungen, um informierte Zustimmung wenn nötig zu gewähren, – Kann

informierte Zustimmung eingeholt werden? Muss sie eingeholt werden? – waren hier pro Stationenbereich die Öffentlichkeitsdimensionen zwischen weit öffentlich, bedingt öffentlich, bedingt privat und stark privat sowie die namentliche bzw. anonyme Kennung der Akteure zwischen namentlich-persönlich, pseudonym-identifizierbar, pseudonym-nicht identifizierbar und anonym heuristisch zu bestimmen. Darüber hinaus war als zusätzlicher Faktor der Grad an Interaktion von Feld und Beobachter zwischen synchron wechselseitig, asynchron wechselseitig, einseitig und nicht vorhanden einzustufen.

Das Entscheidungsraaster und die Entscheidungsheuristik, exemplarisch für den bedingt privaten Stationenbereich, gestaltete sich entsprechend wie in Abbildung 3 dargestellt. Ein Beispiel für diesen Stationenbereich wäre die Episode

Abbildung 3: Heuristisches forschungsethisches Entscheidungsraaster: Anonymität vs. informierte Zustimmung



Quelle: Eigene Darstellung

einer Wikipedia-internen Diskussion, bei der angemeldete und unangemeldete Autoren die Station der Diskussionsseite eines Autors gebrauchten, um die Berechtigung einer Benutzersperre wegen des Vorwurfs, ein Editor habe Einträge vandalisiert, zu debattieren.

Zusammenfassend wurden folgende forschungsethische Entscheidungen getroffen: Es wurde eine offene Beobachtung durchgeführt, insoweit die Plattform ein Offenbaren des beobachtenden Teilnehmers und seiner Ziele zuließ. Das hieß, als Pseudonym wurde der Klarname gewählt und auf der entsprechend verlinkten Nutzerseite von Christianpentzold fanden sich Informationen zu dem Forschungsvorhaben und es wurde die Möglichkeit gegeben, per E-Mail auch außerhalb von Wikipedia in Kontakt zu treten. Jeder Beitrag in Wikipedia wurde angemeldet getätigt und jeder Beitrag zu Diskussionsseiten signiert, sodass die Aktivitäten der Accounts nachverfolgt werden konnten. Wenn Interaktionen mit angemeldeten Nutzern stattfanden, wurden diese auf das Forschungsvorhaben aufmerksam gemacht und den angefragten Autoren wurde das Angebot gemacht, den Kontakt per Mail oder Telefon fortzusetzen.⁸

5 Ausblick: Forschungsethische Zusicherung unter forschungspraktischer Unsicherheit

Die forschungsethischen Verfahrensentscheidungen der teilnehmenden Beobachtung an Wikipedia konzentrierten sich im ersten Problemfeld Öffentlichkeit vs. Anonymität auf den Umgang mit den öffentlichen Aktivitätsaufzeichnungen und öffentlichen Selbstauskünften der Autoren, der Identifizierbarkeit aller im Wiki gemachten Edits und des Umgangs der Nutzer mit Publizität zum einen und den lizenzrechtlichen Anforderungen nach Autorennennung und den methodischen Kriterien der Authentizität des empirischen Materials sowie der Validität seiner Interpretation zum anderen. Im zweiten Problemfeld Anonymität vs. informierte

8 Als Interviewpartner wurden zudem nur volljährige Teilnehmer ausgewählt. Sie wurden vorab über den Zweck der Studie, die Art der Auswertung und die Publikationsart informiert, ihnen wurde das transkribierte Interview zugesandt und das Zugänglichmachen der Ergebnisse zugesichert. Die Pseudonyme der interviewten Wikipedianer wurden nochmals maskiert.

Zustimmung galt es wiederum, einerseits die unzuverlässige Ansprechbarkeit pseudonymer bzw. anonymer Nutzer mit der Forderung nach dem Einholen informierter Zustimmung andererseits forschungspraktisch auszubalancieren.

Mit diesen Problemfeldern und den darauf erfolgten Entscheidungen wurde versucht, das teilnehmende Beobachten auf der Online-Plattform Wikipedia und im plattformunabhängigen Handlungsraum der Wikipedianer methodisch viabel und forschungsethisch plausibel zu gestalten. Dieses Vorgehen kann aber nicht alle möglichen Unsicherheiten, unter denen forschungsethische Integrität hergestellt und zugesichert werden sollte, ausräumen. Vielmehr verweist die Studie auf drei allgemeine Probleme, die nicht durch forschungspraktische Vorkehrungen innerhalb eines Untersuchungsrahmens aufgelöst werden können.

Die *erste Frage der mediatisierten Forschung* betrifft das Problem, *wie Vertraulichkeit und Anonymität überhaupt zugesichert werden können, wenn Daten und Datenauswertung nicht mehr souverän zu regeln sind*. Trifft die Diagnose des zunehmenden medialen Durchdringens alltäglicher und professioneller Lebenswelten zu, dann wachsen dadurch wohl die empirischen Felder, die durch kommunikationswissenschaftliche Ansätze erschlossen und erklärt werden können (Krotz & Hepp, 2012). Zugleich aber sind die hier verfügbaren Daten in steigendem Maße vermittelt und von Dritten wie den meist kommerziellen Anbietern von Plattformen und Kommunikations- und Informationsdiensten die Facebook, Twitter oder den zahlreichen Apps generiert, sodass sie auch, aber nicht exklusiv, von kommunikationswissenschaftlichen Untersuchungen genutzt werden können. Im Fall der Untersuchung online-medialer Plattformen wird dieses Thema umso virulenter, je stärker die Betreiber der Plattformen selber ein Interesse verfolgen, digitale Methoden voranzubringen, um Nutzerdaten zu generieren und auszuwerten (boyd & Crawford, 2012).

Mit der *zweiten Frage nach den potenziellen Kontexten und potenzierten Analyseoptionen* wird problematisch, *wie die Grenzen voraussehbaren Schadens bestimmt werden können und ob mit ihnen die forschungsethische Verantwortung endet*, gerade wenn sozial- und speziell kommunikationswissenschaftliche Forschung nur einer unter anderen Sektoren mit analytisch-beschreibendem Interesse an digitaler Kommunikation ist (Savage, 2009). In dieser Hinsicht generiert akademische Forschung als eine Instanz unter anderen, darunter Plattformbetreibern, Netzbetreibern, Informationsdienstleistern und staatlichen Agenturen, Daten und Interpretationen. Diese können – gerade wenn sie digitalisiert offen gemacht werden – unter

Umständen auch für andere, mit der ursprünglichen Forschungsabsicht kongruente oder konträre, Zwecke gebraucht werden.

Schließlich ist für die *dritte Frage der Dynamik forschungsethischer Grenzen angesichts analytischer Neuerungen* zu klären, *inwiefern methodische und technologische Innovationen forschungsethische Probleme lösen können, ohne neue zu schaffen*. Dem Wandel der Gegenstände und Felder sozialwissenschaftlicher Forschung durch Digitalisierung, Telekommunikationsnetzwerke und mobile Endgeräte begegnen Digital Methods damit, dass sie sich gerade diese Entwicklungen mit neuen Verfahren und Instrumenten zunutze machen wollen (Rogers, 2013). Während aber die Nützlichkeit und Verlässlichkeit der so bereitgestellten Softwaretools und Abläufe schon mehrfach erwiesen wurde, steht ihre forschungsethische Bewertung noch am Anfang.

Dr. Christian Pentzold ist wissenschaftlicher Mitarbeiter am Institut für Medienforschung und Post-Doc am DFG-Graduiertenkolleg „Crossworlds. Connecting Virtual and Real Social Worlds“ der Technischen Universität Chemnitz sowie assoziierter Forscher am Alexander von Humboldt Institut für Internet und Gesellschaft in Berlin

Quellenverzeichnis

- Albers, M. (2005). *Informationelle Selbstbestimmung*. Baden-Baden: Nomos.
- ASA American Sociological Association (1999). *Code of Ethics*. Abgerufen von <http://www.asanet.org/about/ethics.cfm>
- Ayers, P., Matthews, C., & Yates, B. (2008). *How Wikipedia Works. And how you can be part of it*. San Francisco: No Starch Press.
- Basset, E., & O’Riordan, K. (2002). Ethics of Internet Research. Contesting the human subjects research model. *Ethics and Information Technology*, 4(3), 233-247.
- Baym, N., & boyd, d. (2012). Socially Mediated Publicness. *Journal of Broadcasting and Electronic Media*, 56(3), 320-329. doi: 10.1080/08838151.2012.705200
- Benkler, Y. (2006). *The Wealth of Networks*. New Haven, CT: Yale University Press.
- Boellstorff, T., Nardi, B., Pearce, C., & Taylor, T.L. (2012). *Ethnography and Virtual Worlds. A Handbook of Method*. Princeton, NJ: Princeton University Press.
- boyd, d., & Crawford, K. (2012). Critical Questions for Big Data. *Information, Communication & Society*, 15(5), 662-679. doi: 10.1080/1369118X.2012.678878

- Buchanan, E., & Ess, C. (2009). Internet Research Ethics and the Institutional Review Board: Current practices and issues. *Computers and Society*, 39(3), 43-49. doi: 10.1145/1713066.1713069
- Coleman, G. (2013). *Coding Freedom: The ethics and aesthetics of hacking*. Princeton: Princeton University Press.
- DGPuK Deutsche Gesellschaft für Publizistik und Kommunikationswissenschaft (1999). *Erklärung der Deutschen Gesellschaft für Publizistik und Kommunikationswissenschaft zur Sicherung guter wissenschaftlicher Praxis*. Abgerufen von <http://www.dgpuk.de/uber-die-dgpuk/ethikerklarung/>
- DGOF Deutsche Gesellschaft für Online-Forschung (2007). *Richtlinien für Online-Befragungen*. Abgerufen von http://rat-marktforschung.de/fileadmin/user_upload/pdf/R08_RDMS.pdf
- DGS Deutsche Gesellschaft für Soziologie (2014). *Ethik-Kodex der Deutschen Gesellschaft für Soziologie und des Berufsverbandes deutscher Soziologinnen und Soziologen*. Abgerufen von <http://www.sociologie.de/index.php?id=19>
- van Dijck, J. (2013). *The Culture of Connectivity*. Oxford: Oxford University Press.
- van Dijk, Z. (2010). *Wikipedia. Wie Sie zur freien Enzyklopädie beitragen*. München: Open Source Press.
- Ess, C. (2006). Ethics and the use of the Internet in social science research. In A. Joinson, T. Postmes, & U.-D. Reips (Hrsg.), *Oxford Handbook of Internet Psychology* (S. 487-503). Oxford: Oxford University Press.
- Ess, C. (2009). *Digital Media Ethics*. Cambridge: Polity.
- Ess, C., & AoIR Ethics Working Committee (2002). *Ethical decision-making and Internet research. Recommendations from the AoIR Ethics Working Committee*. Abgerufen von <http://aoir.org/reports/ethics.pdf>
- Eynon, R., Fry, J., & Schroeder, R. (2008). The Ethics of Internet research. In N. Fielding, R. Lee, & G. Blank (Hrsg.), *The Sage Handbook of Internet Research* (S. 23-41). Thousand Oaks: Sage.
- Fiebig, H. (2005). *Wikipedia. Das Buch*. Berlin: Zenodot.
- Giddens, A. (1997). *Die Konstitution der Gesellschaft. Grundzüge einer Theorie der Strukturierung*. Frankfurt a. M., New York: Campus.
- Gillespie, T. (2010). The Politics of "Platforms". *New Media & Society*, 12(3), 347-364. doi: 10.1177/1461444809342738
- Glaser, B., & Strauss, A. (1967/2005). *Grounded Theory. Strategien qualitativer Forschung*. Bern: Huber.

- Goffman, E. (1983). The Interaction Order. *American Sociological Review*, 48(1), 1-17.
- Gola, P., & Schomerus, R. (2010). *BDSG Bundesdatenschutzgesetz. Kommentar*. München: C.H. Beck.
- Hammersley, M., & Atkinson, P. (2007). *Ethnography: Principles in practice*. London: Routledge.
- Heider, D., & Massanari, A. (Hrsg.) (2012). *Digital Ethics. Research and Practice*. New York et al.: Peter Lang.
- Hopf, C. (2005). Forschungsethik und qualitative Forschung. In U. Flick, E. von Kardoff, & I. Steinke (Hrsg.), *Qualitative Forschung. Ein Handbuch* (S. 589-600). Reinbek bei Hamburg: Rowohlt.
- Hudson, J., & Bruckman, A. (2004). "Go away": Participant objections to being studied and the ethics of chatroom research. *The Information Society*, 20, 127-139. doi: 10.1080/01972240490423030
- ICA International Communication Association (o.J.). *Ethics Statement*. Abgerufen von http://www.icahdq.org/about_ica/ethics.asp
- ICC/ESOMAR European Society for Opinion and Market Research (2011). *ESOMAR Guideline for Online Research*. Abgerufen von http://www.esomar.org/uploads/public/knowledge-and-standards/codes-and-guidelines/ESOMAR_Guideline-for-online-research.pdf
- Kozinets, R. (2010). *Netnography. Doing Ethnographic Research Online*. London: Sage.
- Kraut, R., Olson, J., Banaji, M., Bruckman, A., Cohen, J., & Couper, M. (2004). Psychological Research Online: Report of Board of Scientific Affairs' Advisory Group on the Conduct of Research on the Internet. *American Psychologist*, 59(2), 105-117. doi: 10.1037/0003-066X.59.2.105
- Krotz, F. (2005). *Neue Theorien entwickeln. Eine Einführung in die Grounded Theory, die heuristische Sozialforschung und die Ethnographie anhand von Beispielen aus der Kommunikationsforschung*. Köln: Herbert von Halem.
- Krotz, F., & Hepp, A. (Hrsg.) (2012). *Mediatisierte Welten. Forschungsfelder und Beschreibungsansätze*. Wiesbaden: VS Verlag.
- Mahrt, M., & Scharrow, M. (2013). The Value of Big Data for Digital Media Research. *Journal of Broadcasting and Electronic Media*, 57(1), 20-33. doi: 10.1080/08838151.2012.761700
- Markham, A., Buchanan, E., & AoIR Ethics Working Committee (2012). *Ethical decision-making and Internet Research 2.0: Recommendations from the AoIR Ethics Working Committee*. Abgerufen von <http://aoir.org/reports/ethics2.pdf>

- McKee, H., & Porter, J. (2008). *The Ethics of Digital Writing Research: A Rhetorical Approach*. New York: Peter Lang.
- Nissenbaum, H. (2011). *Privacy in Context. Technology, Policy, and the Integrity of Social Life*. Palo Alto: Stanford University Press.
- Pentzold, C. (2007). *Wikipedia: Diskussionsraum und Informationsspeicher im neuen Netz*. München: Reinhard Fischer.
- Pentzold, C. (2013). *Die Praxis der Online-Kooperation. Praktiken und Institutionen online-medialer Zusammenarbeit*. Unveröff. Dissertation, Philosophische Fakultät, Technische Universität Chemnitz.
- Rogers, R. (2013). *Digital Methods*. Cambridge, MA: MIT Press.
- Roth, W.-M. (2004). Qualitative Forschung und Ethik. *Forum Qualitative Sozialforschung*, 5(2). urn: urn:nbn:de:0114-fqs040275
- Savage, M. (2009). Contemporary sociology and the challenge of descriptive assemblage. *European Journal of Social Theory*, 12(1), 155-174. doi: 10.1177/1368431008099650
- Schmidt, J. (2009). *Das neue Netz*. Konstanz: UVK.
- Schmidt, J. (2012). Persönliche Öffentlichkeiten und informationelle Selbstbestimmung im Social Web. In J. Schmidt & T. Weichert (Hrsg.), *Datenschutz. Grundlagen, Entwicklungen und Kontroversen* (S. 215-225). Bonn: Bundeszentrale für politische Bildung.
- Schmidt, J., & Weichert, T. (Hrsg.) (2012). *Datenschutz. Grundlagen, Entwicklungen, Kontroversen*. Bonn: Bundeszentrale für politische Bildung.
- Spoerri, A. (2007). What is Popular on Wikipedia and Why? *First Monday*, 12(4). Abgerufen von <http://firstmonday.org/ojs/index.php/fm/article/view/1765/1645>
- Stegbauer, C. (2009). *Wikipedia. Das Rätsel der Kooperation*. Wiesbaden: VS Verlag.
- Stöcklin, N. (2010). *Wikipedia clever nutzen - in Schule und Beruf*. Zürich: Orell Füssli.
- Sveningsson-Elm, M. (2009). How Do Various Notions of Privacy Influence Decisions in Qualitative Internet Research? In A. Markham & N. Baym (Hrsg.), *Internet Inquiry* (S. 69-87). London: Sage.
- Tilley, L., & Woodthorpe, K. (2011). Is it the end for anonymity as we know it? A critical examination of the ethical principle of anonymity in the context of 21st century demands on the qualitative researcher. *Qualitative Research*, 11(2), 197-212. doi: 10.1177/1468794110394073

- von Unger, H., Narimani, P., & M'Bayo, R. (Hrsg.) (2014). *Forschungsethik in der qualitativen Forschung. Reflexivität, Perspektiven, Positionen*. Wiesbaden: Springer VS.
- Walther, J. (2002). Research ethics in Internet-enabled research: Human subjects issues and methodological myopia. *Ethics and Information Technology*, 4(3), 205-216. doi: 10.1023/A:1021368426115
- Waskul, D. & Douglass, M. (1996). Considering the electronic participant: Some polemical observations on the ethics of on-line research. *The Information Society*, 12(2), 129-140. doi: 10.1080/713856142
- WMF Data Retention Policy (2008). Abgerufen von http://wikimediafoundation.org/wiki/Data_retention_policy
- WMF Privacy Policy (2011). Abgerufen von http://meta.wikimedia.org/wiki/Privacy_policy
- WMF:Terms of Use (2012). Abgerufen von http://wikimediafoundation.org/wiki/Terms_of_Use
- Wikipedia:Ethically researching Wikipedia (2012). Abgerufen von http://en.wikipedia.org/wiki/Wikipedia:Ethically_researching_Wikipedia
- Wikipedia:Statistics (2014). Abgerufen von <http://en.wikipedia.org/wiki/Special:Statistics>
- Wikipedia:Statistik (2014). Abgerufen von <http://de.wikipedia.org/wiki/Spezial:Statistik>

Empfohlene Zitierung: Ausserhofer, J. (2015). „Die Methode liegt im Code“: Routinen und digitale Methoden im Datenjournalismus. In A. Maireder, J. Ausserhofer, C. Schumann, & M. Taddicken (Hrsg.), *Digitale Methoden in der Kommunikationswissenschaft* (S. 87-111). doi: 10.17174/dcr.v2.5

Zusammenfassung: Digitale Methoden haben in den letzten Jahren nicht nur in die Sozial- und Geisteswissenschaften Einzug gehalten. Auch der Journalismus hat seine Verfahren digitalisiert und seine Recherchepraktiken sowie seine Darstellungsformen entsprechend angepasst. Immer öfter bedienen sich JournalistInnen für ihre Berichterstattung des Werkzeugkastens der digitalen Methoden. Am deutlichsten zeigt sich das am Datenjournalismus, der seit einem Jahrzehnt einen enormen Aufschwung erfahren hat. Dieser Beitrag beleuchtet digitale Methoden im Datenjournalismus. Zu diesem Zweck werden hier zunächst aus den Science and Technology Studies und der Digital-Methods-Forschung heraus die Grundzüge einer datenjournalistischen Methodologie entwickelt. Auf Grundlage von Redaktionsbesuchen und elf semistrukturierten Interviews mit PraktikerInnen werden im Anschluss konkrete Arbeitsabläufe und Akteure in datenjournalistischen Methoden beschrieben. Es wird erläutert, wie datenjournalistische Kollaborationen im Newsroom funktionieren, wie sich dort Routinen etablieren und welche Technologien dabei vermitteln. Abschließend werden Empfehlungen zur Weiterentwicklung digitaler Methoden in der Kommunikationswissenschaft und im Datenjournalismus formuliert.

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

Julian Ausserhofer

„Die Methode liegt im Code“

Routinen und digitale Methoden im Datenjournalismus

1 Einleitung

Am 15. April 2013 explodierten zwei Bomben in der Zielgeraden des Marathons von Boston. Drei Menschen verloren ihr Leben, mehr als 260 weitere wurden teils schwer verletzt. Während vor Ort Sicherheits- und Rettungsmaßnahmen eingeleitet wurden, machte sich die Polizei an die Aufklärung der Tat. Auch ReporterInnen und Laien versuchten zur Aufklärung beizutragen und sammelten auf Online-Plattformen sämtliche verfügbaren Informationen zum Attentat. Sie analysierten Fotos und Videos, transkribierten den Polizeifunk, glichen die Ergebnisse mit Datenbanken, Karten und Social Networks ab. Nachdem die Polizei die Namen von zwei Verdächtigen veröffentlicht hatte, dauerte es nicht lange bis auch deren Social-Media-Profile öffentlich bekannt waren (Madrigal, 2013). Viele Nachrichtenorganisationen zitierten in der Folge die Tweets von Dzhokhar Tsarnaev, dem jüngeren der zwei Brüder (z.B. Preston & Roston, 2013). Die Medienhäuser behandelten Twitter als eine weitere Textquelle neben vielen anderen. Wenige andere Nachrichtenseiten wählten einen alternativen Ansatz: Um mehr über den Verdächtigen und dessen Lebensumstände zu erfahren, beschäftigten sie sich mit den Metadaten seines Twitter-Profiles. So extrapolierte die Website *Quartz* etwa die Zeitpunkte, zu denen der mutmaßliche Attentäter im Laufe von

eineinhalb Jahren seine rund tausend Tweets abgesendet hatte (Seward et al., 2013). Das Resultat: Eine Aufschlüsselung nach Wochentagen und Stunden über Tsarnaevs durchschnittliche Twitter-Aktivität, woraus die Autoren ableiteten, wann er für gewöhnlich schlafen würde.¹ Die Nachrichtenseite *Digg* veröffentlichte wenige Tage später eine umfassende Analyse von Tsarnaevs Twitternetzwerk und des *#FreeJahar*-Hashtags, unter dem die Freilassung und Schuld von Tsarnaev diskutiert wurden (Lotan, 2013). Mit Hilfe einer Clusteranalyse wurden Gruppen identifiziert, mit denen Tsarnaev auf Twitter verbunden war. Ausgewählte anonymisierte Zitate aus deren Timelines komplettierten das Bild über einen Informationskanal, den der mutmaßliche Attentäter wohl stark verwendet hat.

Derartige „computerunterstützte Explorationen im Journalismus“ (Gynnild, 2014, S. 3) haben in den vergangenen Jahren stark zugenommen. Der experimentelle Einsatz von Algorithmen, Daten und sozialwissenschaftlichen Methoden in der Nachrichtenproduktion – von der Recherche über die Datenanalyse bis hin zur Visualisierung – wird im deutschsprachigen Raum vor allem unter dem Begriff Datenjournalismus diskutiert (für eine Begriffstypologie siehe Coddington, 2014).² Wie die eingangs beschriebenen Twitter-Analysen beispielhaft demonstrieren, werden im Datenjournalismus immer wieder Daten verwendet, die von Webplattformen generiert werden. Auch wenn es nicht so titulierte wird, werden auch „digitale Methoden“ eingesetzt: DatenjournalistInnen untersuchen wiederkehrend Themen, Akteure oder Strukturen von Online-Kommunikation. In leitenden Fragestellungen wie auch in eingesetzten Methoden zeigen sich starke

- 1 Der Vollständigkeit halber soll nicht unerwähnt bleiben, dass die Autoren später im Artikel derartige Recherchen durchaus kritisch hinterfragen: „Where it was once only reporters and the police who dug up information about people of interest, a whole nation is at it today. [...] But those caveats are almost beside the point. We don't know what we think we know because these digital details don't connect the dots; they merely draw the dots. They offer trivia but not insight.“ (Seward et al., 2013)
- 2 Neben der Bezeichnung „Datenjournalismus“ haben sich je nach Medium, Region oder disziplinärem Hintergrund auch andere Begriffe etabliert: Data-driven journalism, precision journalism, computational journalism, database journalism, structured journalism und computational journalism sind nur einige der vielen Ausdrücke, die im Umlauf sind. Mit graduellen Unterschieden benennen sie die gleichen Tätigkeiten. Aus Gründen der Kürze und Einfachheit verwende ich im Folgenden auch immer wieder die Bezeichnung „DatenjournalistIn“, wohlwissend dass dieser Begriff so wie auch Datenjournalismus Unschärfen beinhaltet und es „den“ Datenjournalisten nicht gibt.

Parallelen zur sozialwissenschaftlichen Internetforschung. Aber ähnlich wie beim „Geschwisterpaar“ empirischer Journalismus und empirische Sozialwissenschaft (Haas, 1999) gibt es auch deutliche Unterschiede zwischen Datenjournalismus und sozialwissenschaftlicher Internetforschung. Das wechselseitige Verständnis des jeweils anderen Metiers kann indes fruchtbare Impulse zur Verbesserung des eigenen Metiers liefern (Rühl, 1981, zitiert nach Haas, 1999, S. 111).

Vor diesem Hintergrund beschreibt dieser Aufsatz Routinen, Arbeitsabläufe und Technologien im Datenjournalismus. Für die Kommunikationswissenschaft will dieser Text ein „Blick über den Tellerrand“ sein: Eine Analyse der digitalen Methoden des Datenjournalismus soll der sozialwissenschaftlichen Internetforschung alternative Herangehensweisen zeigen. Für Datenjournalismus-Interessierte bietet der Beitrag eine strukturierte Beschreibung der Arbeitsabläufe in datenjournalistischen Projekten (Abschnitt 5). Bevor das passiert, soll jedoch noch erörtert werden, was unter „digitalen Methoden“ verstanden wird und wie sich diese erforschen lassen (Abschnitt 2). Auch wird dargelegt, was die Forschung bislang zu datenjournalistischen Routinen und Methoden weiß (Abschnitt 3) und wie für diesen Forschungsbeitrag vorgegangen wurde (Abschnitt 4).

2 Methodologischer Zugriff: Digitale Methoden, ihre Verteilung und Erforschung

Wenn im Folgenden von digitalen Methoden die Rede ist, dann bezieht sich das auf den Entwurf von Richard Rogers (2009, 2010, 2013) und die von ihm gegründete *Digital Methods Initiative* (DMI) an der Universität Amsterdam. Formuliert als Methodologie für die Sozial- und Geisteswissenschaften unterscheidet Rogers zwischen digitalisierten und digitalen Methoden: Erstere umfassen ursprünglich analoge Forschungsansätze, die für das Internet adaptiert wurden – Ethnografien oder Fragebögen etwa. Zumeist beschäftigen sich Forschungsprojekte, die diese Methoden zur Grundlage haben, mit netzkulturellen Phänomenen oder sie nutzen das Web schlicht als Kommunikationsmedium im Forschungsprozess. Digitale Methoden hingegen sind Teil des „Computational Turn“ (Berry, 2011), sie sind „nativ digital“ (R. Rogers, 2014), sie wurden also speziell für Online-Forschungsumgebungen entwickelt und haben kein direktes analoges Vorbild. Auch die Forschungsperspektive ist bei digitalen Methoden eine andere: Das Web wird

nicht als virtueller Raum gesehen, dessen Kultur durch Forschung durchdrungen wird; die Analyse von Online-Umgebungen soll vielmehr dazu beitragen, Aspekte gesamtgesellschaftlicher Kommunikation und Kultur zu verstehen: „Rather, in studying the online, we make and ground findings about society and culture with the Internet. Thus, the Internet is a research site where one can ground findings about reality.” (R. Rogers, 2010, S. 243)

So beschäftigen sich Digital-Methods-Forschungsprojekte etwa mit der Rolle von Bots³ in der Wissensakkumulation bei Wikipedia (Niederer & van Dijck, 2010), den Plattformen der holländischen Blogosphäre (Weltevrede & Helmond, 2012) oder den Online-Kontroversen zum Klimawandel (Niederer, 2013). Die Forschungsinteressen sind durch einen starken Fokus auf Plattformen und deren digitale Objekte wie Hyperlinks, Likes oder Tags gekennzeichnet. Automatisierungen wie etwa das „Crawlen“ von Netzwerken und Websites oder das „Scrapen“⁴ von Social-Media-Inhalten sind typische Datenbeschaffungsvorgänge. Die Ergebnispräsentationen sind geprägt von Visualisierungen wie Netzwerken, Wortwolken oder Zeitstrahlen.

Unter welchen methodologischen Prämissen lassen sich digitale Methoden erforschen? Nortje Marres (2012), die starke Verbindungen zur *Digital Methods Initiative* hat, argumentiert unter Rekurs auf unter anderem Bruno Latour (1988) und John Law (2004), dass Methoden ein verteiltes Unterfangen darstellen. Anschließend an die Tradition der Akteur-Netzwerk-Theorie (ANT) sieht sie Methoden als gemeinsame Entwicklung von ForscherInnen, Forschungssubjekten und -objekten sowie Technologien. Neues Wissen und Innovation in Forschungsprozessen werden durch komplexe Aushandlungsprozesse arbeitsteilig zwischen ebendiesen Akteuren entwickelt. Die Digitalisierung, so Marres, habe sozialwissenschaftliche Methoden stark verändert und etwa Online-Plattformen als wichtige Akteure mit Forschungskapazität etabliert. Weitere Agenten digitaler Methoden seien neben digitalen Objekten (R. Rogers, 2009) unter anderem auch Datenbanken, Designagenturen, Algorithmen, IT-Unternehmen oder BürgerwissenschaftlerInnen.

- 3 Ein Bot ist eine Softwareanwendung, die automatische einfache, repetitive Aufgaben im Internet ausführt.
- 4 Crawler und Scraper sind Skripte, die ganze Webseiten bzw. Datenbanken oder Teile davon abgreifen, auf eine lokale Repräsentanz kopieren und für die weitere Analyse aufbereiten.

Führt man die von Marres aufgebrachte Methodologie konsequent weiter, dann stellt sich vor dem Hintergrund dieses Aufsatzes die Frage, inwieweit man die Handlungsmacht von Agenten, die an digitalen Methoden im Datenjournalismus beteiligt sind, konsequent operationalisieren und damit erforschen kann. Es liegt nahe, auf die Aushandlungsprozesse und die daraus resultierende verbundene Arbeitsteilung zwischen den einzelnen Akteuren zu fokussieren. Um die Handlungsmacht der involvierten Plattformen und Software-Agenten zu bestimmen, eignen sich Forschungsansätze wie jener der Software Studies (Manovich, 2013) oder der digitalen Methoden selbst – unter der Voraussetzung, dass eine selbst-reflexive Offenlegung⁵ des Vorgehens stattfindet. Zur Analyse der physischen Situationen in den Redaktionen bieten sich neben Newsroom-Ethnografien (z.B. Domingo & Paterson, 2011) Forschungsprogramme wie die ANT (Latour, 2005) oder die von Robert Schmidt (2012) entwickelte Soziologie der Praktiken an, die verschiedene Stränge sozialwissenschaftlicher Forschung inkorporiert. Schmidt vereint für seinen praxeologischen Entwurf Ansätze der Ethnomethodologie, der Interaktionsanalyse, der ANT sowie Bourdieus ethnologischen Feldstudien. Im Mittelpunkt von Schmidts Ansatz stehen die Beobachtbarkeit sozialer Praktiken sowie die Körperlichkeit und Materialität bei Interaktionen.

Im Ergebnisteil versucht diese Arbeit, diese Perspektiven auf digitale Methoden und die Möglichkeit ihrer Erforschung in die Praxis umzusetzen und insbesondere die Bandbreite und die Eigenschaften der Akteure, die an datenjournalistischen Methoden beteiligt sind, zu ergründen. Empirisch wird neben Beobachtungsprotokollen auf Interviewmaterial zurückgegriffen.

3 Forschungsstand: Datenjournalistische Routinen

Die Dichotomie von Euphorie und Skepsis, die regelmäßig in der Frühphase von gesellschaftlichen Debatten über neue technologiegetriebene Praktiken auftritt (Münch & Schmidt, 2005), ist auch beim Datenjournalismus festzustellen;

5 Die selbstreflexive Offenlegung umfasst die Transparenzmachung von Entscheidungen und Interpretationen im Forschungsprozess sowie die Veröffentlichung von Forschungsartefakten wie Bibliografien, Notizen, Algorithmen und rohen sowie prozessierten Daten (siehe dazu Ausserhofer, 2015; Rieder & Röhl, 2012).

wobei in der (halb-)akademischen Literatur die optimistische Sichtweise überwiegt (Anderson, 2013) – ganz ähnlich wie bei den Diskursen zu partizipativem Journalismus (Borger et al., 2013). Frühe Veröffentlichungen zu Datenjournalismus weisen insbesondere auf die sich durch die neue Technologie eröffnenden Potenziale für den Journalismus und die Demokratie im Allgemeinen hin (exemplarisch: Hamilton & Turner, 2009; Cohen et al. 2011; Flew et al., 2010).

Eine ähnliche Perspektive zeigte bereits Philip Meyer (2002/1973), der aufgrund seiner journalistischen Arbeiten und des Referenzhandbuchs „Precision Journalism“ als einer der Pioniere von Datenjournalismus gilt. Er beschreibt, wie JournalistInnen Statistik und andere quantitative sozialwissenschaftliche Verfahren für die Berichterstattung verwenden können. Darüber hinaus liefert er erstmals ein Workflow-Modell, das den Ablauf datengetriebener journalistischer Projekte operationalisiert. Auch andere haben derartige Schemata entwickelt (z.B. Bradshaw, 2011a, 2011b; S. Rogers, 2011; Lorenz, 2010; Weinacht & Spiller, 2014), in denen sie Methoden und Prozesse dokumentieren und visualisieren. Diese Schemata sind unterschiedlich detailliert und haben auch verschiedene Schwerpunkte, erklären im Kern jedoch folgende Prozesse: a) Datenbeschaffung und -bereinigung, b) Storytelling, Entwicklung und Visualisierung, c) Veröffentlichung und Distribution.

Mit den Routinen und Methoden in Nachrichtenorganisationen (neben anderen Aspekten) beschäftigen sich auch einige jüngere, vor allem auf qualitativer Forschung aufbauende Veröffentlichungen. Es sind meist Fallstudien zu einzelnen Redaktionen oder Projekten (Royal, 2012; Parasie & Dagiral, 2013; Weber & Rall, 2013; Young & Hermida, 2014), oder Analysen der Situation von Datenjournalismus in einem Land oder einer Sprachregion (Weber & Rall, 2013; Applgren & Nygren, 2014; De Maeyer et al., 2014; Fink & Anderson, 2014; Karlsen & Stavelin, 2014; Weinacht & Spiller, 2014). In den Arbeiten zeigt sich, dass zwar so manche Unterschiede in den analysierten Häusern und Ländern (USA, Belgien, Schweden, Norwegen, Deutschland) bestehen – Stichwort Hierarchie und Organisationskultur –, im Hinblick auf Epistemologie, Praktiken und technologische Entwicklungsstandards aber durchaus viele Ähnlichkeiten⁶ bestehen: So existieren

6 Die Homogenität in den Ergebnissen in den bisherigen Forschung hängt vermutlich auch mit der Erhebungsmethode „Interview“ zusammen, in der es die Tendenz zu erwünschten Antworten gibt, die wiederum vom normativen transnationalen Diskurs über Datenjournalismus beeinflusst sind (Gynnild, 2014; De Maeyer et al., 2014).

einerseits einige wenige datenjournalistische Leuchtturmprojekte und -redaktionen, die auch häufig im Blickpunkt der Analysen stehen, und andererseits viele (Individuen in) Organisationen, die mit Ressourcenknappheit kämpfen.

Diese Dichotomie wird auch in der Art der realisierten Projekte offenbar – zum einen gibt es jenen Datenjournalismus, der von einer Einzelperson im „Tagesgeschäft“ produziert wird, zum anderen Investigativprojekte, mit denen Teams in Arbeitsteilung oft wochen- oder monatelang beschäftigt sind (Fink & Anderson, 2014). Ähnlich wie im „traditionellen“ Investigativjournalismus bilden sich dabei die interdisziplinären Teams je nach Projektanforderung neu, oft auch unter Einbeziehung von externen Dienstleistern. Deadlines treten in den Hintergrund, im Vordergrund stehen die Vollständigkeit des Beitrags und der Wunsch, mit jedem Projekt das technologische Potenzial etwas weiter auszuloten (Royal, 2012). Dementsprechend gelten viele DatenjournalistInnen in ihren Redaktionen als technologische Avantgarde, als „Nerds“: Sie zeigen Affinität zu Transparenz-Ideologien aus der IT-Kultur und versuchen Methodologien wie Hacking in die Redaktionen zu tragen (Parasie & Dagiral, 2013; Howard, 2014; Weinacht & Spiller, 2014). Diese Arbeitskulturen sind mit jenen im „traditionellen“ Journalismus nicht völlig kompatibel, weshalb sich Redaktionen durchaus schwer damit tun, Routinen zu etablieren (Karlsen & Stavelin, 2014).

4 Forschungsfokus und -design

Ausgehend von den obigen Ausführungen rund um digitale Methoden und deren Erforschung sowie um datenjournalistische Routinen sollen hier das Forschungsinteresse und -design spezifiziert werden. In Fortführung der methodologischen Argumentation und Literaturdiskussion ergeben sich folgende Forschungsfragen, die in dieser Reihenfolge im Ergebnisteil beantwortet werden:

- *FF1: Wie funktioniert Zusammenarbeit bei datenjournalistischen Projekten?*
- *FF2: Wie sehen datenjournalistische Routinen und Arbeitsabläufe aus?*
- *FF3: Welche Akteure (im Sinne der Akteur-Netzwerk-Theorie) beteiligen sich an digitalen datenjournalistischen Methoden?*
- *FF4: Wie werden digitale datenjournalistische Methoden stabilisiert?*

Um diese Forschungsfragen zu beantworten, wurden zwischen Ende August 2013 und Mitte Mai 2014 semi-strukturierte Interviews mit elf Personen durchgeführt, die überwiegend einer Tätigkeit nachgingen, die mit Datenjournalismus assoziiert wird.⁷ Diese übten sie entweder in Redaktionen (acht der elf Befragten) oder für spezialisierte Datenjournalismus-Agenturen (3) aus – bei insgesamt sieben unterschiedlichen Organisationen.⁸ Sie wurden um ein Interview gebeten, weil sie in der Berichterstattung über Datenjournalismus immer wieder genannt worden waren oder weil sie sich auf Twitter unter dem einschlägigen Hashtag „#ddj“ geäußert hatten und sich in ihrem dortigen Profil entsprechend darstellten. Zehn der Interviewten arbeiteten im deutschsprachigen Raum, eine Person in Großbritannien.⁹

Die Interviews wurden nach einem schwach strukturierten Leitfaden face-to-face in den Redaktionen bzw. den jeweiligen Arbeitsumgebungen aufgezeichnet.¹⁰ Zusätzlich zu den Interviews wurden nach den Besuchen Beobachtungsprotokolle verfasst, die Aspekte wie räumliche Ordnung, Gesten und Materialien in interpersonellen Interaktionen dokumentierten. Die durchschnittliche Gesprächsdauer betrug eine Stunde und zwei Minuten. Zu Beginn wurden die Interviewten nach ihrer Ausbildung, ihrem beruflichen Werdegang und ihrem Einstieg in den Datenjournalismus gefragt. Anschließend rückten die Akteure, Abläufe und Methoden bei datenjournalistischen Projekten in den Mittelpunkt des Gesprächs. Der gesamte Prozess, von Themenfindung über Produktionsbedingungen bis hin zur Dissemination datenjournalistischer Projekte wurde im Detail durchgegangen. Den Abschluss bildeten Reflexionen über digitale Methoden, die eigene

7 Die Befragten waren (in alphabetischer Reihenfolge): Mona Chalabi, Christina Elmer, Florian Gossy, Markus Hametner, Thomas Jöchler, Nicolaus Kayser-Bril, Mirko Lorenz, Lorenz Matzat, Maximilian Schäfer, Julius Tröger und Sascha Venohr.

8 Zum Zeitpunkt der Interviews arbeiteten die Personen bei den bzw. für die folgenden Organisationen (in alphabetischer Reihenfolge): Berliner Morgenpost, The Guardian, Journalism++, Open Data City, Der Spiegel, Der Standard, Die Zeit.

9 Wie groß die Dynamik in der Branche ist, zeigt sich am Umstand, dass mit Herbst 2014, zum Zeitpunkt der Niederschrift dieses Textes, wenige Wochen nach dem letzten Interview, mindestens sechs der Interviewten zwar noch in derselben Branche aber bereits in anderen Abteilungen oder bei anderen Organisationen tätig sind.

10 Ein Interview wurde per Skype geführt, zwei Personen wurden an anderen Orten interviewt.

Praxis und die allgemeine Situation des Datenjournalismus. Insbesondere wurde dabei auf die Nähe zur Wissenschaft, auf die Objektivität und auf die Ausbildungssituation von DatenjournalistInnen eingegangen.

Die Interviews wurden transkribiert und anschließend mit den Beobachtungsprotokollen in ein Programm zur computerunterstützten qualitativen Auswertung, den TAMS Analyzer (Weinstein, 2006), importiert. Die Analyse orientierte sich an den Konzepten der Grounded Theory (Berg, 2001; Glaser & Strauss, 1967; Strauss & Corbin, 1990) und wurde auf Basis der Transkripte entwickelt. Die Codes der Analyse wurden in einem iterativen Prozess auf die Gesamtheit des Materials angewendet.

5 Ergebnisse

„Wie haltet ihr eure Methoden für später fest?

„Was meinst du damit?“

„Wie dokumentiert ihr etwa eure Visualisierung oder Datenakquise?“

„Das ist bei uns als Code vorhanden.“

Aus einem Interview mit einem Datenjournalisten (I1)¹¹

5.1 Interdisziplinäre Kollaborationen

Die Produktion eines datenjournalistischen Stücks ist in der Regel ein kollektives Unterfangen, erklärte die Mehrzahl der Interviewten. Viele unterschiedliche Akteure helfen mit, bis ein Projekt erscheint. „Den Datenjournalisten“ oder „die Datenjournalistin“ gibt es in den wenigsten Organisationen. Gewöhnlich werden datenjournalistische Stücke arbeitsteilig von SpezialistInnen produziert. Ihre Expertise liegt in Programmierung, Statistik, Software-Projektmanagement, interaktiver Grafik, nutzerzentriertem Design, Kartographie, Fact-Checking oder in einem spezifischen Themengebiet. Viele der Interviewten sind ExpertInnen in mehreren Domänen. Mit geringer externer Unterstützung können sie allein datenjournalistische Stücke veröffentlichen. Zentral dafür sind das Beherrschen

11 Die Zitate wurden der Schriftsprache angepasst, inhaltlich jedoch nicht verändert. Zwei der Interviews wurden auf Englisch geführt – die Zitate wurden übersetzt. Auch wurden die Namen der Interviewten durch Nummern ersetzt.

dreier Kompetenzen: Datenanalyse bzw. Programmierung, Design sowie Formulierungsstärke. Mängel in Visualisierungsfähigkeiten können in vielen Fällen durch den Einsatz webbasierter Tools kompensiert werden.

Die finanziellen Probleme, von denen viele Medienhäuser betroffen sind, beeinträchtigen auch die Personalpolitik im Datenjournalismus. So wie im Textjournalismus ist es im Datenjournalismus nichts Ungewöhnliches, für bestimmte Aufgaben externe Freischaffende zu beauftragen. Nur ressourcenstärkere Nachrichtenorganisationen leisten es sich, Vollzeitkräfte oder sogar Teams zu beschäftigen, die sich allein um datenjournalistische Projekte kümmern. Viele, die sich in Redaktionen an journalistischen Datenprojekten beteiligen, interessieren sich privat für Informationsvisualisierung oder Datenbanken. Sie sind ReporterInnen oder RedakteurInnen, häufig auch MitarbeiterInnen der Grafik-, Archiv- oder Entwicklungsabteilung, die sich innerhalb der Organisation ein neues Aufgabefeld suchen und einen Teil ihrer Arbeitszeit auf einschlägige Projekte verwenden. Je nach Ressourcen und Projektanforderungen beteiligen sie sich in den interdisziplinären Teams, die an den Projekten arbeiten. Die Leitung in diesen Teams übernimmt meist eine Person mit journalistischem Profil, die die Arbeit dann zwischen sich selbst und den Verantwortlichen für Datenanalyse, Programmierung, Visualisierung und dem Fachressort, dem die Geschichte thematisch zugeordnet ist, verteilt.

Insofern nehmen Personen, die überwiegend einer mit Datenjournalismus assoziierten Tätigkeit nachgehen, in der Aufgabenverteilung in ihren Organisationen eine selten klar definierte Funktion ein: Einerseits sind sie selbstständige ReporterInnen und RedakteurInnen, die selbst recherchierte Projekte publizieren, andererseits unterstützen sie RedakteurInnen, die keine ausgeprägte Datenkompetenz haben, bei ihren Veröffentlichungen. Sie sehen sich sowohl als Teil der Redaktion als auch als Teil der Supportdienste. RedakteurInnen „suchen regelmäßig die fachliche Beratung [bei uns]“, erzählt ein Datenjournalist (17). Und weiter: „Es ist mittlerweile so, dass wir in der Redaktion etabliert sind, dass wir uns als geschätzten Partner sehen, der Arbeit nicht wegmoderiert. [...] [Die Fachredakteure] wissen: Wenn wir sagen, etwas geht nicht, dann geht es wirklich nicht.“ Zu dieser Rolle zwischen Redaktionsmitglied und Dienstleister passt auch die Tatsache, dass DatenjournalistInnen interne Weiterbildungen durchführen, in denen sie die technischen Kompetenzen von FachredakteurInnen verbessern und ihnen Tools vermitteln, mit Hilfe derer sie ohne externe Hilfe Datenvisualisierungen anfertigen können.

Interessierte treffen sich in internen Fortbildungen genauso wie auf einschlägigen, klassisch abgehaltenen Konferenzen. Darüber hinaus werden auch informelle Treffen mit Gleichgesinnten organisiert – sowohl hausintern als auch extern. Eine besondere Rolle spielen Formate wie die Veranstaltungen der internationalen Graswurzelbewegung *Hacks/Hackers*, bei der „Hacks“, also JournalistInnen, mit „Hackern“¹² zusammenkommen, um sich auszutauschen und zusammenzuarbeiten (Lewis & Usher, 2014). Ein Reporter (I8), der sich hauptsächlich mit Datenprojekten beschäftigt und auf einem solchen Treffen ein Projekt vorgestellt hat, berichtet von seinen Erfahrungen im Anschluss an sein Referat: „Dort kommen die entsprechenden ‚Nerdfragen‘, die du sonst nicht bekommst. Und ich erzähle dann genau, wie wir das gemacht haben, helfe dem einen oder anderen auch damit.“ Die gegenseitige Unterstützung ist im Datenjournalismus laut den Interviewten stark ausgeprägt – auch über Medienunternehmen hinweg. Selbst MitarbeiterInnen von Unternehmen, die miteinander in Konkurrenz stehen, würden sich gegenseitig helfen. Das gute Klima der Kollaboration betonen alle Befragten in ihren Aussagen. Sie unterstützen einander etwa bei technischen Problemen oder inhaltlichen Fragen. „In dieser Szene geht das sehr schnell. Man bringt eine Frage auf Twitter ein und dann antworten einem drei Koryphäen aus den USA oder von wo auch immer“, erklärt ein Interviewter (I7). Twitter wird übereinstimmend als ein zentraler Disseminations- und Feedbackkanal genannt. Die Beteiligten organisieren ihre Kommunikation unter anderem durch die Verwendung des Hashtags „#ddj“, das für „data-driven journalism“ steht. Auch kommunizieren viele Akteure miteinander per Telefon, Skype oder E-Mail. Praktiken wie etwa das Einbetten fremder Inhalte oder die Übernahme von Visualisierungsmethoden oder von Stories mit lokalisiertem Inhalt sind üblich.

5.2 Routinen und Arbeitsabläufe

Als große Herausforderung nannte ein Großteil der Interviewten die Einführung von Routinen in die Produktion der datenjournalistischer Stücke. Viele

12 Hacker in diesem Kontext bezeichnet nicht Personen, die Sicherheitslücken in Computersystemen ausnützen, sondern technologieaffine Personen, die mit ihren Programmierkenntnissen spielerisch gesellschaftliche Probleme lösen wollen.

der Projekte haben experimentellen Charakter – selbst in Organisationen, die bereits seit mehreren Jahren Datenjournalismus betreiben. Die Installation von Routinen wird nicht zuletzt durch die Heterogenität der Projekte erschwert. Diese zeigt sich an drei Aspekten: Erstens sind Datenjournalismus-Verantwortliche in vielen Organisationen sowohl für tagesaktuelle als auch für langfristig angelegte Projekte zuständig. Da diese Projekte je nach Realisierungshorizont unterschiedliche Herangehensweisen erfordern, ist es schwer, Routinen zu etablieren. Zweitens macht das Themenspektrum, das von einer einzelnen Person oder einem Team behandelt wird, nicht vor Ressortgrenzen Halt: Mit Datenjournalismus betraute Personen arbeiten gewöhnlich an Fragestellungen und heterogenen Daten aus allen Themenbereichen einer Redaktion – von Außenpolitik über Sport bis zu Chronik. Drittens setzen viele Projekte den Einsatz neuer Tools, Software oder Visualisierungsbibliotheken voraus. „Man hat keine Strategie für die nächsten Jahre [...], man kann online nur die nächsten zwei Wochen planen und ‚wurstet sich durch‘ [bringt sich behelfsmäßig durch]“, resümiert ein Journalist (18).

Auch wenn kaum datenjournalistische Routinen etabliert werden, so können trotzdem abgeschlossene *Praktiken* identifiziert werden, die miteinander in Beziehung stehen: *Initiiert* werden datenjournalistische Projekte durch selbstständige Ideen und Recherchen sowie durch redaktionsinterne oder externe Anregungen und Einflüsse. Mehrere der Interviewten nennen die Veröffentlichung neuer Daten durch statistische Dienste als zentrales Ausgangsmoment ihrer Projekte. Die Datenerfassung passiert in komplexen Projekten oftmals automatisiert mit Hilfe von Scrapern oder Crawlern. Auch werden Daten auf „Vorrat“ gesammelt, obwohl zum Zeitpunkt der Akquise noch nicht klar ist, inwieweit sich die Daten für eine Projektrealisierung eignen. Ein freischaffender Datenjournalist (I1) erläutert den Hintergrund:

„Am besten sammelst du einmal alles ein, was es gibt. Du kannst später immer noch filtern. Der Scraper läuft dann, du lässt ihn, wenn man so will, ‚reifen‘. Nach einem Monat oder zwei guckst du in die Daten und schaust, ob du Muster findest.“

Das Finden von Mustern ist Teil einer „explorative[n] Phase, in der man erst einmal schaut: Was sind das überhaupt für Daten? In welchem Zustand sind sie? Was für ein Aufwand ist es, sich damit zu beschäftigen? Welche Form kann das

am Ende annehmen?“, erklärt ein anderer Interviewter (I3). Weitere Praktiken, die ebenso am Beginn eines Projekts stehen, sind die Erstreinigung der Daten, die Formulierung von Hypothesen und das Herausbilden eines Narrativs. Nicht jedes Projekt enthält zwangsläufig jede Praxis – und auch die Reihenfolge unterscheidet sich von mal zu mal. In einigen Projekten etwa verzichten die ProduzentInnen auf das Formulieren von Hypothesen oder das Reinigen der Daten.¹³

Die *Explorationsphase* ist geprägt durch Experimentieren und Sondieren. Zitat eines Datenjournalisten (I8): „Es passiert ständig: Wir haben irgendwelche Ideen, werfen Datensätze rein und dabei kommt gar nichts raus.“ „Es geht viel um herumspielen“, charakterisiert ein anderer (I10) diese Phase. Mit der Formulierung eines Narrativs geht gewöhnlicherweise die Entwicklung der Datenvisualisierung einher. Dabei werden Skizzen per Hand aufs Papier gebracht, bei komplexeren Projekten werden auch so genannte Wireframes, also am Computer gezeichnete Entwürfe, angefertigt. Diese sollen dabei helfen, die Sicht der EndanwenderInnen von Anfang an in die Konzeption zu integrieren. Sofern mehr als eine Person an einem Projekt arbeitet, ist in diesem Produktionsabschnitt die Abstimmung zwischen den einzelnen Teammitgliedern besonders intensiv. Allerdings gibt es kaum formalisierte Besprechungen, dafür aber starken informellen Austausch: Indem die Teammitglieder auf Abbildungen deuten, Skizzen anfertigen oder unter Zuhilfenahme von Gestik ihre Ideen artikulieren, aktualisieren sie wechselseitig ihre Vorstellungen über das Projekt. Unterstützend kommen in den Organisationen verschiedene physische und elektronische Projektkoordinationstools wie Pinnwände, die den Projektstatus abbilden, gemeinsame Chatkanäle in IRC oder Skype, Wikis, oder Mailinglisten zum Einsatz.

So werden datenjournalistische Projekte in Arbeitsteilung und intensiver gegenseitiger Abstimmung weiterentwickelt, bis sie kurz vor der Finalisierung stehen. Es folgen interne Tests, Feedbackschleifen mit den Ideen- bzw. den Auftraggebern. Ein wichtiges Element, das enorme Ressourcen benötigt, ist die Adaption der Visualisierungen für verschiedene Plattformen, Betriebssysteme und Browser. Gerade in Projekten, bei der NutzerInnen mit den Inhalten interagieren können, sie also die Fakten erst durch weitere Eingaben erfassen können, stel-

13 Der im Englischen gängige Ausdruck „Data-driven Journalism“ impliziert die Anwendung von theorieinduzierten und hypothesenformulierenden Methoden (siehe auch Mahrt in diesem Band).

len unterschiedliche Darstellungs- und Interaktionstechnologien die Entwickle-
rInnen vor große Herausforderungen. Schließlich spielt der Veröffentlichungs-
zeitpunkt gerade bei aufwändigeren Projekten eine wichtige Rolle, geben die
Interviewten übereinstimmend an. Neben der prominenten Platzierung auf der
Startseite eines Mediums setzen sie im „Rollout“ intensiv auf die Distributions-
effekte sozialer Medien.

In der *Evaluierungsphase* wird die Qualität von Unternehmungen von den
PraktikerInnen nicht zuletzt danach beurteilt, inwieweit sie die Grenzen des
technisch Machbaren ausloten. Mit anderen Worten: Die Qualität eines Projekts
misst sich nicht allein am Inhalt bzw. dessen erfolgreicher Vermittlung an das
Publikum, sondern auch am Arbeitsaufwand, am Komplexitätsgrad der einge-
setzten Daten und Technologien sowie an der Weiterentwicklung von Visuali-
sierungsansätzen.

Was beeinträchtigt die Güte eines Projekts? So gut wie alle Befragten nennen
Ressourcenknappheit, insbesondere den Mangel an Zeit und Geld, an erster Stel-
le. Weiters fehlt es in sehr spezialisierten Projekten zuweilen an technischer Ex-
pertise. Exemplarisch hierfür die Aussage einer Redakteurin (12):

„Das ist so eine Zwickmühle im tagesaktuellen Online-Journalismus –, dass das, was
wir datenjournalistisch machen, in einer perfekten Welt natürlich viel länger und
noch reflektierter und optisch opulenter daherkommen könnte, wir uns das aber
oft einfach nicht leisten können.“

Die Knappheit an Ressourcen führt unter anderem dazu, dass die Nachberei-
tung der Projekte in vielen Fällen nicht so passiert, wie es sich die PraktikerInnen
wünschen würden. Insbesondere betrifft das die Veröffentlichung der Daten, der
selbst entwickelten Programme und Werkstattberichte, die von allen Interview-
ten als wichtige Elemente am Ende des datenjournalistischen Workflows genannt
werden. Damit soll die eigene Arbeit transparent und nachvollziehbar gemacht
werden sowie eventuell gemachte Fehler nach der Veröffentlichung korrigiert
werden. Auch wenn DatenjournalistInnen diese selbst gesteckten Ziele nicht im-
mer erreichen, halten sie die Idee sehr hoch, möglichst viel Transparenz über
das eigene Schaffen herzustellen. „Wir machen alles open source, weil wir sehen
nicht, warum es anders sein sollte“, erklärt ein Gründer einer Datenjournalismu-
sagentur (14) seine Motive.

5.3 Akteure digitaler datenjournalistischer Methoden

In den Beschreibungen der Arbeitsabläufe im vorhergehenden Abschnitt wurden bereits zahlreiche Akteure erwähnt, die an datenjournalistischen Methoden beteiligt sind. Die Mehrheit der Interviewten schildert, dass Software-Komponenten im gesamten Prozess zentrale Rollen einnehmen. In der Phase der Datenakquise sind es etwa so genannte „digitale Briefkästen“. Dabei handelt es sich um Webformulare, über die eine Organisation verschlüsselt und anonymisiert Daten empfangen kann. Auch Crawler und Scraper kommen zum Einsatz. Nicht zuletzt sind auch Schnittstellen zu Datenbanken zentrale Agenten der datenjournalistischen Methode, kontrollieren sie nicht nur den Zugang zu strukturierten Informationen, sondern formatieren gemeinsam mit den Datenbanken selbst und deren Nutzungsbedingungen auch die Weiterverarbeitung der Daten vor (Ausserhofer, 2015; Bodle, 2011; Bucher, 2013).

Weitere Hard- und Software-Agenten beeinflussen die lokale Datenaufbereitung und -visualisierung: In den Interviews wurden unter anderem Tabellenkalkulationsprogramme und Visualisierungswerkzeuge genannt. Die DatenjournalistInnen würden die eigenen Vorstellungen ständig anpassen, weil die Software-Anwendungen nur bestimmte Operationen zulassen. Durch Programmieren könne man diese Restriktionen jedoch weitestgehend überwinden. Ein Datenjournalist (I10) erläutert: „Die Software gibt vor, wie Geschichten auszusehen haben. Zugleich ist es ein großer Vorteil, dass wir im Newsroom einen Entwickler haben, der mit uns eigene Lösungen baut, damit wir uns nach den Daten richten können und nicht nach der Software.“

Daten werden auf lokalen Servern oder in der „Cloud“ abgelegt. Darüber hinaus dienen die Server auch als „Hosts“ für Datenscraper und -crawler sowie News-Apps. Äußerst wesentlich erscheinen Hosting-Dienste wie *GitHub*, mit Hilfe derer einerseits kollaborativ Programmcode entwickelt wird, andererseits Daten veröffentlicht werden. Immer wieder stören lokale Sicherheitsvorkehrungen und Kompatibilitätsprobleme die Arbeit an Projekten, zum Beispiel wenn der Laptop, auf dem die Visualisierung entwickelt wird, bestimmte Programme nicht ausführen kann oder wegen Sicherheitsbedenken von den lokalen Servern ausgesperrt wird – alltägliche Situationen, wie sie mehrfach geschildert wurden.

Eine nicht unwesentliche Rolle spielt auch die räumliche Position der DatenjournalistInnen in der Organisation. Aus den Aussagen der Interviewten und den

Beobachtungen während der Redaktionsbesuche lässt sich keine eindeutige Tendenz ablesen – auch weil die räumliche Situation in jeder Organisation unterschiedlich ist. Manche Akteure waren mitten im Newsroom stationiert, andere in der Entwicklungsabteilung, dritte waren ohne zugeteilten Arbeitsplatz, saßen mit ihrem Laptop an freien Tischen oder besetzten eigens für Recherche vorgesehene Räume. In den Agenturen dominierten Gemeinschaftsbüros.

Neben der physischen Position in der Organisation während des Entwicklungsprozesses ist auch der Platz der Veröffentlichung fertiger Projekte auf der Website von großem Belang für die Interviewten. Dies wird gewöhnlich von leitenden Angestellten entschieden. Eine prominente Platzierung auf der Start- oder Ressortseite, zu Zeiten, in denen die Seite oft aufgerufen wird, interpretieren DatenjournalistInnen als Wertschätzung der Redaktionsleitung gegenüber der geleisteten Arbeit. Hinzu kommt, dass eine sichtbarere Positionierung eines Beitrags die Chance auf viele Seitenaufrufe erhöht – und Seitenaufrufe sind neben Verweildauer und Besucherzahl eine zentrale Metrik, anhand derer der Erfolg datenjournalistischer Projekte gemessen wird.

Zunehmend werden traditionelle Web-Aufmerksamkeitsmetriken wie Verweildauer durch andere quantitative Indikatoren aus Sozialen Medien ergänzt. Im Datenjournalismus sind Social-Web-Währungen wie „Shares“, „Likes“, „Fans“ oder „Follower“ doppelt relevant: einerseits bei der Veröffentlichung fertiger Projekte als Maßstab für deren Erfolg, andererseits bei Projekten, die sich inhaltlich mit dem Web beschäftigten, als Datenquelle für Analysen und Visualisierungen. Mit letzterem beschäftigt sich der folgende Abschnitt: mit datenjournalistischen Projekten, die ähnlichen Fragen wie kommunikationswissenschaftliche Internetforschung nachgehen.

5.4 Stabilisierung digitaler datenjournalistischer Methoden

Automatisierungen, Fokus auf digitale Objekte und die Analysen von Web-Plattformen: Das sind einige zentrale Merkmale von *Digital Methods*, wie sie in Abschnitt zwei besprochen wurden. Diese Prinzipien sind auch vielen datenjournalistischen Projekten inhärent. So gibt es zahlreiche journalistische Analysen, die die Kommunikationsinhalte, das Verhalten oder die Netzwerke von NutzerInnen untersuchen. Das Themenspektrum ist dabei ähnlich breit wie in der kommuni-

kationswissenschaftlichen Internetforschung und reicht von Mappings von Nutzerinhalten bei Naturkatastrophen bis hin zur Analyse politischer und persönlicher Öffentlichkeiten. Aufgrund des relativ einfachen Datenzugangs greifen viele Analysen auf Social-Media-Dienste und hier insbesondere auf Twitter zurück.

Bei größeren Nachrichtenereignissen werden datenjournalistische Digital-Methods-Untersuchungen immer wieder herangezogen, um die Kommunikationsintensität oder Meinungen eines Teils der Bevölkerung zu erheben. Ein Datenjournalist (17) erklärt, warum er Social-Media-Daten für seine Berichterstattung heranzieht: „Jetzt haben wir die Möglichkeit, zumindest ein Gefühl dafür zu bekommen, wie tausende Menschen zu einem Thema denken.“ Beispielhaft für diese Herangehensweise ist das Projekt „Die Freunde von Pegida“ (Biermann et al., 2015), bei dem die Anhängerbasis der Facebook-Seite der ausländerfeindlichen Dresdner Organisation *Pegida* analysiert wurde. Nachdem der Ertrag klassischer soziologischer Erhebungen bislang „dünn“ sei, wandten sich die Autoren der Datenbasis der Fansseite zu, untersuchten etwa deren Wachstum, deren angegebene demografischen Besonderheiten und Vorlieben.¹⁴ Dass diese Social-Media-Öffentlichkeiten nicht repräsentativ sind und mit ihrer Konstruktion auch Politik gemacht wird, ist den befragten ExpertInnen dabei durchaus bewusst. Die Interviews zeigten, dass insbesondere aufgrund des häufig vorherrschenden Ressourcenmangels Hoffnungen in Automatisierungen gesetzt werden: Einige der Interviewten würden gerne Scripts und Scraper so konfigurieren, dass sie manuelle Arbeit verringern, und die Veröffentlichung schneller bzw. annähernd in Echtzeit (etwa bei Social-Media-Daten) passieren kann.

Viele journalistische Digital-Methods-Projekte beschäftigen sich auch mit den persönlichen Öffentlichkeiten von Personen, die im Medieninteresse stehen. Exemplarisch seien die eingangs beschriebene Netzwerkanalyse des mutmaßlichen Boston-Marathon-Attentäters (Lotan, 2013) sowie die Analysen über die Twitteraktivität von Berliner (Tröger, 2012) und Schweizer (Thelitz et al., 2014) Abgeordneten angeführt. Ein gut dokumentiertes Beispiel für digitale Methoden im Datenjournalismus ist eine Untersuchung über die Verbreitung von Falschmeldungen auf Twitter während der Ausschreitungen in London im Sommer 2011 (Dant & Richards,

14 Diese Daten konnten ohne großen Aufwand per Schnittstelle von Facebook ausgelesen werden.

2011; Guardian Interactive Team et al., 2011). Für dieses Projekt arbeitete der Guardian mit drei britischen Universitäten zusammen. Insgesamt wurden 2,6 Millionen Tweets mit spezifischen Hashtags gesammelt, gefiltert und ein Teilkorpus davon manuell codiert. Mit Hilfe einer zeitbasierten interaktiven Visualisierung konnten die AutorInnen Momente aufzeigen, in denen Gerüchte auf Twitter verbreitet und entkräftet wurden (Procter et al., 2013).

6 Resümee

Dieser Aufsatz beleuchtete die Rolle von digitalen Methoden im Datenjournalismus, die mit Hilfe von Interviews untersucht wurden. In einem ersten Schritt wurden unter Rückgriff auf unter anderem Rogers (2013), Marres (2012), Latour (2005) und Schmidt (2012) die Grundzüge einer Methodologie formuliert, mithilfe derer die Anwendung von Digitalen Methoden im Journalismus untersucht werden sollte. Im Zentrum stand dabei die Frage, wie sich datenjournalistische Routinen manifestieren und welche Praktiken und heterogenen Akteure dabei vermitteln. Aufbauend auf qualitativen Analysen von elf semistrukturierten Interviews und Beobachtungsprotokollen aus Redaktionsbesuchen konnten einerseits neue Perspektiven auf Methoden und Routinen im Datenjournalismus gegeben werden, andererseits auch Arbeiten besprochen werden, in denen Daten aus dem Web den Ausgangspunkt für datenjournalistische Werke darstellen.

Datenjournalismus stellt sich als ein dynamisches Feld dar, das seine Routinen und Methoden in den letzten Jahren stark verändert und an die Digitalisierung angepasst hat. Was kann die digitale Kommunikationsforschung von den Entwicklungen des Datenjournalismus lernen?¹⁵ Potenzial gibt es sowohl im Entwicklungsprozess als auch in der Ergebnispräsentation. Wie die Resultate dieser Arbeit darlegen, ist die enge Zusammenarbeit zwischen Personen mit unterschiedlichen Hintergründen (Journalismus, Design, Programmierung, etc.) zentral für das Gelingen eines Projekts. Die Einbeziehung sämtlicher Akteure von Beginn an, kurze Feedbackschleifen und die Anlehnung an Methoden der itera-

15 Mit der Frage, wie Journalismus von Digital Methods profitieren kann, beschäftigen sich Jonathan Gray und Liliana Bounegru (2014).

tiv-inkrementellen Software-Entwicklung ermöglichen es, im Datenjournalismus rasch zu präsentablen Ergebnissen zu kommen. Klar ist zugleich, dass in der Internetforschung in vielen Fällen andere Tempi notwendig sind, weil Wissenschaft andere Ansprüche an Validität stellt.

Im Bereich der Ergebnispräsentation liegt Potenzial für die Kommunikationswissenschaft in den interaktiven Darstellungsformen des Datenjournalismus: Viele Projekte zeichnen sich durch innovative interaktive Visualisierungen aus, über die Narrative vermittelt werden. In anderen Projekten wiederum können sich die RezipientInnen in den Ergebnissen wiederfinden und selbst einen Narrativ entwickeln. Auch die digitale Kommunikationsforschung kann Ähnliches in der Dissemination leisten oder Kooperationen mit Datenjournalismus-Teams in Redaktionen suchen. In den Vergangenheit gab es einige erfolgreiche Kollaborationsprojekte, von denen sowohl Journalismus als auch Wissenschaft profitiert haben. Ein Beispiel ist das oben diskutierte „Reading the Riots“-Projekt des *Guardian*.

Hinsichtlich der Methodentransparenz und der Offenlegung von Algorithmen kann Datenjournalismus durchaus Vorbild für die Kommunikationswissenschaft sein. Profitieren könnte sie hier darüber hinaus auch von den Digital Humanities und der Open-Science-Bewegung, die sich schon viele Jahre mit diesen Themen auseinandersetzen (Ausserhofer, 2015).

Auch wäre es sinnvoll, wenn der Datenjournalismus die Standardisierung seiner digitalen Methoden weiter vorantreibt. Langsam kristallisieren sich etwa neue Erzählformen heraus. Eine Vereinheitlichung von digitalen Methoden und Technologien ist wünschenswert, um die längerfristige Zugänglichkeit zu gewährleisten und Archivierung zu ermöglichen.

Mag. (FH) Julian Ausserhofer ist wissenschaftlicher Mitarbeiter am Institut für Journalismus und Public Relations der FH Joanneum Graz und Doktorand am Institut für Publizistik- und Kommunikationswissenschaft der Universität Wien

Quellenverzeichnis

Anderson, C. W. (2013). Towards a Sociology of Computational and Algorithmic Journalism. *New Media & Society*, 15(7), 1005-1021. doi: 10.1177/1461444812465137

- Appelgren, E., & Nygren, G. (2014). Data Journalism in Sweden: Introducing New Methods and Genres of Journalism into “Old” Organizations. *Digital Journalism*, 2(3), 394-405. doi: 10.1080/21670811.2014.884344
- Ausserhofer, J. (2015). Die Datenbank verdient die Hauptrolle: Aspekte einer Methodologie für Open Digital Humanities. In S. Eichhorn, B. Oberreither, M. Rauchenbacher, I. Schwentner, & K. Serles (Hrsg.), *Aufgehoben? Speicherorte, -diskurse und -medien*. Würzburg: Königshausen & Neumann.
- Berg, B. L. (2001). *Qualitative Research Methods for the Social Sciences*. Boston: Allyn and Bacon.
- Berry, D. M. (2011). The Computational Turn: Thinking About the Digital Humanities. *Culture Machine*, 12, 1-22. Abgerufen von <http://www.culturemachine.net/index.php/cm/article/viewArticle/440>
- Biermann, K., Blickle, P., & Venohr, S. (2015, 5. Februar). Dürfen wir vorstellen: Die Freunde von Pegida. *Die Zeit*. Abgerufen von <http://www.zeit.de/gesellschaft/zeitgeschehen/2015-02/wer-ist-pegida-facebook-daten/>
- Bodle, R. (2011). Regimes of Sharing: Open APIs, Interoperability, and Facebook. *Information, Communication & Society*, 14(3), 320-337. doi: 10.1080/1369118X.2010.542825
- Borger, M., van Hoof, A., Costera Meijer, I., & Sanders, J. (2013). Constructing Participatory Journalism as a Scholarly Object. *Digital Journalism*, 1(1), 117-134. doi: 10.1080/21670811.2012.740267
- Bounegru, L., & Gray, J. (2014, September). Mapping Issues with the Web: An Introduction to Digital Methods. Abgerufen von <http://www.slideshare.net/jwyg/mapping-issues-with-the-web-an-introduction-to-digital-methods>
- Bucher, T. (2013). Objects of Intense Feeling: The Case of the Twitter API. *Computational Culture*, (3). Abgerufen von <http://computationalculture.net/article/objects-of-intense-feeling-the-case-of-the-twitter-api>
- Bradshaw, P. (2011a, 7. Juli). The inverted pyramid of data journalism. Abgerufen von <http://onlinejournalismblog.com/2011/07/07/the-inverted-pyramid-of-data-journalism/>
- Bradshaw, P. (2011b, 13. Juli). 6 ways of communicating data journalism (The inverted pyramid of data journalism part 2). Abgerufen von <http://onlinejournalismblog.com/2011/07/13/the-inverted-pyramid-of-data-journalism-part-2-6-ways-of-communicating-data-journalism/>

- Coddington, M. (2014). Clarifying Journalism's Quantitative Turn: A Typology for Evaluating Data Journalism, Computational Journalism, and Computer-Assisted Reporting. *Digital Journalism*, Online First. doi: 10.1080/21670811.2014.976400
- Cohen, S., Hamilton, J. T., & Turner, F. (2011). Computational Journalism. *Communications of the ACM*, 54(10), 66-71. doi: 10.1145/2001269.2001288
- Dant, A., & Richards, J. (2011, 8. Dezember). Behind the Rumours: How We Built Our Twitter Riots Interactive. *The Guardian*. Abgerufen von <http://www.guardian.co.uk/news/datablog/2011/dec/08/twitter-riots-interactive>
- De Maeyer, J., Libert, M., Domingo, D., Heinderyckx, F., & Cam, F. L. (2014). Waiting for Data Journalism: A Qualitative Assessment of the Anecdotal Take-Up of Data Journalism in French-Speaking Belgium. *Digital Journalism*, Online First. doi: 10.1080/21670811.2014.976415
- Domingo, D., & Paterson, C. (Hrsg.) (2011). *Making Online News Volume 2: Newsroom Ethnographies in the Second Decade of Internet Journalism*. New York: Peter Lang.
- Flew, T., Daniel, A., & Spurgeon, C. L. (2010). The Promise of Computational Journalism. In K. McCallum (Hrsg.), *Media, Democracy and Change: Refereed Proceedings of the Australian and New Zealand Communication Association Annual Conference* (S. 1-19). Canberra. Abgerufen von <http://eprints.qut.edu.au/39649/>
- Glaser, B., & Strauss, A. (1967). *The Discovery of Grounded Theory: Strategies for Qualitative Research*. Chicago: Aldine.
- Fink, K., & Anderson, C. W. (2014). Data Journalism in the United States: Beyond the "usual suspects." *Journalism Studies*, Online First. doi: 10.1080/1461670X.2014.939852
- Guardian Interactive Team, Procter, R., Vis, F., & Voss, A. (2011, 7. Dezember). How Riot Rumours Spread on Twitter. *The Guardian*. Abgerufen von <http://www.theguardian.com/uk/interactive/2011/dec/07/london-riots-twitter>
- Gynnild, A. (2014). Journalism Innovation Leads to Innovation Journalism: The Impact of Computational Exploration on Changing Mindsets. *Journalism*, 15(6), 713-730. doi: 10.1177/1464884913486393
- Haas, H. (1999). *Empirischer Journalismus: Verfahren zur Erkundung gesellschaftlicher Wirklichkeit*. Wien: Böhlau.
- Hamilton, J. T., & Turner, F. (2009). Accountability Through Algorithm: Developing the Field of Computational Journalism. Presented at the Behavioral Sciences Summer Workshop. Abgerufen von <http://dewitt.sanford.duke.edu/wp-content/uploads/2011/12/About-3-Research-B-cj-1-finalreport.pdf>

- Howard, A. B. (2014). *The Art and Science of Data-Driven Journalism*. Tow Center for Digital Journalism. Abgerufen von <http://towcenter.org/wp-content/uploads/2014/05/Tow-Center-Data-Driven-Journalism.pdf>
- Karlsen, J., & Stavelin, E. (2014). Computational Journalism in Norwegian Newsrooms. *Journalism Practice*, 8(1), 34-48. doi: 10.1080/17512786.2013.813190
- Latour, B. (1988). *The Pasteurization of France*. (A. Sheridan & J. Law, Übers.). Cambridge: Harvard University Press.
- Latour, B. (2005). *Reassembling the Social: An Introduction to Actor-Network-Theory*. New York: Oxford University Press.
- Law, J. (2004). *After Method: Mess in Social Science Research*. London: Routledge.
- Lewis, S. C., & Usher, N. (2014). Code, Collaboration, and the Future of Journalism: A Case Study of the Hacks/Hackers Global Network. *Digital Journalism*, 2(3), 383-393. doi: 10.1080/21670811.2014.895504
- Lotan, G. (2013, 25. April). When Your Twitter Friend Turns Out To Be The Boston Bomber. *Digg*. Abgerufen von <http://digg.com/originals/dzhokhar-tsarnaev-twitter-map>
- Lorenz, M. (2010). Status and Outlook for Data-Driven Journalism. In European Journalism Centre (Hrg.), *Data-Driven Journalism: What Is There to Learn?* (S. 8-17). Amsterdam.
- Madrigal, A. C. (2013, 19. April). #BostonBombing: The Anatomy of a Misinformation Disaster. *The Atlantic*. Abgerufen von <http://www.theatlantic.com/technology/archive/2013/04/-bostonbombing-the-anatomy-of-a-misinformation-disaster/275155/>
- Manovich, L. (2013). *Software Takes Command*. New York: Bloomsbury.
- Marres, N. (2012). The Redistribution of Methods: On Intervention in Digital Social Research, Broadly Conceived. *The Sociological Review*, 60(S1), 139-165. doi: 10.1111/j.1467-954X.2012.02121.x
- Meyer, P. (2002/1973). *Precision Journalism: A Reporter's Introduction to Social Science Methods* (4. Auflage). Oxford: Rowman & Littlefield.
- Münch, R., & Schmidt, J. (2005). Medien und sozialer Wandel. In M. Jäckel (Hrsg.), *Mediensoziologie: Grundfragen und Forschungsfelder* (S. 201-218). Wiesbaden: VS Verlag.
- Niederer, S., & van Dijck, J. (2010). Wisdom of the Crowd or Technicity of Content? Wikipedia as a Sociotechnical System. *New Media & Society*, 12(8), 1368-1387. doi: 10.1177/1461444810365297

- Niederer, S. (2013). "Global Warming Is Not a Crisis!": Studying Climate Change Skepticism on the Web. *NECSUS. European Journal of Media Studies*, 2(1), 83-112. doi: 10.5117/NECSUS2013.1.NIED
- Parasie, S., & Dagiral, E. (2013). Data-Driven Journalism and the Public Good: "Computer-Assisted-Reporters" and "Programmer-Journalists" in Chicago. *New Media & Society*, 15(6), 853-871. doi: 10.1177/1461444812463345
- Preston, J., & Roston, M. (2013, 20. April). A Closer Look at the Bombing Suspect's Twitter Account. *New York Times*. Abgerufen von <http://thelede.blogs.nytimes.com/2013/04/20/dzhokhar-tsarnaevs-jahar-twitter-account-prompts-scrutiny/>
- Procter, R., Vis, F., & Voss, A. (2013). Reading the Riots on Twitter: Methodological Innovation for the Analysis of Big Data. *International Journal of Social Research Methodology*, 16(3), 197-214. doi: 10.1080/13645579.2013.774172
- Rieder, B., & Röhle, T. (2012). Digital Methods: Five Challenges. In D. M. Berry (Hrsg.), *Understanding Digital Humanities* (S. 67-84). Basingstoke: Palgrave Macmillan.
- Rogers, R. (2009). *The End of the Virtual: Digital Methods*. Amsterdam: Vossiuspers UvA.
- Rogers, R. (2010). Internet Research: The Question of Method – A Keynote Address from the YouTube and the 2008 Election Cycle in the United States Conference. *Journal of Information Technology & Politics*, 7, 241-260. doi: 10.1080/19331681003753438
- Rogers, R. (2013). *Digital Methods*. Cambridge: MIT Press.
- Rogers, R. (2014). Political Research in the Digital Age. *International Public Policy Review*, 8(1), 73-87.
- Rogers, S. (2011, 7. April). Data Journalism Broken down: What We Do to the Data Before You See It. Abgerufen von <http://www.theguardian.com/news/datablog/2011/apr/07/data-journalism-workflow>
- Royal, C. (2012). The Journalist as Programmer: A Case Study of the New York Times Interactive News Technology Department. *#ISOJ The Official Research Journal of the International Symposium on Online Journalism*, 2(1). Abgerufen von <https://online.journalism.utexas.edu/2010/papers/Royal10.pdf>
- Rühl, M. (1981). Journalismus und Wissenschaft: Anmerkungen zu ihrem Wirklichkeitsverständnis. *Rundfunk und Fernsehen*, 2(3), 211-222.

- Schmidt, R. (2012). *Soziologie der Praktiken: Konzeptionelle Studien und empirische Analysen*. Berlin: Suhrkamp.
- Seward, Z. M., Mirani, L., & King, R. (2013, 20. April). We know when Dzhokhar Tsarnaev sleeps. *Quartz*. Abgerufen von <http://qz.com/76442/we-know-when-dzhokhar-tsarnaev-sleeps/>
- Strauss, A., & Corbin, J. M. (1990). *Basics of Qualitative Research: Grounded Theory Procedures and Techniques*. Thousand Oaks: Sage.
- Thelitz, N., Kohli, A., Fichter, A., Kobler, S., Preusse, T., Schudel, B., ... Interactive Things. (2014, 8. September). Ungleiches Gezwitscher aus dem Parlament. *Neue Zürcher Zeitung*. Abgerufen von <http://www.nzz.ch/schweiz/ungleiches-gezwitscher-aus-dem-parlament-1.18378048>
- Tröger, J. (2012, 11. Oktober). Wie die Berliner Abgeordneten Twitter nutzen. *Berliner Morgenpost*. Abgerufen von <http://www.morgenpost.de/berlin-aktuell/article109759023/Wie-die-Berliner-Abgeordneten-Twitter-nutzen.html>
- Weber, W., & Rall, H. (2013). "We Are Journalists": Production Practices, Attitudes and a Case Study of the New York Times Newsroom. In W. Weber, M. Burmester, & R. Tille (Hrsg.), *Interaktive Infografiken* (S. 161-172). Berlin: Springer. doi: 10.1007/978-3-642-15453-9_9
- Weinacht, S., & Spiller, R. (2014). Datenjournalismus in Deutschland: Eine explorative Untersuchung zu Rollenbildern von Datenjournalisten. *Publizistik*, 59(4), 411-433. doi: 10.1007/s11616-014-0213-5
- Weinstein, M. (2006). TAMS Analyzer Anthropology as Cultural Critique in a Digital Age. *Social Science Computer Review*, 24(1), 68-77. doi: 10.1177/0894439305281496
- Weltevrede, E., & Helmond, A. (2012). Where Do Bloggers Blog? Platform Transitions Within the Historical Dutch Blogosphere. *First Monday*, 17(2). doi: 10.5210/fm.v17i2.3775
- Young, M. L., & Hermida, A. (2014). From Mr. and Mrs. Outlier to Central Tendencies: Computational Journalism and Crime Reporting at the Los Angeles Times. *Digital Journalism*, Online First. doi: 10.1080/21670811.2014.976409

Danksagung

Die Forschung, die dieser Arbeit zugrunde liegt, wurde ermöglicht durch Mittel der Universität Wien, der FH Joanneum sowie der Österreichischen Forschungsförderungsgesellschaft (FFG) (Projekt „Visual Analytics in Data-driven Journalism“ (VALiD), Projektnummer: 845598, Fördergeber: Bundesministerium für Verkehr, Innovation und Technologie). Bedanken möchte ich mich bei meinen InterviewpartnerInnen sowie Wolfgang Aigner, Robert Gutounig, Axel Maireder, Sarah Matiasek, Alice Miller, Wiebke Loosen, Alexander Rind, Elena Rudowsky und Christina Schumann. Dank gebührt darüber hinaus dem Alexander von Humboldt Institut für Internet und Gesellschaft in Berlin sowie dem Center for Science, Technology, Medicine and Society an der University of California, Berkeley, wo Teile der Forschung und Textproduktion stattgefunden haben.

II STRUKTURANALYSEN

Empfohlene Zitierung: Maireder, A., & Schlögl, S. (2015). Twitter-Öffentlichkeiten: Identifikation und Interpretation der Strukturen von Follower-Netzwerken. In A. Maireder, J. Ausserhofer, C. Schumann, & M. Taddicken (Hrsg.), *Digitale Methoden in der Kommunikationswissenschaft* (S. 115-139). doi: 10.17174/dcr.v2.6

Zusammenfassung: Der Beitrag nimmt die Struktur der Follower-Verknüpfungen zwischen TwitternutzerInnen in den Blick, die sich an bestimmten Themen und Diskursen beteiligen. Verdichtungen in diesen Netzwerken verstehen die Autoren als *(Teil-)öffentlichkeiten*, deren Analyse Einblicke in die Partizipationsstrukturen in öffentlichen Kommunikationsprozessen ermöglicht. Anhand zweier Beispiele aus dem Feld der politischen Kommunikation wird die Vorgehensweise bei der Datensammlung, -aufbereitung und -analyse dargelegt und vor dem Hintergrund des methodologischen Konzepts diskutiert. Die Ergebnisse der beiden Fallstudien werden anschließend verglichen und daraus Faktoren für die Beschreibung von Twitter-Communities in politischen Kontexten abgeleitet. Der Beitrag schließt mit einem Ausblick auf weiterführende Analysen und Anwendungsfelder von Clusteranalysen in Twitter-Follower-Netzwerken.

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

Axel Maireder & Stephan Schlögl

Twitter-Öffentlichkeiten: Identifikation und Interpretation der Strukturen von Follower-Netzwerken

1 Einleitung

In den öffentlichen Debatten zu Social Media und Politik – und zuweilen auch in der Forschung – spielen strukturelle Kennzahlen wie die Anzahl von Freunden und Likes, von Views und Shares, von Followern oder Tweets mit einem bestimmten Hashtag eine zentrale Rolle. Hohe Werte werden von den Akteuren stolz kommuniziert und von Journalismus und Öffentlichkeit als Indikatoren für die Relevanz von Personen, Organisationen und Themen herangezogen. Diese ‚Economy of Metrification‘ (Gerlitz & Helmond, 2013) ermöglicht Orientierung in den sonst weitgehend unüberschaubaren Kommunikationsräumen der sozialen Medien. Für Forschung zu den Strukturen und Dynamiken öffentlicher Kommunikationsprozesse liefern die Kennzahlen aber nur oberflächliche Anhaltspunkte, da sie die Komplexität der ineinander verwobenen Kommunikationsräume außerordentlich stark reduzieren. In unserem Beitrag schlagen wir einen Ansatz abseits der Analyse der Standardmetriken vor, bei dem wir jene sozio-technischen Strukturen in den Blick nehmen, die den Konversations- und Diffusionsdynamiken in diesen Kanälen – hier am Beispiel Twitter – zu Grunde liegen.

Twitter ist neben den Massenmedien zu einem der wichtigsten Kanäle für öffentliche Kommunikation avanciert. Politische Akteure und BürgerInnen tauschen Nachrichten sowie Meinungen aus und verhandeln die Bedeutung von Ereignissen, wobei sich Stimmungen und Standpunkte entwickeln und verbreiten. Die strukturelle Basis für die individuelle Praxis auf der Mikroebene und die Konversationsdynamiken auf der Mesoebene der (politischen) Kommunikationsprozesse auf Twitter sind Verknüpfungen zwischen Accounts, welche die NutzerInnen durch ihre individuelle Auswahl von ‚Followees‘ herstellen. Das Aggregat dieser Verknüpfungen stellt sich als Netzwerk dar, in dem NutzerInnen (Accounts) als Knoten und ihre informationellen Beziehungen (Kontakte) als Kanten erscheinen. Diese soziotechnische Makroebene nimmt dieser Beitrag in den Blick. Ausgehend von einem strukturtheoretischen Öffentlichkeitskonzept stellen wir Methoden vor, mit denen spezifische Ausschnitte dieser Netzwerke als (Teil-)Öffentlichkeiten („publics“) verstanden und analysiert werden können. Anschließend zeigen wir, dass die so gewonnenen Erkenntnisse eine sinnvolle Ausgangsbasis für weiterführende Beschreibungen und vergleichende Analysen darstellen können.

Der Aufsatz fokussiert auf die Darstellung unseres Forschungsdesigns und der dahinterliegenden methodologischen Annahmen. Zur Illustration ziehen wir zwei konkrete Fallbeispiele heran, anhand derer wir Potenziale und Herausforderungen unseres Ansatzes diskutieren. Einerseits untersuchen wir die Beteiligten am Hashtag #aufschrei, der Ende Jänner 2013 dazu benutzt wurde, persönliche Erfahrungen mit sexueller Belästigung im Alltag publik zu machen. Innerhalb weniger Stunden nach Etablierung des Hashtags hatten ihn schon mehrere tausend NutzerInnen verwendet und damit in weiterer Folge eine breite Debatte rund um Sexismus im Alltag ausgelöst. Wir fokussieren auf jene NutzerInnen, die den Hashtag in den ersten 24 Stunden verwendeten. Mit Hilfe unseres Ansatzes können wir zeigen, welche Nutzerkreise sich an den Diskussionen beteiligten und wie sich in der Dynamik dieser Beteiligungen die Verbreitung und Bedeutung des Hashtags entwickelte.

In einem zweiten Fallbeispiel analysieren wir die Nutzerstruktur rund um österreichische Innenpolitik. Dabei greifen wir auf Daten aus einem weitaus längeren Zeitraum – 28 Wochen – zurück, um zu zeigen, wie die (innen-)politische ‚Twittersphäre‘ in Österreich strukturiert ist und wie diese Struktur mit Medienpräferenzen der NutzerInnen zusammenhängt. Abschließend fassen wir zusammen, wie der Ansatz zur Analyse von Twitter-Öffentlichkeiten kommunikationstheoretisch eingeordnet werden kann.

2 Theorie

2.1 Öffentlichkeiten

Unsere Perspektive auf Twitterkommunikation ist öffentlichkeitstheoretisch, wobei wir nicht normativ sondern strukturtheoretisch argumentieren. Wir beziehen uns vorerst dennoch auf Habermas (1998), wenn wir Öffentlichkeit als hoch komplexes Netzwerk geographisch, kulturell, thematisch oder funktional differenzierbare Kommunikationssphären verstehen, die vielfach miteinander verknüpft und ineinander verschränkt sind. In diesen Kommunikationssphären tauschen sich Menschen jeweils zu bestimmten Interessen und Themen aus. Dieses Gemeinsame kann vielfältig gelagert und auch – aber bei weitem nicht nur – an politischen Kategorien ausgerichtet sein. So kann die Kommunikation von Fußballfans über ein Spiel ebenso als eine solche Sphäre verstanden werden, wie jene der Anhängerschaft einer politischen Idee. Im englischsprachigen Fachdiskurs wird auf diese interessensgeleiteten Kommunikationszusammenhänge mit dem Begriff ‚publics‘ verwiesen, während im deutschsprachigen Raum vielfach der Terminus ‚Teilöffentlichkeit‘ genutzt wird – auch, um sich von ‚Öffentlichkeit‘ als Konzept für eine gesamtgesellschaftliche Kommunikationssphäre wie bei Habermas (1998) oder in der Systemtheorie (z.B. bei Luhmann, 2004; Kohring, 1997) abzugrenzen.

Öffentlichkeiten sind hier und im Folgenden als Kommunikationszusammenhänge zu verstehen, die sich zwischen Menschen (oder Akteuren) rund um Probleme und Themen entwickeln, die für alle Beteiligten relevant sind. Gruning & Hunt (1984) betonen dabei den Unterschied zwischen heterogenen ‚Massen‘, deren TeilnehmerInnen nichts gemeinsam haben müssen, und Öffentlichkeiten, deren TeilnehmerInnen im Hinblick auf das Interesse an einem bestimmten Thema oder Problem homogen sind – aber nicht unbedingt darüber hinaus. Als solche umfassen Öffentlichkeiten Kommunikationszusammenhänge, an denen jeweils unterschiedlich viele Menschen beteiligt sind, und so können wir Öffentlichkeiten durchaus spezifische ‚Größen‘ oder ‚Reichweiten‘ zuschreiben (Emirbayer & Sheller, 1999). Auch Hepp, Brüggemann, Königslöw, Lingenberg und Möller (2012) nutzen eine solche Raummetapher, wenn sie von Öffentlichkeiten als kommunikativen Verdichtungen sprechen, die „durch die Grenzen dieser Verdichtungen sozialräumlich voneinander“ unterschieden werden können (S.

22). Bei der Diskussion von Twitternetzwerken weiter unten werden wir insbesondere an diese Vorstellungen von Öffentlichkeiten als Räume und Verdichtungen anschließen.

Vorerst sei noch darauf hingewiesen, dass Öffentlichkeiten auf unterschiedlichste Art und Weise miteinander verknüpft und ineinander verschränkt sind, insofern als Aussagen, Meinungen und Themen in unterschiedlichen Sphären verhandelt werden und von Sphäre zu Sphäre ‚wandern‘ können; und zwar sowohl zwischen ‚gleichrangigen‘ thematisch konzentrierten Sphären, als auch von ‚kleineren‘ hin zu ‚größeren‘ Öffentlichkeiten (Benkler, 2007; Anderson, 2010). Als breitenwirksamste und somit effektivste Form zur ‚Herstellung‘ größerer Öffentlichkeiten wird Massenkommunikation verstanden; die unmittelbarste Verknüpfung unterschiedlicher Sphären ist jedoch die Kommunikation einzelner Individuen. Denn jeder Mensch ist, abhängig von Lebensumfeld und Interessenlage, Teil verschiedener Öffentlichkeiten, die er unterschiedlich intensiv beobachtet und an denen er sich unterschiedlich aktiv beteiligt. Themen können in bestimmten Kommunikationszusammenhängen aufgenommen und in anderen weiterkommuniziert und -verhandelt werden. Für gesellschaftliche Kommunikation *en gros* sind diese und andere Verbindungen unerlässlich, denn nur wenn sich die einzelnen Teile der Gesellschaft untereinander verständigen, können sie zu gemeinsamen Vorstellungen über die Welt gelangen (Dahlgren, 2005).

2.2 Netzöffentlichkeit und Twitternetzwerke

Das Internet ermöglicht neue Formen der Organisation gesellschaftlicher Kommunikation und damit vielfältiger ‚Sphären‘ öffentlicher Kommunikation rund um spezifische Interessen und Themen. Da diese in ebenso vielfältiger Art und Weise direkt und indirekt miteinander verknüpft und ineinander verflochten sind, kann das Internet als „umfassendes Öffentlichkeitsforum“ verstanden werden, „das alle Öffentlichkeitssphären in sich vereint“ (Schweiger & Weihermüller 2008, S. 545). Neben dem hohen Grad an sozialer und inhaltlicher Vernetzung zeichnet sich diese *Netzöffentlichkeit* auch durch die Möglichkeit zur breiten sozialen Inklusion aus, da sie sehr niederschwellige Formen kommunikativer Beteiligung bereithält (Benkler, 2007) – nicht nur, aber insbesondere über Social-Media-Anwendungen wie Twitter.

Twitter und andere ‚Social Networks‘ sind grundlegend als Netze strukturiert. Bei Twitter wählen NutzerInnen Accounts, deren Mitteilungen sie erhalten wollen, sie ‚folgen‘ ihnen und werden so ihre ‚Followees‘. Gleichzeitig können dieselben oder andere NutzerInnen ihnen folgen, als sogenannte ‚Follower‘. Aus einer Makroperspektive lässt sich so ein Netzwerk aus Accounts als Knoten und Followerbeziehungen als Kanten beobachten. Die Position der Accounts in diesem Netzwerk reguliert den Erhalt von Tweets in den individuell strukturierten ‚Newsfeeds‘ und die kommunikativen Potenziale der NutzerInnen. Je mehr Verbindungen (also gemeinsame Followers und Followees) zwei NutzerInnen haben, desto näher sind sie einander im Netzwerk und je stärker diese kommunikative Verknüpfung, umso eher werden sie von denselben Mitteilungen erreicht, kommunizieren zu denselben Themen und nehmen gemeinsam an Konversationen teil. Netzwerkteile mit besonders vielen Verknüpfungen können als kommunikative Verdichtungen interpretiert werden, in denen der Austausch zwischen den Beteiligten höher ist als in anderen Teilen des Netzwerks.

Wir verstehen diese Verdichtungen als ‚Öffentlichkeiten‘ im oben skizzierten Sinne. Da wir NutzerInnen bei der Auswahl von Followees ein – wie auch immer geartetes – Interesse an den Mitteilungen dieser Accounts unterstellen können, zeigen sich in Twitter-Netzwerkstrukturen Interessenskonstellationen. Diese Konstellationen sind nicht nur Ergebnis eines sozialen Kommunikationsprozesses, sondern prägen auch die Diffusions- und Themensetzung auf Twitter grundlegend: Sie sind zugleich Ausdruck von Interessenzusammenhängen als auch grundlegende sozio-technische Kommunikationsstruktur für die weitere Entwicklung dieser Zusammenhänge. Dabei kann jeder einzelne Account mit (theoretischem) Blick auf das umfassende Gesamtnetzwerk Teil unterschiedlicher Twitteröffentlichkeiten sein, da NutzerInnen nicht nur einzelne sondern vielfältige Interessen haben. Die relative Positionierung eines Accounts gegenüber anderen *innerhalb* der Beteiligten eines bestimmten Diskurses (Themas, Interessensgebietes) jedoch, so die Annahme, zeigt die Interessenskonstellationen rund um diesen Diskurs. Da sich Followerstrukturen nur langsam verändern, sind diese Strukturen als relativ stabile thematisch zentrierte Kommunikationszusammenhänge zwischen Akteuren der Öffentlichkeit zu deuten, als ‚Öffentlichkeiten‘ im Sinne unserer Definition.

Mit dem Fokus auf Followerbeziehungen gehen wir einen anderen Weg als jene Netzwerkstudien zu Twitter-Öffentlichkeiten, die sich primär auf @mentions als

Artefakte interpersonaler Kommunikation (Ausserhofer & Maireder, 2013; Larsson & Moe, 2011), Retweets als Diffusionsmarkierungen (siehe Bruns in diesem Band) oder Hashtags als Themenmarkierungen (Bruns & Burgess, 2011; Bastos et al., 2013) stützen. @mention- und Retweet-Netzwerke aggregieren digitale Repräsentationen einzelner Handlungen (@mentions, Retweets) in jeweils spezifischen Kommunikationssituationen, um Konversations- oder Diffusionsstrukturen zu zeigen (Bruns & Moe, 2014). Es sind flüchtige Netzwerke, die nur in und durch die Aggregation bestehen und als solche für die Nutzung selbst keine Relevanz entfalten. Ähnliches gilt für Netzwerke aus der Ko-Nutzung von Hashtags, die sich für die Analyse gesellschaftlicher Diskursstrukturen eignen. Followernetzwerke hingegen sind nicht nur vergleichsweise stabil, sondern werden für die NutzerInnen laufend beobachtungs- und handlungsrelevant. Sie beeinflussen die Handlungsoptionen der NutzerInnen, strukturieren die individuellen Beobachtungsfelder (die Newsfeeds) und liegen den Konversations- und Diffusionsdynamiken auf Twitter zu Grunde. Sie repräsentieren also nicht ‚nur‘ gesellschaftliche Kommunikationsstrukturen, sondern prägen die Kommunikationsdynamiken selbst massiv mit. Durch diese Dualität der Struktur von Followernetzwerken erscheint ihre Analyse besonders lohnend, insbesondere für Forschung zur Struktur und Strukturierung von Netzöffentlichkeit.

3 Forschungsdesign

Twitter hat für sozialwissenschaftliche Forschung den immensen Vorteil, dass Elemente und Strukturen der Kommunikation standardisiert vorliegen. Tweets sind auf 140 Zeichen begrenzt, es gibt nur wenige syntaktisch-funktionale Merkmale (z.B. Replies, Hashtags), die Netzwerkstruktur beruht auf qualitativ gleichwertigen Followerbeziehungen und die Diffusionsprozesse werden im Gegensatz zu anderen Social-Media-Systemen (wie bspw. Facebook) bisher nicht durch anbieterseitige Selektionsroutinen beeinflusst. Zudem ist nicht nur die Kommunikation auf Twitter hochgradig transparent, das Unternehmen stellt auch Schnittstellen (APIs) zur Verfügung, um Informationen standardisiert abzurufen. Auch wenn diese APIs primär für Entwickler von Drittsoftware gedacht sind, vielfältige Nutzungseinschränkungen aufweisen und gelegentlich Unklarheiten bei der Vollständigkeit von Datensätzen auftauchen (Gaffney & Pusch-

mann, 2014), erscheinen die so gewonnen Daten relativ zuverlässig und können für quantitative Analysen herangezogen werden.

In unserem fünfstufigen Forschungsprozess greifen wir anfangs zweifach auf Twitter APIs als Informationsquellen zurück: (1) Wir sammeln Mitteilungen zu bestimmten Themenbereichen und rufen (2) anschließend Informationen zu den AutorInnen dieser Tweets – insbesondere zu ihren Followees – ab. Auf dieser Basis konstruieren wir (3) Followernetzwerke und führen Clusteranalysen durch. Danach werden (4) die gefundenen Cluster durch eine Inhaltsanalyse der Nutzerprofile beschrieben und dienen (5) als Ausgangspunkt für weiterführende Analysen zu Zusammenhängen zwischen Netzwerkstruktur und Kommunikationsdynamiken oder -inhalten. Im Folgenden werden diese Schritte besprochen und methodologisch reflektiert.

3.1 *Definition eines Themenfeldes und Erhebung der Tweets*

Da wir an den Netzwerken von NutzerInnen interessiert sind, die sich an Konversationen zu einem bestimmten Thema oder Themenfeld beteiligen, müssen in einem ersten Schritt Tweets zu diesen Themen gesammelt und die Autorennaccounts identifiziert werden. Dazu nutzen wir die Search oder Streaming APIs von Twitter, welche die Suche nach Schlagworten und Schlagwortkombinationen in der nahen Vergangenheit oder im laufenden Twitterstrom sowie die Speicherung der so gefundenen Tweets ermöglicht. Bei der Definition von Suchbegriffen gehen wir davon aus, dass die unter Beobachtung stehenden Themenfelder durch spezifische Begriffe und Begriffskombinationen eindeutig repräsentiert werden. Die Verwendung dieser Begriffe durch NutzerInnen bedeutet so eine Beteiligung dieser NutzerInnen an der thematisch fokussierten Kommunikation.

Im Gegensatz zu Analysen, die von spezifischen Nutzergruppen ausgehen (z.B. PolitikerInnen, JournalistenInnen; wie bei Golbeck et al., 2010; Jürgens & Jungheer, 2011) ist unsere Methode offen und sozial inklusiv: JedeR NutzerIn ist potenziell Teil der Grundgesamtheit, wenn wir unter Grundgesamtheit alle an einem Thema beteiligten NutzerInnen verstehen. Auf diese Weise kommen wir dem im niederschweligen Zugang zu Twitter begründeten hohen Beteiligungspotenzial nach: JedeR NutzerIn hat die gleichen technischen – aber natürlich nicht sozialen – Voraussetzungen für eine Teilnahme an Twitterkommunikation. So ist in

unserem Ansatz nicht die soziale Position eines Akteurs von Relevanz, sondern einzig und allein seine durch die Schlagwörter repräsentierte Beteiligung an den Kommunikationszusammenhängen.

Für unser erstes Fallbeispiel, den Diskurs zu Sexismus im Alltag im Jänner 2013, griffen wir auf den genutzten Hashtag #aufschrei als Stichwort zurück. Da die Diskussion unter diesem Begriff begonnen und geführt wurde, ist das Thema mit dem Hashtag hinlänglich abgedeckt. Für die Analyse waren wir besonders daran interessiert, wie die Debatte auf Twitter entstanden ist und welche NutzerInnen in den ersten 24 Stunden daran teilgenommen haben. Da die Entstehung der Debatte nicht vorhersehbar war, mussten die Tweets im Nachhinein über die Twiter Search-API erhoben werden. Auf diese Weise konnten 24.643 Tweets von insgesamt 8.126 NutzerInnen aufgefunden werden. Die Problematik im Umgang mit der Twitter-Search-API besteht darin, dass Ergebnisse laut Angaben von Twitter „auf Relevanz und nicht auf Vollständigkeit ausgelegt sind“, wobei der Plattformbetreiber nicht offenlegt, an welchen Kriterien sich Relevanz misst. Unsere Erfahrung zeigt jedoch, dass diese Probleme primär bei großen Datensätzen schlagend werden, die Nahe an API-Limits (z.B. 1 Prozent aller aktuell abgesetzten Tweets) kommen. Für unsere Untersuchung ist dies nur beschränkt problematisch, da wir vergleichsweise wenige Tweets gezogen haben, und weil wir nicht an den Tweets selbst sondern an den NutzerInnen interessiert sind (zu diesem Beispiel auch Maireder & Schlögl, 2014).

Für das zweite Beispiel, die innenpolitische Diskussion in Österreich im Vorfeld der Nationalratswahlen 2013, haben wir uns für einen Suchstring aus Parteienamen (BZÖ, FPÖ, Grüne, ÖVP, SPÖ, Team Stronach) und SpitzenkandidatInnen entschieden. Wir gehen dabei davon aus, dass NutzerInnen, die sich zu österreichischer Innenpolitik austauschen, diese Begriffe in einem längeren Zeitraum mehrfach nennen. Die Begriffe sind den gesuchten Parteien und Personen eindeutig zuordbar, einzig das Datenmaterial zu „Grüne“ musste ob der vielfältigen Verwendung des Begriffs manuell gesichtet und selektiert werden. Die siebenmonatige Erhebungsperiode – 14. Jänner bis zum 14. August 2013 – erlaubte nicht nur die NutzerInnen zu identifizieren, sondern auch aktive und weniger aktive NutzerInnen zu unterscheiden. So konnten wir aus den 11.776 AutorInnen der 149.691 Tweets im Erhebungszeitraum die aktivsten Accounts auswählen. Kriterium war dabei, dass

1 Twitter Developers: dev.twitter.com/docs/using-search

ein Account zumindest zehn Tweets mit den erwähnten Schlagwörtern gesendet hatte. Dies war bei 1.952 NutzerInnen der Fall (zu diesem Beispiel auch Maireder, Schlögl, & Schütz, 2013; Maireder & Schlögl, 2014).

3.2 Erhebung der Nutzerinformationen

Der nächste Schritt ist die Erhebung der Nutzerinformationen. Dies sind zum einen die inhaltlichen Informationen zum Account (z.B. Description, Language, Location), zum anderen die Followee-Listen (Friends) als Basis für die Netzwerkanalyse. Wir greifen dazu mit einem leicht veränderten Python Skript (Russel 2011, S. 380) auf die Twitter-REST-API zu und rufen die Daten an Hand der zuvor auf Basis des Nutzernamens erhobenen User-IDs ab. Dies ist ein sehr langwieriger Prozess, da Twitter die Menge der über die Schnittstelle abrufbaren Information zeitlich begrenzt. Zum Erhebungszeitpunkt konnte beispielsweise nur 60 Nutzer-IDs pro Stunde abgefragt werden und die Anzahl der bei einer solchen Anfrage abrufbaren Followee-IDs lag bei 5.000. Zudem sind die Informationen nicht für alle NutzerInnen verfügbar, da die Anfrage mit Accountnamen erfolgt, sich diese Accountnamen aber zwischen den beiden Erhebungszeitpunkten (Abruf des Tweets und Abruf der Nutzerinformationen) verändert haben können, oder Accounts gelöscht oder geschützt wurden. In unseren beiden Beispielen war die Erhebung für #aufschrei in 97 Prozent und für die Accounts der Österreichischen Innenpolitik in 95 Prozent der Fälle erfolgreich. Zur Minimierung dieser Probleme ist eine zur Sammlung der Tweets zeitnahe Erhebung der Nutzerinformationen zu empfehlen.

3.3 Netzwerkanalyse und strukturelle Identifikation von Öffentlichkeiten

Die erhobenen Followee-Informationen sind relationale Daten, die sich nach Zusammenführung als Basis für die Analyse als Netzwerk mit Accounts als Knoten und Followerbeziehungen als Kanten eignen. Die Erhebung grundlegender Netzwerk-Metriken (z.B. durchschnittlicher Grad, durchschnittliche Pfadlänge, Dichte) erlaubt einen ersten quantitativen Blick auf die Struktur der Netzwerke, für unsere Analyse sind wir jedoch insbesondere an der inneren Struktur des Netzwerks interessiert. Die Netzwerkanalyse unterscheidet zwischen *emic* und

etic groups: Erstere basieren auf netzwerkexogenen Kategorien (in unserem Fall z.B. die Zuordnung der Accounts zu Berufsgruppen, Geschlecht, etc.), zweite auf Strukturmerkmalen des Netzwerks selbst (Kadushin, 2012, S. 45). Dies sind beispielsweise dichte Stellen im Netzwerk, entsprechend unserer Operationalisierung von Öffentlichkeiten als kommunikative Verdichtungen. Solche Verdichtungen können mittels Clusterverfahren identifiziert werden, wobei zahlreiche Vorgehensweisen zur Auswahl stehen.

Die Konzepte und Methoden der (sozialen) Netzwerkanalyse entstehen an der Schnittstelle zwischen Mathematik und Sozialwissenschaft. SozialwissenschaftlerInnen, die mit dieser Methode arbeiten, sehen sich dementsprechend oft mit der Frage konfrontiert, ob die jeweiligen Maßzahlen und Optionen, welche die Netzwerkanalysesoftware zur Verfügung stellt, die (oft) soziologischen Konzepte, die aus der Theorie mitgebracht werden, tatsächlich abbilden können. Die Auswahl eines geeigneten Algorithmus zur Identifikation von Netzwerkclustern stellt eine schwierige und mitunter folgenschwere Entscheidung im Forschungsprozess dar. Grund dafür sind die meist divergierenden Forschungsziele zwischen SozialwissenschaftlerInnen einerseits, und den MathematikerInnen und ComputerwissenschaftlerInnen, die diese Art von Algorithmen entwickeln, andererseits. Schnelligkeit und die Anwendbarkeit auf besonders große Netzwerke sind wichtige Kriterien im Bereich der Computerwissenschaften, die für SozialwissenschaftlerInnen jedoch nicht alleine ausschlaggebend sind. Hier steht die Nachvollziehbarkeit des Algorithmus und seine methodologische Anschlussfähigkeit im Vordergrund, die es erlaubt, ihn als Operationalisierung eines sozialwissenschaftlichen Konzeptes zu nutzen.

Wir haben uns für die Analyse der Followernetzwerke für den von Pons und Latapy (2005) vorgeschlagenen ‚Walktrap‘-Algorithmus entschieden, der Knoten in einem agglomerativen Verfahren iterativ zu Gruppen zusammenführt. Die Rechenoperationen beruhen dabei auf der Annahme, dass zufällige gewählte Pfade in einem Netzwerk (‚random walks‘) insbesondere zwischen Knoten der gleichen Community verlaufen. Durch die Berechnung dieser Pfade ergibt sich ein Maß struktureller Ähnlichkeit durch das einzelne Knoten zu Communities aggregiert werden. Für die Wahl dieses Algorithmus waren insbesondere drei Punkte ausschlaggebend:

- Das Wichtigste dieser Kriterien ist die Anbindung an die theoretischen Überlegungen. Wir verstehen Öffentlichkeiten wie angemerkt „als Räume

verdichteter Kommunikation“ und der von Pons und Latapy vorgeschlagen Algorithmus operationalisiert diese Definition angemessen, da von Communities als Gruppen von Knoten gesprochen wird, die miteinander hochgradig verknüpft sind, aber wenige Links zu anderen Knoten aufweisen.

- Ein weiteres Kriterium ist, dass Communities bei unserem explorativen Design im Vorhinein nicht bekannt sein können. In verschiedenen Anwendungen der Netzwerkanalyse ist die Anzahl der Gruppen im Vorhinein festzulegen und die Aufgabe des Algorithmus besteht darin, nach bestimmten Kriterien die beste Zuordnung zu diesen zu finden (Newman & Girvan, 2004). Bei unserem Zugang richten wir die Analyse an den Ergebnissen des Algorithmus aus. Wir suchen demnach nach Communities, die sich einzig und allein aus der Netzwerkstruktur ergeben und setzen weder Anzahl noch Größe der resultierenden Communities voraus.
- Nicht zuletzt spielten auch weniger inhaltliche als forschungspragmatische Gründe eine Rolle bei der Entscheidung. Die Größe und Dichte des Netzwerkes (ganz besonders im Fall Aufschrei) machten die Anwendung klassischer Community-Algorithmen (z.B. basierend auf betweenness centrality) zu rechenintensiv, auch die Schnelligkeit des Algorithmus wurde damit zum Kriterium. Walktrap hat sich als besonders effizienter und schneller Algorithmus herausgestellt, der Community-Strukturen in Netzwerken wie den unseren auf modernen Standardrechnern innerhalb weniger Minuten findet.

3.4 Inhaltliche Beschreibung der Öffentlichkeiten

Nach der strukturellen Identifikation der Netzwerkverdichtungen führen wir manuelle Analysen der Nutzerinformationen durch, um die Communities inhaltlich zu beschreiben. Dabei gehen wir davon aus, dass die NutzerInnen innerhalb der Cluster Gemeinsamkeiten auf Basis geteilter Interessen aufweisen, die über die Nutzerbeschreibungen ersichtlich werden. Diese Profilbeschreibungen sind kurze, oft nur Schlagwörter enthaltende Textfelder, in denen die NutzerInnen sich selbst und/oder ihre Tweetpraxis beschreiben. Wir sichten also diese In-

formationen für alle NutzerInnen eines Clusters und suchen in einem induktiven Verfahren nach Gemeinsamkeiten, welche die Ausrichtung der Community anzeigen. Wie wir im Ergebnisteil zeigen, können diese Beschreibungsmuster ganz unterschiedlich ausgerichtet sein, wobei neben der Nutzerbeschreibung zuweilen auch die Ortsangabe (Location) der NutzerInnen Zusammenhänge offenbart – in diesen Fällen geographische. Diese Vorgangsweise verlangt den Forschenden bestimmte Grundkenntnisse über die jeweiligen Interessensgebiete ab, weil letztere in den Profilbeschreibungen auf verschiedenste Weise formuliert und ausgedrückt werden können.

Bei diesem Vorgehen beachten wir die ungleiche Position (Zentralität) der NutzerInnen in den einzelnen Clustern. Die meisten Clusteralgorithmen ordnen Knoten einer einzigen konkreten Nummer und damit einer Community zu. Diese Zuordnung kann jedoch eine starke Vereinfachung darstellen, da viele NutzerInnen möglicherweise zwischen zwei oder mehreren Communities stehen, im Endeffekt jedoch nur einer einzigen zugeteilt werden können. So beachten wir die Accounts an den ‚Rändern‘ der Cluster weniger als jene, die im Zentrum stehen. Dies sind die NutzerInnen, denen die integrativste Rolle innerhalb der Community zukommt, da ihnen viele andere in der Community folgen. In der Analyse berücksichtigen wir dies, indem wir mit den zentralsten Accounts der Communities (nach Indegree = Anzahl der eingehenden Verbindungen) beginnen und die Analyse so lange fortsetzen, bis sich eine theoretische Sättigung ergibt. Wie wir im Ergebnisteil zeigen, konnten wir so – bis auf einen Cluster im Netzwerk der DiskutantInnen rund um österreichische Innenpolitik – alle Communities beschreiben und ihre Ausrichtung auf Basis dieser Beschreibung interpretieren. Zu beachten ist dabei, dass die Ergebnisse dieser Analysen Beschreibungen von Aggregaten sind und nicht ohne weiteres auf jeden einzelnen Knoten rückgeführt werden können.

3.5 Weiterführende Analysen

Die identifizierten Netzwerkcluster können einen Ausgangspunkt für weiterführende Analysen darstellen, welche die Follower-Strukturen zur zeitlichen Entwicklung und inhaltlichen Dynamik von Diskussionen oder der erweiterten Vernetzungsstruktur in Bezug setzen. Dabei verstehen wir die Zuordnung der NutzerInnen zu einem Cluster als spezifisches Merkmal dieser NutzerInnen und

untersuchen Zusammenhänge zu Merkmalen der von ihnen gesendeten Mitteilungen. In den vorliegenden Fallstudien war dies einerseits die inhaltliche Ausrichtung der Tweets in ihrem Bezug zur #aufschrei-Diskussion, die wir für eine nach Stunden des Tages geschichtete Zufallsstichprobe (N=5.840) manuell codiert haben (siehe Maireder & Schlögl, 2014). Im Fall des Österreich-Politik-Netzwerks lag das Interesse darin, Unterschiede der Communities in ihrer externen Verlinkungspraxis zu beschreiben.

4 Ergebnisse

4.1 Netzwerkmetriken

In unseren beiden Beispielnetzwerken zeigt sich die Vernetzungsdichte in spezifischen Ausschnitten von Twitters Kommunikationsnetzwerk: JedeR NutzerIn hat durchschnittlich 44,3 (#aufschrei) bzw. 105,1 (Österreichische Politik) Verbindungen zu anderen NutzerInnen (average degree) und die durchschnittliche Entfernung zweier NutzerInnen (average path length) im Netzwerk beträgt 3,0 bzw. 2,3 Knoten. Trotz der jeweils hunderttausenden Kanten sind die Netzwerke im Vergleich zu einem vollständigen Netzwerk eher lose geknüpft, die Dichte beträgt 0,005 bzw. 0,057. Dabei überrascht nicht, dass das Netzwerk der Beteiligten an der spontan entstandenen und deutschlandweit geführten #aufschrei-Diskussion deutlich loser ist als die über einen langen Zeitraum erhobene Twitteröffentlichkeit zu österreichischer Innenpolitik (siehe Tabelle 1).

Tabelle 1: Metriken aus der Netzwerkanalyse der beiden Beispiele

	Knoten	Kanten	Ø Grad	Ø Pfadlänge	Dichte	Modularität
	Nodes	Edges	Average degree	Average path length	Density	Modularity
#aufschrei	8.179	362.024	44,3	3,0	0,005	0,38
AT Politik	1.852	194.614	105,1	2,3	0,057	0,11

Die Clusteranalyse zeigt die innere Struktur dieser Netzwerke, wobei wir für #aufschrei acht und für das Politik-Netzwerk 13 größere Communities identifiziert haben, die zwischen 146 und 2.114 bzw. 21 und 497 NutzerInnen haben. Das Maß für die Ausprägung der Cluster, die Modularität (modularity), ist in beiden Fällen relativ gering, die Verbindungen zwischen NutzerInnen verschiedener Cluster also relativ stark. Nichtsdestotrotz sind die Communities auf Basis spezifischer Merkmale in den Nutzerinformationen weitgehend homogen, wie die qualitative Analyse zeigt.

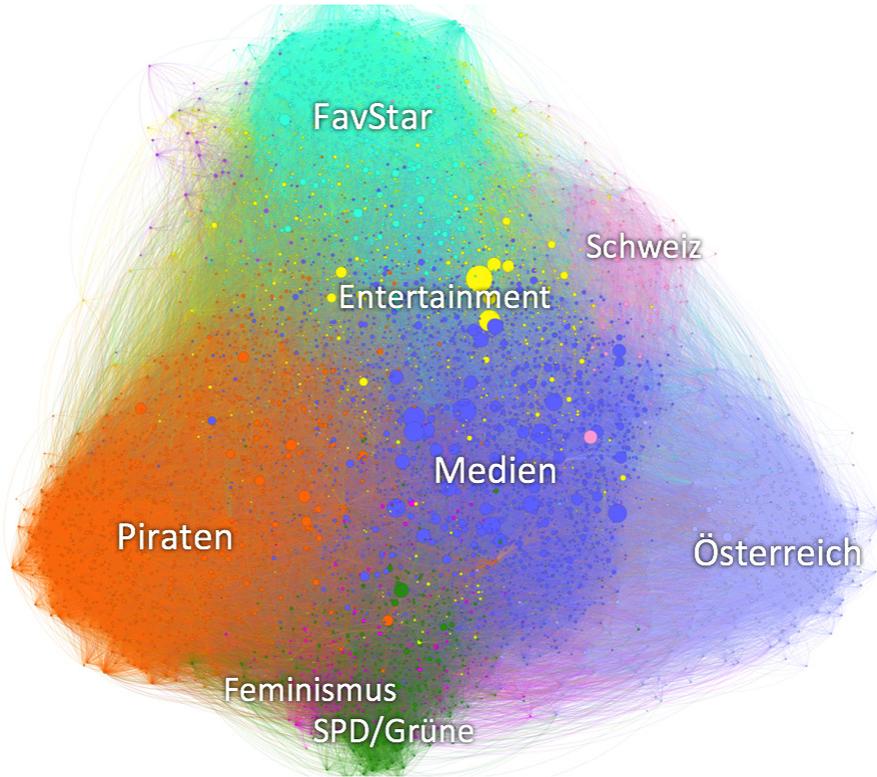
4.2 *Faktoren für die Beschreibung der Communities*

Im Folgenden beschreiben wir einige Communities aus den Fallbeispielen #aufschrei und österreichische Innenpolitik beispielhaft. Es geht hierbei hauptsächlich darum, aufzuzeigen, wie unterschiedlich die Interessenshintergründe gelagert sein können. Des Weiteren veranschaulichen wir, auf welche Arten sich Öffentlichkeiten auf Twitter strukturieren. Abbildungen 1 und 2 zeigen die Netzwerke, wobei die Farbe die Zugehörigkeit zur Community widerspiegelt und die Größe der Knoten den Indegree-Wert im Netzwerk. Die Community-Strukturen, die sich mit unserer Herangehensweise abbilden lassen, sind freilich vor allem davon abhängig, in welchem Kontext die Daten erhoben wurden. In unseren Fällen lassen sich Communities beider Netzwerke in fünf verschiedene Kategorien zusammenfassen:

Partei-Communities

In beiden Netzwerken finden wir Communities, die sich parteipolitisch verorten lassen. Als Indikatoren finden wir Verweise auf eine Partei (z.B. SPÖ) oder Vorfeldorganisationen (z.B. Sozialistische Jugend) oder die Deklaration als AnhängerIn einer bestimmten Partei-Ideologie (z.B. als ‚Sozialdemokrat‘). Im Fallbeispiel #aufschrei waren zwei parteipolitisch verortbare Communities zu identifizieren: eine ist stark von SPD und Grün-PolitikerInnen dominiert und eine andere der Piratenpartei zuzuordnen. Im Beispiel der österreichischen Innenpolitik finden wir erwartungsgemäß eine ganze Reihe solcher Communities: die Cluster ‚Sozialdemokratie‘ (SPÖ-nahe), ‚Konservative‘ (ÖVP-nahe), ‚Freiheitliche‘ (FPÖ-nahe) und ‚Piraten‘. Zudem finden wir eine kleine Community, deren Mitglieder sich häufig als Liberale deklarieren oder sich der neuen liberalen Partei

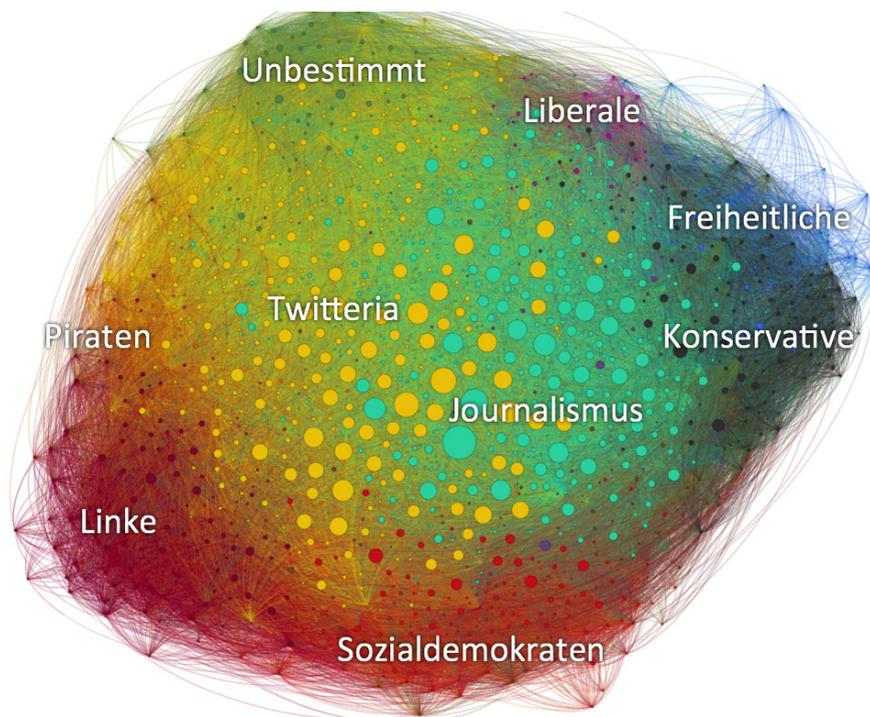
Abbildung 1: Followernetzwerk #aufschrei. Knoten=6.402, Kanten=343.445, Größe=Indegree, Farbe=Community



Quelle: Eigene Darstellung; Details zu den Clustern siehe Maireder & Schlägl, 2014

NEOS und dem Liberalen Forum zuschreiben. Die österreichischen Grünen sind die einzige der größeren Parteien, die keine als solche erkennbare Community im Netzwerk bildet. GrünpolitikerInnen und grünaffine NutzerInnen finden sich stattdessen in einer Community, die wir ‚Twitteria‘ genannt haben und weiter unten beschreiben.

Abbildung 2: Followernetzwerk Österreichische Innenpolitik. Knoten = 1.638, Kanten = 191.045, Größe=Indegree, Farbe=Community



Quelle: Eigene Darstellung; Details zu den Clustern siehe Maireder et al., 2013; Maireder & Schlögl, 2014

Politische Communities ohne Partei-Zusammenhang

Einige Communities beziehen sich zwar stark auf (gesellschafts-)politische Kategorien, lassen sich jedoch nicht ohne weiteres in die deutsche oder österreichische Parteienlandschaft einordnen. Im Netzwerk der #aufschrei-Beteiligten finden wir einen Cluster, dessen Accounts in ihren Profilbeschreibungen über-

aus oft auf Frauenpolitik, Gleichberechtigung und Gender Mainstreaming Bezug nehmen. Wir haben ihn ‚Feminismus‘ getauft. In der österreichischen politischen Twittersphäre finden wir eine Community, deren NutzerInnen auf Themen wie Antirassismus und Antifaschismus verweisen. Ebenso starke Verweise finden wir auf Positionen innerhalb der österreichischen Hochschülerschaft und damit der stark links geprägten studentischen Universitätspolitik. Aus anderen Studien wissen wir, dass viele NutzerInnen dieses Clusters ihre Twitternutzung im Zuge der Wiener Studierendenproteste 2009 aufgenommen haben und das die Beteiligten Twitter bis heute vielfach als Koordinations- und Mobilisierungskanal für politischen Aktivismus nutzen (Maireder & Schwarzenegger, 2012)

Medien & Entertainment

In beiden Untersuchungen gruppieren sich Accounts von Medienunternehmen und JournalistInnen in einem Cluster. Zusätzlich finden sich in diesen Clustern zahlreiche Accounts, die primär diesen Medienaccounts folgen, und die wir entsprechend als NutzerInnen verstehen können, die im Kontext unserer Analysen primär an journalistischen Inhalten interessiert sind. Entsprechend stark sind in diesen Clustern die Unterschiede in der Zentralität der Accounts. Dies gilt insbesondere auch für den ‚Entertainment‘-Cluster, den wir im #aufschrei-Netzwerk identifizieren konnten. Hier stehen ‚Celebrities‘ (MusikerInnen, SportlerInnen, Entertainer) im Mittelpunkt des Interesses.

Geographisch verortbare Communities

Geographisch verortbare Communities finden sich in beiden Fallstudien, jedoch auf unterschiedlichem Niveau. #aufschrei stellte einen ursprünglich deutschen Diskurs dar, jedoch finden wir auch TeilnehmerInnen aus Österreich und der Schweiz, die den Hashtag innerhalb der ersten 24 Stunden verwendet haben. Diese NutzerInnen erscheinen in der Netzwerkanalyse als eigene Communities, wobei vor allem die Twitter-interne Profilinformation „location“ in Kombination mit den Profilbeschreibungen hilft, diesen Zusammenhang zu verstehen. Bei den NutzerInnen, die sich an dem Diskurs zur österreichischen Innenpolitik beteiligt haben, können wir regionale Communities identifizieren, nämlich Nutzergruppen aus den Bundesländern Tirol und Kärnten.

Praxisgemeinschaften

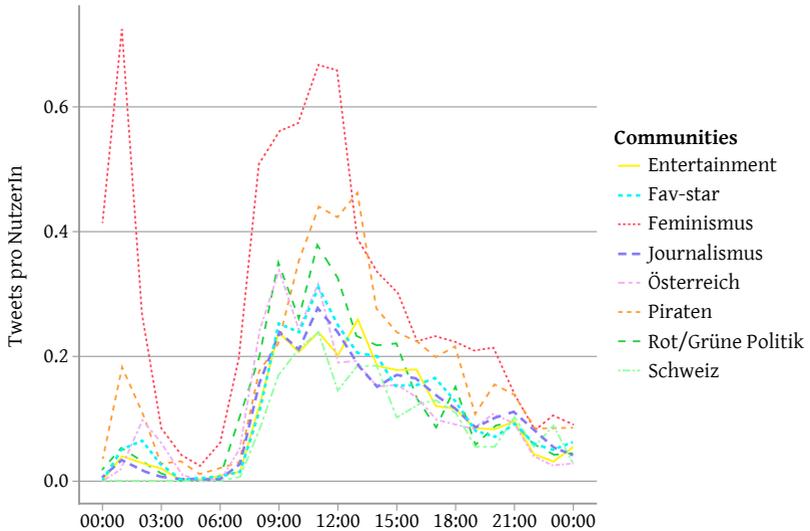
Unter Praxisgemeinschaften verstehen wir Communities, die primär über eine spezifische gemeinsame Kommunikationspraxis definiert werden können (vgl. Konzept der ‚Communities of Practice‘, z.B. bei Lave & Wenger, 1991). Eine solche Community ist der im Follower Netzwerk der #aufschrei-Debatte identifizierte ‚Favstar‘-Cluster. Ein überwiegender Teil dieser NutzerInnen nimmt in irgendeiner Weise (oft mit Link) im Profil Bezug auf die Website favstar.fm, bei der TwitternutzerInnen Punkte für originelle Tweets vergeben und so ihre ‚Favstars‘ wählen (dazu genauer bei Paßmann et. al, 2013). Auch die Community mit dem Namen *Twitteria* im Österreichbeispiel lässt sich (mit Einschränkungen) als Praxisgemeinschaft verstehen. Es handelt sich dabei um eine Gruppe von NutzerInnen, die zu den *Early Adopters* des Mediums in Österreich gehören und sich als Gruppe primär über ihre geteilte Nutzung von Twitter definieren. Gleichzeitig gehören sie zu den aktivsten aller NutzerInnen in Österreich (Ausserhofer & Maireder, 2013; Maireder, 2014).

4.3 *Weiterführende Analysen*

Die strukturelle und inhaltliche Identifizierung von Communities in Follower Netzwerken ermöglicht Einblick in die grundlegenden sozio-technischen Kommunikationsstrukturen ausgewählter Nutzergruppen, die für weiterführende Analysen herangezogen werden können. Dabei geht es um die Verknüpfung der Strukturmerkmale mit grundlegenden Metriken oder inhaltlichen Kriterien. Beim Fallbeispiel #aufschrei konnten wir die Ergebnisse nutzen, um die Entstehung der Debatte genauer zu rekonstruieren und die Beteiligung der verschiedenen Nutzergruppen in der Zeit mit der inhaltlichen Dynamik der Debatte zusammenführen (Maireder & Schlögl, 2014). Bei den Beteiligten an den innenpolitischen Diskussionen können wir unter anderem zeigen, dass die Communities jeweils verschiedene Nachrichtenmedien präferieren, gemessen an der Anzahl der Hyperlinks aus Twitter zu Nachrichtenwebsites.

Abbildung 3 zeigt die Aktivität zum Hashtag #aufschrei der jeweiligen Communities im Zeitverlauf der ersten 24 Stunden. Die Linien bilden die durchschnittliche Anzahl an Tweets pro NutzerInnen und Stunde ab. Dabei wird zwar deutlich,

Abbildung 3: #aufschrei-Tweets pro Minute nach Communities in den ersten 24 Stunden



dass die Aktivität aller Cluster einem ähnlichen Trend folgt, die Quantität der Beteiligung aber stark differiert. Die Feminismus-Community ist erwartungsgemäß über beinahe den ganzen Tag hinweg die weitaus Aktivste. Tatsächlich handelt es sich bei den NutzerInnen dieser Community auch um jene, die den Hashtag und damit die Debatte ins Leben gerufen haben. Dies geschah zwischen null und ein Uhr morgens und der Hashtag wurde besonders von dieser Gruppe über die Nacht stark weitergetragen. Noch in der Nacht schlossen sich auch NutzerInnen aus der Community *Piraten* an und schließlich waren es auch letztere, die zumindest zu einem Zeitpunkt (13:00) noch mehr als die Community Feminismus zum Hashtag twitterten. Die beiden ausländischen Communities unterscheiden sich stark von einander. Während die österreichische Community noch in der Nacht an der Debatte teilnimmt und schließlich auch am Morgen des nächsten Tages zu den aktivsten gehört, erreicht die Debatte die Community Schweiz erst im Laufe des Vormittages und ihre NutzerInnen gehören durchweg zu den weniger aktiven (Maireder & Schlögl, 2014)

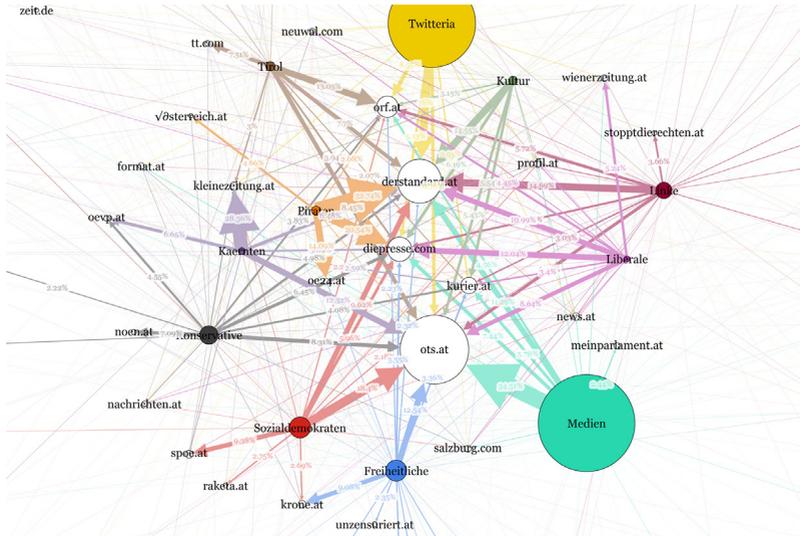
Die Inhaltsanalyse der Tweets zeigt, dass die Communities nicht nur in unterschiedlichem Maße aktiv sind, sondern sich auch inhaltlich auf unterschiedliche Weise zu dem Thema äußern: In über 40 Prozent aller Tweets (inklusive Retweets) aus der Community ‚Feminismus‘ werden persönliche Erfahrungen mit sexueller Belästigung geschildert, wobei diese Zahl auch Retweets anderer NutzerInnen beinhaltet. Es handelt sich dabei um jene Art von Tweets, für die der Hashtag #aufschrei eigentlich ins Leben gerufen wurde. Bei allen anderen Communities liegt der Anteil dieser Tweets zwischen 13 Prozent und 23 Prozent Prozent.

Wie sich im Zeitverlauf sehen lässt, machten persönliche Erfahrungen in den ersten Stunden bis zu 60 Prozent aller Miteilungen aus. Im Laufe des Nachmittags des Folgetages waren dies nur mehr zwischen 8 Prozent und 21 Prozent, stattdessen nahmen Beiträge zur Relevanz der Aufschrei-Diskussion an sich zu. Recht starke Unterschiede ergeben sich zudem bei Tweets, in denen der Debatte gegenüber ein Zuspriechen ausgedrückt wurde, beispielsweise durch Aufrufe, Aufschrei ernst zu nehmen, oder die Unterstreichung der Relevanz des Problems Sexismus. Tweets dieser Art machten einen wichtigen Teil bei den Communities *SPD/Grüne* und *Schweiz* aus (in beiden Fällen über 20 Prozent aller Tweets), spielten aber beispielsweise bei der *FavStar*-Community mit nur 9 Prozent eine untergeordnete Rolle. Dafür waren dort vergleichsweise viele kritische und satirische Kommentare zur Diskussion zu verzeichnen.

Abbildung 4 zeigt nun für das Fallbeispiel der politischen Netzwerke in Österreich, auf welche Online-Medien die NutzerInnen aus den jeweiligen Clustern in ihren Tweets verweisen. Die Beschriftung der Kanten weist aus, wie viel Prozent aller Links aus einer Community auf die jeweilige Domain verlinken. Der Originaltextservice der APA (Austria Presse Agentur), über den politische Akteure Presseaussendungen verschicken, dominiert das Bild. Die meisten Communities verweisen zu einem erheblichen Anteil (über 10 Prozent) auf diesen Service, welcher Presseaussendungen im Originaltext zur Verfügung stellt.

Besonders bei den parteipolitisch dominierten Communities überrascht dieses Ergebnis wenig, weil die Parteiangehörigen meist selbst über Twitter auf ihre Presseaussendungen auf dem Portal verweisen. Auf die beiden großen österreichischen Qualitätszeitungen *derstandard.at* und *diepresse.com* wird von den politisch zuordenbaren Communities in unterschiedlicher Weise verwiesen. Konservative und Liberale verweisen mit rund 2 Prozentpunkten häufiger auf die traditionell bürgerlich-liberale Presse als auf den Standard, während Lin-

Abbildung 4: Hyperlinks von NutzerInnen auf Online-Medien aggregiert auf Communities



Quelle: Eigene Darstellung

ke, Sozialdemokratie und Piraten stärker auf die den sozial-liberalen Standard verlinken. Zudem zeigen sich für mehrere Communities Besonderheiten in der Linkstruktur. Die Freiheitlichen Communities (in der Abbildung zusammengekommen) sind die einzigen, die zu einem erheblichen Anteil (9 Prozent) auf die Kronenzeitung (kronen.at) verweisen. Weitere communityspezifische Domains sind die kleinezeitung.at (Kärntner regionale Tageszeitung) mit 28 Prozent für Kärnten, tt.com (Tiroler Tageszeitung) mit 8 Prozent für Tirol, spoe.at (Homepage der Sozialdemokratischen Partei) mit 9 Prozent für die Community Sozialdemokratie und schließlich noen.at (Niederösterreichische Nachrichten) 7 Prozent und oewp.at (Homepage der österreichischen Volkspartei) mit 5 Prozent für den Cluster Konservative.

5 Conclusio

Wir haben ein Verfahren vorgestellt, mit dem sich Verdichtungen in Twitter-Followernetzwerken identifizieren und beschreiben lassen. Dabei nehmen wir jeweils die Netzwerke jener NutzerInnen in den Blick, die sich an der Kommunikation zu bestimmten Themen beteiligen. Diese Communities verstehen wir als ‚Öffentlichkeiten‘ im Sinne relativ stabiler Kommunikationszusammenhänge rund um geteilte Interessen. Dabei ist zu beachten, dass wir immer nur spezifische Ausschnitte der Netzwerke vor Augen haben, die sich aus der Auswahl der NutzerInnen als Beteiligte an thematisch fokussierter Kommunikation ergeben. Die gefundenen Communities spiegeln so die Follower Strukturen aus einer ganz spezifischen Perspektive wieder, die auch nur aus dieser Perspektive heraus interpretiert werden können. In unseren Fällen waren dies ein kurzer und intensiver sozio-politischer Diskurs einerseits und eine laufende, innenpolitische Auseinandersetzung andererseits. Da NutzerInnen – auch auf Twitter – immer Teil verschiedener kleinerer und größerer Öffentlichkeiten sind, können sich ein- und dieselben Accounts aus anderen Perspektiven auch in gänzlich anderen Communities wiederfinden.

Ein Beispiel dafür zeigt sich auch im Vergleich der hier vorgestellten Analysen, da sich einige der NutzerInnen aus unterschiedlichen Communities der österreichischen innenpolitischen Diskussion im #aufschrei-Netzwerk in einem gemeinsamen Cluster wiederfinden – Österreich eben. Bei Twitter-Followernetzwerken haben wir es so mit Vernetzungsstrukturen zu tun, die je nach Blickweite der Analyse unterschiedliche Muster zeigen. Die Vorstellung verschiedener voneinander abgrenzbarer ‚Ebenen‘ von Öffentlichkeit, wie prominent von Gerhards & Neidhard (1993) ausgearbeitet, erscheint so in Sozialen Online-Netzwerken obsolet. Öffentlichkeiten auf Twitter sind vielmehr als komplexe kommunikative Konstellationen zu verstehen, deren Strukturen sich je nach Beobachtungsperspektive anders darlegen.

Dennoch – oder auch genau deswegen – ist die Analyse von Followernetzwerken auf Twitter ein lohnendes Unterfangen. Denn die auf der Makroebene beobachteten Strukturen sind sowohl Ausdruck der Nutzungspraxis auf der Mikroebene – der Wahl von Informationsbeziehungen durch die NutzerInnen – als auch strukturelle Basis für kommunikative Dynamiken im Netzwerk. So ermöglicht unser Analyseverfahren die Exploration sozialer Strukturen in der Repräsen-

tation durch Twitterkommunikation, die vor dem Hintergrund der jeweiligen Kommunikationskontexte interpretiert und für weitere Analysen herangezogen werden können. Da kleinteilig strukturierte Kommunikationsnetzwerke wie jenes von Twitter öffentliche Kommunikationsprozesse immer stärker prägen, verstehen wir Strukturanalysen dieser Netzwerke als außerordentlich erkenntnisreiche Erweiterung des sozialwissenschaftlichen Methodensets zur Erforschung kommunikativer Praxis und der Dynamik öffentlicher Kommunikation.

Dr. Axel Maireder ist wissenschaftlicher Leiter des Social Media Intelligence Center der GfK und war bis Ende 2014 Mitarbeiter am Institut für Publizistik- und Kommunikationswissenschaft der Universität Wien

Mag. Stephan Schlögl ist wissenschaftlicher Mitarbeiter am Social Media Intelligence Center der GfK und promoviert an der Universität Wien im Fach Internationale Entwicklung

Quellenverzeichnis

- Anderson, C. W. (2010). Journalistic Networks and the Diffusion of Local News: The Brief, Happy News Life of the “Francisville Four.” *Political Communication*, 27(3), 289-309. doi: 10.1080/10584609.2010.496710
- Ausserhofer, J., & Maireder, A. (2013). National Politics on Twitter. *Information, Communication & Society*, 16(3), 291-314. doi: 10.1080/1369118X.2012.756050
- Bastos, M. T., Raimundo, R. L. G., & Travitzki, R. (2013). Gatekeeping Twitter: message diffusion in political hashtags. *Media, Culture & Society*, 35(2), 260-270. doi: 10.1177/0163443712467594
- Benkler, Y. (2007). *The wealth of networks: How social production transforms markets and freedom*. New Haven London: Yale University Press.
- Bruns, A., & Burgess, J. E. (2011). The Use of Twitter Hashtags in the Formation of Ad Hoc Publics. Reykjavik. Abgerufen von [http://snurb.info/files/2011/The Use of Twitter Hashtags in the Formation of Ad Hoc Publics \(final\).pdf](http://snurb.info/files/2011/The_Use_of_Twitter_Hashtags_in_the_Formation_of_Ad_Hoc_Publics_(final).pdf)
- Bruns, A., & Moe, H. (2014). Structural layers of communication on Twitter. In A. Bruns, K. Weller, J. Burgess, M. Mahrt, & C. Puschmann (Hrsg.), *Twitter and Society* (S. 15-28). New York, NY: Peter Lang.

- Dahlgren, P. (2005). The Internet, Public Spheres, and Political Communication: Dispersion and Deliberation. *Political Communication*, 22(2), 147-162. doi: 10.1080/10584600590933160
- Emirbayer, M., & Sheller, M. (1999). Publics in history. *Theory and Society*, 28(1), 143-197. Abgerufen von <http://www.jstor.org/stable/3108509>
- Gaffney, D., & Puschmann, C. (2014). Data collection on Twitter. In K. Weller, A. Bruns, J. Burgess, M. Mahrt, & C. Puschmann (Hrsg.), *Twitter and Society* (S. 55-67). New York, NY: Lang.
- Gerhards, J., & Neidhart, F. (1993). Strukturen und Funktionen moderner Öffentlichkeit. In W. R. Langenbucher (Hrsg.), *Politische Kommunikation* (S. 52-88). Wien: Braumüller.
- Gerlitz, C., & Helmond, A. (2013). The like economy: Social buttons and the data-intensive web. *New Media & Society*, 15(8), 1348-1365. doi: 10.1177/1461444812472322
- Golbeck, J., Grimes, J. M., & Rogers, A. (2010). Twitter use by the U.S. Congress. *Journal of the American Society for Information Science and Technology*, 61(8), 1612-1621. doi: 10.1002/asi.21344
- Gruning, J. E., & Hunt, T. T. (1984). *Managing Public Relations*. New York: Holt, Rinehart & Winston.
- Habermas, J. (1998). *Faktizität und Geltung*. Frankfurt am Main: Suhrkamp.
- Hepp, A., Brüggemann, M., Königslöw, K. K., Lingenberg, S., & Möller, J. (2012). Ein Beschreibungs- und Erklärungsansatz: Transnationale Öffentlichkeit und politische Diskurskulturen in Europa. In A. Hepp, M. Brüggemann, K. Kleinen-von Königslöw, S. Lingenberg, & J. Möller (Hrsg.), *Politische Diskurskulturen in Europa* (S. 21-47). Wiesbaden: VS Verlag.
- Jürgens, P., & Jungheer, A. (2011). Wahlkampf vom Sofa aus: Twitter im Bundestagswahlkampf 2009. In J. Schweitzer & S. Albrecht (Hrsg.), *Das Internet im Wahlkampf* (S. 201-225). Wiesbaden: VS Verlag.
- Kadushin, C. (2012). *Understanding social networks: Theories, concepts, and findings*. Oxford: Oxford University Press.
- Kohring, M. (1997). *Die Funktion des Wissenschaftsjournalismus: Ein systemtheoretischer Entwurf*. Opladen: Westdeutscher Verlag.
- Larsson, A. O., & Moe, H. (2011). Studying political microblogging: Twitter users in the 2010 Swedish election campaign. *New Media & Society*, 14(5), 729-747. doi: 10.1177/1461444811422894

- Lave, J., & Wenger, E. (1991). *Situated Learning. Legitimate peripheral participation*. Cambridge: University of Cambridge Press.
- Luhmann, N. (2004). *Die Realität der Massenmedien*. Wiesbaden: VS Verlag.
- Maireder, A. (2014). Ein Tweet: Zur Struktur von Netzöffentlichkeit am Beispiel Twitter. In H. Ortner, D. Pfurtscheller, M. Rizzolli, & A. Wiesinger (Hrsg.), *Datenflut und Informationskanäle* (S. 55-70). Innsbruck: iup.
- Maireder, A., Schlögl, S., & Schütz, A. (2013). Twitter-Politik 2013. Parteien, Medien, Netzwerke. Forschungsbericht. Universität Wien.
- Maireder, A., & Schwarzenegger, C. (2012). A Movement of connected Individuals. *Information, Communication & Society*, 15(2), 171-195. doi: 10.1080/1369118X.2011.589908
- Maireder, A., & Schlögl, S. (2014). 24 hours of an #outcry: The networked publics of a socio-political debate. *European Journal of Communication*, 20(6), 687-702. doi: 10.1177/0267323114545710
- Newman, M. E. J., & Girvan, M. (2004). Finding and evaluating community structure in networks. *Physical Review E*, 69(2), 026113. doi: 10.1103/PhysRevE.69.026113
- Paßmann, J., Boeschoten T., & Schäfer M. (2013). The Gift of the Gab: Retweet cartels and gift economies on Twitter. In A. Bruns, K. Weller, J. Burgess, M. Mahrt, & C. Puschmann (Hrsg.), *Twitter and Society* (S. 331-344). New York, NY: Peter Lang.
- Pons, P., & Latapy, M. (2005). Computing communities in large networks using random walks. *Computer and Information Sciences-ISCIS*, 3733, 284-293. Physics and Society; Disordered Systems and Neural Networks; Statistical Mechanics. doi: 10.1007/11569596_31
- Russell, M. A. (2011). *Mining the Social Web: Analyzing Data from Facebook, Twitter, LinkedIn, and Other Social Media Sites*. Beijing: O'Reilly Media.
- Schweiger, W., & Weihermüller, M. (2008). Öffentliche Meinung als Online-Diskurs – ein neuer empirischer Zugang. *Publizistik*, 53(4), 535-559. doi: 10.1007/PL0002234

Empfohlene Zitierung: Bruns, A., & Sauter, T. (2015). Anatomie eines Trending Topics: Methodische Ansätze zur Visualisierung von Retweet-Ketten. In A. Maireder, J. Ausserhofer, C. Schumann, & M. Taddicken (Hrsg.), *Digitale Methoden in der Kommunikationswissenschaft* (S. 141-161). doi: 10.17174/dcr.v2.7

Zusammenfassung: Social Media dienen immer häufiger als Disseminationsmechanismen für Medieninhalte. Auf Twitter ermöglicht besonders die Retweet-Funktion den schnellen und weitläufigen Transfer von Nachrichten. In diesem Beitrag etablieren neue methodische Ansätze zur Erfassung, Visualisierung und Analyse von Retweet-Ketten. Insbesondere heben wir hervor, wie bestehende Netzwerkanalysemethoden ergänzt werden können, um den Ablauf der Weiterleitung sowohl temporal als auch spatial zu erfassen. Unsere Fallstudie demonstriert die Verbreitung des Videoclips einer am 9. Oktober 2012 spontan gehaltenen Wutrede der australischen Premierministerin Julia Gillard, in der sie Oppositionsführer Tony Abbott als Frauenhasser brandmarkte. Durch die Erfassung von Hintergrunddaten zu den jeweiligen NutzerInnen, die sich an der Weiterleitung des Videoclips beteiligten, erstellen wir ein detailliertes Bild des Disseminationsablaufs im vorliegenden Fall. So lassen sich die wichtigsten AkteurInnen und der Ablauf der Weiterleitung darstellen und analysieren. Daraus entstehen Einblicke in die allgemeinen Verbreitungsmuster von Nachrichten auf Twitter.

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

Axel Bruns & Theresa Sauter

Anatomie eines Trending Topics

Methodische Ansätze zur Visualisierung von Retweet-Ketten

1 Netzwerkanalyse und -visualisierung

Soziale Netzwerkplattformen und Microbloggingdienste gewinnen als Medien für Nachrichtenkonsum und -verbreitung zunehmend an Bedeutung. Statistiken aus den USA zeigen, dass die Hälfte aller erwachsenen Facebook- und Twitter-NutzerInnen regelmäßig Nachrichten über diese Seiten beziehen. Bei Reddit sind es sogar über 60 Prozent (Matsa & Mitchell, 2014). Des Weiteren beteiligen sich die NutzerInnen dieser Seiten zunehmend an der Distribution von Medieninhalten (Maireder, 2013). Woong Yun & Park (2011, S. 202) bezeichnen Facebook und andere soziale Netzwerkseiten als „major news distributing outlets“. Der strukturelle Aufbau von Nutzerverbindungen auf Twitter (durch nicht zwingend reziproke Follower-Verbindungen) als „meso-level follower network“ (Bruns & Moe, 2014, S. 22) macht die Microbloggingplattform zu einem besonders effektiven Disseminationsmedium für „breaking news“ und Gerüchte. Auch Kwak, Lee, Park, & Moon (2010) weisen darauf hin, dass Twitter weniger zur Mitteilung alltäglicher persönlicher Begebenheiten als zu Verbreitung von Nachrichten genutzt wird. In Krisen- und Katastrophensituation (Bruns & Burgess, 2014; Mendoza, Poblete, & Castillo, 2010; Palen, Starbird, Vieweg, & Hughes, 2010) und während entscheidender politischer Ereignisse (Hermida, 2014; Lotan, Gra-

eff, Ananny, Gaffney, Pearce, & boyd, 2011; Burns & Eltham, 2009) hat sich Twitter in den letzten Jahren als besonders schneller Verbreitungsmechanismus erwiesen, der oft sogar den herkömmlichen Nachrichtenmedien zuvorkommt (Hermida, Lewis, & Zamith, 2014; Bruns & Burgess, 2011).

Twitter und andere soziale Netzwerkseiten spielen bei der Verbreitung von aktuellen Ereignissen also eine immer präsentere Rolle. Die Dynamik der Informationsverbreitung durch digitale Netzwerkstrukturen ist jedoch noch relativ unerforscht. Zu welchem Zeitpunkt wird ein Thema von wem aufgegriffen? Wer leitet eine Nachricht wann an wen weiter? Welche Rolle spielen einzelne KommunikationsteilnehmerInnen, bestehende digitale Communities oder die technischen Grundlagen der einzelnen Netzwerkplattformen bei der Verbreitung von Nachrichten? In diesem Kapitel erörtern wir anhand des Beispiels eines auf Twitter weitergeleiteten Videos zu einem aktuellen sozialpolitischen Thema in Australien einige neue Methoden für die dynamische Darstellung und Analyse von Kommunikationsabläufen. Unsere Methode verbindet zeitliche und räumliche Analyseansätze und eröffnet neue Einblicke in die Verbreitung von Nachrichten in digitalen Netzwerken.

1.1 *Welche Twitternetzwerke?*

Das Nachzeichnen der Verbreitung aktueller Nachrichten über soziale Netzwerkplattformen wie Twitter eröffnet ForscherInnen einzigartige Einblicke in die Dynamik sozialer, kultureller und politischer Diskurse, Öffentlichkeiten und Kontroversen. Webberley, Allen und Whitaker (2011) bekunden, dass die Untersuchung von Retweets es ForscherInnen ermöglicht, zu analysieren, wie Informationen durch die soziale Struktur des Twitternetzwerkes verbreitet werden, um anhand dessen zu deuten, welche Art von Information die größte Wahrscheinlichkeit hat, verbreitet zu werden. Die Entwicklung analytischer Methoden für die Untersuchung solcher Verbreitungsdynamiken und insbesondere auch die Anwendung solcher Methoden auf konkrete Fragestellungen in den Medien- und Kommunikationswissenschaften steht jedoch noch am Anfang. Die Analyse der zeitlichen Dynamik von Kommunikationsvolumina gehört inzwischen zwar zum Standardrepertoire der Twitterforschung (Bruns &

Stieglitz, 2014) und ermöglicht es, wichtige Momente in der Entwicklung eines gegebenen Diskussionsthemas zeitlich genau zu bestimmen. Netzwerkanalysen und -visualisierungen zur Verbreitung von Nachrichten und politischen Diskussionen auf Twitter beschränken sich jedoch zumeist auf die Untersuchung statischer Gesamtnetzwerke auf Basis von Replies und/oder Retweets über einen kompletten Untersuchungszeitraum und ignorieren somit die zeitliche Dimension. Eine solche Beschränkung erfolgt dabei größtenteils aus rein forschungspraktischen Gründen.

Die Analyse derartiger Gesamtnetzwerke (wie z.B. bei Maireder & Ausserhofer, 2014; Paßmann, Boeschoten, & Schäfer, 2014; Nuernbergk, 2013; Nuernbergk & Neubarth, 2013) ist zweifelsohne höchst interessant und wichtig, weil sie im Rückblick auf das zu untersuchende Ereignis die insgesamt aktivsten, zentralsten oder wichtigsten DiskussionsteilnehmerInnen identifiziert. Dadurch ist eine genauere Untersuchung der Kommunikationsaktivitäten dieser führenden NutzerInnen sowie des Echos der weiteren Teilnehmerschar möglich. Jedoch sind die langfristig zentralsten TeilnehmerInnen nicht notwendigerweise *ab origine* bei der Etablierung eines Diskussionsthemas (oder auch etwaiger zugehöriger Hashtags) federführend. Stattdessen handelt es sich bei ihnen mitunter um bereits einschlägig in der Netzgemeinde bekannte NutzerInnen, die einem Thema, das in kleinerem Kreise bereits einiges an Fahrt aufgenommen hat, durch ihre Teilnahme zu einem späteren Zeitpunkt weitere Prominenz vermitteln. Dies war zum Beispiel bei der unter dem Hashtag #aufschrei bekanntgewordenen Twitterkampagne gegen alltäglichen Sexismus der Fall, als eine zunächst eher überschaubare Diskussion persönlicher Erfahrungen durch die Teilnahme prominenter NetzaktivistInnen und JournalistInnen deutschlandweit bekannt wurde (siehe Maireder & Schlögl, 2014, sowie in diesem Band; Eckert, 2013).

Eine dynamische Netzwerkanalyse hingegen ermöglicht es, die zu verschiedenen Zeitpunkten jeweils führenden TeilnehmerInnen zu identifizieren und so die Formation der Interessentenschar zu verfolgen und dadurch die Evolution der Diskussion nachzuvollziehen. Dabei muss in erster Linie zwischen drei verschiedenen Netzwerkformen unterschieden werden, die sich an den von Bruns & Moe (2014) beschriebenen drei Kommunikationsebenen der Twitterplattform orientieren. Hierbei werden die Mikroebene der Kommunikation mittels Mentions und Replies, die Mesoebene der miteinander verwobenen persönlichen Followernetz-

werke einzelner NutzerInnen, und die Makroebene spontan oder geplant entstehender Hashtags unterschieden:

1. Das durch Replies und Retweets aufgespannte Netzwerk zwischen aktiven DiskussionsteilnehmerInnen zeigt, wer mit wem interagiert und identifiziert so für jeden Zeitpunkt, um welche TeilnehmerInnen herum der Austausch gerade am stärksten zentriert ist bzw. wessen Beiträge aktuell die größte Anzahl neu hinzukommender aktiver TeilnehmerInnen anziehen. Ein derartiges Netzwerk existiert nach dem oben erwähnten Dreiebenenschema der Twitterkommunikation auf der Mikroebene, da es sich hier in erster Linie um direkte aber öffentliche Interaktionen zwischen aktiven TeilnehmerInnen handelt. Diese Art dynamisches Netzwerk lässt sich daher auch durch die von Bruns (2012) aufgezeigten Visualisierungsmethoden darstellen.
2. Das zuvor beschriebene Netzwerk aktiver TeilnehmerInnen an einem Thema besteht jedoch auch vor dem Hintergrund der Mesoebene bereits existierender Followernetzwerke. Die Interaktionen zwischen aktiven TeilnehmerInnen an einem Thema sind daher zumeist nicht nur für diese TeilnehmerInnen sichtbar, sondern auch für ihre jeweiligen Follower, sodass die Gesamtzahl der passiven TeilnehmerInnen um einige Größenordnungen höher liegen kann als die der aktiven TeilnehmerInnen. Eine Darstellung dieses grundlegenden Followernetzwerks, auf welche dann die aktive Verbreitung relevanter Tweets abgebildet wird, ermöglicht es zudem die Reichweite eines Themas einzuschätzen: Ist es lediglich in einem schon vorher eng vernetzten Teil der gesamten Twittersphäre aktiv, oder erreicht es ein wirklich breites Twitterpublikum? Die Analyse einer solchen hybriden Überlagerung zweier Netzwerke (des Mikronetzwerks direkter Replies und Retweets und des Mesonetzwerks bestehender Followerverbindungen) setzt jedoch die Verfügbarkeit von Followerdaten für alle aktiven TeilnehmerInnen voraus und erfordert besonders bei größeren Ereignissen und Netzwerken einen hohen forschungstechnischen Aufwand.¹

1 Streng genommen müsste zudem auch die Dynamik der Followernetzwerke selbst berücksichtigt werden, welche etwa durch Methoden wie den von Meeder, Karrer, Sayedi, Ravi, Borgs, & Chayes (2011) sowie Bruns, Woodford, & Sadkowsky (2014b)

3. Zumindest theoretisch könnte zudem auch noch das Gesamtnetzwerk aller TwitternutzerInnen untersucht werden, die einem ereignisrelevanten Hashtag folgen, um so die Netzwerkdynamik auf der Makroebene der Twitterkommunikation abzubilden. Dadurch würde berücksichtigt, dass TwitternutzerInnen einem Thema mit Hilfe von Suchfunktionen auf der Twitter-Website oder in einschlägigen Applikationen auch dann folgen können, wenn sie oder die NutzerInnen, denen sie im Rahmen ihre alltäglichen Nutzungspraxis folgen, nicht selbst aktiv am Thema beteiligt sind. Zu einer verlässlichen Einschätzung dieses Zusatzpublikums auf der Makroebene jedoch fehlt der Twitterforschung die notwendige Datengrundlage, da über das Twitter-API keine Daten dazu zur Verfügung stehen, welche NutzerInnen einem bestimmten Hashtag folgen. In der weiteren Diskussion beschränken wir uns daher auf die ersten beiden Forschungsszenarien.

1.2 Welche Datengrundlagen?

Im Prinzip sind die für die Netzwerkanalyse und -visualisierung von Twitterdiskussionen benötigten Datensätze relativ leicht über das Twitter Application Programming Interface (API) zu erfassen. Falls ein einschlägiger Hashtag oder eindeutige Begriffe existieren, können die entsprechenden Tweets mit Hilfe von Tools wie *yourTwrapperkeeper* (Gaffney & Puschmann, 2014) erfasst und archiviert werden. Gleiches gilt für Tweets, die Links zu bestimmten URLs bzw. Domains beinhalten. Jedoch erlauben viele solcher Tools nur die Erfassung laufender Ereignisse über das Twitter Streaming API und sind bei der Erfassung älterer Inhalte durch die Funktionalität des Twitter Search API stark beschränkt (Gaffney & Puschmann, 2014). Dies ist besonders für Forschungsansätze problematisch, die an der Frühphase der Entstehung von Diskussionen interessiert sind – die sich also zum Beispiel dafür interessieren, welche TeilnehmerInnen als Erste Links auf ein später viral weit verbreitetes Video weitergeleitet haben, oder welche NutzerInnen schon vor dem Entstehen eines Hashtags wie #aufschrei an

beschriebenen Ansatz erfasst werden kann, da sich das Followernetzwerk einzelner DiskussionsteilnehmerInnen während oder sogar wegen der aktuellen Diskussion weiter verändert haben mag.

der später für den Hashtag verantwortlichen Twitterdebatte beteiligt waren. Solche Daten aus der Frühzeit eines kommunikativen Ereignisses sind oft nur dann verfügbar, wenn sie zufällig schon durch laufende Erfassungen mitgesammelt oder nachträglich von kommerziellen Datenanbietern wie DataSift erworben wurden. Im unten behandelten Fallbeispiel etwa war es uns möglich, die benötigten Daten zur Verbreitung einer bestimmten URL aus einem größeren Datensatz zu extrahieren, in dem fortlaufend alle Tweets mit Links zu der übergeordneten Domain erfasst werden.

Von diesen Einschränkungen abgesehen ist es in Folge relativ einfach, aus einem Datensatz von Tweets zu einem bestimmten kommunikativen Ereignis (die jeweils etwa von den Forschern bestimmte Hashtags, Kennworte, oder URLs beinhalten) die relevanten dynamischen Netzwerkinformationen zu extrahieren. Daten und Metadaten zu jedem Tweet beinhalten jeweils den Nutzernamen und die numerische Nutzer-ID des Absenders sowie einen sekundengenauen Time-stamp. Der Tweettext selbst enthält – soweit es sich um einen an eine/n andere/n TeilnehmerIn gerichteten Tweet handelt – eine Adressierung im Format *@user*, wobei zusätzlich noch zwischen echten Replies und Retweets (in Formaten wie *RT @user*, *MT @user* usw.) unterschieden werden kann (Bruns & Stieglitz, 2014). Für jeden Tweet lassen sich daher die für die dynamische Netzwerkvisualisierung zwingend notwendigen Merkmale *Absender*, *Zeitpunkt* sowie gegebenenfalls *Empfänger* bestimmen.²

Diese Datengrundlage ist für die Analyse und Visualisierung der ersten oben angeführten Netzwerkvariante bereits ausreichend. Für die zweite Variante hingegen fehlen noch die Hintergrundinformationen zu den bereits existierenden Followernetzwerken der am aktuellen Thema teilnehmenden NutzerInnen. Auch

2 Selbst wenn *Empfänger* nicht gesetzt sein sollte – wenn es sich also um einen nicht spezifisch adressierten Tweet an ein unbestimmtes Publikum handelt –, ist es mitunter dennoch sinnvoll, einen solchen Tweet mit in die Netzwerkanalyse aufzunehmen, da es sich hier zum Beispiel um den Ausgangspunkt einer späteren Kette von Retweets handeln könnte. Anfangs würde ein solcher Tweet somit als vom weiteren Netzwerk isolierter einzelner Knotenpunkt erscheinen, doch zu späterem Zeitpunkt greifen weitere NutzerInnen aus dem bestehenden Followernetzwerks des Absenders den Tweet eventuell auf, retweeten ihn, und verbreiten ihn somit weiter. In der Netzwerkanalyse würde der zuvor isolierte Knotenpunkt damit nachträglich mit anderen Knoten verbunden.

diese lassen sich zumindest im Prinzip über die Twitter-API erfassen, wobei wegen der technischen Beschränkungen der API besonders bei großen Netzwerken oder einer Vielzahl von teilnehmenden NutzernInnen viel Geduld nötig ist.

Bei der Erfassung und Auswertung der Followernetzwerke muss zudem beachtet werden, dass die Daten aus der Twitter-API jeweils nur den aktuellen Stand eines Followernetzwerks zeigen. Zwar ist es durch verschiedene Methoden (Meeder et al., 2011; Bruns et al., 2014b) möglich, den Stand des Netzwerks zu einem beliebigen gegebenen Zeitpunkt – also etwa dem Zeitpunkt des zu untersuchenden kommunikativen Ereignisses – zumindest annähernd zu extrapolieren und somit erst später hinzugekommene Follower von der Netzwerkanalyse auszuschließen, doch ist auch diese Korrektur nicht in der Lage, etwaige frühere Follower, die die Verbindung mittlerweile wieder gelöst haben, zu erkennen. Das einem konkreten, durch Replies und Retweets realisierten kommunikativen Ereignis zugrundeliegende Followernetzwerk kann daher grundsätzlich immer nur annähernd ermittelt werden, ist jedoch für die weitere Analyse trotzdem nützlich.

Unser Beitrag veranschaulicht diese neuen methodischen Ansätze zur Analyse von Entwicklung und Verbreitung digitaler Diskussionen. Anhand einer Fallstudie zur Verbreitung eines Videos auf Twitter haben wir dynamische Darstellungen der Disseminationsstrukturen entwickelt, die beide der oben aufgezeigten Netzwerkanalysevarianten illustrieren. Solche Visualisierungen machen es möglich, die wichtigsten AkteurInnen in der Verbreitung eines Themas und den dynamischen Ablauf nachzuzeichnen. Die Anwendung einer solchen Visualisierung bietet Einblicke in die Entstehung globaler Kommunikationsmuster und öffentlichen Meinungsaustauschs. Dadurch können einige wichtige Fragen beantwortet werden: Wer verbreitet Material? Wer hat am meisten Tragkraft? In welche Bereiche breitet sich ein Thema aus? Und wie lange bleibt es tatsächlich aktuell und interessant?

Neben dem Einblick in die Disseminationsdynamik des vorliegenden Falles ermöglicht die Analyse dieses Medienereignisses auch allgemeine Erkenntnisse zu den Mustern der Nachrichtenverbreitung auf Twitter. Da die nutzergesteuerte Weiterleitung von Links zu aktuellen Nachrichten eine immer größere Rolle im Informationskonsum der NetznutzerInnen darstellt, ist die Untersuchung auch von allgemeiner Bedeutung.

2 Fallstudie – Der Fall Gillard

2.1 *Das Video*

Am 9. Oktober 2012 geriet die damalige australische Premierministerin Julia Gillard während einer Parlamentssitzung in Rage über die sexistische Attitüde des Oppositionsführers Tony Abbott und der allgegenwärtigen Frauenfeindlichkeit in der australischen Politik. Abbott hatte den Rücktritt des Präsidenten des australischen Unterhauses, Peter Slipper, gefordert, nachdem dieser in einen Sexskandal mit einem Ex-Mitarbeiter verwickelt worden war und seine E-Mails mit teilweise vulgären Inhalten öffentlich geworden waren. Abbott, dessen eigene Vorstellungen zu Frauen und Familie schon seit Langem in der öffentlichen Kritik standen, beschuldigte dabei auch Gillard, mit ihrer Unterstützung für Slipper einem Frauenfeind Tür und Tor geöffnet zu haben. In einer flammenden, aus dem Stegreif gehaltenen fünfzehnminütigen Rede lehnte Gillard es vehement ab, sich von Abbott zum Thema Frauenfeindlichkeit belehren zu lassen und bezeichnete ihn selbst als Sexisten.

Der öffentlich-rechtliche australische Sender ABC veröffentlichte sofort ein Video der Rede auf seiner Website. Auch auf YouTube und anderen Plattformen erschienen Kopien und Ausschnitte. Von dort verbreitete sich das Video zuerst in Australien und nach Tagesanbruch in anderen Zeitzonen weltweit. Innerhalb eines Tages wurde das Video über 300.000 mal aufgerufen. Nach etwa einem halben Jahr hatte sich diese Zahl bei YouTube bereits auf über zweieinhalb Millionen Aufrufe gesteigert. Auf Facebook verbreiteten sich Best-of-Ausschnitte der Rede und auf Twitter wurde das Thema unmittelbar zum „Trending Topic“. Auch in den traditionellen Medien hinterließ das Thema ein internationales Echo (z.B. Farr, Jones, Hudson, & Caldwell, 2012; Huffington Post, 2012; Lester, 2012; Telegraph, 2012).

2.2 *Die Daten*

Im Rahmen eines bestehenden Forschungsprojektes, das zur Erstellung des Australian Twitter News Index (ATNIX) fortwährend sämtliche über Twitter verbreitete und auf einschlägige australische Leitmedien verweisende Links erfasst (Bruns, Harrington, & Highfield, 2013a), wurden auch die Tweets zum

Gillard-Video bei der ABC von Anfang an von uns gesammelt. Dies ermöglicht es uns, beginnend mit den ersten Tweets am 9. Oktober 2012, die Weiterverbreitung des Videos auf Twitter zu verfolgen. Wir berücksichtigen dabei nur die Weiterverbreitung des Videolinks. Weitere Anschlusskommunikation, die von der Verbreitung des Videos angestoßen wurde, aber den Link selbst in ihren Tweets nicht mehr enthält, wurde nicht miteinbezogen. Ein Forschungsansatz, der weniger an der viralen Verbreitung als an der Entwicklung eines *Ad-Hoc*-Diskussionsnetzwerks um das Thema interessiert wäre, müsste solche linkfreien Anschlussweets selbstverständlich miteinbeziehen, könnte dann aber ebenfalls die hier vorgestellten Analyse- und Visualisierungsmethoden nutzen. Dieselben Methoden lassen sich problemlos auch zur Analyse anderer Inhalte verwenden, solange die relevanten Daten verfügbar sind.

Die vom ATNIX-Projekt erfassten Daten sind für die Visualisierung des aktiven Teilnehmernetzwerks (Mikroebene) nach dem ersten der oben beschriebenen Typen ausreichend. Zusätzlich verfügen wir – über ein weiteres Forschungsprojekt – über die Followerdaten für etwa 1,05 Millionen australische TwitternutzerInnen (Stand Frühjahr 2012), sodass wir das passive Teilnehmernetzwerk bzw. die Sichtbarkeit von Links zu dem Gillard-Video in verschiedenen Bereichen der australischen Twittersphäre approximieren können (Bruns, Burgess, & Highfield, 2014a).

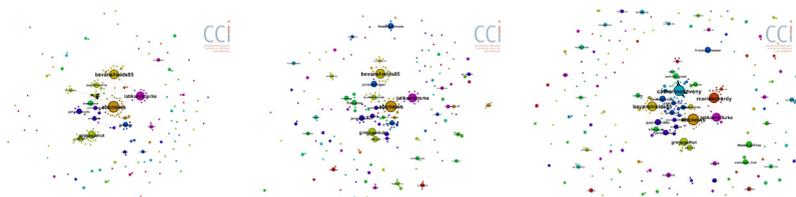
2.3 Netzwerkvisualisierungen und -analysen

Zur Visualisierung dieser Netzwerke benutzen wir den von der Open-Source-Software Gephi angebotenen Algorithmus Force Atlas 2 (Jacomy, Heymann, Venturini, & Bastian, 2011), nachdem wir die nach der oben beschriebenen Extrahierungsmethode erfassten Tweetdaten im Format *Absender, Zeitpunkt, Empfänger* als Netzwerkverbindungen (Edges) eingelesen haben. Gephi ermöglicht die Einbeziehung einer Zeitschiene als dritte bzw. vierte Dimension bei der Visualisierung. Im einfachsten Fall werden dabei bestimmte Anfangs- und Endwerte für den numerisch als Unix-Timestamp (in Sekunden seit dem 1.1.1970 00:00 Uhr) angegebenen Absendezeitpunkt jedes Tweets eingestellt und auf diese Weise Tweets vor bzw. nach dem gewünschten Zeitraum ausgeschlossen. In der Netzwerkvisualisierung stellt sich dies dann so dar, dass je nach eingestelltem Zeitraum bestimmte Verbindungen zwischen den Netzwerkknoten erscheinen oder

verschwinden. Eine graduelle Verschiebung der Anfangs- und/oder Endwerte innerhalb des zu untersuchenden Gesamtzeitraums ermöglicht es nun, die jeweils aktiven Teilnetzwerke zu untersuchen und ihre derzeit führenden TeilnehmerInnen zu identifizieren. Videoaufnahmen der Netzwerkdynamik, wie sie Gephi während dieser Verschiebung darstellt, sind ebenfalls möglich.³

Wir visualisieren dabei zunächst nur das Netzwerk aktiver Diskussionsteilnehmer. Abbildung 1 zeigt die Entwicklung des Verbreitungsnetzwerks für das Gilard-Video der ABC über die ersten Stunden (nach 15 Minuten, sowie nach einer bzw. drei Stunden). Die Netzwerkknotten in der Visualisierung repräsentieren je einen Account, und jede Verbindung zwischen ihnen einen oder mehrere Replies und Retweets.

Abbildung 1: Netzwerk aktiver Teilnehmer nach 15 Minuten, 1 Stunde und 3 Stunden

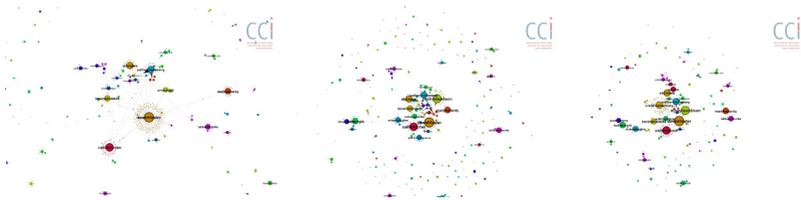


Quelle: Eigene Darstellung

Nach den ersten fünf Stunden erweiterte sich die Verbreitung des Videos außerhalb Australiens, verlangsamte sich jedoch auch – zum Teil, weil nun auch Versionen des Videos auf YouTube sowie auf den Seiten einschlägiger Nachrichtenorgane in anderen Ländern kursierten, sodass das ABC-Video nicht mehr die einzige häufig weitergeleitete Variante war. Somit sind in Abbildung 2 (Netzwerk nach fünf, zehn, und zwanzig Stunden) die Zeitschritte schneller und die Teilnehmerschar größer.

3 Für die aktuelle Fallstudie, siehe <http://youtu.be/4V81s5sAiQM> (0-5 Stunden) sowie <http://youtu.be/-1hAdHO1a08> (5-75 Stunden nach dem ersten Tweet).

Abbildung 2: Netzwerk aktiver Teilnehmer nach 5, 10 und 20 Stunden



Quelle: Eigene Darstellung

Diese Visualisierung veranschaulicht sowohl die Disseminationsdynamik des ABC-Videos, als auch das Netzwerk der AkteurInnen, die das Video verbreiteten. Gephi teilt jedem Knoten eine Größe zu, welche die Anzahl der erhaltenen Retweets und Replies anzeigt. Anhand dieser Darstellung lässt sich leicht ermitteln, wer die wichtigsten AkteurInnen in der Verbreitung des Videos sind und in welchen Beziehungen sie zu einander stehen – wobei wir die Wichtigkeit einzelner TeilnehmerInnen im vorliegenden Fall daran messen, wie oft ihr Tweet weitergeleitet wurde und welche Rolle ihr Account als Knotenpunkt bei der Verbreitung des Videos spielte. Somit können die in diesem Fall einflussreichsten AkteurInnen genauer untersucht und die Dynamik der Retweet- und Diskussionskette analysiert werden.

Um die Verbreitung unseres Gillard-Videos zu erörtern, untersuchten wir die größten Knotenpunkte genauer. Hierzu verwendeten wir die öffentlich einsehbaren Twitterprofile der AkteurInnen und (falls vorhanden) die in den Profilbeschreibungen verlinkten Websites, um Steckbriefe der jeweiligen VerbreiterInnen zu erstellen. Dies führte zu einer Einteilung der wichtigsten AkteurInnen während der ersten fünf Stunden in die Kategorien: *JournalistInnen* (10 NutzerInnen), *FeministInnen* (8), *Privatpersonen* (5), *AktivistInnen* (5), *KomödiantInnen* (3), und *CartoonistInnen* (1). In den folgenden 70 Stunden hingegen verschob sich das Profil der führenden TeilnehmerInnen: *FeministInnen* (6), *JournalistInnen* (4), *MusikerInnen* (3), *KomödiantInnen* (2), *Privatpersonen* und *PolitikerInnen* (jeweils 1).

Des Weiteren führten wir anhand der Twitterprofile eine Zuordnung nach geografischen Angaben durch. So ermittelten wir, dass die wichtigsten AkteurInnen in der Verbreitung des Videos in den ersten fünf Stunden überwiegend *AustralierInnen* waren (29 NutzerInnen), sowie jeweils ein/e *NeuseeländerIn* und *EngländerIn*.

In den folgenden 70 Stunden ergab sich dagegen eine weitaus internationalere Zusammensetzung: *EngländerInnen* und *AustralierInnen* (jeweils 7), *US-AmerikanerInnen* (2), und ein/e *KanadierIn* fanden sich nun unter den einflussreichsten VerbreiterInnen des Videos. Diese geografische Verschiebung ist mit einiger Sicherheit auch auf verschiedene Zeitzonen zurückzuführen: Die führende Stellung australischer TeilnehmerInnen in den ersten Stunden hängt damit zusammen, dass die Rede selbst am frühen Nachmittag Ortszeit gehalten wurde – und somit einige Stunden vor Tagesbeginn in Europa. Retweets aus Großbritannien begannen in großem Maße erst nach dem Arbeitsbeginn einiger auf Twitter bekannter und einflussreicher englischer JournalistInnen. Die längerfristige Verschiebung weg von JournalistInnen und hin zu anderen Personen des öffentlichen Lebens (wie etwa der Komödiantin Julia Morris oder der Sängerin Missy Higgins) hingegen kann als Zeichen der Wandlung des Themas von einem primär politischen Ereignis hin zu einer gesellschaftlichen *Cause Célèbre* gesehen werden. Eine solche Wandlung ist auch im nationalen und internationalen Medienecho auf Gillards Rede erkennbar, denn Wochen und Monaten nach der Rede zeigte sich eine mediale Aufmerksamkeit weit über konventionelle politische Medien hinaus.

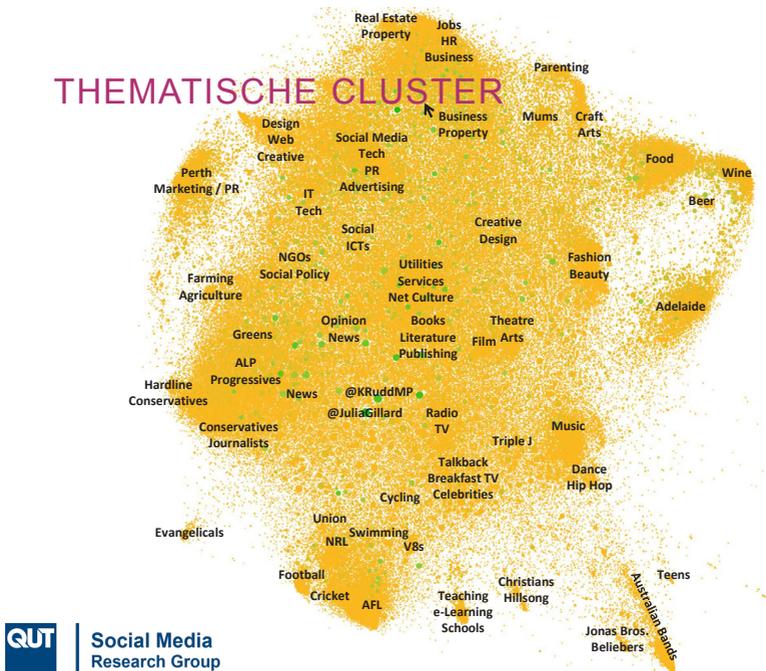
Im zweiten Teil unserer Analyse bilden wir nun das aktive Tweetnetzwerk zum Gillard-Video auf ein darunterliegendes Netzwerk bestehender Followerverbindungen in der australischen Twittersphäre ab. Zur Vermeidung langwieriger Neuerfassungen dieses Netzwerks benutzen wir hierzu bereits zuvor erfasste Netzwerkdaten für 1,05 Millionen australische TwitternutzerInnen. Dieser Ansatz ermöglicht es uns, die hier beschriebene Methode zu demonstrieren, zieht allerdings die folgenden Einschränkungen nach sich, welche nur durch eine Neuerfassung des aktuellen Followernetzwerks umgangen werden könnten:

- Das bis zum Frühjahr 2012 erfasste Followernetzwerk stellt nur ein Teilnetzwerk der bis September 2013 von uns erfassten gesamten australischen Twittersphäre von 2,8 Millionen Accounts dar, bietet jedoch eine für unsere Zwecke ausreichende Approximation der Struktur des Gesamtnetzwerks.
- Neue NutzerInnen und Followerverbindungen, die zwischen der Netzwerkerfassung bis zum Frühjahr 2012 und der Rede von Gillard im Oktober 2012 entstanden, sind im Followernetzwerk nicht enthalten.

- Durch die Beschränkung auf die australische Twittersphäre werden TwitternutzerInnen außerhalb Australiens von der Analyse ausgeschlossen, die den Link zum Video ebenfalls weiterleiteten.

Diese Einschränkungen sind im vorliegenden Fall jedoch vertretbar, da es uns hier um die Illustration der Methodik und nicht um eine detaillierte Analyse der Videoverbreitung selbst geht.

Abbildung 3: Netzwerkdarstellung der Followerverbindungen in der australischen Twittersphäre, Forschungsstand Frühjahr 2012



Quelle: Eigene Darstellung; Dargestellt sind die 120,000 am stärksten vernetzten australischen Twitteraccounts. Jeder Knotenpunkt repräsentiert einen Account, und Knoten sind mit Hilfe des Force-Atlas-2-Algorithmus so zueinander positioniert, dass sich besonders stark untereinander vernetzte Knotengruppen zu Netzwerkclustern zusammenschließen.

Die der Visualisierung zugrundeliegende Darstellung der australischen Twittersphäre in Abbildung 3 zeigt die Existenz einiger intern besonders eng verbundener Netzwerkcluster, die – wie eine Untersuchung der jeweils führenden Mitglieder dieser Cluster zeigt – zumeist thematisch bedingt sind. So existieren Cluster um Themen wie Politik, Wirtschaft, Sport, die teilweise wiederum nach verschiedenen Spielarten in einige Teilcluster getrennt werden können (zur weiteren Erläuterung siehe Bruns et al., 2014a).

Soweit es sich bei den TeilnehmerInnen an der Weiterleitung des Gillard-Videos als Link in Form von Tweets, Replies und Retweets um australische NutzerInnen handelt, die von unserer vorigen Visualisierung der australischen Twittersphäre erfasst worden sind, lassen sich diese TeilnehmerInnen nun auf

Abbildung 4: Verbreitung des ABC-Videos durch die australische Twittersphäre nach 15 Minuten, 30 Minuten, 1 Stunde und 5 Stunden



Quelle: Eigene Darstellung

dieser Gesamtkarte verorten. Darüber hinaus bietet der Datensatz zur Videoweiterleitung in Verbindung mit dieser Karte die Möglichkeit, den Verlauf der Verbreitung des Videos innerhalb des australischen Twitternetzwerks zu verfolgen. Falls ein uns bekannter australischer Twitteraccount den Tweet eines anderen uns bekannten australischen Twitteraccounts weiterleitet (also den ursprünglichen Sender im weiterleitenden Tweet erwähnt – zum Beispiel in der Form „RT @user ...“ – bzw. die von Twitter selbst bereitgestellte Retweet-Funktion zur Weiterleitung benutzt), so sind sowohl der/die ursprüngliche SenderIn als auch der/die weiterleitende NutzerIn auf der Gesamtkarte zu finden. Unabhängig von der durch die Struktur des Followernetzwerks vorgegebenen Netzwerkstruktur ist es daher möglich, die durch solche weiterleitenden Tweets zumindest temporär geknüpften Verbindungen zwischen einzelnen TeilnehmerInnen in die bestehende Karte einzutragen – in Abbildung 4 sind diese in roter Farbe gekennzeichnet.

Besonders am Anfang der Weiterverbreitung sticht dabei ein offizieller Account der ABC, @abcnews, heraus, dessen Tweet mit einem Link zum Video von einigen wenigen Followern früh weitergeleitet wurde. Jedoch erscheinen auch hier schon einige andere WeiterleiterInnen, die den Link unabhängig vom @abcnews-Account gefunden und in einem Tweet weiterverbreitet haben. Innerhalb der ersten 30 Minuten erfolgen weitere Retweets besonders im stark etablierten Cluster, der primär an Nachrichten und Politik interessierten australischen TwitternutzerInnen (auf der Darstellung zentral links), während andere große Teile der australischen Twittersphäre zunächst unbeteiligt blieben. Erst mit Hilfe eines weitgehend feministisch orientierten Clusters von TwitternutzerInnen (auf der Karte zentral gelegen) sowie von anderen TeilnehmerInnen in verschiedenen Netzwerkclustern ergab sich nach einigen Stunden eine sehr viel weitere Sichtbarkeit des Videos im Gesamtnetzwerk. Hierbei muss natürlich zusätzlich noch berücksichtigt werden, dass um jede/n der in Abbildung 4 markierten TeilnehmerInnen herum ein Egonetzwerk von Follower-Accounts existiert, welche somit alle den Link gesehen haben könnten, so sie zu diesem Zeitraum ihren Twitterfeed verfolgt haben.

3 Diskussion

Die Verbindung dieser quantitativen und qualitativen Methoden zur Untersuchung einer Retweet- und Reply-Kette, die dynamische Visualisierung

durch Gephi, durch die die wichtigsten AkteurInnen identifiziert werden können, die Untersuchung und Kategorisierung dieser AkteurInnen und die Lokalisierung der AkteurInnen auf einer bestehenden thematischen Karte der australischen Twittersphäre ermöglicht die Analyse digitaler Kommunikationsabläufe und die Thematisierung aktueller kultureller, sozialer und politischer Kontexte.

Durch die dynamische Netzwerkvisualisierung der Retweet-Kette und die in Folge durchgeführte Erstellung von Steckbriefen der wichtigsten AkteurInnen lässt sich nachverfolgen, wie sich die Gillard-Rede von einem lokalen Ereignis zu einem internationalem Medienereignis entwickelte. Erste Aufmerksamkeit bekam das Video zunächst hauptsächlich von australischen JournalistInnen und „Politjunkies“ (Coleman, 2003). Dann verbreiteten FeministInnen und KomödiantInnen das Video weiter im australischen Netzwerk. Schließlich fanden internationale JournalistInnen, AktivistInnen, und Medienpersönlichkeiten Interesse an der Geschichte, und disseminierten das Video (sowie seine Entsprechungen auf YouTube und anderen Plattformen) auf weltweiter Basis. Die virale Verbreitung allerdings – und auch dies soll erwähnt sein – ist nicht alleine ein Verdienst dieser besonders sichtbaren NutzerInnen, sondern insbesondere auch der Vielzahl „normaler“ TeilnehmerInnen, deren Retweets en masse überhaupt erst für die große Sichtbarkeit dieser führenden NutzerInnen gesorgt haben.

Zeitonenbedingt erreichte die Geschichte erst nach Ablauf der ersten vier bis fünf Stunden erste JournalistInnen in England und verbreitete sich dann über die nächsten Tage auch hier wieder hauptsächlich durch FeministInnen und „Celebrities“ – KomödiantInnen, MusikerInnen, PolitikerInnen. Die Vernetzung der AkteurInnen und das Interesse an dem Thema lassen sich anhand der historischen Verbindungen zwischen Australien und England und der großen australischen Expat-Community in England erklären. Des Weiteren teilen die beiden Commonwealth-Länder ähnliche politische Systeme und eine Medienvernetzung – der *Guardian Australia* ist zum Beispiel eine Dependance des britischen *Guardian*.

Die Weiterleitung des Videos im Fall Gillard folgte einer „Long Tail“-zu-„Fat Head“-Dynamik: Anfangs wurde das Video von vielen einzelnen Verlinkungen durch diverse TeilnehmerInnen verbreitet, die nicht immer in direkter Verbindung miteinander stehen. Später werden fast nur noch die Tweets der bekanntesten und einflussreichsten AkteurInnen geretweetet. Dies demonstriert das Phänomen des „preferential attachment“ (Barabási & Albert, 1999), in dem die Tweets bereits sichtbarer TeilnehmerInnen bevorzugt weitergeleitet werden

und diese dadurch ihre Sichtbarkeit weiter erhöhen können. Somit ist das Netzwerk zuerst relativ flach und später zunehmend stark strukturiert, wobei die Tweets der größten Opinion Leaders am langlebigsten sind. Einzelne AkteurInnen wurden zudem zu wichtigen Knotenpunkten in der Überleitung des Themas in neue Netzwerke, zum Beispiel die britische Guardian-Journalistin Dawn Foster, die gegen Ende der ersten intensiven Fünf-Stunden-Phase auftaucht und die Nachricht in das englische Twitternetzwerk von JournalistInnen und FeministInnen einführt.

Die Dynamik des Weiterleitungsnetzwerkes zeigt, wie sich das Thema von einem von JournalistInnen verbreiteten „breaking news“-Phänomen zu einem längerfristiges Medien- und Politikthema entwickelte. Das internationale Medienecho, das dem Twitter-Lauffeuer folgte, entwickelt das Thema von einem politischen Spektakel in eine anhaltende, über Australien hinausreichende Diskussion über Sexismus und Misogynie in der Politik und in anderen Bereichen des öffentlichen Lebens.

4 Abschluss

Die massive Verbreitung bestimmter Inhalte, wie etwa des Gillard-Videos, durch neue digitale Medien bietet uns eine Gelegenheit zu analysieren, wie sich Informationen im globalen Netz verbreiten, und zu erörtern, was uns dies zu modernen Kommunikationsabläufen und -inhalten zu sagen hat. Neue Forschungsansätze wie die von uns oben aufgeführten ermöglichen es, diese digitalen Kommunikationsabläufe zu analysieren und somit aktuelle kulturelle, soziale und politische Kontexte zu thematisieren. Dabei ergeben sich neue Einsichten in die Rolle einzelner KommunikationsteilnehmerInnen, zum Beispiel als „opinion leaders“ oder als Brücken zwischen diversen thematischen Communities im Netz, sowie in die Dynamik solcher Verbreitungsmechanismen und ihre Abhängigkeit von bereits existierenden Netzwerkstrukturen oder den technischen Grundlagen einzelner Netzwerkplattformen.

Besonders die Dynamik der Informationsverbreitung im Netzwerk bleibt dabei noch weitgehend unerforscht. Die Verbindung räumlicher und zeitlicher Analysen – und somit die Erschließung der Zeit als einer zusätzlichen Dimension bei der Netzwerkanalyse und -visualisierung – steckt in der aktuellen Social-Media-For-

schung, insbesondere aus der Perspektive der Medien- und Kommunikationswissenschaft, noch in den Kinderschuhen. Ein methodischer Ansatz, wie wir ihn hier vorgestellt und demonstriert haben, benötigt jedoch nur einen geringen technischen Mehraufwand und ermöglicht im Gegenzug eine ganze Reihe neuer Analysen.

Prof. Dr. Axel Bruns ist Professor am ARC Centre of Excellence for Creative Industries and Innovation, Queensland University of Technology, Brisbane, Australien

Dr. Theresa Sauter ist wissenschaftliche Mitarbeiterin am ARC Centre of Excellence for Creative Industries and Innovation, Queensland University of Technology, Brisbane, Australien

Quellenverzeichnis

- Barabási, A.-L., & Albert, R. (1999). Emergence of Scaling in Random Networks. *Science*, 286(5439), 509-512. doi: 10.1126/science.286.5439.509
- Bruns, A. (2012). How Long Is a Tweet? Mapping Dynamic Conversation Networks on Twitter Using Gawk and Gephi. *Information, Communication & Society*, 15(9), 1323-1351. doi: 10.1080/17512786.2012.663610
- Bruns, A., & Burgess, J. (2011). The use of Twitter hashtags in the formation of ad hoc publics. European Consortium for Political Research conference, Reykjavík, 25.-27. August 2011. Abgerufen von <http://eprints.qut.edu.au/46515/>
- Bruns, A., & Burgess, J. (2014). Crisis Communication in Natural Disasters: The Queensland Floods and Christchurch Earthquakes. In K. Weller, A. Bruns, J. Burgess, M. Mahrt, & C. Puschmann (Hrsg.), *Twitter and Society* (S. 373-384). New York, NY: Peter Lang.
- Bruns, A., Burgess, J., & Highfield, T. (2014a). A 'Big Data' Approach to Mapping the Australian Twittersphere. In P. L. Arthur & K. Bode (Hrsg.), *Advancing Digital Humanities* (S. 113-129). Basingstoke: Palgrave Macmillan.
- Bruns, A., Harrington, S., & Highfield, T. (2013a). Sharing the News: Dissemination of Links to Australian News Sites on Twitter. In J. Gordon, P. Rowinski, & G. Stewart (Hrsg.), *Br(e)aking the News* (S. 181-210). New York: Peter Lang.
- Bruns A., & Moe, H. (2014). Structural Layers of Communication on Twitter. In K. Weller, A. Bruns, J. Burgess, M. Mahrt, & C. Puschmann (Hrsg.), *Twitter and Society* (S. 15-28). New York, NY: Peter Lang.

- Bruns, A., & Stieglitz, S. (2014). Metrics for Understanding Communication on Twitter. In K. Weller, A. Bruns, J. Burgess, M. Mahrt, & C. Puschmann (Hrsg.), *Twitter and Society* (S. 69-82). New York, NY: Peter Lang.
- Bruns, A., Woodford, D., & Sackowsky, T. (2014b). Towards a Methodology for Examining Twitter Follower Accession. *First Monday*, 19(4). doi: 10.5210/fm.v19i4.5211
- Burns, A., & Eltham, B. (2009). Twitter free Iran: An evaluation of Twitter's role in public diplomacy and information operations in Iran's 2009 election crisis. *Record of the Communications Policy & Research Forum 2009* (S. 298-310). Abgerufen von http://www.networkinsight.org/publications/record_of_the_2009_cprf.html/group/16
- Coleman, S. (2003). A Tale of Two Houses: The House of Commons, the Big Brother House and the People at Home. *Parliamentary Affairs*, 56(4), 733-758.
- Eckert, S. (2013). #aufschrei/#outcry: Solidarity and Aggression in a Twitter Debate over Sexual Harassment. The Association of Internet Researchers conference, Denver, 23.-26. Oktober 2013.
- Farr, M., Jones, G., Hudson, P., & Caldwell, A. (2012). Prime Minister Julia Gillard Blasts Tony Abbott in Extraordinary Scenes in Parliament. *The Australian*, 9. Oktober 2012. Abgerufen von <http://www.theaustralian.com.au/news/deputy-opposition-leader-julie-bishop-says-peter-slipper-cannot-remain-as-speaker/story-e6frg6n6-1226491792937>
- Gaffney, D., & Puschmann, C. (2014). Data Collection on Twitter. In K. Weller, A. Bruns, J. Burgess, M. Mahrt, & C. Puschmann (Hrsg.), *Twitter and Society* (S. 55-68). New York, NY: Peter Lang.
- Hermida, A. (2014). Twitter as an Ambient News Network. In K. Weller, A. Bruns, J. Burgess, M. Mahrt, & C. Puschmann (Hrsg.), *Twitter and Society* (S. 359-372). New York, NY: Peter Lang.
- Hermida, A., Lewis, S. C., & Zamith, R. (2014). Sourcing the Arab Spring: A case study of Andy Carvin's sources on Twitter during the Tunisian and Egyptian revolutions. *Journal of Computer-Mediated Communication* 19(3), 479-499. doi: 10.1111/jcc4.12074
- Huffington Post (9. Oktober 2012). Julia Gillard, Australia Prime Minister, Launches Blistering Attack on Sexism during Parliament Speech. *Huffington Post*. Abgerufen von http://www.huffingtonpost.com/2012/10/09/julia-gillard-sexism-speech_n_1952231.html

- Jacomy, M., Heymann, S., Venturini, T., & Bastian, M. (2011). ForceAtlas2, a Graph Layout Algorithm for Handy Network Visualization. Abgerufen von http://www.medialab.sciences-po.fr/publications/Jacomy_Heymann_Venturini-Force_Atlas2.pdf
- Kwak, H., Lee, C., Park, H., & Moon, S. (2010). What is Twitter, a social network or a news media? *19th International Conference on World Wide Web - WWW '10* (S. 591). Raleigh, North Carolina, USA.
- Lester, A. (9. Oktober 2012). Ladylike: Julia Gillard's Misogyny Speech. The New Yorker. Abgerufen von <http://www.newyorker.com/online/blogs/newsdesk/2012/10/julia-gillards-misogyny-speech.html>
- Lotan, G., Graeff, E., Ananny, M., Gaffney, D., Pearce, I., & boyd, d. (2011). The revolutions were tweeted: Information flows during the 2011 Tunisian and Egyptian Revolutions. *International Journal of Communication*, 5, 1375-1405. Abgerufen von <http://ijoc.org/ojs/index.php/ijoc/article/view/1246>
- Maireder, A. (2013). Framing von Nachrichten in Social Media. In O. Jandura, A. Fahr, & H.-B. Brosius (Hrsg.), *Theorieanpassungen in der digitalen Medienwelt* (S. 191-206). Baden Baden: Nomos.
- Maireder, A., & Ausserhofer, J. (2014). Political Discourses on Twitter: Networking Topics, Objects, and People. In K. Weller, A. Bruns, J. Burgess, M. Mahrt, & C. Puschmann (Hrsg.), *Twitter and Society* (S. 305-318). New York, NY: Peter Lang.
- Maireder, A., & Schlögl, S. (2014). 24 Hours of an #outcry: The Networked Publics of a Socio-Political Debate. *European Journal of Communication*, 20(6), 687-702. doi: 10.1177/0267323114545710
- Matsa, K. E., & Mitchell, A. (26. März 2014). 8 Key Takeaways about Social Media and News. Pew Research Journalism Project. Abgerufen von <http://www.journalism.org/2014/03/26/8-key-takeaways-about-social-media-and-news/>
- Meeder, B., Karrer, B., Sayedi, A., Ravi, R., Borgs, C., & Chayes, J. (2011). We Know Who You Followed Last Summer: Inferring Social Link Creation Times in Twitter. *Proceedings of the 20th International Conference on World Wide Web - WWW '11* (S. 517-526). doi: 10.1145/1963405.1963479
- Mendoza, M., Poblete, B., & Castillo, C. (2010). Twitter under crisis: Can we trust what we RT? *Proceedings of the First Workshop on Social Media Analytics - SOMA '10* (S. 71-79). Abgerufen von http://snap.stanford.edu/soma2010/papers/soma2010_11.pdf

- Nuernbergk, C. (2013). *Anschlusskommunikation in der Netzwerköffentlichkeit*. Baden-Baden: Nomos.
- Nuernbergk, C., & Neubarth, J. (2013). Networked Politics on Twitter: Social Media Use by German Politicians during the Federal Election 2013. *14th Annual Conference of the Association of Internet Researchers - IR 14: Resistance and Appropriation*, Denver (CO), 23.-26. Oktober 2015.
- Palen, L., Starbird, K., Vieweg, S., & Hughes, A. (2010). Twitter-based information distribution during the 2009 Red River Valley flood threat. *Bulletin of the American Society for Information Science and Technology*, 36(5), 13-17. doi: 10.1002/bult.2010.1720360505
- Paßmann, J., Boeschoten, T., & Schäfer, M. T. (2014). The Gift of the Gab: Retweet Cartels and Gift Economics on Twitter. In K. Weller, A. Bruns, J. Burgess, M. Mahrt, & C. Puschmann (Hrsg.), *Twitter and Society* (S. 331-344). New York, NY: Peter Lang.
- The Telegraph (9. Oktober 2012). Australian PM Julia Gillard's Furious Attack on Tony Abbott. *The Telegraph*. Abgerufen von <http://www.telegraph.co.uk/news/worldnews/australiaandthepacific/australia/9597488/Australian-PM-Julia-Gillards-furious-attack-on-Tony-Abbott.html>
- Webberley, W., Allen, S., & Whitaker, R. (2011). Retweeting: A study of message-forwarding in Twitter. *IEEE, 2011 Workshop on Mobile and Online Social Networks* (S. 13-18). Mailand, 6.-8. September 2011. doi: 10.1109/MOSN.2011.6060787
- Woong Yun, G., & Park, S.-Y. (2011). Selective Posting: Willingness to post a message online. *Journal of Computer-Mediated Communication*, 16(2), 201-227. doi: 10.1111/j.1083-6101.2010.01533.x

Empfohlene Zitierung: Grimm, J., & Grill, C. (2015). Politische Online-Pluriversen. Aggregatanalyse der Twitter-Kommunikation zur Salzburger Landtagswahl 2013. In A. Maireder, J. Ausserhofer, C. Schumann, & M. Taddicken (Hrsg.), *Digitale Methoden in der Kommunikationswissenschaft* (S. 163-195). doi: 10.17174/dcr.v2.8

Zusammenfassung: Methodische Herausforderungen der Twitterforschung betreffen unter anderem den Datenumfang, die Subjektivität der Tweet-Auswahl, die Datenreduktion und das Einbettungsproblem bezüglich ausgegrenzter Kommunikationszusammenhänge. In diesem Aufsatz wird ein Verfahren der *Twitter-Aggregatanalyse* (TAA) vorgeschlagen, das die Willkür des Vorgehens reduziert, auf kommunikative Einbettungen fokussiert und an den Agenda-Setting-Approach bei der Analyse politischer Online- und Offline-Kommunikation anschlussfähig ist. Die neu entwickelte TAA-Methode kombiniert automatisierte Selektion mit manueller Reduktion und konventioneller Inhaltsanalyse und generiert auf dieser Grundlage Tweet-Aggregate, so genannte *Twitter-Universen*. Durch die Variation der Selektoren und die standardisierten Transformationen der Twitter-Universen bietet die TAA vielfältige Vergleichsmöglichkeiten im Hinblick auf die Themenprioritäten von Politikern, Journalisten und Laien, die sich im Sinne wechselseitiger Beeinflussung interpretieren lassen. Die Leistungsfähigkeit der TAA und deren Interpretationsrelevanz werden am Beispiel der Salzburger Landtagswahlen 2013 demonstriert. Besonderes Augenmerk gilt der Einbettung der Tweets in Gesprächszusammenhänge sowie der Rolle von Tweets ohne Gesprächsresonanz. Des Weiteren werden Themen und Themenhandlungsweisen mit und ohne Politikerbeteiligung ermittelt, die auf Defizite in der Qualität politischer Online-Kommunikation verweisen.

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

Jürgen Grimm & Christiane Grill

Politische Online-Pluriversen

Aggregatanalyse der Twitter-Kommunikation
zur Salzburger Landtagswahl 2013

1 Einleitung

Auf das enorme Wachstum der Online-Kommunikation in *Social Media* und darin insbesondere der *Microblogging*-Dienste reagieren Parteien und Spitzenkandidaten in Wahlkampfzeiten immer stärker mit planmäßiger politischer PR im Netz (Schmitt-Beck & Mackenrodt, 2010; Thimm, Dang-Anh & Einspänner, 2011; Thimm, Einspänner & Dang-Anh, 2012). Trotzdem gibt es bislang kaum generalisierbare Erkenntnisse zur besonderen Funktions- und Wirkungsweise von Online-Medien, die sich für ein strategisches Agenda Building nutzen ließen. Insbesondere die Themenabhängigkeit der Kommunikationsdynamiken innerhalb versus außerhalb des Netzes wird überwiegend am Erfolg einzelner berühmter Persönlichkeiten (Plehwe, 2008; Moorstedt, 2008) oder einzelnen Themenkarrieren festgemacht (siehe die Metaanalyse bei Weller, 2014). Im Folgenden versuchen wir am Beispiel der Salzburger Landtagswahlen 2013 Möglichkeiten der Aggregatanalyse aufzuzeigen, mit deren Hilfe Twitter-Kommunikation von einem stärker systematischen Standpunkt aus betrachtet und mit Offline-Kommunikation verglichen werden kann.

Ziel des methodischen Konzepts ist es, abgegrenzte Tweet-Sammlungen zu bilden, die durch semantische und zeitliche Selektoren sowie Akteursspezifizierungen definiert und mit Hilfe einer Inhaltsanalyse auf Themenprioritäten und Themenbehandlungsweisen untersucht werden können. Die jeweiligen thematischen Besonderheiten der Tweet-Aggregate lassen sich dann zur Funktionsbestimmung einzelner Teilbereiche der Twitter-Kommunikation wie auch zur Relationierung und Abgrenzung gegenüber Inhalt und Art der Thematisierungen in der politischen Offline-Kommunikation nutzen.

Methodisch geht es uns *erstens* darum, Twitter-Analysen operativ so zu strukturieren, dass die generierten Daten einer kontrollierten theoretischen Reflexion zugänglich werden. *Zweitens* ist die Vorgehensweise darauf ausgerichtet, die Anschlussfähigkeit der Twitterforschung an die traditionelle Massenkommunikationsforschung zu erhöhen.

Der Agenda-Setting-Approach eignet sich für dieses Vorhaben insofern, als er den Zusammenhang von interpersonaler Kommunikation und auf Massenmedien basierter Kommunikation in den Mittelpunkt der Analyse von Medienwirkungen stellt und dabei alternativ einzelne Themenkarrieren oder ganze Themenhierarchien untersucht. McCombs & Shaw (1972) haben bereits in ihrer klassischen Studie das Aggregatverfahren bei der Untersuchung von Agenda-Setting-Effekten ausgearbeitet, indem sie die massenmedial verbreiteten Themen den Themenprioritäten des Elektorats gegenüberstellten (McCombs, 2004; Schenk, 2007). Im Unterschied zur strikten Bindung an individuell rezipierte Medieninhalte erfasst das Aggregatverfahren nicht nur direkte Rezeptionseffekte, sondern auch indirekte Medienwirkungen, die zum Beispiel durch Meinungsführer vermittelt werden (Grimm et al., 2008). Im Aggregatansatz der Agenda-Setting-Forschung werden Themenhierarchien im Bereich der Medienkommunikation sowie die Gesprächsaktivitäten und die Wichtigkeitszuschreibungen im Publikum ermittelt und auf der Zeitachse korrelativ verknüpft. In dieses Vorgehen fügen sich zeitlich geordnete Themenhierarchien in diversen Segmenten der Online-Kommunikation nahtlos ein.

Dazu ist es allerdings erforderlich, über den üblichen akteurszentrierten Ansatz innerhalb der Twitterforschung (Ausserhofer & Maireder, 2013; Cheong & Cheong, 2011; Dang, Einspänner & Thimm, 2013; Maireder & Schwarzenegger, 2012; mit Ausnahme von Bruns & Stieglitz, 2012; Highfield, 2012) hinauszugehen und mit einem themenzentrierten Konzept zu verbinden.

In diesem Aufsatz entwerfen wir ein Modell der *Twitter-Aggregatanalyse* (TAA), das sich auf Themen und Themenbehandlungsweisen in genau definierten Teilbereichen der Online-Kommunikation konzentriert. Im zweiten Teil wenden wir die TAA dann auf die Salzburger Landtagswahlen 2013 an, um die Leistungsfähigkeit des Verfahrens zu demonstrieren. Bevor wir aber zur Modellkonstruktion kommen, skizzieren wir zunächst jene methodischen Probleme, die durch die TAA gelöst bzw. einer Lösung näher gebracht werden sollen.

2 Methodische Ausgangslage: Problemdefinition

Die Analyse von Twitter-Daten stellt die Forschung vor besondere methodische Probleme, die eine Systematisierung erschweren und bislang ein Hindernis für die Verknüpfung von Twitterforschung mit klassischen kommunikationswissenschaftlichen Ansätzen darstellen. Die Herausforderungen betreffen vor allem die Quantität der Textmenge, die Subjektivität des Vorgehens und Schwierigkeiten der Zusammenhangsanalyse, die aus der Zerlegung in Codiereinheiten resultieren. Bei Twitter-Analysen werden methodische Probleme der Inhaltsanalyse in einer Weise verschärft, die mit den etablierten Verfahren alleine nicht gelöst werden können.

1. *Umfangsproblem:* Generell erfährt Twitter-Kommunikation ein rasantes Wachstum, das mit herkömmlichen Erfassungsmethoden nicht zu bewältigen ist. Es stellt sich die Frage, inwieweit automatisierte Verfahren herangezogen werden können, ohne den forschungsökonomischen Gewinn durch Validitätsverluste zu bezahlen.
2. *Subjektivitätsproblem:* Die Auswahl der zu analysierenden Tweets ist von Suchkriterien abhängig, für die bislang keine Modelle zur Verfügung stehen. Twitter-Samples weisen je nach Konstruktion eine hohe Variabilität und Willkürlichkeit auf.
3. *Reduktionsproblem:* In den gesammelten Tweets finden sich häufig Beiträge, die zwar formale Suchkriterien aus einer Wortliste (z. B. „Salzburg“) erfüllen, aber inhaltlich nicht mit den Untersuchungszielen einer Wahlkampf-

analyse (z. B. touristische Hinweise auf Salzburg) übereinstimmen und daher maschinell und/oder konventionell identifiziert werden müssen.

4. *Einbettungsproblem*: Generell gehen bei der Sammlung von Einzeltweets Kommunikationszusammenhänge verloren, sei es, dass diese außerhalb des Erfassungsbereichs von Suchbegriffen liegen, oder dass sie innerhalb des Erfassungsbereichs Vernetzungen mit anderen Tweets vernachlässigen. Da die Twitter-Kommunikation durch komplexe Vernetzungen gekennzeichnet ist, müssen Strategien entwickelt werden, um den Atomisierungseffekt von Tweet-Sammlungen zu kompensieren.

Einzelfallanalysen zu bestimmten Themen werden dem Umfangs- und Subjektivitätsproblem nur unzureichend gerecht, da die Auswahl relativ willkürlich und extrem selektiv erfolgt. Auch dann wenn die Vernetzung via Hashtag und/oder Retweet bei monothematischen Tweet-Ketten berücksichtigt wird (Ausserhofer & Maireder, 2013), bleiben die Vernetzungsbefunde auf die gerade betrachtete Kette beschränkt. Der explorative Wert von Einzelfallanalysen ist unbestritten, zur Hypothesenprüfung in Verbindung mit quantitativen Methoden der Medienwirkungsforschung eignen sie sich aber kaum. Hashtags und Retweets bieten immerhin die Möglichkeit zu einer quantifizierenden Analyse (Bruns & Stieglitz, 2012), die Aussagen über den Grad der Schlagwort-Vernetzung und der Resonanz in Einzeltweet-Sammlungen und einzelnen Tweet-Ketten gestattet. Allerdings werden dabei zwei wesentliche Vernetzungsformen nicht erfasst:

Erstens sind Tweets oftmals in Gesprächszusammenhänge eingebunden, die den semantischen Raum der gerade betrachteten einzelnen Mitteilung diskursiv erweitern. In manchen Fällen ist die Kenntnis der Gesprächszusammenhänge erforderlich, um die Bedeutung des ursprünglichen Tweets verstehen zu können (z. B. die Bedeutung des Tweets „Dieser Hype ist unerträglich“. Welcher ist gemeint?). Die Gesprächsaktivität kann aber auch die ursprüngliche Bedeutung erweitern und ganz andere Themen mit ins Spiel bringen.

Zweitens sind Tweets intertextuell vernetzt, auch und gerade in den Bereich der Offline-Kommunikation hinein. Wenn ein Tweet einen gedruckten Zeitungsartikel zitiert, so wird ein Teil der Offline-Kommunikation in den Online-Bereich „importiert“. Ein Verweisartikel, der „nur“ online präsentiert wurde, generiert

ebenfalls einen semantischen Raum der Intertextualität, da er von einem Medium ins Netz gestellt wurde, das selbst auch eine Printausgabe produziert.

Die Frage stellt sich, wie dem Gesprächscharakter der Tweets in Aggregatanalysen Rechnung getragen und der Übergang zur Offline-Kommunikation einbezogen werden kann. Von der Berücksichtigung interner und externer Vernetzungen hängt ab, ob und inwieweit die Befunde zu Online-Aggregaten mit den Befunden aus Offline-Analysen (z. B. der gedruckten Zeitungsausgaben oder der Nachrichtensendungen im Fernsehen) im Sinne einer erweiterten Agenda-Setting-Forschung kompatibel sind. Das Einbettungsproblem wird in der TAA durch Ausschöpfung der Reply-Funktion innerhalb des Twitter-Tools bzw. durch die Verlinkung mit Texten der von Printmedien und/oder ihren Online-Nachrichtendiensten gelöst, die jeweils eigene Untereinheiten der Twitter-Analyse bilden.

Schließlich wird zu zeigen sein, inwieweit aggregatanalytische Verfahren auch einen methodischen Fortschritt bezüglich des Umfangs-, Subjektivitäts- und Reduktionsproblems ermöglichen.

3 Forschungsleitende Annahmen und Ziele

Unsere Strategie zur Integration der Twitterforschung in den Agenda-Setting-Ansatz besteht darin, transparente und klar definierte Tweet-Aggregate – sogenannte „Twitter-Universen“ – zu bilden und diese mit Offline-Aggregaten zu vergleichen. Die Aggregate können durch zeitliche, sprachliche und/oder semantische Indikatoren definiert werden und sind nicht zwangsläufig an einzelne Akteure bzw. Akteursgruppen gebunden. Im Unterschied zu singulären Tweet-Ketten oder rein personenbezogenen Tweet-Sammlungen wird dabei mit Hilfe von Suchbegriffen ein Analyseraum definiert, der mit anderen Analyseräumen (z. B. der Berichterstattung nationaler Offline-Medien wie Zeitung und Fernsehen) auf der Zeitachse in einen kausalen Zusammenhang gebracht werden kann. „Twitter-Universen“ sind soziale Konstruktionen in Abhängigkeit von den verwendeten Konstruktionsprinzipien. Einen Vorteil sehen wir in der Flexibilität des Vorgehens, das die Konstruktion einer Vielzahl kriteriengebundener Universen mit Hilfe von Wortlisten oder kontextsensitiver Auswahlkriterien gestattet, vorausgesetzt ein Ableiten in Beliebigkeit wird vermieden und die Intersubjektivität bleibt gewahrt.

Minimale Voraussetzungen hierfür sind:

- Die Tweet-Auswahl erfolgt transparent anhand definierter Selektoren (z. B. Hashtags oder Schlüsselbegriffe),
- weitere analytische Transformationen der Daten (wie der Ausschluss irrelevanter Tweets oder die Einbeziehung von Tweets, die mit der ursprünglichen Tweet-Sammlung gesprächsvernetzt sind) erfolgt logisch stringent und
- die Interpretation wird an die verwendeten Konstruktionsprinzipien rückgebunden (inklusive der Reflexion auf die Begrenztheit des damit verbundenen Aussageanspruchs).

Unter diesen Bedingungen können durch Variation der Selektoren und Konstruktionsvorschriften Kaskaden von *Twitter-Universen* gebildet und die Veränderung von Themenhierarchien, Themenbehandlungsweisen und Verknüpfungstechniken genau beobachtet werden. Mit dem Terminus „Universum“ akzentuieren wir die Geschlossenheit des Aggregats, das mit den Kriterien der Selektion und Konstruktion, die es bilden, korrespondiert. Darauf aufbauend lassen sich dann weitere Universen ableiten, sei es durch Teilmengenbildung (Binnenstruktur), sei es durch Erweiterung der Selektoren (angelagerte externe Struktur).

Die TAA zielt generell auf Differenzen, die sich durch unterschiedliche Aggregatbildungen im Analyse-Output ergeben. Die primäre Aggregierungslogik bestimmt die Basisinterpretation der Daten. Vom ursprünglichen Universum abgeleitete Derivat-Universen werden entsprechend ihrer Differenzlogik interpretiert. So können etwa durch Gesprächsvernetzung erweiterte Twitter-Universen herangezogen werden, um den Einfluss des Gesprächscharakters im Ursprungsuniversum zu verstehen. In einer ersten Anwendung der TAA geschieht dies im Hinblick auf die Twitter-Kommunikation zur Salzburger Landtagswahl 2013. Wir konstruieren zunächst mit Hilfe zeitlicher, thematischer und personaler Selektionskriterien ein Twitter-Universum, um dieses „Ursprungsuniversum“ sodann durch Einbeziehung vernetzter Gesprächs-Tweets zu modifizieren.

Aus den resultierenden Differenzen zwischen den Universen erwarten wir Aufschluss über die Funktion der Gesprächsvernetzung innerhalb der Twitter-Kommunikation zum Beispiel im Hinblick auf bestimmte Themen und Themenbehandlungsweisen: Welche Themen weisen eine hohe Resonanzqualität auf, die sich in zahlreichen kommunikativen Anschlusshandlungen manifestiert?

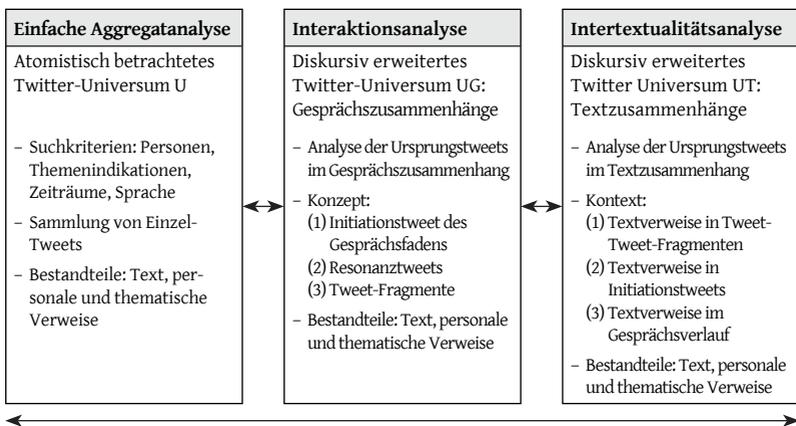
Auch ist von Interesse, welche Themen zwar gesetzt werden, aber sofort „versanden“, also ohne Resonanz im Twitter-Universum bleiben.

Themenbezogene Ergebnisse lassen sich in dreierlei Weise zur Agenda-Setting-Analyse nutzen: In Verbindung mit Akteurszuschreibungen werden *erstens* online-immanente Agenda-Setting-Prozesse (von Journalisten induzierte Thematisierungen) und Agenda-Building-Prozesse (von Politikern initiierte Themen) rekonstruiert. Unter Berücksichtigung von Verweisen auf Texte der Massenkommunikation in den Tweets geraten *zweitens* „individualisierte“ Schnittstellen zur Offline-Kommunikation in den Blick. Schließlich können *drittens* die durch Gesprächsresonanzen gewichteten Themenhierarchien der Online-Kommunikation mit der im selben Zeitraum vorfindlichen Themenhierarchie in der Offline-Massenkommunikation parallelisiert und durch cross-lagged Korrelationen und Strukturgleichungen (Reinecke, 2005; Hoyle, 2012) auf kausale Zusammenhänge geprüft werden.

4 Modell der Twitter-Aggregatanalyse (TAA)

Basierend auf unseren Annahmen und Zielstellungen lässt sich folgendes dreigliedriges Modell der Twitter-Aggregatanalyse aufstellen (siehe Abbildung 1):

Abbildung 1: Modell der Twitter-Aggregatanalyse (TAA)



In der TAA werden drei Untersuchungsebenen unterschieden, denen je eigene Erkenntnisinteressen zugeordnet sind: a) einfache Aggregatanalyse, b) Interaktionsanalyse und c) Intertextualitätsanalyse.

Auf der ersten Ebene, der *einfachen Aggregatanalyse* wird eine durch Selektion und Reduktion gebildete Gesamtheit von Tweets auf Häufigkeiten von Vernetzungsparametern wie Hashtags und Retweets sowie von Textindikatoren (z. B. für Themen und Themenbehandlungsweisen) hin untersucht. Die einfache Aggregatanalyse befasst sich ausschließlich mit dem „Ursprungsuniversum“ U, das durch Selektion und Reduktion definiert und von der Vernetzung der „atomistisch“ betrachteten Tweets innerhalb und außerhalb abstrahiert ist. Auf dieser Analyseebene lassen sich bereits interessante Vergleiche zwischen verschiedenen U-Universen anstellen, wenn diese auf unterschiedlichen Zeiträumen und/oder unterschiedlichen Akteursgruppen wie Journalisten und Politikern basieren. Diskrete Twitter-Aggregate, die überschneidungsfrei koexistieren, heißen „Paralleluniversen“. Erkenntnisinteresse ist die Ermittlung quantitativer Konvergenzen und Divergenzen zwischen den Universen.

Die *Interaktionsanalyse* unterscheidet sich von der einfachen Aggregatanalyse dadurch, dass U als *Ursprungsuniversum* figuriert, von dem ein durch die Anwendung der Reply-Funktion modifiziertes und erweitertes *Gesprächsuniversum* UG abgeleitet wird. UG enthält mehr Tweets als U, da auch verknüpfte Tweets enthalten sind, die außerhalb von U liegen. Innerhalb von UG wird danach unterschieden, ob Tweets einen Gesprächsfaden initiieren, auf einen Initiationstweet reagieren oder aber ein unverbundenes Fragmenttweet repräsentieren. UG bildet somit den durch Gespräche strukturierten Interaktionsraum verbundener und unverbundener Einzeltweets. Die Differenzen zwischen UG und U beziehen sich zum Beispiel auf Veränderungen der formalen und inhaltlichen Kommunikationsqualität (Abnahme semantischer Fernverknüpfung durch Hashtags, Verschiebungen der thematischen Schwerpunkte etc.), die sich aus der Berücksichtigung interner und externer Gesprächsnetzungen ergeben. Sie charakterisieren die Einbettung von U in einen Gesprächszusammenhang.

Im Rahmen der Themenanalyse repräsentiert „UG minus U“ den *semantischen Gesprächsdruck von außen*, der offenbart, welche Themen im Twitter-Ursprungsuniversum durch angelagerte Gespräche zusätzlich gepusht wurden. Das heißt, der Erweiterungsraum enthält genau die Themen, die durch Gesprächsakteure in das Ursprungsuniversum (z. B. alle Tweets mit Bezug zur Salzburger Land-

tagswahl) hereingetragen wurden, obwohl sie nicht das primäre Selektions- und Relevanz-Kriterium der Samplebildung erfüllen. Wenn etwa das Thema „soziale Sicherheit“ im erweiterten Gesprächsuniversum an Bedeutung gewinnt, so lässt sich hieraus ableiten, dass es den Gesprächsaktivisten mehr am Herzen liegt als den Durchschnittsakteuren im Ursprungsuniversum. Umgekehrt verweisen Themen, die in UG bzw. in „UG minus U“ weniger häufig vorkommen als in U, auf spezifische Kommunikationsinhalte des Wahlkampfs, die außerhalb des Wahlkampfs weniger bedeutsam sind.

Die Unterscheidung zwischen Initiations-, Resonanz- und Fragmenttweets erlaubt es überdies, die *Binnenstruktur* der Twitter-Kommunikation in UG einer genaueren Analyse zu unterziehen. Welche Akteure initiieren Gesprächsverläufe, welche folgen ihnen? Welche Themen induzieren einen großen Resonanzraum, welche bleiben isoliert und folgenlos? Hieraus sind differenzierte Einsichten in Agenda-Setting- und Agenda-Building-Prozesse innerhalb der politischen Online-Kommunikation zu gewinnen, die Erfolge und Misserfolge des Themen-Spinnings offenbaren. Außerdem kann die Rolle von Laien bei der Initiation und viralen Verbreitung von Themen genauer bestimmt werden.

Auf der dritten Analyseebene ist UT das um vernetzte Texte (z. B. der Massenmedien) erweiterte Ursprungsuniversum (Intertextualitätsraum), auf das sich die *Intertextualitätsanalyse* bezieht. Es kann in analoger Weise wie UG zur Differenzbildung herangezogen werden, wobei „UT minus U“ bei der Analyse verlinkter Texte von Presse und Fernsehen den *externen semantischen Druck der Massenmedien* repräsentiert, sofern diese im Twitter-Universum Spuren hinterlassen. Wenn UT inklusive der verlinkten Texte der Massenmedien in höherem Maße das Thema „politisches System und Finanzskandal“ enthält als dies im Ursprungsuniversum der Fall ist, dann hat die Berichterstattung zu diesem Thema an den Rändern des Wahlkampf einen viralen Spin erzeugt, der über den im Kernbereich realisierten Spin hinausgeht. Umgekehrt verweist eine gesunkene Bedeutung des Themas „soziale Sicherheit“ in UT darauf, dass die Massenmedien hier nur wenig zur Themendynamik bei Twitter beigetragen haben (womöglich im Unterschied zu den Gesprächsaktivisten, siehe oben). Neben massenmedialen können aber auch andere Textsorten (z. B. Blog-Posts) berücksichtigt und in analoger Weise als *externer semantischer Druck von Blogs* interpretiert werden.

Quer zu Twitter-Universen U, UG, und UT sind weitere Untergliederungen zum Beispiel nach Person, Partei und Zeitraum möglich, die dann für die Akteurs-

bzw. für die Analyse im Zeitverlauf zur Verfügung stehen. Selbstverständlich können innerhalb der Twitter-Universen auch Einzelfallanalysen durchgeführt werden, die den quantitativen Ertrag um eine qualitative Komponente erweitern. Umgekehrt profitiert die Einzelfallanalyse von der Einbindung in die Twitter-Aggregatanalyse insofern, als sie in einem klar definierten Referenzrahmen stattfindet und die Ergebnisse mit den quantitativen Befunden interpretativ verknüpft werden können.

5 Architektur der Twitter-Universen

Im Folgenden werden die Selektions- und Konstruktionsschritte bei der Twitter-Analyse zur Salzburger Landtagswahlen skizziert, die wir als paradigmatischen Anwendungsfall der TAA sehen und uns Anlass zu einer grundsätzlichen Darstellung des Vorgehens bietet.

Zur Verminderung des Subjektivitäts-Problems ist *Transparenz* unerlässlich. Daher ist uns die genaue Dokumentation der Analyseschritte wichtiger als das Postulat eines vermeintlich einzig richtigen Vorgehens. Wir gehen davon aus, dass es immer alternative Varianten der Selektion und Konstruktion von Twitter-Universen gibt und fokussieren auf die *Forschungslogik der mit der Wahl des Vorgehens verbundenen Interpretationsperspektiven* inklusive deren Grenzen. Ein logisch verbundenes Set von Twitter-Universen, das an ein bestimmtes Ursprungsuniversum rückgebunden ist, heißt „Pluriversum“, wobei die Mehrzahl der Vokabel im Titel des Aufsatzes die Pluralität der Konstruktionen doppelt unterstreicht. Ein *Pluriversum* besteht aus teilweise überlappenden Universen, die durch Binnendifferenzierung bzw. durch Erweiterungsregeln bezüglich des Ursprungsuniversums entstehen. Innerhalb des TAA-Modells sind UG und UT typische Twitter-Pluriversen mit Bezug auf ein bestimmtes Ursprungsuniversum.

Dem *Umfangsproblem* begegnen wir mit der Aufstellung von Selektionskriterien und der Festlegung von Selektionsschritten, die zur Abfrage der Twitter-Datenbank genutzt werden. Auch wenn Twitter teilweise schwer durchschaubare Limitierungen in den angefragten Output einbaut, erscheint dies zurzeit als die beste Strategie, um dem Anspruch auf Transparenz, Systematik und Umfangsbewältigung zu genügen. Zur Sammlung der Tweets bezüglich der Salzburger Land-

tagswahl wurde zunächst eine Akteursliste aus österreichischen Twitter-Usern erstellt, die auf folgenden Daten basiert:

- a) „Heavy Twitter-User“, die auf *politometer.at*, *socialmediaradar.at* und *ots.at/twitterlist* gelistet sind und den Akteurstypen Politiker, Partei, Journalist, Medium und Laie zugeordnet sind,
- b) Nennung des Twitter-Accounts auf der Webseite einer der österreichischen Parteien (Akteurstyp Partei und Politiker) bzw. eines Mediums (Medium, Journalist) und schließlich
- c) Referenz zu Partei (Politiker), Medium (Journalist) oder politischer Orientierung (politisch interessierter Laie) im Twitter-Profil.

Die daraus resultierende Liste umfasste 364 Twitter-Accounts und gliedert sich in die Typen (1) *Politiker* (Person oder Partei), (2) *Journalist* (Person oder Medium) und (3) *Laie*. Mithilfe des *Twitter Search Application Programming Interface* (API) haben wir von diesen 364 Twitter-Usern alle Tweets im Zeitraum der Salzburger Landtagswahl (vom 29.4. bis 6.5.2013) ausgelesen und gefiltert. Das ergab ein Gesamtsample von 34.001 Tweets. Aufgrund der Datenmengen wurden für die weitere Analyse nur der 4. Mai (Samstag, Tag vor der Wahl) und der 5. Mai (Sonntag, Tag der Wahl) ausgewählt (N Samstag = 2.496 Tweets, N Sonntag = 5.240 Tweets, N Sa+So = 7.736 Tweets), um in exemplarischer Weise die Umstellungen des Twitter-Diskurses von der Endphase des Wahlkampfes zur kommunikativen Begleitung der Wahl beobachten zu können. Da nur Tweets mit Bezug zu einem Politikfeld- bzw. einem Wahlbezug als relevant für politische Kommunikation erachtet wurden, musste die Ausgangssammlung von Tweets um nicht-relevante reduziert werden. Das *Reduktionsproblem* wurde konventionell mit menschlichen CodiererInnen gelöst, die mit einem Relevanz-Kriterienkatalog (Beispiele für Wahlbezug und Vorgabe der Politikfeld-Kategorien siehe unten) die maschinell erzeugte Liste an Tweets bearbeiteten.

5.1 *Bildung eines Ursprungsuniversums aus der Ausgangssammlung von Tweets*

Zur Konstruktion des Ursprungsuniversums wurde die Ausgangssammlung an Tweets (N=7.736) nach den folgenden Relevanzkriterien reduziert: Das

Twitter-Ursprungsuniversum U^p ($N = 1.357$) besteht aus Politiker-Tweets mit semantisch explizitem Politikfeld- bzw. Wahlbezug. Folglich besteht U^p aus Tweets, die von Politikern, Parteien, politischen Organisationen versendet (aktiv) bzw. Tweets, die an diese adressiert wurden (passiv), sofern sie zugleich einen Bezug zu einem der definierten Politikfelder und/oder den auf die Landtagswahl bezogenen Kontextualisierungen (z. B. eine Spekulation über den Wahlausgang) aufweisen. Die Akteurstypen *Journalist* und *Laie* kommen in dieser Sammlung vor, aber nur, wenn sie von *Politikern* adressiert wurden bzw. sie Politiker adressiert haben. Es gibt demnach fünf dyadische Kombinationen: Politiker—Politiker, Politiker—Journalist, Journalist—Politiker, Politiker—Laie, Laie—Politiker. Inhaltsleere und nicht-politische Tweets („Ich habe Fotos auf Facebook hochgeladen“, „Die Sonne scheint heute früh über Wien“) wurden entsprechend rigoros aussortiert.

5.2 Bildung eines Paralleluniversums aus Nicht-Politiker-Tweets

Reine Nicht-Politiker-Tweets, bei denen Politiker weder als Sender noch als Adressat vorkommen, bilden in der Tweet-Ausgangssammlung ein *Paralleluniversum* zu U^p , nämlich U^{np} . Dieses besteht aus Tweets, die von Journalisten, Medien und Laien versendet und dabei ausschließlich auch an diese adressiert wurden. Auch hier gelten zusätzlich die beiden thematischen Relevanzkriterien: (1) Bezug zu einem der definierten Politikfelder und/oder (2) Landtagswahl-Kontextualisierung (N von $U^{np} = 1.220$ Tweets, N von $U^p + U^{np} = 2.577$ Tweets). Aussortiert wurden wiederum inhaltsleere bzw. nicht-politische Tweets.

5.3 Pluriversum der Gesprächsinteraktionen für Politiker-Tweets

Eine wichtige inhaltsanalytische Innovation der TAA stellt die Interaktionsanalyse der Gespräche (kurz: die *Gesprächsanalyse*) dar. Diese wurde exemplarisch für die Politiker-Tweets im Universum U^p durchgeführt. Twitter bietet den Usern über die Reply-Funktion die Möglichkeit, auf Tweets anderer User zu reagieren, wodurch sich ein Gespräch entwickeln kann. Da die Extraktion der Tweets mittels der Search API jedoch keine Rückschlüsse auf potentielle Einbettungen in Gesprä-

chen zulässt, wurden die Gesprächsfäden manuell über die Twitter-Suchfunktion verfolgt. Im Rahmen dieses Vorgangs wurden die einzelnen Tweets bezüglich ihrer Funktion in einem Gespräch gekennzeichnet. Dabei wurden auch Tweets, die bis dato noch nicht in U^P existierten, hinzugefügt.

Für die Gesprächsanalyse ergeben sich somit folgende Teilmengen an Tweets:

Pluriversum der Gesprächsanalyse

- U^P = Ursprungsuniversum aktiver und passiver Politiker-Tweets, N=1.357
 G_{UP} = Durch Gespräch vernetzte Tweets innerhalb von U^P , N=367
 F_{UP} = Durch Gespräch nicht vernetzte Fragmenttweets in U^P , N=164¹
 Gadd = Tweets, die durch die Rekonstruktion der Gespräche neu hinzukamen (angelagerte Gesprächstweets außerhalb von U^P), N=171

$$\begin{aligned} \text{Gesprächsuniversum:} & \quad UG (N = 1.528) = U^P + \text{Gadd} \\ \text{Teilmenge der Gesprächstweets:} & \quad G_{UG} (N = 538) = G_{UP} + \text{Gadd} \end{aligned}$$

Die Gesprächsanalyse bezieht sich auf den direkten interpersonellen Interaktionsraum auf Twitter, der durch weitere Vernetzungsoperatoren wie zum Beispiel Hashtags überlagert wird. Im Rahmen der Gesprächsanalyse kann etwa ermittelt werden, ob semantische Fernvernetzungen via Hashtag in den gesprächsvernetzten Teilen oder aber zur Kompensation interpersoneller Vernetzungsdefizite in den fragmentierten Teilen von UG eine größere Rolle spielen.

Zusätzlich lassen sich im Gesprächsuniversum noch folgende Sub-Unterscheidungen treffen, die für die Indexbildung im Rahmen der Interaktionsanalyse wichtig sind: (1) *Initiationstweets* (lösen Gespräche aus) und (2) *Resonanztweets* (beantworten Tweets).

$$\begin{aligned} I_{UG} &= \text{Initiationstweets, die aus } UG \text{ stammen und Gespräche in } UG \text{ auslösen, } N=132 \\ R_{UG} &= \text{Resonanztweets, die aus } UG \text{ stammen und auf Tweets aus } UG \text{ antworten, } N=405 \end{aligned}$$

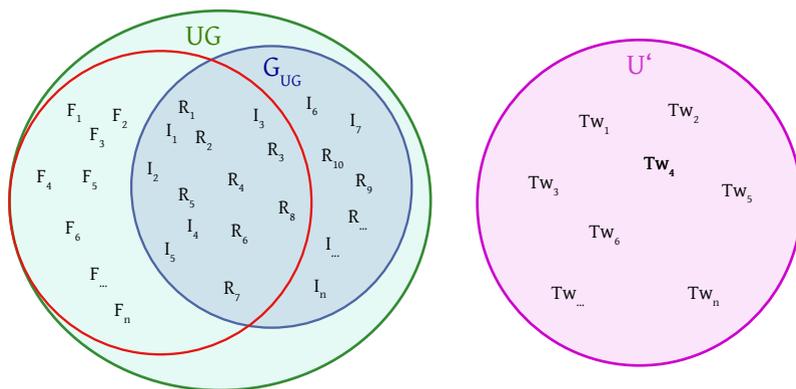
1 826 Tweets des Ursprungsuniversums konnten aufgrund des Analysezeitpunkts nicht mehr zugeordnet werden. Daher addieren sich Gesprächstweets und Fragmenttweets nicht zu U^P .

Bildet man die Differenz zwischen den Initiations- und Resonanztweets, ergibt sich ein Maß für den Initiations- bzw. den Resonanz-Bias von Themen, Politikern und Parteien. Ein Übermaß von Initiationstweets über Resonanztweets innerhalb des Themenbereichs *Umweltpolitik* zeigt an, dass dem Initiationsaufwand von Twitter-Aktivisten zu diesem Thema keine gleichermaßen hohe Resonanz entspricht. Umgekehrt indizieren resonanzstarke Themen, dass die kommunikative Dynamik im Twitter-Universum besonders hoch ausfiel, ohne dass dazu ein erhöhter Initiationsaufwand getrieben werden musste. In analoger Weise kann die Bilanz von Initiation und Resonanz zur Charakterisierung von Politikern und Parteien dienen.

Abbildung 2 veranschaulicht die teilweise überlappenden, teilweise separierten Twitter-Universen der TAA, wobei UG ein komplexes Pluriversum der Interaktionsanalyse definiert und U versus U' Paralleluniversen – in unserem Falle gebildet aus Politiker- und Nicht-Politiker-Tweets – repräsentieren.

Der Überschneidungsbereich in UG ist die *Zone verdichteter Interaktion*. Hier lässt sich zum Beispiel ermitteln, welche Themen und Akteure im direkten Twit-

Abbildung 2: Twitter-Pluriversum der Interaktionsanalyse und Paralleluniversum U'



- F_{1-n} = Fragmenttweets in U und UG,
- I_{1-n} = Initiationstweets in UG,
- R_{1-n} = Resonanztweets in UG,
- $R+I$ = Teilmenge gesprächsnetzter Tweets in UG ($=G_{UG}$)

Tw_{1-n} = Tweets im Paralleluniversum U'

ter-Gespräch dominieren. Die Erweiterung von U zu UG markiert die *durch Interaktion verursachte Ausdehnung von U*, die wir als *diskursive Dynamik* im Sinne eines eingehenden oder ausgehenden Gesprächsdrucks interpretieren und wiederum auf Themen und Akteure beziehen. Und schließlich erbringt die Analyse der Paralleluniversen U und U' (in unserem Fall U^P und U^{NP}) Erkenntnisse zur Differenz, durch die sich Stilmerkmale, Themenschwerpunkte und Themenbehandlungsweisen in den Politiker- und Nicht-Politiker-Tweets unterscheiden. Die *einfache Aggregatanalyse* kommt derweil mit Differenzierungen auf der Zeitachse und nach Akteursgruppen aus, die wir aus Anschaulichkeitsgründen nicht in die Grafik integriert haben, mit der wir nun aber die praktische Analyse beginnen.

Im Folgenden präsentieren wir ausgewählte Ergebnisse der Inhaltsanalyse, die wir gemäß der TAA und der angegebenen Architektur der Twitter-Universen am Beispiel der Landtagswahlen in Salzburg 2013 erzielten.

6 Inhaltsanalyse der Twitter-Kommunikation zur Salzburger Landtagswahl

Die Salzburger Landtagswahl 2013 stellte für die österreichischen Parteien insofern eine Sondersituation dar, als die mehrjährige Dominanz der Sozialdemokraten (SPÖ) aufgrund eines Finanzskandals gefährdet schien. Sowohl alteingesessene als auch junge, aufsteigende Parteien sahen ihre Chancen auf einen Wahlerfolg. Ziel unserer Untersuchung ist *erstens* die Rekonstruktion von Themen und Themenbehandlungsweisen, die zur Analyse von Agenda-Setting- und Agenda-Building-Prozessen herangezogen werden und eine Erklärung des negativen Wahlausgangs für die SPÖ liefern können. *Zweitens* wollen wir exemplarisch die Leistungsfähigkeit der TAA und die Interpretationsfähigkeit der generierten Daten im Hinblick auf grundlegende Fragen der politischen Online-Kommunikation demonstrieren. Wir konzentrieren uns dabei auf formale Vernetzungsparameter (@, #, http, retweets), um etwa Kommunikationsstile in den Twitter-Universen von Politikern und Nicht-Politikern zu unterscheiden, sowie auf die Themenhierarchien im Rahmen des Gesprächs-Pluriversums und der Sub-Universen von Initiations- und Resonanztweets. Die Vorführung der Intertextualitätsanalyse muss aus Platzgründen unterbleiben. Sie folgt aber im Wesentlichen der Logik der Interaktionsanalyse.

Die Auszählung der Vernetzungsparameter erfolgte maschinell. Der themenbezogene Teil wurde mit Hilfe einer konventionellen Inhaltsanalyse (IA) durchgeführt, deren Kategoriensystem sich bereits bei vorangegangenen Wahlen in Österreich bewährt hat (Grimm et al., 2008). Das Codebuch gliedert sich in vier Teile: *Erstens* werden grundlegende Charakteristika der Tweets wie etwa der Typ des Users (u. a. Politiker, Journalist, Laie), Anzahl und Wortlaut der im Tweet vorkommenden Hashtags oder Anzahl und Name der Adressaten codiert. Darauf aufbauend wurde *zweitens* der Inhalt der Tweets kategorisiert. Neben dem Politikfeld (u. a. Außenpolitik, innere Sicherheit, Umweltpolitik) wurde hier auch die Politikebene (regional, national, EU, global) und die Art der Themenbehandlung (problem- versus lösungsorientiert) codiert. *Drittens* wurden die singulären Tweets bezüglich ihrer Kontextualisierung zu einer Wahl im Sinne eines Bezugs zu einem Wahlkampf, Wahlinformationen oder dem Ausgang der Wahl erfasst. Im vierten und letzten Schritt wurden die Tweets hinsichtlich ihrer Einbettung in Gespräche codiert.

Die Intercoder-Reliabilität war durchwegs gut. Die formalen Vernetzungsparameter wie Hashtags und Retweets wurden maschinell ausgezählt. In den für die Interpretation der Agenda-Setting- und Agenda-Building-Prozesse wichtigen themenbezogenen Kategorien ist für die Variable *Politikfeld* $R=0,74$ bzw. für die Variable *Themenbehandlungsweise* $R=0,82$.²

6.1 *Stilmerkmale der Twitter-Kommunikation*

Die Elemente der Adressierung anderer User mittels @, der Thematisierung mittels #, der Einbindung von Hyperlinks und Retweets (d. h. der Weiterverbreitung eines Tweets eines anderen Users) und zählen zu den Kernelementen von Twitter-Kommunikation. Tabelle 1 stellt den Einsatz dieser Twitter-Metriken im Universum der Politiker-Tweets (U^P) dem Paralleluniversum der Tweets ohne Politikerbeteiligung (U^{NP}) gegenüber. Die Prozentangaben beziehen sich auf die Anzahl der Tweets, die die jeweiligen Stilmerkmale aufweisen, gemessen an allen Tweets im jeweiligen Universum (=100%).

2 Die Reliabilitätskoeffizienten wurden nach Früh (1981) in der erweiterten Version von Grimm (1989) berechnet.

Tabelle 1: Der Einsatz von Adressierungen, Hashtags, Hyperlinks und Retweets in den Paralleluniversen mit und ohne Politiker-Beteiligung

	U ^P %	U ^{NP} %	Differenz d%
@	80,0	73,5	6,4
#	42,3	35,3	7,0
http	26,9	36,6	-9,7
retweets	57,3	64,7	-7,4

U^P = Universum der Politiker-Tweets, U^{NP} = Paralleluniversum der Nicht-Politiker-Tweets. Angaben in Prozent. Berechnungsgrundlage sind N (U^P) = 1.357 Tweets und N (U^{NP}) = 1.220 Tweets.

Die Tweets in den Paralleluniversen unterscheiden sich zunächst dadurch, dass im Falle einer Beteiligung von Politikern verstärkt direkte User-Adressierungen enthalten sind (d% = 6,4) und die Verschlagwortung von Themen mittels Hashtags eine größere Rolle spielt (d% = 7,0). Im Gegensatz dazu werden in Tweets, die ausschließlich Nicht-Politiker erwähnen, häufiger als im Politiker-Universum Verbindungen zu Online-Quellen (d% = -9,7) hergestellt. Auch retweeteten Nicht-Politiker häufiger untereinander, als dies der Fall ist, wenn Politiker die Kommunikation mitgestalten. Wir verstehen dies als einen Hinweis darauf, dass der Eingriff von Politikern in die Twitter-Kommunikation vor allem thematisch motiviert ist (vermutlich zum Zwecke des Agenda-Building). Demgegenüber trägt die ausgelagerte Journalisten-Laien-Kommunikation diskursiven Charakter und bezieht eher auch externe Online-Informationsquellen mit ein, die zumindest teilweise ihren Ursprung in den klassischen Massenmedien haben dürften. Die geringeren Retweets unter Politiker-Beteiligung deuten im Übrigen auf kommunikative Zurückhaltung potenzieller Gesprächspartner hin, die aus der wahrgenommenen parteipolitischen Instrumentalisierung oder aus dem Durchschauen einzelner Manöver der Themen-Promotion erklärt werden können.

Im nächsten Schritt beobachten wir die Veränderungen des Kommunikationsstils, die aus der Erweiterung des Politiker-Ursprungsuniversums U^P zum Gesprächsuniversum U^G resultieren. Tabelle 2 zeigt, dass im diskursiv erweiterten

Interaktionsraum thematische Verschlagwortung und Verlinkungen mit externen Informationsquellen wie auch Adressierungen und Retweets geringer enthalten sind als im Ursprungsuniversum. Offensichtlich verlieren diese an Bedeutung in der direkten Interaktion unter den Usern. Im Umkehrschluss lässt dies vermuten, dass die Twitter-spezifischen Metriken primär der Fernvernetzung dienen, während im direkten Gespräch semantische (#, http) und pragmatische Operatoren (@, retweets) eher verzichtbar erscheinen.

Tabelle 2: Der Einsatz von Adressierungen, Hashtags, Hyperlinks und Retweets im Ursprungsuniversum aus Politiker-Tweets und dem erweiterten Gesprächsuniversum

	U^P %	UG %	Differenz d%
@	80,0	79,3	0,7
#	42,3	41,0	1,3
http	26,9	24,7	2,2
retweets	57,3	55,4	1,9

U^P = Ursprungsuniversum der Politiker-Tweets, UG = Gesprächsuniversum. Angaben in Prozent. Berechnungsgrundlage sind N (U^P) = 1.357 Tweets und N (UG) = 1.528 Tweets.

Die Differenz in der Funktionalität der Metriken zeigt sich noch deutlicher, wenn wir die Fragmenttweets den gesprächsvernetzten Tweets gegenüberstellen (siehe Tabelle 3). Unter Bedingungen der Gesprächsvernetzung werden um 25 Prozent weniger externe Wissensquellen und um 11,8 Prozent weniger thematische Hashtags eingesetzt, als dies bei Fragmenttweets der Fall ist. Dies unterstreicht die Bedeutung der Metriken für die semantische Fernkommunikation. Man könnte aber auch kritischer formulieren: Die Verwendung von Hashtags und WWW-Verweisen findet (häufig) kein entsprechendes Echo in der direkten Twitter-Interaktion. Aus Tabelle 3 ist überdies ersichtlich, dass die pragmatischen Vernetzungsoperatoren – vor allem die persönlichen Adressierungen durch @ – auch im gesprächsvernetzten Teil der Twitter-Kommunikation noch eine erhebliche Rolle spielen.

Tabelle 3: Der Einsatz von Adressierungen, Hashtags, Hyperlinks und Retweets bei fragmentierten und gesprächsvernetzten Tweets

	F_{UP} %	G_{UG} %	Differenz d%
@	74,4	74,5	-0,1
#	44,5	32,7	11,8
http	38,4	13,4	25,0
retweets	65,2	50,9	14,3

F_{UP} = Fragmenttweets (=F_{UG}), G_{UG} = Gesprächstweets. Angaben in Prozent. Berechnungsgrundlage sind $N(F_{UP}) = 164$ Tweets und $N(G_{UG}) = 538$ Tweets.

Immerhin ist ja auch die Gesprächsvernetzung selbst eine pragmatische Verlinkungsoperation.

Aus Tabelle 4 geht hervor, dass bei Initiationstweets, die ein Gespräch auslösen, deutlich mehr semantische Verlinkungsoperatoren (#, http) verwendet werden als in Resonanztweets, die auf Gesprächsinitiativen reagieren. Bei den pragmatischen Operatoren der Adressierung und Weiterversendung (@, retweets) ist es umgekehrt.

Tabelle 4: Der Einsatz von Adressierungen, Hashtags, Hyperlinks und Retweets bei der Initiierung von Gesprächen und kommunikativen Anschlusshandlungen in Gesprächen

	I_{UG} %	R_{UG} %	Differenz d%
@	31,1	88,9	-57,8
#	54,5	25,7	28,9
http	23,5	10,1	13,4
retweets	47,7	52,1	-4,4

I_{UG} = Initiationstweets, R_{UG} = Resonanztweets. Angaben in Prozent. Berechnungsgrundlage sind $N(I_{UG}) = 132$ Tweets und $N(R_{UG}) = 405$ Tweets.

Die Initiierung von Gesprächen scheint also von einem Bemühen getragen zu sein, die semantische Basis für kommunikative Anschlusshandlungen zu schaffen, indem der thematische Fokus und die Wissensfundierungen betont werden. Die weitere Gesprächsdynamik ist dann dadurch gekennzeichnet, den begonnenen Gesprächsfaden durch Adressierungen und Versendungen zu bewahren bzw. weiterzuentwickeln.

Hier ist nicht der Raum, um differenziertere Analysen etwa in Bezug auf die Rolle von Politikern, Journalisten und Laien bei der Initiierung und Dynamisierung von Gesprächsverläufen anzustellen. Es muss als bisheriges Resümee der *einfachen Aggregatanalyse* und der *erweiterten Interaktionsanalyse* der Hinweis genügen, dass sich die Verwendung der Twitter-typischen Metriken unterschiedlich in den Sub-Universen abbildet und die gefundenen Differenzen plausibel aus den jeweiligen funktionalen Erfordernissen erklärt werden können. Der Vergleich der Paralleluniversen aus Politiker- und Nicht-Politiker-Tweets zeigte markante Unterschiede des Kommunikationsstils. Vor allem Politiker scheinen die Twitter-Kommunikation aus Gründen des Agenda-Building zu instrumentalisieren, auf das die Twitter-Gemeinde freilich mit kommunikativer Zurückhaltung reagiert. Im Twitter-Gesprächspluriversum unter Politiker-Beteiligung wurde eine Tendenz zur Aufspaltung der Rollen nach Gesprächsinitiierung einerseits und Weiterverbreitung von Botschaften andererseits festgestellt. Eine klare Rollenzuweisung nach Akteurstypen lässt sich allerdings bislang nicht belegen. Manche Politiker gerieren sich als Gesprächsinitiatoren, andere ziehen es vor, sich in bereits bestehende Gesprächsverläufe einzuklinken.

Welche Themenschwerpunkte beherrschten nun die politische Kommunikation zur Salzburger Landtagswahl auf Twitter? Lassen sich Zusammenhänge zwischen Themenprioritäten und Wahlausgang erkennen?

6.2 Themen in den Twitter-Universen

Bei jedem Tweet wurden die Themenvariablen *Politikfeld* und *Wahlbezug* codiert. Als „Politikfelder“ gelten Handlungsbereiche der Politik, wie sie sich zum Beispiel in der Organisationsstruktur von Ministerien widerspiegeln. Politische Handlungen umfassen idealtypisch eine Situations- oder Problemdefinition und optional ein Lösungskonzept. Durch die Beachtung des Handlungsaspekts bei der Kategori-

sierung der Twitter-Themen können diese über Debatten-Referenzen hinaus auch auf die Inhalte politischer Entscheidungen projiziert werden. „Wahlbezüge“ sind demgegenüber ereigniszentrierte Thematisierungen. Sie lassen sich nicht in die Hierarchie von *Politikfeldern* integrieren und werden daher gesondert ausgewiesen. Als minimale Relevanzbedingung für die Konstruktion des Ursprungsuniversums galt, dass Tweets zumindest eine der beiden Themenbedingungen erfüllen. Dementsprechend gibt es Tweets, die nur Politikfelder, nur den Wahlkampf oder aber beides ansprechen. Die Prozentangaben in Tabelle 5 beziehen sich auf die Anzahl der Tweets, die im jeweiligen (Teil-)Universum gegeben sind (=100%).

Es fällt auf, dass alle Twitter-Universen mehrheitlich Bezüge zur Wahl und weniger zu Politikfeldern aufweisen. Offenbar ist der Ereignischarakter der Wahl ein genereller Fokus der Twitter-Kommunikation, nicht nur am Wahltag, dort freilich umso mehr. Wir vermuten, dass die Twitter-Kommunikation generell ereignisorientiert veranlasst, dynamisiert und erst sekundär mit inhaltlichen Politik-Themen aufgeladen wird, die dem jeweiligen Ereignis zugeordnet werden können.

Tabelle 5: Politikfelder und Wahlbezüge nach Twitter-Universen

	Tweets gesamt	Tweets mit Politikfeld- bezug	Tweets mit Politikfeld- bezug in %	Tweets mit Wahlbezug	Tweets mit Wahlbezug in %
U^P	1.357	366	27,0	1.064	78,4
F_{UP}	164	85	51,8	96	58,5
G_{UP}	367	99	27,0	285	77,7
UG	1.528	386	25,3	1.224	80,1
I_{UG}	132	26	19,7	117	88,6
R_{UG}	405	92	22,7	328	81,0
U^{NP}	1.220	329	27,0	915	75,0

U^P = Ursprungsuniversum der Politiker-Tweets, F_{UP} = Fragmenttweets,
 G_{UP} = Gesprächstweets, UG = Gesprächsuniversum, I_{UG} = Initiationstweets,
 R_{UG} = Resonanztweets, U^{NP} = Paralleluniversum der Nicht-Politiker-Tweets.
 Angaben in Prozent. Basis der Prozentuierung sind die
 korrespondierenden „Tweets gesamt“ des jeweiligen Universum.

Beachtenswert ist außerdem, dass Fragmenttweets, also Tweets ohne Gesprächszusammenhang, den stärksten Bezug zu Politikfeldern aufweisen (51,8%). Dies belegt, dass die Politikfeldbezüge häufig intentional gesetzt werden (z. B. von Spin Doctors der Parteien oder von engagierten Journalisten und Laien), ohne dass ihnen kommunikative Anschlusshandlungen folgen. Dennoch gibt es hinreichend viele Tweets mit Politikfeldbezügen, die erfolgreich Gesprächsfäden generieren. Immerhin beinhalten 19,7 Prozent der Initiationstweets thematische Einlassungen zu Politikfeldern. Im Resonanzraum der Gespräche sind die Politikfeldbezüge sogar noch stärker ausgeprägt (22,7%). Hieraus folgt, dass gut 20 Prozent aller erfolgreich konstituierten Gesprächszusammenhänge im erweiterten Interaktionsraum sich mit praxisrelevanten politischen Themen befassen.

Im Hinblick auf den Wahlbezug ergibt sich ein umgekehrter Zusammenhang. Im erweiterten Gesprächsuniversum sind die ohnehin sehr häufigen Wahlbezüge gegenüber dem Ursprungsuniversum noch zahlreicher. Die Gesprächsdynamik der politischen Twitter-Kommunikation gibt der Ereigniszentrierung zusätzlichen Raum, während sie die thematischen Bezüge reduziert. Wir schließen daraus, dass eine der besonderen Qualitäten der Twitter-Kommunikation, sofern sie Gesprächszusammenhänge konstituiert, darin zu bestehen scheint, die Ereigniszentrierung der politischen Kommunikation zu verstärken. Der höchste Prozentsatz an Wahlbezügen ist in Initiationstweets enthalten (88,6%). Dies belegt eindrücklich das intentional erfolgreiche Moment der Ereigniszentrierung innerhalb der Twitter-Kommunikation. Demgegenüber sind in Resonanztweets um 7,6 Prozent weniger Wahlbezüge enthalten. Die Gesprächs-Follower sind demnach nicht in gleichem Umfang bereit, ereigniszentrierte Gespräche zu verlängern, wie es der Erfolgsquote gelungener Gesprächsinitiationen entspräche. Wir sehen darin ein Maß dafür, die Fortsetzung von Gesprächsfäden zu begrenzen. Offenbar dynamisiert die Ereigniszentrierung die Twitter-Kommunikation, erhöht aber zugleich auch ihr Verfallsdatum.

Tabelle 6 enthält die relativen Anteile einzelner Politikfelder gemessen an den Tweets mit Politikfeldbezug. Im Unterschied zu den Tabellen oben, die als Prozentuierungsbasis jeweils alle Tweets im jeweiligen (Teil-)Universum verwenden, haben wir aus Gründen der Anschaulichkeit das Thematisierungsaufkommen als Basis für 100 Prozent verwendet, so wie es auch in der traditionellen Agenda-Setting-Forschung üblich ist. Das atomistisch konstruierte Universum der Politiker-Tweets U^p wird dominiert von den Politikfeldern „politisches System“ (29,7%),

Tabelle 6: Bezüge zu Politikfeldern im Ursprungs- und erweiterten Gesprächsuniversum

$N_U=366$ $N_{UG}=386$ Politikfeld	U^P %	UG %	Differenz d%
Politisches System	29,7	30,1	-0,4
Öffentliche Verwaltung und staatliche Unternehmen	1,6	1,6	0,0
Innere Sicherheit	8,0	7,8	0,2
Frauenpolitik	4,3	4,1	0,1
Rechtspolitik	3,5	3,4	0,1
Außenpolitik	7,2	7,5	-0,3
Europäische Integration	3,5	3,4	0,1
Verteidigungspolitik	1,3	1,3	0,0
Finanzpolitik	4,5	4,7	-0,1
Wirtschaftspolitik	2,4	2,3	0,1
Land- und Forstwirtschaftspolitik	4,0	3,9	0,1
Verkehrs- & Technologiepolitik	0,3	0,3	0,0
Umweltpolitik	10,7	10,4	0,3
Bildungspolitik	4,0	3,9	0,1
Universitäts- und Forschungspolitik	5,6	5,7	-0,1
Kulturpolitik	3,2	3,4	-0,2
Arbeits- und Beschäftigungspolitik	1,3	1,3	0,0
Soziale Sicherheit	2,1	2,6	-0,5
Gesundheitspolitik	2,1	2,1	0,1
Familien- und Jugendpolitik	0,5	0,5	0,0
% Summe	100,0	100,0	0,0

*U^P = Ursprungsuniversum aktiver und passiver Politiker-Tweets (N=1.357),
UG = Gesprächsuniversum (N=1.528). Basis für Prozentuierung: Tweets mit
Bezug zu Politikfeld in U^P N=366, Tweets mit Bezug zu Politikfeld in UG N=386.*

„Umweltpolitik“ (10,7%), „innere Sicherheit“ (8,0%) und „Außenpolitik“ (7,2%). Im Thema „politisches System“ spiegelt sich neben Fragen des Wahlmodus und der Korruption vor allem das Thema Finanzskandal in der öffentlichen Verwaltung wider, das, verursacht durch eine SPÖ-Ministerin, zu einer Reform der Finanzverwaltung herausforderte. Das Thema „Finanzpolitik“ selbst figuriert mit 4,5 Prozent auf dem sechsten Platz in der Themenhierarchie.

Eine höhere Beachtung im erweiterten Gesprächsuniversum zeigt sich vor allem bei den Politikfeldern „soziale Sicherheit“³, „politisches System“ und „Außenpolitik“. Die Differenzen sind aufgrund der geringen Fallzahl zwar numerisch klein, sie lassen sich aber plausibel mit dem Wahlausgang verknüpfen. Das mit Abstand dominierende Thema „politisches System“ zeigt den zweithöchsten Zuwachs. Überwiegend gewinnen gerade die quantitativ dominierenden Politikfelder durch die angelagerte Gesprächsvernetzung in U_G an Bedeutung. Dies kann als Indikator für Themenkonvergenz zwischen den Tweet-Aussendungen insgesamt und den gesprächsvernetzten Tweets verstanden werden. Die externe Gesprächsdynamik verstärkt demnach saliente Themen und repliziert die Themenrangfolge in hohem Maß.⁴ Das gilt auch für das wahlentscheidende Thema „Finanzpolitik“, bei dem ebenfalls ein, wenn auch schwacher Zugewinn im Gesprächsuniversum zu verzeichnen ist. Offenbar hat das Thema über die Summe der Einlassungen von Politikern, Journalisten und Laien hinaus eine Kommunikationsdynamik entfaltet, die zur SPÖ-Niederlage beitrug.

Es gibt allerdings auch interessante Ausnahmen von der Regel einer gesprächsverstärkten und gesprächsverstärkenden Themenkonvergenz. So weist das Thema „soziale Sicherheit“ mit 0,5 Prozent den höchsten gesprächsbedingten Bedeutungsgewinn auf, rangiert selbst aber mit 2,1 Prozent im unteren Mittelfeld der Themen-Rangreihe. Wir deuten dies als Hinweis auf ein Input-Defizit der

- 3 Die Prozentsatz-Differenzen des erweiterten Gesprächsuniversums U_G gegenüber U sind abhängig vom relativen Zuwachs an Tweets, der im vorliegenden Fall gering ausfällt. Dies bedeutet zunächst, dass der äußere Gesprächsdruck auf U eher klein ist. Ungeachtet dessen können und müssen die gefundenen Unterschiede zwischen den Prozentsatzdifferenzen als themenspezifischer Einfluss der Gesprächskomponente innerhalb der Twitter-Kommunikation bewertet werden.
- 4 Berechnet man die Betragssumme der Abweichungen zwischen U^P und U_G ergibt sich ein Maß für divergente Thematisierungen in den gesprächsvernetzten Teilen des Twitter-Universums. Diese sind mit 3,0 im vorliegenden Fall sehr niedrig.

Twitter-Kommunikation, das durch Gesprächsaktivitäten kompensiert wurde. Das Thema hatte offenbar eine größere Relevanz, als es an den rein quantitativen Einlassungen im Twitter-Universum abzulesen ist. Möglicherweise erklärt dieses SPÖ-affine Thema, warum die Partei aufgrund des Finanzskandals zwar verloren hat, die Niederlage aber unerwartet gering ausfiel.

Weitere divergente Politikfelder, nun aber in umgekehrter Richtung, sind „Umweltpolitik“ und „innere Sicherheit“. In beiden Fällen sind die relativen Anteile im Ursprungsuniversum höher als im erweiterten Gesprächsuniversum. Bei diesen Themen zeigen sich Widerstände innerhalb der Twitter-Kommunikation gegenüber den hohen (oder allzu hohen) Input-Leistungen interessierter Kommunikatoren. Es mag kein Zufall sein, dass dies ausgerechnet die beiden Politikfelder betrifft, die mit den Grünen bzw. mit der FPÖ assoziiert sind, die – aus vermutlich unterschiedlichen Gründen – Probleme hatten, ihre Themen „viral“ durch Gesprächskanäle zu verbreiten.

Der direkte Vergleich der Politikfelder in Fragment- und Gesprächstweets bestätigt im Wesentlichen die Befunde des U-UG-Vergleichs. Die Unterschiede zwischen „typischen“ Gesprächsthemen, die kommunikative Anschlusshandlungen provozieren, und fragmentierten thematischen Einlassungen ohne Resonanz im Gesprächsraum treten hier sogar noch deutlicher in Erscheinung. Aus Platzgründen verzichten wir auf die Präsentation detaillierter Befunde ebenso wie die Befunde zur Unterscheidung von Initiationstweets und Resonanztweets, die allesamt, wenn auch je spezifisch, diskursive Dynamiken der politischen Twitter-Kommunikation abbilden.

Die Themenanalyse abschließend fragen wir danach, in welchem Verhältnis die Themenprioritäten in den Paralleluniversen der Politiker- und Nicht-Politiker-Tweets zueinander stehen. Das Ausmaß der Diskrepanzen gibt Auskunft über die Kluft, die zwischen dem politischen Diskurs mit den Profis und der Debatte außerhalb der berufspolitischen Zone besteht. Thematische Defizite des Profidiskurses auf Twitter sind vor allem bei den Themen „Verteidigungspolitik“ ($d\%=-11,1\%$), „Umweltpolitik“ ($d\%=-7,9\%$), und „Finanzpolitik“ ($d\%=-6,9\%$) belegt. Diese Themen spielen in Nicht-Politiker-Kreisen unter den Twitter-Usern eine größere Rolle, sei es, weil sie in ideologisch abgeschirmten Nischen diskutiert werden, sei es, weil die Majorität der Berufspolitiker sich einer Thematisierung auf Twitter widersetzt. Umgekehrt sind Themen mit einem Überhang im politischen Profidiskurs wie „innere Sicherheit“ ($d\%=7,3$), „Außenpolitik“

(d%=6,2%) und „Universitäts- und Forschungspolitik“ (d%=5,7%) Ausdruck dafür, dass die Themen-Promotion auf Zurückhaltung der Adressaten stößt. Bei diesen Politikfeldern dringt die Kommunikation der Profis auf Twitter kaum zu den Nicht-Politikern durch. Insgesamt finden sich Hinweise auf eine nicht unerhebliche *Themenkluft im politischen Twitter-Diskurs*. Immerhin summieren sich die Divergenzen zwischen Profi- und Journalisten-Laien-Tweets auf eine Betragssumme von 73,4 – weit höher als die thematischen Unterschiede zwischen den anderen Teil-Universen. Ein Vergleich mit Aggregaten der politischen Offline-Kommunikation müsste zeigen, ob die Themenkluft nach Ausmaß und Inhalt spezifisch für politische Online-Kommunikation ist oder aber eine grundsätzliche Spaltung in der politischen Diskurskultur darstellt.

6.3 *Problem- und lösungsorientierte Themenbehandlungsweisen in den Twitter-Universen*

Politikfeld-Thematisierungen können erfolgen, um ein Problem zu kennzeichnen und Handlungsdruck zu erzeugen. Sie können auch Lösungsvorschläge beinhalten oder aber beides: die Problemdefinition und die Lösung gemeinsam präsentieren. Aus der Art der Themenbehandlung lassen sich Rückschlüsse über die diskursive Qualität der politischen Twitter-Kommunikation gewinnen. Tabelle 7 gibt einen Überblick über die Art der Themenbehandlung, die zu jeder Politikfeldnennung im Twitteruniversum von Politikern- und Nicht-Politikern erhoben wurde. Außerdem werden die Veränderungen der Themenbehandlung, die sich am Wahltag gegenüber dem Wahlkampftag ergaben, dokumentiert. Aus der Tabelle geht hervor, dass der Profi-Nonprofi-Gap nicht nur die Themenschwerpunkte, sondern auch die Art der Themenbehandlung auf Twitter betrifft. Während die Berufspolitiker Themenfelder im Wahlkampf nur selten mit Problem- oder Lösungsbezug (35,5% der Tweets mit Politikfeldbezug) präsentieren, stehen bei den Nicht-Politikern die Problem-Lösungsperspektiven im Vordergrund (65,4%). Wir sehen darin einen Mangel des Profidiskurses auf Twitter.

Interessant ist die Verteilung nach Fragment- und Gesprächstweets. Die neutrale Präsentation ohne besonderen Modus der Themenbehandlung ist im fragmentierten Bereich um 5,5 Prozent höher ausgeprägt als im durch Gespräche vernetzten Teil des Twitter-Universums. Daraus lässt sich ableiten, dass die Twit-

Tabelle 7: Art der Themenbehandlung in den Twitter-Universen

	Tweets mit Politikfeldbezug	Wahlkampftag				Wahltag			
		kein Modus	problemorientiert	lösungsorientiert	problem- und lösungsorientiert	kein Modus	problemorientiert	lösungsorientiert	problem- und lösungsorientiert
		%	%	%	%	%	%	%	%
U^P	366	65,5	19,6	4,3	10,6	79,4	10,7	4,6	5,3
F_{UP}	85	69,5	22,0	5,1	3,4	96,2	3,8	0,0	0,0
G_{UP}	99	77,4	14,5	3,2	4,8	77,1	11,4	2,9	8,6
U^{NP}	329	24,6	57,3	11,7	6,4	48,1	34,2	3,8	13,9

U^P = Ursprungsuniversum der Politiker-Tweets, F_{UP} = Fragmenttweets, G_{UP} = Gesprächstweets, UG = Gesprächsuniversum. I_{UG} = Initiationstweets, R_{UG} = Resonanztweets, U^{NP} = Paralleluniversum der Nicht-Politiker-Tweets. Angaben in Prozent. Basis der Prozentuierung sind die korrespondierenden „Tweets mit Politikfeldbezug“ des jeweiligen Teil-Universums.

tergespräche von der Diskussion um Problem- und Lösungsperspektiven angetrieben werden. Sieht man diesen Befund im Zusammenhang des grundsätzlichen Mangels an problem- und lösungsorientierten Formen des politischen Diskurses von Berufspolitikern auf Twitter, dann ergibt sich ein weiteres Defizitmerkmal demokratischer Kommunikationskultur. Offenbar müssen die Problem-Lösungsperspektiven von den Nicht-Politikern über die Gesprächskanäle „zwangsweise“ in die Twitter-Diskussion eingebracht werden, weil sie von den Politikern eher ignoriert werden. Gerade die Laien tragen zur problem- und lösungszentrierten Debatte bei, die als *conditio sine qua non* einer nachhaltigen Diskursqualität in der Demokratie gelten muss.

7 Fazit und Ausblick

Die Twitter-Aggregatanalyse (TAA) erbrachte im Anwendungsfall der Salzburger Landtagswahlen durchweg plausibel interpretierbare Befunde, die zur Er-

klärung des Wahlausgangs ebenso wie zur Aufhellung besonderer Qualitäten der Twitter-Kommunikation beitragen. So werden gesprächsverstärkt auf Twitter genau diejenigen Themen gepusht, die wie der „Finanzskandal“ zur Niederlage der SPÖ führten. Als bedeutsam für das Funktionieren deliberativer Demokratie (Habermas 1999) können die Twitter-Einlassungen von Nicht-Politikern gelten, die teilweise andere Themenschwerpunkte setzen und vor allem die Lösungsorientierung mehr als die Politikprofis in den Mittelpunkt ihrer Kommunikationsaktivitäten stellen.

In methodischer Hinsicht hat sich die TAA bei der Lösung des Umfangs-, Subjektivitäts-, Reduktions- und Einbettungsproblems insofern bewährt, als sie gekennzeichnet ist durch:

- automatisierte Vorselektion kombiniert mit manueller Reduktion,
- systematische Variation von Twitter-Universen und Kontrolle der Differenzen,
- Standardisierung von Vergleichsprozeduren,
- Einbeziehung des Gesprächscharakters der Twitter-Kommunikation,
- Berücksichtigung der Intertextualität innerhalb und außerhalb des Internet.

Die Anwendung automatisierter Sampling-Techniken und die Standardisierung der Analyse ist eine wesentliche Voraussetzung, um den Herausforderungen der Datenmengen und den Ansprüchen auf Nicht-Willkürlichkeit des Vorgehens gerecht zu werden. Bei der Analyse der formalen Vernetzungsparameter zeigte sich überdies, dass die TAA prinzipiell auch mit automatisierter Inhaltsanalyse sinnvoll kombiniert werden kann, um die Vernetzung der Tweets in Teilaspekten zu quantifizieren und im Rahmen einer Kommunikationsstil-Analyse fruchtbar werden zu lassen. Im Kernbereich der Themenanalyse sehen wir allerdings bislang noch keine Möglichkeit, auf konventionelle Inhaltsanalysen zu verzichten. Eine Schwierigkeit stellen die häufig telegrammartigen oder sogar elliptischen Aussagen in Tweets dar, die sich erst im Zusammenhang mit Gesprächsketten dem Codierer bzw. der Codiererin ihrem Sinn nach erschließen. Diese die Codierung komplizierende Schwierigkeit kann mit Hilfe von Wortlisten, wie sie in automatisierten Inhaltsanalysen üblich sind, bis heute nicht befriedigend gelöst werden. Dies schließt nicht aus, dass in Zukunft dieser Bereich bei Anwendung entsprechender Validierungsverfahren (z. B. Vergleich zwischen automatisierter und konventioneller Inhaltsanalyse) ausgedehnt werden kann und sollte.

Als außerordentlich fruchtbar hat sich die *Interaktionsanalyse* von Gesprächen innerhalb der TAA erwiesen. So ergaben sich themenbezogene Initiationshandlungen und Resonanzen, die sich als prozessuales Agenda-Setting innerhalb der Twitter-Kommunikation deuten lassen. Durch Akteursspezifizierungen nach Politikern, Journalisten und Laien können die Dynamiken weitergehend ausgeleuchtet werden, was im vorliegenden auf die Erläuterung der Methode und die Konstruktion der Twitter-Universen ausgerichteten Aufsatz aus Platzgründen entfiel. Hinreichend deutlich wurde bereits, dass Politiker beim Agenda-Building im Netz typische Defizite vor allem im Hinblick auf die Lösungsorientierung bei der Themenbehandlung aufweisen, die sie von den stärker auf Lösungen ausgerichteten Laien unterscheidet. Weitere Forschung muss zeigen, inwieweit es zivilgesellschaftlichen Akteuren gelingt, zur politischen Diskurskultur produktiv beizutragen.

Von der *Intertextualitätsanalyse* versprechen wir uns Aufschlüsse über die Schnittstelle von Offline- und Online-Kommunikation, um die Rolle der Massenmedien für die Twitter-Kommunikation besser zu verstehen. Diese kann und sollte zu einer Input-Output-Analyse in beide Richtungen ausgebaut werden, um die Rückkopplungen der Online-Kommunikation auf die massenmediale Offline-Kommunikation zu explorieren. Und schließlich lässt sich die TAA auch ohne Weiteres im Rahmen umfassender Agenda-Setting-Studien nutzen, die bislang auf die klassischen Massenmedien fokussiert bleiben. Die Aggregatbildung in methodisch kontrollierter Weise zu leisten ist der Schlüssel zur Vernetzung der Twitterforschung mit den Offline-Analysen der Massenkommunikationsforschung.

Vor unreflektierten Aggregierungen sei allerdings ebenso gewarnt wie vor einer unkritischen Automatisierungseuphorie. Mahrt und Scharnow (2014) weisen zu Recht darauf hin, dass die Twitterforschung sich zuweilen in allzu experimentierfreudigen Methoden-Applikationen und Innovationsforderungen verliert. Die methodischen Standards der empirischen Sozialforschung wie Transparenz, Intersubjektivität und Reproduzierbarkeit der Ergebnisse gelten selbstverständlich auch im Bereich der Twitterforschung. Und auch die größten Datensammlungen im Ausmaß von Big Data unterliegen spezifischen Restriktionen, die gerade im Bereich von API generierten Tweet-Listen durch kommerziell ausgerichtete Anbieter (boyd & Crawford 2012) verursacht werden. Dies alles muss bei der Dateninterpretation mitbedacht werden, die auf unabsehbare Zeit auf methodisch-kritische Reflexion und empirische Kontroll-Untersuchungen mit Hilfe

konventioneller manueller Verfahren angewiesen bleibt. Reliabilitätsprüfungen und Validitäts-Checks sind der Lackmus-Test, an dem sich die wissenschaftliche Qualität der Twitterforschung erweist. Im Hinblick auf die Validität der TAA könnte dies u. a. dadurch geschehen, dass man nacheinander verschiedene Twitter-Universen in zeitlich erstreckte Untersuchungsdesigns t_1, t_2, \dots, t_n integriert und sie mit Paralleluniversen im Offline-Bereich verknüpft, z. B. mit Themenhierarchien in den Massenmedien und Themenhierarchien im Elektorat zu den Zeitpunkten t_1, t_2, \dots, t_n . Mit Hilfe von Strukturgleichungsmodellen kann man dann die besten Online-Prädiktoren für Agenda-Setting-Effekte im Elektorat ermitteln, unter Berücksichtigung der Wechselwirkung mit den Offline-Medien, versteht sich. Auf dieser Grundlage wiederum lassen sich Rückschlüsse auf die Konstruktion von Twitter-Universen ziehen, die die Prognosevalidität verbessern.

Eine Eliminierung von Subjektivität ist hier allerdings auch auf diesem Weg nicht vollumfänglich möglich, aber doch eine Ausdehnung der intersubjektiven Zone. Im Übrigen geht es methodisch gar nicht um die Ausmerzung, sondern um die Kontrolle der Subjektivität, die eine wesentliche Ressource für Verstehensleistungen darstellt. Auch die formale Strenge von Differenzmatrizen innerhalb der TAA kommt nicht ohne ein erhebliches Maß an Verstehensleistung aus. Verstanden werden müssen nicht nur die zuweilen kryptischen Einzeltweets, sondern auch die Gesamtheit einer Themenhierarchie, deren Bedeutung für den Wahlausgang sich ohne hinreichende analytische Abstraktion und interpretative Kompetenz kaum erschließt.

Eine Stärke der TAA sehen wir darin, dass systematische Variationen von Twitter-Universen eine Gesamtschau von Datenmengen in Abhängigkeit von Selektions- und Konstruktionsprinzipien ermöglichen und ganzheitliches Denken kombiniert mit methodischer Strenge erfordern und fördern. Die Differenzlogik mildert überdies die Störvarianz, die durch nicht erkannte Einflüsse bei der Bereitstellung der Twitter-Daten entsteht. Denn es gibt bislang keinen Grund zu der Annahme, dass der durch Anbieter-Algorithmen erzeugte Bias in den Daten die politischen Twitter-Universen in unterschiedlicher Weise affiziert. Die Differenzen bleiben davon relativ unberührt unabhängig davon, ob sie durch konventionelle Inhaltsanalyse oder durch automatische Verfahren gewonnen wurden (King & Lowe 2003). Sie können jedenfalls als „robuster“ im Sinne eines Konzepts der Begrenzung von Irrtumswahrscheinlichkeiten gelten als dies für einfache Häufigkeiten innerhalb ein und desselben Aggregats der Fall sein mag.

Die TAA ist sowohl deduktiv als auch induktiv einsetzbar. Letzteres erscheint uns im Stadium der Exploration und Hypothesengenerierung für die Twitterforschung ein wesentlicher Vorzug zu sein. Ohne induktive Komponente bliebe die TAA eine oberflächliche Registratur von Quantitäten, die das Erkenntnispotenzial in Anbetracht von *Pluriversen der politischen Online-Kommunikation* nur unvollkommen ausschöpft. Gefragt ist daher eine methodisch kontrollierte *Hermeneutik der Aggregatanalyse*, offen für mathematische Formalisierungen ebenso wie für verständnisfördernde Interpretationshypothesen – sei es im Interesse ganzheitlicher Betrachtungen, sei es, um weiterer, deduktiv ausgerichteter Forschung zum Ausgangspunkt zu dienen.

Prof. Dr. Jürgen Grimm ist Professor am Institut für Publizistik- und Kommunikationswissenschaft der Universität Wien

Christiane Grill, MMag. ist wissenschaftliche Mitarbeiterin und Prae-doc am Institut für Publizistik- und Kommunikationswissenschaft der Universität Wien

Quellenverzeichnis

- Ausserhofer, J., & Maireder, A. (2013). National politics on Twitter. Structures and topics of a networked public sphere. *Information, Communication & Society*, 16(3), 291-314. doi: 10.1080/1369118X.2012.756050
- boyd, d., & Crawford, K. (2012). Critical questions for big data. *Information, Communication & Society*, 15(5), 662-679. doi: 10.1080/1369118X.2012.678878
- Bruns, A., & Stieglitz, S. (2012). Quantitative approaches to comparing communication patterns on Twitter. *Journal of Technology in Human Services*, 30(3-4), 160-185. doi: 10.1080/15228835.2012.744249
- Cheong, F., & Cheong, C. (2011). Social media data mining: a social network analysis of tweets during the 2010-2011 Australian floods. *Proceeding of PACIS 2011*. Abgerufen von <http://aisel.aisnet.org/pacis2011/46>
- Früh, W. (1981). *Inhaltsanalyse. Theorie und Praxis*. München: Ölschläger
- Grimm, J. (1989). Inhaltsanalyse. In J. Bellers & W. Woyke (Hrsg.), *Methoden zur Erforschung der internationalen Beziehungen* (S.170-180). Opladen: Westdeutscher Verlag.

- Grimm, J., Schmidt, P., Kneidinger, B., & Winkelnkemper, P. (2008). Individualisierung oder Aggregation? Zur Anwendung von Strukturgleichungsmodellen im Rahmen der Agenda Setting-Forschung. *Medien Journal*, 2, 7-52.
- Habermas, J. (1999/1996). *Die Einbeziehung des Anderen. Studien zur politischen Theorie*. Frankfurt am Main: Suhrkamp.
- Highfield, T. (2012). Talking of many things: using topical networks to study discussions in social media. *Journal of Technology in Human Services*, 30(3-4), 204-218. doi: 10.1080/15228835.2012.746894
- Hoyle, R. H. (2012). *Handbook of structural equation modelling*. New York: Guilford Press.
- King, G., & Lowe, W. (2003). An automated information extraction tool for international conflict data with performance as good as human coders: A rare events evaluation design. *International Organization*, 57(3), 617-642. doi: 10.1017/S0020818303573064
- Mahrt, M., & Scharkow, M. (2014). Der Wert von Big Data für die Erforschung digitaler Medien. In R. Reichert (Hrsg.), *Big Data. Analysen zum digitalen Wandel von Wissen, Macht und Ökonomie* (S. 221-237). Bielefeld: transcript Verlag.
- Maireder, A., & Schwarzenegger, C. (2012). A movement of connected individuals. *Information, Communication & Society*, 15(2), 171-195. doi: 10.1080/1369118X.2011.589908
- McCombs, M. E. (2004). *Setting the agenda. The mass media and public opinion*. Cambridge: Polity.
- McCombs, M. E., & Shaw, D. L. (1972). Agenda Setting. *The Public Opinion Quarterly*, 36(2), 177-187. doi: 10.1086/267990
- Moorstedt, T. (2008). *Jeffersons Erben. Wie die digitalen Medien die Politik verändern*. Frankfurt am Main: Suhrkamp.
- Plehwe, K. (2008). *Von der Botschaft zur Bewegung. Die 10 Erfolgsstrategien des Barack Obama*. Hamburg: Hanseatic Lighthouse.
- Reinecke, J. (2005). *Strukturgleichungsmodelle in den Sozialwissenschaften*. München, Wien: Oldenbourg.
- Schenk, M. (2007). *Medienwirkungsforschung* (3., überarb. Aufl.). Tübingen: Mohr.
- Schmitt-Beck, R., & Mackenrodt, C. (2010). Social networks and mass media as mobilizers and demobilizers: A study of turnout at a German local election. *Electoral Studies*, 29(3), 392-404. doi: 10.1016/j.electstud.2010.03.011

- Thimm, C., Dang-Anh, M., & Einspänner, J. (2011). Diskurssystem Twitter: Semiotische und handlungstheoretische Perspektiven. In M. Anastasiadis & C. Thimm (Hrsg.), *Social Media - Theorie und Praxis digitaler Sozialität* (S. 178-223). Frankfurt am Main: Peter Lang.
- Thimm, C., Einspänner, J., & Dang-Anh, M. (2012). Twitter als Wahlkampfmedium: Modellierung und Analyse politischer Social-Media-Nutzung. *Publizistik*, 57(3), 293-314. doi: 10.1007/s11616-012-0156-7
- Weller, K. (2014). Twitter und Wahlen. Zwischen 140 Zeichen und Milliarden von Tweets. In R. Reichert (Hrsg.), *Big Data. Analysen zum digitalen Wandel von Wissen, Macht und Ökonomie* (S. 239-257). Bielefeld: transcript Verlag.

Empfohlene Zitierung: Sattelberger, F., & Seufert, W. (2015). Lassen sich mit Social-Media-Analysen die Absatzprognosen für Vertrauensgüter verbessern? Wechselwirkungen bei der Kommunikation über Filme vor dem Kinostart. In A. Maireder, J. Ausserhofer, C. Schumann, & M. Taddicken (Hrsg.), *Digitale Methoden in der Kommunikationswissenschaft* (S. 197-221). doi: 10.17174/dcr.v2.9

Zusammenfassung: Die vorliegende Untersuchung präsentiert am Beispiel des deutschen Kinomarkts die Ergebnisse eines Screening-Ansatzes verschiedener Social-Media-Metriken und -Kanäle. Ziel der Studie ist es zum einen, jene Kommunikationskanäle zu identifizieren, die eine signifikante Absatzwirkung aufweisen. Weiterhin werden die Wechselwirkungen zwischen dem Kommunikationsumfang einzelner Kommunikationskanäle im Zeitverlauf untersucht. Zu diesem Zweck werden autokorrelative sowie kreuzverzögerte Effekte für einen Zeitraum von zwölf Wochen vor dem Kinostart synchron modelliert. Bislang werden Kommunikationsprozesse vor der Markteinführung in Erfolgsfaktorenmodellen für Medienprodukte – ebenso wie in der Diffusionsforschung generell – vernachlässigt. Der Ansatz erlaubt es, Pfadabhängigkeiten im zeitlichen Verlauf unterschiedlicher Kommunikationsangebote zu untersuchen und gibt einen Einblick in deren Wachstumsprozess. Durch Reduktion auf relevante Kommunikationsindikatoren und deren Ausprägung zu frühen Zeitpunkten lässt sich Social-Media-Kommunikation als zusätzliche erklärende Variable in Prognosemodelle für Medienprodukte und andere Vertrauensgüter integrieren.

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

Felix Sattelberger & Wolfgang Seufert

Lassen sich mit Social-Media-Analysen die Absatzprognosen für Vertrauensgüter verbessern?

Wechselwirkungen bei der Kommunikation über Filme vor dem Kinostart

1 Problemstellung

Erfahrungs- und Vertrauensgüter weisen für den Käufer vor der Nutzung ein hohes Maß an Informationsunsicherheit auf. Dies gilt auch für alle Unterhaltungsmedien mit Unikatcharakter (z.B. Romane, Spielfilme, Tonträger), die überwiegend als *paid content* angeboten werden. So kann im Kinomarkt, mit dem sich dieser Aufsatz beschäftigt, beispielsweise erst nach der Rezeption beurteilt werden, inwiefern ein Kinofilm die Erwartungen des Rezipienten erfüllt hat. In derartigen Entscheidungssituationen mit asymmetrischer Informationsverteilung zwischen Verkäufer und Käufer über die Produktqualität unterscheidet die Ökonomie zwei grundlegende Strategien zum Abbau des Informationsdefizits der Nachfrager: Das Qualitäts-*Signalling* durch Anbieter (z.B. den Aufbau von Medienmarken) und das Qualitäts-*Screening* durch Rezipienten. Innerhalb der Screening-Strategie sind Prozesse interpersonaler sowie computervermittelter interpersonaler Kommunikation von zentraler Bedeutung. In der ökonomischen

Diffusionsforschung (u.a. Rogers, 1962; Dogruel, 2013), die sich mit der Verbreitung von Produktinnovationen als einer besonderen Klasse von Vertrauensgütern beschäftigt, wird beispielsweise der Mundpropaganda (*word-of-mouth*, WOM) und damit der interpersonalen Kommunikation zwischen Konsumenten ein stärkerer Einfluss auf die Diffusionsgeschwindigkeit zugesprochen als den Marketing-Maßnahmen der Anbieter.

Die interpersonale Kommunikation zwischen den Rezipienten wird im Gegensatz zu den Marketing-Aktivitäten der Filmanbieter im Rahmen von Erfolgsfaktoren- und Absatz-Prognosemodellen für Kinofilme bislang nur eingeschränkt berücksichtigt. Wenn dies der Fall ist, dann bezieht man in der Regel lediglich WOM-Effekte aus der Phase nach dem Kinostart ein (Duan, Gu, & Whinston, 2008; Hennig-Thurau, Houston, & Walsh, 2007; Dellarocas, Zhang, & Awad, 2007; Chang & Ki, 2005). Nur die Untersuchung von Liu (2006) bildet eine Ausnahme, wobei dort der Zeitraum vor dem Kinostart als Aggregat des Kommunikationsaufkommens bis zu der Woche direkt vor dem Kinostart in das Modell eingeht. Die Auswahlkriterien für die verwendeten Kommunikationsindikatoren werden dabei meist nicht explizit dargelegt oder begründet. Ein systematischer Vergleich sowie eine strukturierte Bewertung verschiedener Indikatoren im Zeitverlauf finden ebenfalls nicht statt.

Die vorliegende Untersuchung präsentiert am Beispiel des deutschen Kinomarkts die Ergebnisse eines Screening-Ansatzes zur Beobachtung computervermittelter *institutioneller* und *interpersonal-öffentlicher Kommunikation* mithilfe verschiedener Social-Media-Metriken. Ziel dieser Studie ist es, diejenigen Kommunikationsindikatoren zu identifizieren und im Zeitverlauf in ihren Wechselbeziehungen zu untersuchen, für die in Einzelanalysen eine signifikante Absatzwirkung festgestellt werden kann. Zu diesem Zweck werden autokorrelative sowie kreuzverzögerte Effekte für einen Zeitraum von bis zu zwölf Wochen vor dem Kinostart synchron modelliert.

Institutionalisierte Kommunikationsprozesse umfassen hier im Internet verbreitete Kritiken auf spezialisierten Kinofilmportalen sowie auf Online-Ablegern klassischer Medienunternehmen. Der Begriff *interpersonal-öffentliche Kommunikation* soll den Besonderheiten im Internet verbreiteter Nutzerkommunikation Rechnung tragen, bei der es zu einer Verschmelzung zwischen interpersonaler und Massenkommunikation kommt (Haas, Keyling, & Brosius, 2010, S. 249). Sie umfasst hier die Kommunikation in sozialen Netzwerkplattformen über Kino-

filme. Der Ansatz erlaubt Pfadabhängigkeiten zwischen der Kommunikation auf unterschiedlichen Kanälen zu untersuchen und gibt einen Einblick in den Wachstumsprozess einzelner handlungsrelevanter Kommunikationsprozesse sowie deren Wechselwirkungen. Starke Autokorrelationen verdeutlichen hierbei, dass ein Indikator durch dessen vorangegangene Ausprägungen determiniert ist. In ähnlicher Weise erlauben kreuzverzögerte Effekte eine Schätzung durch die früheren Ausprägungen der anderen Kommunikationsvariablen. Liegen entsprechende Zusammenhänge vor, kann für eine zuverlässige Modellierung auf frühere Zeitpunkte sowie eine geringere Anzahl an Variablen zurückgegriffen werden. Dies reduziert wiederum den Aufwand, um beispielsweise im Rahmen eines Web-Monitoring-Ansatzes eine Methode zur frühen Erfolgsabschätzung zu etablieren. Der Ansatz erweitert somit bestehende, meist betriebswirtschaftlich geprägte, Nachfragemodelle zu Kinofilmen um die Einflussfaktoren der computervermittelten institutionalisierten und interpersonal-öffentlichen Kommunikation vor dem Kinostart. Bislang wird diese in Erfolgsfaktorenmodellen zu Medienprodukten – ebenso wie in der Diffusionsforschung generell – vernachlässigt.

Die Entscheidungssituation eines Käufers von ihm bis dahin unbekanntem Produkten ähnelt der Entscheidungssituation eines Bürgers bei politischen Wahlen und Abstimmungen. Auch im Bereich der politischen Kommunikation gibt es deshalb Signalling- und Screening-Strategien, um die bestehenden Informationsdefizite zu überwinden. Analog zur Produktkommunikation kann ebenfalls nicht davon ausgegangen werden, dass allein die Anbieterkommunikation (z.B. durch politische Wahlwerbung) oder die von Journalisten aufbereitete institutionelle Kommunikation das Wahlverhalten beeinflussen, sondern dass hierfür auch zusätzliche interpersonale Kommunikationsprozesse (z.B. in Form der Anschlusskommunikation) zwischen Wählern relevant sind (u.a. Sommer, 2010). In der Kommunikationswissenschaft wird diesem Aspekt unter anderem mit dem Meinungsführerkonzept in der Tradition des Two-Step-Flow of Communication Rechnung getragen (u.a. Katz & Lazarsfeld, 1955; Katz, 1957; Lazarsfeld, Berelson, & Gaudet, 1944). Der im Folgenden vorgestellte Ansatz zur Beobachtung von internetbasierter institutioneller und interpersonal-öffentlicher Kommunikation und deren Wechselwirkungen bietet deshalb in methodischer Perspektive nicht nur Transferpotenzial zur Untersuchung des Einflusses von interpersonaler Kommunikation auf den Absatz anderer Vertrauensgüter als Kinofilme, sondern auch für eine Anwendung in der politischen Wahlforschung (siehe Abschnitt 8).

2 Vertrauensgüter in der Kommunikationswissenschaft

Die Informationsökonomie teilt Güter danach ein, inwiefern Rezipienten in der Lage sind, zentrale Charakteristika, etwa Qualität und Nutzen, vor oder nach der Rezeption zu beurteilen. Je nach Ausprägung der Informationsunsicherheit wird in drei Güterarten unterschieden. Während Inspektionsgüter in der Regel Sachleistungen sind, deren qualitative Eigenschaften bereits vor dem Kauf beurteilt werden können, lassen sich Medienprodukte den Erfahrungs- und Vertrauensgütern zuordnen (Kiefer, 2005, S. 141-142). Kinofilme, die den Gegenstand dieser Untersuchung bilden, sind den Erfahrungsgütern zuzuordnen. Die Kinobesucher sind nach der Rezeption in der Lage, den Grad der Bedürfnisbefriedigung zu beurteilen. Dokumentarfilme oder TV-Nachrichten sind dagegen Vertrauensgüter, da ihr Wahrheitsgehalt auch nach der Rezeption nicht beurteilt werden kann. Eine rationale Kosten-/Risiko- und Nutzenabwägung, wie sie beispielsweise dem Konzept des *homo oeconomicus* zugrunde liegt, ist bei Erfahrungs- und Vertrauensgütern wegen dieser unsicheren Informationsbasis ex ante nicht möglich. Dies kann zu Marktversagen führen, sodass bestimmte Mediengüter trotz zahlungsbereiter Nachfrage nicht hergestellt werden (eingeschränkte produktive Effizienz) oder zahlungsbereite Rezipienten keinen Zugang zu bestimmten Mediengütern erhalten (eingeschränkte alloкатive Effizienz; Kiefer, 2005, S. 142).

Um das Informationsdefizit auszugleichen, können Rezipienten gezielt nach Informationen suchen, Expertenmeinungen hinzuziehen oder mit anderen Nutzern kommunizieren. Die Phase vor der Markteinführung ist hierbei entscheidend, da zahlreiche nicht-tagesaktuelle Medienprodukte, wie beispielsweise der Musikmarkt (Asai, 2009, S. 85) oder der Kinofilmmarkt (Sattelberger, 2012, S. 48) zur Einführung den höchsten Absatz aufweisen. Die Absatzmenge nimmt danach exponentiell ab, sodass klassische Effekte interpersonaler Kommunikation nach der Produkteinführung, wie sie im Rahmen der Diffusionstheorie (u.a. Rogers, 1962) angenommen werden, nicht auf diese Medienprodukte anwendbar sind.

3 Interpersonale Kommunikationsprozesse in der Diffusionsforschung

Die im Rahmen der Diffusionsforschung postulierten Dynamiken sind erfolgreich auf technische Medieninnovationen angewendet worden (Dimmick & Wang,

2005). Gemäß der idealtypischen Annahme der Diffusionstheorie folgen Innovationen einem logistischen, „s-förmigen“, Verlauf. Der Wachstumsprozess setzt sich aus einer Kombination innovativer und imitativer Effekte zusammen (Rogers, 1962). Erstere umfassen die objektiven Eigenschaften einer Innovation sowie die Marketingkommunikation, während imitative Wirkungen auf interpersonale Kommunikationsprozesse beruhen. Die Verbreitung des Wissens um die Innovation erfolgt durch „Early adopters“, die ihre Erfahrung weitergeben und dadurch zur Imitation dieses Verhaltens anregen. Diese Dynamik entfaltet einen Selbstläufereffekt, der den exponentiellen Anstieg der Adaptionsrate auslöst (u.a. Bass, 1969; Rogers, 1962; Leibenstein, 1950).

Die Diffusionsforschung impliziert daher ein komplementäres und asynchrones Verhältnis zwischen innovativen und imitativen Effekten sowie zwischen massenmedialen und interpersonale Kommunikationsprozessen. Allerdings ist diese strikte Abgrenzung in der modernen Medienumwelt nicht mehr aufrecht zu erhalten. Wie der Begriff der interpersonal-öffentlichen Kommunikationsprozesse bereits impliziert, verschwimmen die Grenzen zwischen interpersonaler und Massenkommunikation zunehmend. Es resultiert folglich ein Kontinuum zwischen technisch vermittelter Individualkommunikation, über Kommunikation auf öffentlichen Social-Media-Plattformen, die Kriterien interpersonaler Kommunikation aufweisen und dennoch einer potenziell unbegrenzten Leserschaft bereitgestellt werden können, bis hin zu massenmedialen Angeboten.

Die veränderten technischen Rahmenbedingungen dieser Kommunikationsprozesse beeinflussen auch die Dynamik des Diffusionsprozesses. Die Annahme eines komplementären und asynchronen Verhältnisses zwischen innovativen und imitativen Effekten ist für Vertrauensgüter vor der Markteinführung nicht ohne weiteres plausibel. Wie zuvor beschrieben, weisen einige nicht-tagesaktuelle Medienprodukte nach der Markteinführung einen exponentiellen Zerfall in der Nachfrage auf. Während in Diffusionsmodellen das starke Absatzwachstum auf einsetzende interpersonale Kommunikationseffekte zurückgeführt wird, können vergleichbare Dynamiken für Medienprodukte mit Vertrauenseigenschaften in der Phase nach der Markteinführung nicht festgestellt werden (Asai, 2009, S. 85; Sattelberger, 2012, S. 48). Aufgrund der kurzen Lebenszyklen im Primärmarkt (bei Kinofilmen der Kinomarkt) sowie des abnehmenden Absatzpotenzials, kann der grundlegende Diffusionsverlauf nach der Einführung weder anbieter- noch rezipientenseitig grundlegend beeinflusst werden. Bezüglich des Verhältnisses

zwischen interpersonaler und Massenkommunikation sind folglich verschiedene Einflussrichtungen denkbar:

- Bestünde weiterhin, wie von der Diffusionsforschung impliziert, ein komplementäres und asynchrones Verhältnis zwischen interpersonaler und Massenkommunikation, so wäre die Kinofilmnachfrage vor allem durch die Marketingkommunikation des Anbieters sowie die Berichterstattung in Massenmedien vor dem Kinostart determiniert. Interpersonale Kommunikationsprozesse basierend auf der Rezeptionserfahrung nach dem Kinostart hätten hingegen aufgrund des Absatzverlaufs von Kinofilmen keinen entscheidenden Einfluss auf die Nachfrage. Diese Annahmen sind jedoch nicht plausibel, da in diesem Falle der Erfolg eines Kinofilmes ausschließlich von der Höhe des Marketingbudgets abhinge.
- Eine andere Möglichkeit ist, dass vor allem „Early Adaptors“ den Film bereits vor dem Kinostart illegal im Internet sehen, ihre Rezeptionserfahrung kommunizieren und dadurch positiv zum Erfolg des Filmes beitragen (Peukert, Claussen, & Kretschmer, 2012). Vor dem Kinostart sind die Möglichkeiten, den Film legal zu sehen, auf Presse- und Testvorführungen sowie Vorpremieren (Sneak Preview Nights) begrenzt und daher nur einem kleinen Personenkreis zugänglich.
- Wahrscheinlicher sind jedoch Wechselwirkungen zwischen institutionalisierten und interpersonalen Kommunikationseffekten vor dem Kinostart, wobei letztere vor allem auf den Erwartungen der potenziellen Rezipienten beruhen. Diese Kommunikate können in Online-Medien erhoben und für die Analyse nutzbar gemacht werden. Neben dem Einfluss institutionalisierter Kommunikation auf die interpersonale Kommunikation, beispielsweise durch Filmkritiken, ist auch die umgekehrte Wirkungsrichtung möglich, falls bestimmte Filme dann mit höherer Wahrscheinlichkeit rezensiert werden, wenn sich viele potenzielle Rezipienten dafür interessieren und darüber kommunizieren (Clement & Sambeth, 2004, S. 112).

Der vorliegenden Studie liegt diese Annahme einer im Zeitverlauf wechselseitigen Beeinflussung zwischen institutionalisierten und interpersonal-öffentli-

chen Kommunikationsprozessen zugrunde. Um diese Dynamiken untersuchen zu können, wurde ein mehrstufiges Auswertungsverfahren gewählt, wobei im Folgenden vor allem die Analyse der Wechselwirkungen zwischen institutioneller und interpersonaler Kommunikation über Kinofilme detailliert vorgestellt wird.

4 Datenerhebung und Variablenbeschreibung

Der Datensatz umfasst alle Filme mit Erstaufführung in deutschen Kinos zwischen dem 26. Juni und dem 27. Dezember 2012 ($n=234$; $n=216$ vollständige Datensätze). Im Zeitraum zwischen dem 1. Mai und dem 27. Dezember 2012 wurden die entsprechenden Indikatoren zum Umfang institutionalisierter und interpersonal-öffentlicher Kommunikation erhoben. Die geplanten Starttermine wurden den Filmankündigungen durch die *Spitzenorganisation der Filmwirtschaft e.V.* (SPIO; spio.de) und den *Verband der Filmverleiher e.V.* (VdF; vdfkino.de) entnommen und wöchentlich aktualisiert. Für jeden dieser Filme wurden die Daten zu institutionalisierten und interpersonal-öffentlichen Kommunikationsprozessen für die zwölf Wochen vor dem Kinostart wöchentlich zur gleichen Zeit automatisiert als Screenshot erhoben. Während der automatischen Datenerhebung wurde sowohl eine Bilddatei im JPEG-Format als auch eine PDF-Datei gespeichert.

Das Verfahren lässt so potenziell ein maschinelles Auslesen von Textelementen zu, wenngleich dies in der vorliegenden Untersuchung nicht erfolgte. Der Vorteil besteht darin, dass die Seiten exakt so archiviert werden, wie diese für einen Nutzer zum entsprechenden Zeitpunkt angezeigt wurden. Der Ansatz ist dadurch im Vergleich zu einer rückwirkenden Erhebung nicht der Gefahr einer Konfundierung durch anbieterseitige Modifikationen im Zeitverlauf ausgesetzt. Allerdings waren bei der Datenerhebung auch manuelle Bearbeitungsschritte erforderlich, da die Daten nicht in tabellarischer Form exportiert werden konnten. Über stichprobenartige Kontrollen wurde geprüft, ob die entsprechenden Daten korrekt übertragen wurden. Ferner wurden für jede Seite die Social Plugins von Facebook über die bereitgestellte Programmierschnittstelle (API; Application Programming Interface) automatisch erhoben und in tabellarische Form überführt.

Die längsschnittliche Datenerhebung erlaubt eine Analyse der Wachstumsverläufe sowie der Wechselwirkungen zwischen den Indikatoren im Zeitverlauf. Um

die Komplexität der Wachstumskurvenmodelle zu reduzieren und dennoch eine zuverlässige Kurvenschätzung zu ermöglichen, wurden die wöchentlichen Daten zu vier Zeiträumen zusammengefasst: Zwölf und elf Wochen vor dem Kinostart, acht und sieben Wochen vor dem Kinostart, vier und drei Wochen vor dem Kinostart sowie zum Kinostart.

Die Auswahl der Social-Media-Portale selbst erfolgte auf Basis ihrer Reichweite. Bezogen auf die Nutzerzahl in Deutschland ist Facebook Marktführer im Bereich der sozialen Netzwerkplattformen (statista, 2014a), Twitter im Bereich der Micro-Blogging-Plattformen (statista, 2014b) und YouTube im Bereich der Portale für audio-visuelle Inhalte (statista, 2014c). Für Facebook-Fanseiten und YouTube-Trailer wurden jeweils die drei nutzerstärksten filmspezifischen Angebote zum Zeitpunkt zwölf Wochen vor dem Kinostart berücksichtigt, um der Popularität der Filme besser Rechnung tragen zu können. Auf diese Weise war es möglich, Filme mit hohem Kommunikationsaufkommen über mehrere Angebote höher zu gewichten als Filme, für die beispielsweise nur eine Facebook-Fanseite existiert. Gab es zum Zeitpunkt zwölf Wochen vor dem Kinostart weniger als drei Angebote, so wurde für diese Filme wöchentlich geprüft, ob neue Seiten veröffentlicht wurden und entsprechend ergänzt werden konnten. Dadurch konnte der Entstehungs- und Wachstumsprozess der filmspezifischen Kommunikationsprozesse modelliert werden.

Pro Film ergeben sich folglich für Facebook-Fanseiten und YouTube-Trailer jeweils drei Angebote zu zwölf Zeitpunkten (je 36 Datenpunkte). Facebook-Fanseiten oder YouTube-Trailer, die von Filmvermarktern erstellt wurden, wurden hierbei ebenso berücksichtigt, wie von Nutzern erstellte Seiten. Hinzu kommt die wöchentliche Erhebung der Twittermeldungen sowie der filmspezifischen Seiten auf den drei Filmportalen: filmstarts.de, kino.de und moviepilot.de. Diese gingen in aggregierter Form in die Untersuchung ein.

Für die Auswahl der Online-Filmportale wurde die Reichweite, repräsentiert durch die Anzahl der Seitenbesuche pro Monat, herangezogen. Diese Daten werden von der *Informationsgemeinschaft zur Feststellung der Verbreitung von Werbeträgern e.V.* (IVW; ivw.eu) für die bei ihr gemeldeten Online-Angebote ausgewiesen. Folglich werden vor allem Reichweitendaten für kommerzielle, werbefinanzierte Anbieter durch die IVW erhoben. Aufgrund der Vielzahl der jährlich erstaufgeführten Filme in deutschen Kinos ist diese Einschränkung für die vorliegende Untersuchung unkritisch, da vor allem kommerzielle Anbieter über

die Ressourcen verfügen, um für eine große Anzahl an Filmen Bewertungen und Kritiken zu veröffentlichen. Ferner wurden nur Angebote mit inhaltlichem Schwerpunkt auf Kinofilme sowie eigenen redaktionellen Inhalten berücksichtigt. Internetseiten von Kinobetreibern, Videodiensten oder Rundfunkveranstaltern wurden nicht integriert, um die Unabhängigkeit der redaktionellen Beiträge zu gewährleisten.

Die Aktivität auf Facebook-Fanseiten, Online-Filmportalen sowie YouTube-Trailern zum entsprechenden Film wurde als Summenwert operationalisiert, der die Anzahl der Empfehlungen („Likes“), Weiterleitungen („Shares“) und Kommentierungen („Comments“) umfasst (siehe Tabelle 1 auf der nächsten Seite). Die Metriken wurden jeweils für die gesamte Seite, nicht jedoch separat für einzelne Posts auf dieser Seite berücksichtigt. Darüber hinaus gingen die Anzahl der Personen, die auf Facebook-Fanseiten über den Film sprechen, in das Modell mit ein. YouTube weist zudem weitere eigene Kennzahlen aus, die ebenfalls integriert wurden: die Anzahl positiver und negativer Bewertungen, die Anzahl der Nutzerkommentare sowie die Abrufzahlen der Trailer. Ferner wurden die Anzahl der Twittermeldungen zum Film, die Anzahl der Nutzerwertungen sowie die Anzahl der Kommentare auf Online-Filmportalen berücksichtigt. Alle deutschsprachigen Twittermeldungen, die auf den deutschen Filmtitel referenzieren, wurden in die Untersuchung aufgenommen. Irrelevante Meldungen, die sich nicht auf den entsprechenden Film bezogen, mussten manuell bereinigt werden.

Der institutionalisierte Kommunikationsprozess wurde über die Anzahl der Online-Film- und Pressekritiken modelliert. Hierfür wurden die Kritikerbewertungen auf Online-Filmportalen sowie die Pressekritiken, die mit Hilfe des Meta-Mediendienstes filmzeit.de für den deutschsprachigen Raum identifiziert wurden, integriert. Die Kritikerbewertungen setzen sich aus zwei Kriterien zusammen. Neben der Bewertung auf den Filmportalen wurde beurteilt, ob ein Film nur auf einem oder gleichzeitig auf mehreren Portalen rezensiert wurde. Die höchste Wertung von 100 Prozent erhielten nur Filme, die auf allen drei Portalen die beste Wertung erhielten. Für die Pressekritiken ging hingegen nur die Gesamtzahl in die Untersuchung ein.

Die abhängige Variable im Modell ist die Kinobesucherzahl zum Startwochenende. Diese Daten wurden freundlicherweise von der SPIO bereitgestellt und basieren auf den Rückmeldungen der Kinobetreiber in Deutschland an den entsprechenden Spitzenverband.

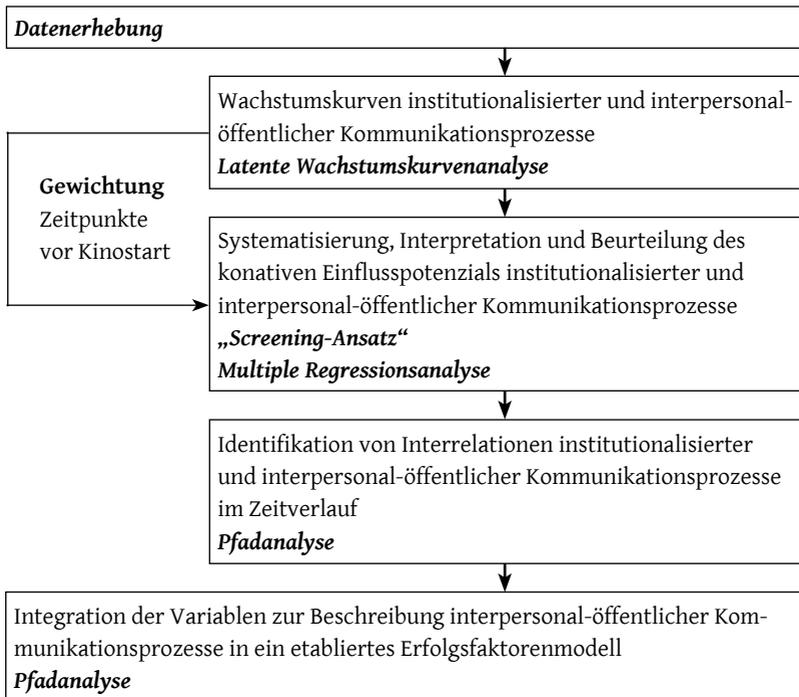
Tabelle 1: Übersicht der Variablen und Datenquellen

Variablenbezeichnung	Kategorie	Datenquelle
Anzahl der Twittermeldungen zum Film	interpersonal-öffentliche Kommunikation	twitter.com
Empfehlungen ("Likes"), negative Wertungen („Dislikes“), Kommentierungen ("Comments") und Seitenaufrufe der YouTube-Trailer	interpersonal-öffentliche Kommunikation	youtube.de
Facebook-Empfehlungen ("Likes"), Weiterleitungen ("Shares") und Kommentierungen ("Comments") auf YouTube-Trailern (Summenwert)	interpersonal-öffentliche Kommunikation	youtube.de facebook.com
Facebook-Empfehlungen („Likes“), Weiterleitungen („Shares“) und Kommentierungen („Comments“) auf Facebook-Fanseiten	interpersonal-öffentliche Kommunikation	facebook.com
Facebook-Empfehlungen ("Likes"), Weiterleitungen ("Shares") und Kommentierungen ("Comments") auf Online-Filmportalen (Summenwert)	interpersonal-öffentliche Kommunikation	filmstarts.de kino.de moviepilot.de facebook.com
Anzahl der Personen, die auf Facebook-Fanseiten über den Film sprechen	interpersonal-öffentliche Kommunikation	facebook.com
Anzahl der Pressekritiken	institutionalisierte Kommunikation	filmzeit.de
Kommentare auf Online-Filmportalen	institutionalisierte Kommunikation	filmstarts.de kino.de moviepilot.de
Kritikerbewertungen auf Online-Filmportalen	institutionalisierte Kommunikation	filmstarts.de kino.de moviepilot.de
Kinobesucherschzahl zum Startwochenende	abhängige Variable	SPIO

5 Übersicht des Gesamtaufbaus der Studie

Im Rahmen der Studie wurde ein komplexer mehrstufiger Aufbau gewählt, der im Folgenden kurz skizziert wird (siehe Abbildung 1). In einem ersten Schritt wird die Dynamik institutionalisierter und interpersonal-öffentlicher Kommunikationsprozesse vor dem Kinostart mittels latenter Wachstumskurvenanalyse modelliert. Das Verfahren zeichnet sich dadurch aus, dass Unterschiede in den Wachstumsfunktionen auf Gruppen- und Individualebene unterschieden werden können (McArdle & Bell, 2000, S. 74). Der gruppenspezifische Wachstumsprozess wird hierbei durch die mittlere Steigung sowie die Regressionsgewichte der einzelnen Zeitpunkte repräsentiert (McArdle & Bell, 2000, S. 101). Für diese Studie stellt die latente

Abbildung 1: Übersicht über die Gesamtkonzeption der Studie



Wachstumskurvenanalyse über die Berechnung gruppenspezifischer Regressionsgewichte ein geeignetes Aggregationsverfahren dar, um die Dynamik der Kommunikationsprozesse vor dem Kinostart später in ein Pfadmodell zur Beurteilung der Wechselwirkungen zwischen diesen Prozessen integrieren zu können.

In einem zweiten Schritt werden die handlungsrelevanten Kommunikationsprozesse identifiziert. Die Anwendung der multiplen Regressionsanalyse, in der die verschiedenen Indikatoren institutioneller und interpersonal-öffentlicher Kommunikation als unabhängige und die Kinobesucherzahl zum Startwochenende jeweils als abhängige Variablen eingehen, erlaubt eine Beurteilung des konativen Wirkungspotenzials der einzelnen Variablen. In die Pfadmodelle gehen folglich nur diejenigen Indikatoren ein, für die in den vorangegangenen Regressionsanalysen ein signifikanter zusätzlicher Erklärungsbeitrag festgestellt werden konnte (siehe Tabelle 2). Das Vorgehen kann als Screening-Ansatz bezeichnet werden, um diejenigen Social-Media-Indikatoren zu ermitteln, welche die Kinobesucherzahl beeinflussen.

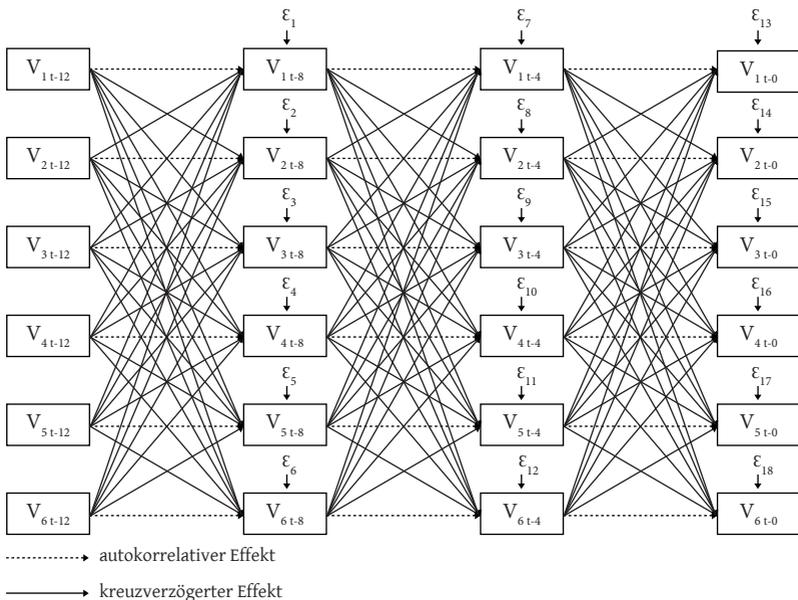
Tabelle 2: Ergebnisse der Regressionsanalyse zur Identifikation der Variablen interpersonal-öffentlicher Kommunikationsprozesse mit signifikantem zusätzlichem Erklärungsbeitrag

Variablenbezeichnung	Standardisierter Regressionskoeffizient β	Variance inflation factor
$R^2_{\text{korr}}=0,72$; $n=216$; $F\text{-Wert}=94,41$ ***		
Facebook-Aktivität auf Online-Filmportalen	0,54***	2,62
Reichweite der YouTube-Trailer	0,68***	7,96
Anzahl der Personen, die auf Facebook-Fan-seiten über den Film sprechen	-0,35***	5,48
Anzahl der Twittermeldungen zum Film	-0,22***	1,67
Anzahl der Pressekritiken	0,22***	1,54
Kritikerbewertungen auf Online-Filmportalen	-0,12**	1,41
Konstante	n.s.	

*p<0,001***; p<0,01**; p<0,05*; n.s. nicht signifikant; AV: Kinobesucherzahl zum Startwochenende*

Basierend auf diesen Erkenntnissen wird in einem dritten Schritt ein Pfadmodell der verschiedenen Zeitpunkte vor dem Kinostart etabliert, um potenzielle Wechselwirkungen zwischen den Indikatoren zu ermitteln (siehe Abbildung 2). Die Analyse erlaubt hierbei eine Differenzierung autokorrelativer und kreuzverzögerter Effekte. Diese geben wiederum an, inwiefern die Ausprägungen einer Variablen zu einem bestimmten Zeitpunkt durch die Werte früherer Phasen oder durch andere Variablen determiniert sind. Zeigen sich entsprechende Effekte, kann dies genutzt werden, um die Komplexität des Prognosemodells zu reduzieren oder die Daten-

Abbildung 2: Pfadmodell autokorrelativer und kreuzverzögerter Effekte institutionalisierter und interpersonal-öffentlicher Kommunikationsprozesse



V1: Facebook-Aktivität auf Online-Filmportalen

V2: Seitenaufrufe der YouTube-Trailer

V3: Anzahl der Personen, die auf Facebook-Fanseiten über den Film sprechen

V4: Anzahl der Twittermeldungen

V5: Anzahl der Pressekritiken

V6: Kritikerbewertungen auf Online-Filmportalen

erhebung auf frühere Perioden zu begrenzen. Dies reduziert den Aufwand bei der Datenerhebung und verlängert gleichzeitig den Prognosezeitraum.

Perspektivisch können diese Erkenntnisse genutzt werden, um die Indikatoren, für die entsprechende konative Wirkungen ermittelt wurden, in etablierte Erfolgsfaktorenmodelle zu integrieren.

6 Interrelationen institutionalisierter und interpersonal-öffentlicher Kommunikationsprozesse im Zeitverlauf

Durch schrittweise multiple Regressionsanalysen konnten sechs Kommunikationsindikatoren identifiziert werden, die einen eigenständigen signifikanten Beitrag zur Varianzaufklärung der Kinobesucherzahlen leisten (siehe Tabelle 2). Dies sind die Facebook-Aktivität (Summe aus der Anzahl der Empfehlungen, Weiterleitungen und Kommentierungen) auf Online-Filmportalen (V1; $\beta_{\text{stand.}}=0,54^{***}$), die Reichweite der YouTube-Trailer (V2; $\beta_{\text{stand.}}=0,68^{***}$), die Anzahl der Personen, die auf Facebook-Fanseiten über den Film sprechen (V3; $\beta_{\text{stand.}}=-0,35^{***}$), die Anzahl der Twittermeldungen zum Film (V4; $\beta_{\text{stand.}}=-0,22^{***}$), die Anzahl der Pressekritiken (V5; $\beta_{\text{stand.}}=0,22^{***}$) sowie die Kritikerbewertungen auf Online-Filmportalen (V6; $\beta_{\text{stand.}}=-0,12^{**}$).

Entgegen der Erwartung weisen die Anzahl der Personen, die auf Facebook-Fanseiten über den Film sprechen ($\beta_{\text{stand.}}=-0,35^{***}$) und die Twittermeldungen ($\beta_{\text{stand.}}=-0,22^{***}$) einen negativen Einfluss auf die Kinobesucherzahl auf. Eine mögliche Erklärung hierfür könnte sein, dass Trailer und Fanseiten vor allem genutzt werden, um die Informationsunsicherheit vor dem Kinostart zu reduzieren. Die starke Nutzung ist dadurch auch auf die Suche nach Erstinformationen zu Filmen zurückzuführen. Nicht alle dieser Filme werden folglich später im Kino betrachtet. Trailer und Fanseiten sind zwar wichtige Informationsquellen, aber keine zuverlässigen Prädiktoren des Filmbesuchs. Kritikerbewertungen auf Online-Filmportalen ($\beta_{\text{stand.}}=-0,12^{**}$) haben ebenfalls einen negativen Einfluss. Dieser kann durch eine hohe Informationsunsicherheit in Kombination mit schlechten Bewertungen erklärt werden. Ein hoher Wert für diese Variable ergibt sich nur, wenn ein Film auf allen drei Portalen gut bewertet wird (siehe Abschnitt 4). Mittels Pfadanalysen wird darauf aufbauend untersucht, ob und wie stark sich die Kommunikationsvariablen im Zeitverlauf gegenseitig beeinflussen. Die Beson-

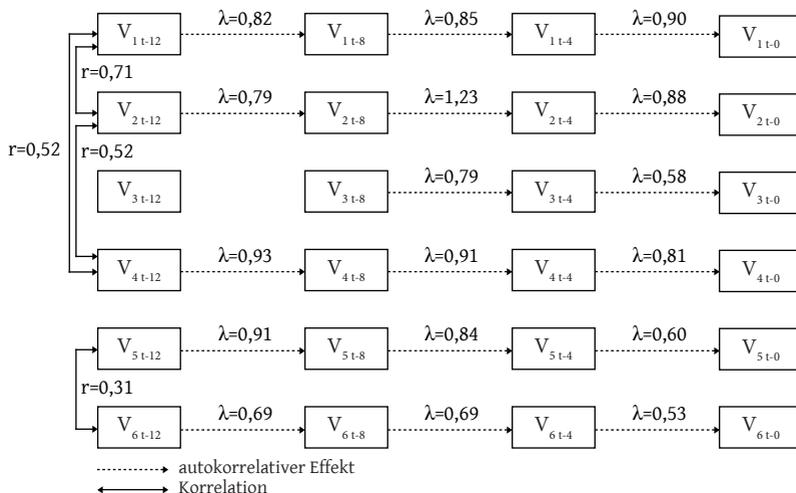
derheit der Pfadanalyse ist die gleichzeitige Berechnung aller Relationen, sodass der eigenständige Erklärungsbeitrag für jede Variable unter Berücksichtigung aller anderen bestimmt werden kann.

Ziel dieser Pfadanalyse ist es, möglichst generalisierbare Erkenntnisse über die Wechselwirkungen der verschiedenen Variablen zu unterschiedlichen Zeitpunkten abzuleiten. Es werden daher nur starke Effekte mit standardisierten Regressionsgewichten von $\lambda \geq 0,2$ und $\lambda \leq -0,2$ (Chin, 1998, S. 11; Bühner & Ziegler, 2009, S. 603) sowie Determinationskoeffizienten von $r \geq 0,30$ (Bühner & Ziegler, 2009, S. 603) dargestellt. Aufgrund der Vielzahl simultan berücksichtigter Relationen werden die Ergebnisse hier in einer schrittweisen Struktur präsentiert. Zunächst wird auf die Gesamtgüte des Modells sowie die Korrelationen zum Startzeitpunkt eingegangen. Anschließend werden die autokorrelativen Beziehungen über alle Zeitpunkte vorgestellt (siehe Abbildung 3). Danach folgt die Darstellung der kreuzverzögerten Beziehungen (siehe Abbildung 4). Abschließend wird zur Übersicht das vollständige Pfadmodell mit allen bedeutsamen autokorrelativen und kreuzverzögerten Zusammenhängen präsentiert (siehe Abbildung 5).

Die Modellanpassung ist mit $NFI=1,0$; $AGFI=1,0$ und $\chi^2/df=0,16$ sehr gut, sodass die beschriebenen Zusammenhänge als robust angesehen werden können. Für einige Variablen institutionalisierter und interpersonal-öffentlicher Kommunikation zeigen sich starke Korrelationen. Am stärksten ist diese zwischen der Facebook-Aktivität auf Online-Portalen sowie den Seitenaufrufen der YouTube-Trailer ($r=0,71$). Ferner zeigt sich ein Zusammenhang zwischen der Facebook-Aktivität auf Online-Portalen und der Anzahl der Mitteilungen zum Film auf Twitter ($r=0,52$) sowie zwischen den Seitenaufrufen der YouTube-Trailer und der Anzahl der Mitteilungen zum Film auf Twitter ($r=0,52$). Interessanterweise lässt sich auch ein Zusammenhang zwischen den Variablen institutionalisierter Kommunikationsprozesse feststellen, welcher darauf hindeutet, dass sich die Anzahl der Pressekritiken sowie die Kritikerwertungen gegenseitig beeinflussen. Dies zeigt, dass Filme, welche auf Online-Ablegern klassischer Medienangebote diskutiert werden, seitens der Filmkritiker auf Online-Filmportalen durchschnittlich häufiger und besser bewertet werden. Nicht alle Filme haben daher die gleiche Wahrscheinlichkeit rezensiert zu werden, was als Indiz für einen Auswahl-Bias der Kritiker angesehen werden kann (Clement & Sambeth, 2004, S. 112).

Generell sind hohe Autokorrelationen feststellbar, sodass sowohl institutionalisierte als auch interpersonal-öffentliche Kommunikationsprozesse in hohem

Abbildung 3: Übersicht der Korrelationen zwischen den Variablen institutionalisierter und interpersonal-öffentlicher Kommunikationsprozesse zum Zeitpunkt t-12 sowie der autokorrelativen Effekte im Zeitverlauf



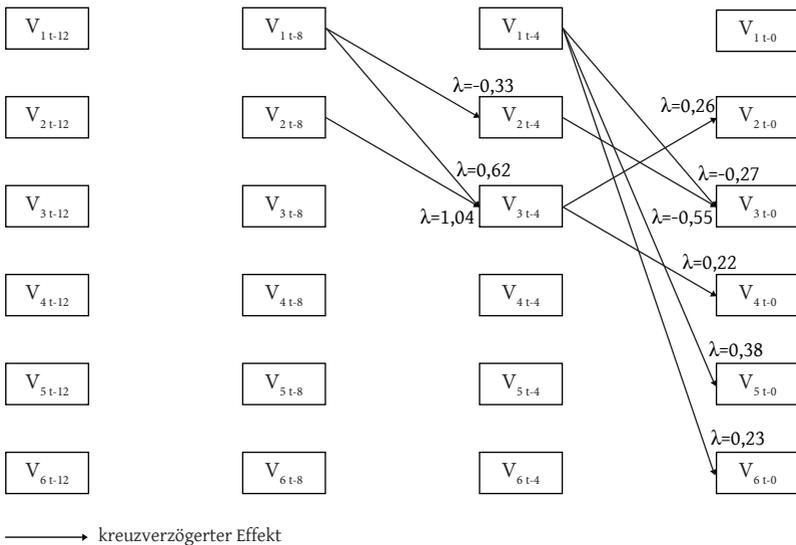
- V1: Facebook-Aktivität auf Online-Filmportalen
- V2: Seitenaufrufe der YouTube-Trailer
- V3: Anzahl der Personen, die auf Facebook-Fanseiten über den Film sprechen
- V4: Anzahl der Twittermeldungen
- V5: Anzahl der Pressekritiken
- V6: Kritikerbewertungen auf Online-Filmportalen

Maße pfadabhängig sind. Im Rahmen von Prognosemodellen ist diese Erkenntnis von großer Bedeutung, da hierdurch frühe Entwicklungen die späteren valide repräsentieren. Um den Aufwand bei der Datenerhebung zu reduzieren und den Prognosezeitraum zu erweitern, ist folglich eine Begrenzung auf frühe Phasen möglich. Eine interessante Entwicklung zeigt sich im Rahmen institutionalisierter Kommunikationsprozesse sowie auf Twitter, da die Pfadabhängigkeiten zum Kinostart hin leicht abnehmen. Dies spricht dafür, dass erst kurz vor dem Kinostart überproportional viele Kritiken oder Mitteilungen für eine breitere Basis an Filmen veröffentlicht werden. Die Pfadabhängigkeit nimmt folglich in Zeitverlauf

ab, da die Varianz im Kommunikationsverhalten zunimmt. Auch die Anzahl der Personen, die auf Facebook-Fanseiten über den Film sprechen, zeigen erst in den beiden Perioden direkt vor dem Kinostart autokorrelative Beziehungen. Im Rahmen der latenten Wachstumskurvenmodellierung wurde für diese Variable ein verzögert einsetzender Wachstumsprozess beobachtet. Dies impliziert, dass sie vor allem durch die anderen Kommunikationsvariablen beeinflusst wird.

Abbildung 4 veranschaulicht diese kreuzverzögerten Effekte. Interessanterweise lassen sich erst zu den letzten beiden Perioden vor dem Kinostart schwache und mittelstarke kreuzverzögerte Effekte feststellen. Die Seitenaufrufe der

Abbildung 4: Übersicht der kreuzverzögerten Zusammenhänge zwischen den Variablen institutionalisierter und interpersonal-öffentlicher Kommunikationsprozesse im Zeitverlauf



- V1: Facebook-Aktivität auf Online-Filmportalen
- V2: Seitenaufrufe der YouTube-Trailer
- V3: Anzahl der Personen, die auf Facebook-Fanseiten über den Film sprechen
- V4: Anzahl der Twittermeldungen
- V5: Anzahl der Pressekritiken
- V6: Kritikerbewertungen auf Online-Filmportalen

YouTube-Trailer beeinflussen die Zahl der Personen, die auf Facebook-Fanseiten über den Film sprechen positiv (t-12 bis -8=0,06; t-8 bis -4=1,04; t-4 bis 0=0,55). Dies bestätigt die vorherige Beobachtung, wonach für die Anzahl der Personen, die auf Facebook-Fanseiten über den Film sprechen, ein zeitlich verzögerter Wachstumsprozess erfolgt. Direkt zum Kinostart ist ein positiver Einfluss dieser Variable auf die Seitenaufrufe der YouTube-Trailer (t-12 bis -8=0,07; t-8 bis -4=0,00; t-4 bis 0=0,26) sowie die Anzahl der Twittermeldungen (t-12 bis -8=0,02; t-8 bis -4=0,01; t-4 bis 0=0,22) feststellbar. Es besteht folglich eine positive Rückkopplung zwischen den Seitenaufrufen der YouTube-Trailer und der Anzahl der Personen, die auf Facebook-Fanseiten über den Film sprechen.

Die Facebook-Aktivität auf Online-Portalen weist hingegen einen negativen Einfluss auf die Anzahl der Personen, die auf Facebook-Fanseiten über den Film sprechen (t-12 bis -8=-0,02; t-8 bis -4=-0,62; t-4 bis 0=-0,27) sowie die Seitenaufrufe der YouTube-Trailer auf (t-12 bis -8=0,1; t-8 bis -4=-0,33; t-4 bis 0=-0,13). Die Facebook-Aktivität auf Online-Filmportalen grenzt sich dadurch von denjenigen Indikatoren interpersonal-öffentlicher Kommunikation ab, die vor allem Reichweite repräsentieren. Die Diskurse auf spezialisierten Online-Portalen beziehen sich folglich schwerpunktmäßig auf andere Filme als diejenigen, welche auf Facebook-Fanseiten oder YouTube-Trailern die größten Reichweiten erzielen. Die Facebook-Aktivität auf Online-Filmportalen könnte daher Eigenschaften von Expertendiskursen aufweisen, an dem sich Presse- und Filmkritiken orientieren.

In der Phase direkt vor dem Kinostart beeinflusst die Facebook-Aktivität auf Online-Filmportalen den Umfang institutionalisierter Kommunikation positiv. Der Effekt auf die Anzahl der Pressekritiken ist hierbei größer (t-12 bis -8=-0,04; t-8 bis -4=-0,04; t-4 bis 0=0,38) als der auf die Kritikerbewertungen auf Online-Filmportalen (t-12 bis -8=-0,07; t-8 bis -4=-0,07; t-4 bis 0=0,23). In Presse- und Filmkritiken werden folglich bevorzugt diejenigen Filme berücksichtigt, welche bereits vor dem Kinostart diskutiert werden. Nicht alle Kinofilme haben demnach die gleiche Wahrscheinlichkeit in Rezensionen berücksichtigt zu werden.

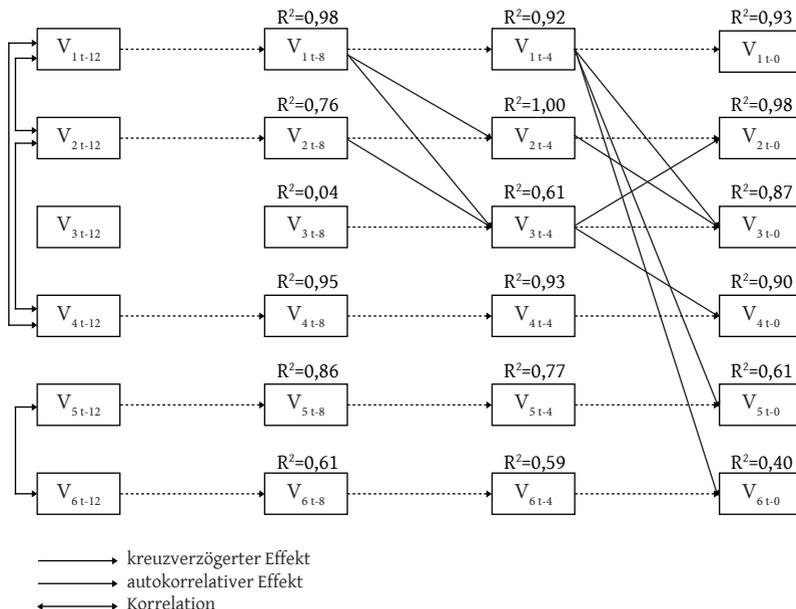
Institutionalisierte Kommunikationsprozesse sind also nicht nur als Erfolgsfaktor anzusehen, sondern werden ihrerseits – zumindest in der Phase direkt vor dem Kinostart – durch die Facebook-Aktivität auf Online-Filmportalen beeinflusst. Die Facebook-Aktivität auf Online-Filmportalen fungiert folglich als Prädiktor institutionalisierter Kommunikationsprozesse. Dieser Argumentati-

on folgend antizipieren die Facebook-Nutzer auf Online-Filmportalen, welche Filme wichtig sind und kommunizieren frühzeitig über diese, während die Bewertung im Rahmen institutionalisierter Kommunikationsprozesse erst später erfolgt. Die Facebook-Aktivität auf Online-Filmportalen setzt daher vor den institutionalisierten Kommunikationsprozessen ein und trägt zu einer Verringerung des Informationsdefizits bei.

7 Kernbefunde und Diskussion

Die vorgestellte Studie ist so konzipiert, dass zunächst eine breite Basis an potenziell einflussreichen Social-Media-Metriken integriert und auf ihr konatives Wirkungspotenzial untersucht wird. Es zeigte sich, dass nur ein Teil dieser Indikatoren einen eigenständigen Einfluss ausübt und dass eine Reihe von Wechselwirkungen bestehen (siehe Abbildung 5). Ein kreuzverzögerter Effekt der Facebook-Aktivität auf Online-Filmportalen auf die Prozesse institutionalisierter Kommunikation bleibt dabei auf die Phase drei und vier Wochen vor dem Kinostart begrenzt. Der negative Einfluss der Facebook-Aktivität auf Online-Filmportalen auf die Anzahl der Personen, die auf Facebook-Fanseiten über den Film sprechen und auf die Seitenaufrufe der YouTube-Trailer ist dagegen im Zeitverlauf vergleichsweise stabil. Zwischen den Seitenaufrufen der YouTube-Trailer und der Anzahl der Personen, die auf Facebook-Fanseiten über den Film sprechen, zeigt sich über den gesamten Zeitraum eine positive Rückkopplung. Die Reichweite beider Kanäle verstärkt sich wechselseitig. Die untersuchten Variablen zum Umfang institutionalisierter und interpersonal-öffentlicher Kommunikation über Kinofilme weisen sehr starke Pfadabhängigkeiten auf. Diese sind vor allem für die institutionalisierte Kommunikation in frühen Phasen des Informationsdiffusionsprozesses höher als in den späteren. Ein Grund hierfür kann darin gesehen werden, dass sich die Kommunikation zunächst vor allem auf die gleichen Filme konzentriert und erst zum Kinostart Rezensionen für eine breitere Basis an Filmen veröffentlicht werden. Zu einzelnen Filmen sind so zu frühen Zeitpunkten überhaupt keine institutionalisierten Kommunikationsangebote verfügbar. Dies führt zu einem Rückgang der Pfadabhängigkeiten zum Kinostart. Die Anteile insgesamt erklärter Varianz durch autokorrelative und kreuzverzögerte Effekte schwanken zum Kinostart zwischen $R^2=0,40$ und $R^2=0,98$.

Abbildung 5: Pfadmodell der autokorrelativen und kreuzverzögerten Effekte institutionalisierter und interpersonal-öffentlicher Kommunikationsprozesse



- V1: Facebook-Aktivität auf Online-Filmportalen
- V2: Seitenaufrufe der YouTube-Trailer
- V3: Anzahl der Personen, die auf Facebook-Fanseiten über den Film sprechen
- V4: Anzahl der Twittermeldungen
- V5: Anzahl der Pressekritiken
- V6: Kritikerbewertungen auf Online-Filmportalen

8 Ausblick: Verbesserte Absatzprognosen und Transferpotenzial für Kommunikationsprozesse im Bereich der politischen Kommunikation

Die entwickelte Screening-Methode erlaubt eine Vorauswahl relevanter Indikatoren und ermöglicht eine Reduktion der zu erhebenden Variablen und Zeitpunkte. Die Ergebnisse für die Kommunikation über Kinofilme zeigen, dass aufgrund starker autokorrelativer Beziehungen eine Beschränkung auf frühe

Phasen im Informationsdiffusionsprozess möglich ist, ohne die Erklärungskraft des Gesamtmodells erheblich einzuschränken. In weiteren Untersuchungen können diese Erkenntnisse über einflussreiche Indikatoren institutionalisierter und interpersonal-öffentlicher Kommunikation sowie deren Wechselwirkungen im Zeitverlauf genutzt werden, um Kommunikationsprozesse über Kinofilme vor der Markteinführung in etablierte Erfolgsfaktoren- und Prognosemodelle zu integrieren. Dadurch kann überprüft werden, inwiefern frühe Kommunikationsprozesse den späteren Erfolg eines Kinofilms beeinflussen und zu welchen Phasen dieser Einfluss am höchsten ist.

Die Identifikation und Beurteilung des konativen Einflusses einzelner Kommunikationskanäle dürfte sich noch verbessern, wenn in vertiefenden Inhaltsanalysen die relevanten Kommunikationsströme nicht nur im Hinblick auf ihren Umfang, sondern auch auf ihre Valenz betrachtet werden. Zudem sollte das Verhältnis von passiven und aktiven Kommunikatoren untersucht werden. Dies erlaubt auch eine Evaluation des Einflusses einzelner am Kommunikationsprozess beteiligter Akteure. Hierbei können in der Kommunikationswissenschaft etablierte Ansätze, wie beispielsweise das Meinungsführerkonzept angewendet werden. Meinungsführerschaft wird in der Regel über die Selbsteinschätzung der Probanden erhoben. An dieser Methode wird kritisiert, dass ein Individuum selbst seinen Einfluss auf andere nicht valide einschätzen kann (Schenk, 2007, S. 351-352, S. 375-376; Rogers & Cartano, 1962, S. 438-439; Katz & Lazarsfeld, 1955, S. 153-154). Mittels netzwerkanalytischer Methoden kann Meinungsführerschaft über die Kommunikatorstruktur direkt aus dem beobachteten Kommunikationsverhalten abgeleitet werden. Die Kommunikatorstruktur kann hierbei über netzwerkanalytische Indikatoren, vor allem die Zentralität und das Prestige von Akteuren, operationalisiert werden, indem die Beziehungen zwischen den Kommunikatoren im Rahmen relationaler Inhaltsanalysen erhoben werden (Adam, 2008, S. 182). Diese Relationen können auf Social-Media-Plattformen beispielsweise über Kommentierungen, Links, Retweets oder aus Verweisen direkt zusammen mit den entsprechenden Inhalten erhoben werden. Der Ansatz erlaubt folglich eine Abschätzung darüber, inwiefern die Kommunikatorstruktur und die Wertung der Aussagen den Kommunikationsverlauf beeinflussen können.

Die generelle Problematik von Social-Media-Analysen, dass private Fanseiten und Diskurse nicht integriert werden konnten, bleibt dennoch bestehen. Diese Informationen würden die Modellgüte sicherlich weiter verbessern, da gerade

private Diskurse als besonders glaubwürdig wahrgenommen werden. Die entsprechenden Seiten werden allerdings selten durch die Nutzer freigegeben. Methodisch können diese Informationen letztlich nur über eine Befragung oder eine Freigabe der individuellen Profile zu Forschungszwecken erschlossen werden. Neben den zu erwartenden Antwortverweigerungen schränkt insbesondere der hohe Aufwand die Anwendbarkeit in der Forschungspraxis ein, sodass entsprechende Untersuchungsdesigns nur für geringe Fallzahlen anwendbar sind. Auch die Nutzung der anbieterseitig bereitgestellten Social-Media-Metriken ist aus methodischer Sicht nicht unbedenklich. So sind über die bereitgestellten Schnittstellen (APIs) meist nicht alle Informationen abrufbar, sondern lediglich die von den jeweiligen Anbietern freigegebenen. Potenziell besteht daher die Gefahr, dass die Datenerhebung durch anbieterseitige Interessen konfundiert wird. Kritisch ist hierbei insbesondere, dass derartige Verzerrungen in der Regel nicht festgestellt werden können. In der vorliegenden Untersuchung wurden daher die Daten aus dem „Frontend“ genutzt, indem die Seiten so archiviert wurden, wie sie einem Nutzer zum entsprechenden Zeitpunkt im Browser angezeigt wurden. Dies führt jedoch zu einem erhöhten manuellen Aufwand, da die Daten zwar automatisch archiviert, aber nicht automatisch in einen Datensatz überführt werden konnten. Über stichprobenartige Kontrollen wurde überprüft, dass die Daten korrekt übertragen wurden.

In methodischer Hinsicht bietet die Studie einen Ansatz konativ bedeutsame Social-Media-Metriken im Zeitverlauf zu identifizieren, zu bewerten und für Prognosemodelle aufzubereiten. Das Vorgehen kann dadurch generell für die Erforschung der Auswirkungen von institutioneller und interpersonal-öffentlicher Kommunikation auf die Nachfrage von Vertrauensgütern, beispielsweise andere Medienprodukte, aber auch im Rahmen der politischen Wahlforschung für Politiker und Parteien, adaptiert werden. In einem ersten Schritt wären hierfür die relevanten institutionellen und interpersonal-öffentlichen Kommunikationskanäle zu bestimmen. Im Vergleich zur Kommunikation über einzelne Medienprodukte ist bei der politischen und insbesondere bei Wahlkampfkommunikation davon auszugehen, dass es in einzelnen Stimmbezirken jeweils unterschiedliche institutionelle Webseiten gibt, die zu beobachten wären. Auch die Suchbegriffe für die Datensammlung müssten entsprechend angepasst werden, insbesondere wenn es sich nicht um Kommunikation über Direktkandidaten, sondern über Parteien handelt. Zumindest bei Bundestagswahlen wäre die Zahl der Stimmbezirke

zur Schätzung latenter Wachstumskurven ausreichend. Hierbei ist interessant, inwiefern die in dieser Studie gewonnenen Erkenntnisse zur relativen Einflussstärke einzelner Social-Media-Metriken und zu deren Wechselwirkungen im Rahmen der Analyse anderer Produkte und Personen mit Vertrauensgutcharakter bestätigt werden können.

Felix Sattelberger, M.A. ist wissenschaftlicher Mitarbeiter am Institut für Kommunikationswissenschaft der Friedrich-Schiller-Universität Jena

Prof. Dr. Wolfgang Seufert ist Professor für Kommunikationswissenschaft mit dem Schwerpunkt Ökonomie und Organisation der Medien am Institut für Kommunikationswissenschaft der Friedrich-Schiller-Universität Jena

Quellenverzeichnis

- Adam, S. (2008). Medieninhalte aus der Netzwerkperspektive. Neue Erkenntnisse durch die Kombination von Inhalts- und Netzwerkanalyse. *Publizistik*, 53(2), 180-199. doi: 10.1007/s11616-008-0074-x
- Asai, S. (2009). Sales Patterns of Hit Music in Japan. *Journal of Media Economics*, 22(2), 81-101. doi: 10.1080/08997760902900106
- Bass, F. M. (1969). A new product growth for model consumer durables. *Management Science*, 15(5), 215-227. doi: 10.1287/mnsc.15.5.215
- Bühner, M., & Ziegler, M. (2009). *Statistik für Psychologen und Sozialwissenschaftler. PS Psychologie*. München: Pearson Studium.
- Chang, B.-H., & Ki, E.-J. (2005). Devising a Practical Model for Predicting Theatrical Movie Success: Focusing on the Experience Good Property. *Journal of Media Economics*, 18(4), 247-269. doi: 10.1207/s15327736me1804_2
- Chin, W. W. (1998). Commentary: Issues and Opinion on Structural Equation Modeling. *Management Information Systems Quarterly*, 22(1), 7-16.
- Clement, M., & Sambeth, F. (2004). Buchkritiker und Bucherfolg: Wie ist der Einfluss wirklich? *MedienWirtschaft*, 1(3), 105-114.
- Dellarocas, C., Zhang, X., & Awad, N. F. (2007). Exploring the value of online product reviews in forecasting sales: The case of motion pictures. *Journal of Interactive Marketing*, 21(4), 23-45. doi: 10.1002/dir.20087

- Dimmick, J., & Wang, T. (2005). Toward an Economic Theory of Media Diffusion Based on the Parameters of the Logistic Growth Equation. *Journal of Media Economics*, 18(4), 233-246. doi: 10.1207/s15327736me1804_1
- Duan, W., Gu, B., & Whinston, A. B. (2008). The dynamics of online word-of-mouth and product sales – an empirical investigation of the movie industry. *Journal of Retailing*, 84(2), 233-242. Abgerufen von <http://ssrn.com/abstract=1101054>
- Dogruel, L. (2013). *Eine kommunikationswissenschaftliche Konzeption von Medieninnovationen. Begriffsverständnis und theoretische Zugänge*. Wiesbaden: Springer VS.
- Haas, A., Keyling, T., & Brosius, H.-B. (2010). Online-Diskussionsforen als Indikator für interpersonale (Offline-)Kommunikation? Methodische Ansätze und Probleme. In N. Jakob, T. Zerback, O. Jandura, & M. Maurer (Hrsg.), *Das Internet als Forschungsinstrument und -gegenstand in der Kommunikationswissenschaft* (S. 246-267). Köln: Herbert von Halem.
- Hennig-Thurau, T., Houston, M. B., & Walsh, G. (2007). Determinants of motion picture box office and profitability: an interrelationship approach. *Review of Managerial Science*, 1(1), 65-92. doi: 10.1007/s11846-007-0003-9
- Katz, E. (1957). The Two-Step Flow of Communication: An up-to-date report on an Hypothesis. *Public Opinion Quarterly*, 21, 61-78. doi: 10.1086/266687
- Katz, E., & Lazarsfeld, P. F. (1955). *Personal influence – the part played by people in the flow of mass communication*. Glencoe: Free Press.
- Kiefer, M. L. (2005). *Medienökonomik: Einführung in eine ökonomische Theorie der Medien* (2., vollst. überarb. Aufl.). München u.a.: Oldenbourg.
- Lazarsfeld, P. F., Berelson, B., & Gaudet, H. (1944). *The people's choice – How the voter makes up his mind in a presidential campaign*. New York: Columbia Press.
- Leibenstein, H. (1950). Bandwagon, Snob, and Veblen Effects in the Theory of Consumers' Demand. *The Quarterly Journal of Economics*, 64(2), 183. doi: 10.2307/1882692
- Liu, Y. (2006). Word-of-Mouth for Movies: Its Dynamics and Impact on Box Office Revenue. *Journal of Marketing*, 70(3), 74-89. doi: 10.1509/jmkg.70.3.74
- McArdle, J. J., & Bell, R. Q. (2000). An introduction to latent growth models for developmental data analysis. In T. D. Little, K. U. Schnabel, & J. Baumert (Hrsg.), *Modeling longitudinal and multilevel data. Practical issues, applied approaches, and specific examples* (S. 69-108). Mahwah, NJ: Lawrence Erlbaum Associates.

- Peukert, C., Claussen, J., & Kretschmer, T. (2012). Piracy and Movie Revenues: Evidence from Megaupload. *SSRN Electronic Journal*. doi: 10.2139/ssrn.2176246
- Rogers, E. M. (1962). *Diffusion of innovations*. Glencoe: Free Press.
- Rogers, E. M., & Cartano, D. G. (1962). Methods of measuring opinion leadership. *Public Opinion Quarterly*, 26, 435-441. doi: 10.1086/267118
- Sattelberger, F. (2012). Veränderte Nachfragedynamik im Markt für Unterhaltungsmedien. Diffusionstheoretische Überlegungen und methodische Anregungen. In J. Müller-Lietzkow (Hrsg.), *Ökonomie, Qualität und Management von Unterhaltungsmedien* (S. 43-61). Baden-Baden: Nomos.
- Schenk, M. (2007). *Medienwirkungsforschung* (3., vollst. überarb. Aufl.). Tübingen: Mohr Siebeck.
- Sommer, D. (2010). *Nachrichten im Gespräch: Wesen und Wirkung von Anschlusskommunikation über Fernsehnachrichten*. Baden-Baden: Nomos.
- statista (2014a). Anzahl der aktiven Nutzer von Facebook in Deutschland. Abgerufen von <http://de.statista.com/statistik/daten/studie/70189/umfrage/nutzer-von-facebook-in-deutschland-seit-2009/>
- statista (2014b). Anzahl aktiver deutschsprachiger Nutzer von Twitter. Abgerufen von <http://de.statista.com/statistik/daten/studie/157936/umfrage/anzahl-deutschsprachiger-nutzer-von-twitter-seit-2009/>
- statista (2014c). Top 20 Videoportale in Deutschland. Abgerufen von <http://de.statista.com/statistik/daten/studie/209190/umfrage/beliebteste-videoportale-in-deutschland/>

III EXPERIMENTE

Empfohlene Zitierung: Hildebrandt, C., Schumann, C., & Wolling, J. (2015). Licht in die Black Box?! Experimente zum Einfluss von Social Signals auf das organische Ranking von Google. In A. Maireder, J. Ausserhofer, C. Schumann, & M. Taddicken (Hrsg.), *Digitale Methoden in der Kommunikationswissenschaft* (S. 225-249). doi: 10.17174/dcr.v2.10

Zusammenfassung: Der Beitrag stellt methodisches Vorgehen und Ergebnisse zweier Experimente mit der Suchmaschine Google vor. Untersucht wurde, ob sogenannte Social Signals einen Einfluss darauf haben, an welcher Stelle der Trefferliste von Google eine Webseite bei einer Suchanfrage positioniert wird. Social Signals sind Links, die durch Empfehlungen (Shares, Likes oder Kommentierungen) in sozialen Netzwerken wie Twitter oder Facebook entstehen. In den Experimenten wurden Testwebseiten angelegt, deren URLs gezielt in den sozialen Netzwerken Facebook, Google+ und Twitter geteilt wurden. Beide Untersuchungen zeigen, dass sich mit kommunikationswissenschaftlichen Methoden Erkenntnisse über die Rankingkriterien von Suchmaschinen generieren lassen: Es konnte nachgewiesen werden, dass sich Social Signals positiv auf die Positionierung von Webseiten auswirken. Deutlich wurde aber auch, dass solche Experimente nur mit fundierten Kenntnissen über die Funktionsweise von Suchmaschinen, technischem Know-How bei der Erstellung geeigneter Webangebote sowie einem erheblichen Aufwand bei der Rekrutierung und Einbindung von Untersuchungsteilnehmern zu realisieren sind.

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

Christin Hildebrandt, Christina Schumann & Jens Wolling

Licht in die Black Box?!

Experimente zum Einfluss von Social Signals
auf das organische Ranking von Google

1 Zur Bedeutung von Suchmaschinen

Suchmaschinen nehmen eine Schlüsselstellung in der digitalen Gesellschaft ein. Sie werden von 83 Prozent der Onliner mindestens einmal wöchentlich verwendet und sind damit die am häufigsten genutzte Anwendung des World Wide Web (van Eimeren & Frees, 2013, S. 363). Suchmaschinen sind von zentraler Bedeutung beim Auffinden von Informationen im Netz und fungieren damit faktisch wie Gatekeeper des Internet, die einen hohen Einfluss auf die Selektion haben (Wolling, 2002; Machill, Neuberger, & Schindler, 2002). Um diese Funktion erfüllen zu können, arbeiten Suchmaschinen mit Rankingkriterien, durch die sie die Relevanz von Webseiten „beurteilen“ (siehe auch Abschnitt 4). Je besser die Webseiten diese Rankingkriterien erfüllen, desto höher und prominenter werden sie den Nutzern in der Ergebnis- bzw. Trefferliste einer Suchanfrage angezeigt. Studien zum Nutzungsverhalten zeigen, dass sich eine gute Platzierung in besonderem Maße darauf auswirkt, ob Nutzer einer Website überhaupt Beachtung schenken (Machill, Neuberger, Schweiger, & Wirth, 2003).

Insbesondere für Webseitenbetreiber mit kommerziellen Interessen ist eine gute Suchmaschinenplatzierung zentral. Aus diesem Grund hat sich die Branche

der Suchmaschinenoptimierung (auch: Search Engine Optimization = SEO) entwickelt. Suchmaschinenoptimierung ist ein Teilbereich des Online-Marketings und bezeichnet verschiedene Maßnahmen, die dazu dienen, Webseiten in den organischen (= nicht gekauften) Ergebnislisten von Suchmaschinen – in erster Linie geht es dabei um Google – auf die vorderen Plätze zu bringen (Düweke & Rabsch, 2012, S. 262). Dazu werden Kenntnisse über die aktuellen Rankingkriterien benötigt, weswegen die Branchenunternehmen kommerzielle Forschungsprojekte initiieren, um die Relevanz der denkbaren Kriterien zu bestimmen. Die Ergebnisse dieser Auftragsforschung sind zumeist nicht öffentlich. Aufgrund der großen Bedeutung von Suchmaschinen auch für den gesellschaftlichen Kommunikationsprozess, sollte dieses Forschungsfeld aber auch aus akademischer Perspektive untersucht werden, damit öffentlich verfügbares Wissen bereitgestellt wird, anhand welcher Kriterien Suchmaschinen über den Zugang zu Informationen entscheiden.

Aufgrund der marktbeherrschenden Stellung in Deutschland (91 Prozent Marktanteil im Februar 2014; Statista, 2014) konzentrierten sich die weiteren Ausführungen ausschließlich auf die Suchmaschine Google.

2 Erkenntnisinteresse und methodische Herausforderung

Die Erstellung von Rankinglisten durch Google basiert auf ähnlichen theoretischen Grundlagen wie die Beurteilung wissenschaftlicher Qualität anhand von Zitierungen: Wissenschaftliche Veröffentlichungen, die häufig zitiert werden, werden als bedeutend eingeschätzt. Übertragen auf das Internet bewerten Suchmaschinen eingehende Links als Empfehlung anderer Webseiten (Erlhofer, 2011, S. 237). Es gilt als gesichert, dass Verlinkungen von externen Webseiten ein bedeutendes Rankingkriterium sind, weil sie als Empfehlung durch externe Akteure gewertet werden.

Durch die weite Verbreitung und intensive Nutzung von sozialen Netzwerken entstehen Linkempfehlungen allerdings zunehmend auch dort. Im Jahr 2010 hat Google bestätigt, dass solche Linkempfehlungen in sozialen Netzwerken – auch Social Signals genannt – eine Rolle bei der Bewertung von Webseiten spielen und somit als Rankingkriterium fungieren (Sullivan, 2010).

An diesem Punkt setzt der Beitrag an und widmet sich der übergeordneten Fragestellung, welchen Einfluss diese Social Signals auf das Ranking von Websei-

ten in den organischen Suchergebnissen von Suchmaschinen haben. Vorgestellt werden Methodik und Ergebnisse zweier Studien, die diese Frage untersucht haben. Methodisch ist dies deshalb herausfordernd, weil sich Algorithmen von Suchmaschinen mit den klassischen sozialwissenschaftlichen Methoden nicht analysieren lassen. Sie sind weder unmittelbar beobachtbar, noch können sie inhaltsanalytisch untersucht werden. Da mit den Codes enorme wirtschaftliche Interessen verknüpft sind und diese daher von den Suchmaschinenbetreibern geheim gehalten werden, erscheint auch eine Befragung der Programmierer wenig erfolgversprechend. Allerdings lassen sich durch experimentelle Manipulationen bis zu einem gewissen Grad Rückschlüsse auf die Funktionsweise der Algorithmen ziehen. Im Fokus des vorliegenden Beitrags stehen die methodischen Herausforderungen, Lösungsansätze und Grenzen, die solche Untersuchungen mit sich bringen. Dazu gibt der Beitrag zunächst eine kurze Einführung in die Funktionsweise von Suchmaschinen und erläutert, was über die Rankingkriterien von Suchmaschinen bereits bekannt ist. Dieses Wissen muss beim Design der nachfolgenden Experimente berücksichtigt werden, um externe wie interne Validität zu sichern.

3 Funktionsweise von algorithmusbasierten Suchmaschinen

Suchmaschinen sind automatisierte Suchdienste, die mithilfe mathematischer Algorithmen die im Internet vorhandenen Seiten erfassen, bewerten, sortieren und zu den Worten der Suchanfrage in Beziehung setzen (Maaß, Skusa, Heß, & Pietsch, 2009).

Google zählt zu den algorithmusbasierten Suchmaschinen, die mit eigener Datenbank arbeiten (für weitere Varianten siehe Schulz, Held, & Laudien, 2005). Entgegen einer häufigen Annahme sucht Google bei einer Suchanfrage nicht das gesamte Internet in Echtzeit nach relevanten Webseiten ab, sondern greift auf die zuvor angelegte Datenbank zurück (Noack, 2010, S. 34). Der Irrtum rührt vermutlich daher, dass Nutzer nur mit der Verarbeitung von Suchanfragen als einer von drei wesentlichen Komponenten des Suchprozesses in Kontakt kommen. Die beiden anderen Komponenten – die Datengewinnung sowie Datenanalyse und -verwaltung (Erlhofer, 2011, S. 170) – laufen hingegen unabhängig von der jeweiligen konkreten Suchanfrage ab. Für eine experimentelle Untersuchung mit

Google ist es insbesondere für den Versuchsaufbau wichtig, die verschiedenen Prozesse zu verstehen. Da alle drei Prozesse von Bedeutung sind, werden sie im Folgenden zum besseren Verständnis der Experimente in ihren Grundzügen skizziert.

Um dem Nutzer passende Suchergebnisse auf seine Suchanfrage liefern zu können, müssen zunächst die notwendigen Daten durch die Suchmaschine beschafft und gesichtet werden (Erlhofer, 2011, S. 170). Dies geschieht während der Phase der *Datengewinnung*, die auch als *Crawling* bezeichnet wird. Dabei kommen spezielle Programme, sogenannte *Crawler* (auch *Spider* oder *Robots* genannt), zum Einsatz, die das WWW kontinuierlich auf neue oder veränderte Dokumente überprüfen (Glöggler, 2008, S. 27). *Crawler* nutzen die Linkstruktur des WWW, um möglichst viele Webseiten ausfindig zu machen: Nachdem ein Dokument erfasst wurde, verfolgt der *Crawler* die darin enthaltenen Links und gelangt so zu weiteren Webseiten, um diese ebenfalls zu erfassen und sich anhand der dort enthaltenen Links wiederum weiter fortzubewegen (Lewandowski, 2005, S. 48).

Die von den *Crawlern* erfassten Daten liegen zunächst in unstrukturierter Form vor. Sie müssen daher im nächsten Schritt, der *Datenanalyse und -verarbeitung*, so aufbereitet werden, dass sie zu einem durchsuchbaren Datensatz werden und möglichst effizient weiterverarbeitet werden können (Bischopnick & Ceyp, 2009, S. 39). Diese Aufbereitung erfolgt in einem mehrstufigen Prozess. Neben den Inhalten werden hierbei unter anderem auch HTML-Strukturinformationen, ausgehende Links oder formale Elemente, wie zum Beispiel das Änderungsdatum, erfasst (Griesbaum, Bekavac, & Ritterberg, 2009, S. 33). Ziel dieser Datenanalyse ist es, relevante Begriffe bzw. Schlagwörter zu identifizieren, die den Inhalt der Seite so präzise wie möglich beschreiben (Erlhofer, 2011, S. 209). Mittels automatisierter Verfahren werden die Seiten schließlich hinsichtlich ihrer Relevanz für einen Suchbegriff mit Hilfe von Rankingkriterien (siehe nächster Abschnitt) bewertet und unter den entsprechenden Schlagworten in einen lokalen Datenbestand, den sogenannten *Index*, aufgenommen (Bischopnick & Ceyp, 2009, S. 23). Bei der letzten der drei Komponenten, der eigentlichen Suchanfrage durch den Nutzer, werden die Ergebnisse dann direkt aus dem *Index* generiert. Dies geschieht durch den *Query-Prozessor*, der über die Eingabeoberfläche die Schnittstelle für den Nutzer mit der Datenbank darstellt. Die Aufgabe des *Query-Prozessors* besteht darin, Dokumente im Datensatz zu finden, die für die Suchanfrage relevant sind und diese in eine Reihenfolge zu bringen (Glöggler, 2008, S. 96). Je höher der

Query-Prozessor die Relevanz der gefundenen Treffer für eine Suchanfrage „einschätzt“, desto besser (= weiter vorne) werden sie in der Ergebnisliste platziert, die dem Nutzer letztlich präsentiert wird.

Die *Bewertungs- bzw. auch Rankingkriterien*, nach denen die Suchmaschine die Trefferliste strukturiert und entscheidet, welche Webseiten in der Ergebnisliste auf eine Suchanfrage weit oben angezeigt werden, stehen im Mittelpunkt dieses Beitrags.

4 Rankingkriterien von Suchmaschinen

Auch wenn die konkrete aktuelle Auswahl und Gewichtung der Rankingkriterien nicht bekannt ist, liegen dennoch grundlegende Kenntnisse zu den Kriterien vor, anhand derer Google die Relevanz einer Webseite bewertet: Einerseits müssen Inhalt, Struktur und Gestaltung der eigentlichen Website bestimmte Qualitätskriterien erfüllen. Hierbei spricht man von *On-Page-Faktoren*. Von Bedeutung sind hier beispielsweise die häufige Aktualisierung der Inhalte sowie die Verankerung bestimmter Schlagworte (= Keywords) auf der Seite (Erlhofer, 2011, S. 398; Bischofnick & Ceyp, 2009; Glöggler, 2008). Andererseits wird von Google auch berücksichtigt, wenn andere Webseitenbetreiber eine Webseite als relevant einschätzen. Externe Relevanzzuschreibungen bemessen sich in erster Linie an den Verlinkungen, die von externen Seiten auf eine Webseite verweisen (= Backlinks). Diese Kriterien werden als *Off-Page-Faktoren* bezeichnet. Dabei kommt es zum einen auf die Quantität dieser Backlinks an; das heißt, je mehr Links auf die eigene Seite verweisen – zumal wenn sie von unterschiedlichen Domains und/oder Servern stammen – desto besser ist dies für die Positionierung der eigenen Seite (Hamacher, 2010, S. 72f). Da dieses Quantitätskriterium allerdings sehr anfällig für Manipulationen durch Suchmaschinenoptimierer ist (Erlhofer, 2011, S. 237), wird insbesondere auch die Qualität der Backlinks durch den Google-Algorithmus berücksichtigt (Glöggler, 2008, S. 179, für genauere Informationen siehe Enge, Spencer, Stricchiola, & Fishkin, 2012, S. 297-298).

Zusammenfassend kann aber festgehalten werden, dass qualitativ hochwertige Backlinks von Seiten kommen, die als vertrauenswürdig gelten (= kein Spam; Erlhofer, 2011, S. 237) und selbst viele Links von wichtigen Seiten erhalten (Koch, 2007, S. 51). Ein Backlink von spiegel.de ist damit bedeutsamer als ein Backlink

irgendeiner unbekanntem, wenig besuchten und wenig verlinkten Seite. Das ist die Grundidee von Googles PageRank-Algorithmus, der die Relevanz eines Dokumentes auf Grundlage der Anzahl und Qualität der Rückverweise von anderen Dokumenten berechnet (Glögler, 2008, S. 83).

Wie eingangs bereits thematisiert, gehen Verlinkungen nicht mehr nur von Webseiten aus, sondern entstehen zunehmend auch in den weit verbreiteten und intensiv genutzten sozialen Netzwerken. Es liegt die Vermutung nahe, dass die Links, die von den Nutzern sozialer Netzwerke geteilt werden, von Google ebenfalls als Empfehlung für eine Seite angesehen werden und somit – ähnlich wie „klassische“ Backlinks – als Rankingkriterium aufgenommen wurden.

5 Social Signals als neues Rankingkriterium?

Linkempfehlungen zu Webseiten, die öffentlich in sozialen Netzwerken geteilt werden, werden als Social Shares bezeichnet (Schiff, 2013, S. 164). Social Shares wiederum sind eine Teilmenge der Social Signals, die definiert sind als „Links, die durch Empfehlungen aus sozialen Netzwerken entstehen (Schiff, 2013, S. 165) und zu denen auch Kommentare, Likes oder Google +1 zählen. Aus forschungsökonomischen Gründen konzentriert sich der vorliegende Beitrag ausschließlich auf die Social Shares in den reichweitenstärksten sozialen Netzwerken Twitter, Google+ und Facebook. Grund für die ausschließliche Fokussierung auf die Social Shares war zum einen die potenzielle Vergleichbarkeit, da es sich um ähnliche Formen der Partizipation (= Teilen eines Links) handelt. Zum anderen lassen sich für diese Teilmenge der Social Signals die stärksten Effekte vermuten, da sie den „herkömmlichen“ Backlinks von Webseiten sehr ähnlich sind. Immer dann, wenn die Nutzer dieser Netzwerke ihre Profile öffentlich machen, sind ihre Aktivitäten nicht nur für andere Nutzer des Netzwerkes sichtbar, sondern können auch von Suchmaschinen gecrawlt werden (Tantau, 2011, S. 84).

Es ist zu vermuten, dass ähnlich wie bei den klassischen Backlinks einerseits die reine Quantität von Social Shares beim Ranking berücksichtigt wird, andererseits aber auch die Qualität der Social Shares (Garner, 2012, S. 27-28; Fischer, 2012, S. 49-50). Um eine Qualitätsbeurteilung vornehmen zu können, benötigen Suchmaschinen Einsicht in die sozialen Verbindungen der Nutzer sozialer Netzwerke (Fischer, 2012, S. 49). Dies geschieht über den Social Graph, der die Bezie-

hungen von Internetnutzern in einem sozialen Netzwerk abbildet (Garner, 2012, S. 27). Durch ihn lässt sich beispielsweise erkennen, wer mit wem befreundet ist und ob es sich bei einem Account um einen echten, inaktiven oder gefakten Account (z.B. Spam) handelt. Insbesondere lässt sich anhand des Social Graphs aber auch der Einfluss eines Users im Netzwerk ablesen, also ob er aktiv ist, selbst Inhalte einstellt oder nur passiv konsumiert (Fischer, 2012, S. 50). Diese Relevanzberechnung von Nutzern geschieht über den sogenannten SocialRank (Schiff, 2013, S. 188). Dieser berechnete Einfluss wird im folgenden Beitrag als „User Authority“ bezeichnet. Welche Faktoren es genau sind, die Einfluss auf die User Authority eines Nutzers haben, ist bisher unbekannt. Enge et al. (2012, S. 349 ff.) haben jedoch einige Faktoren zusammengetragen, anhand derer die User Authority bemessen werden könnte: Die Autoren vermuten beispielsweise, dass Nutzer umso einflussreicher sind, je häufiger sie ihren Status updaten, je mehr Retweets und Reshares ihre Posts oder Tweets aufweisen und je mehr relevante Freunde und Follower sie haben, die wiederum ebenfalls eine große Anzahl an Followern und Freunden mit hoher Authority aufweisen. Je einflussreicher ein User ist, desto relevanter schätzen Suchmaschinen seine Social Shares ein, so die Vermutung.

Für Suchmaschinenoptimierer ist das Wissen um den Einfluss von Social Signals auf die Platzierung einer Webseite in den Trefferlisten von großer Bedeutung. Daher überrascht es nicht, dass bereits versucht wurde, die Bedeutung der Social Signals zu entschlüsseln. Methodisch arbeiten vorliegende Studien entweder mit Expertenbefragungen von Suchmaschinenoptimierern (z.B. SEOmoz, 2011) oder mit Korrelationsanalysen (z.B. Searchmetrics, 2012). Bei letzteren werden für bestimmte Keywords die ersten Rangplätze mit den Eigenschaften der Seiten auf den vorderen Plätzen des Ranking korreliert. Die methodischen Probleme beider Herangehensweisen liegen auf der Hand: Bei der Expertenbefragung ist zum einen nicht klar, wie gesichert das Wissen ist, auf das sich die Experten berufen. Da das Wissen um die Rankingkriterien darüber hinaus zum zentralen Wettbewerbsvorteil von Suchmaschinenoptimierern zählt, ist auch nicht klar, ob die Optimierer in einer Befragung tatsächlich solche Erkenntnisse preisgeben. Korrelationen hingegen lassen keinen gesicherten Rückschluss zu, was Ursache und was Wirkung ist. Die Ergebnisse zeigen also lediglich, dass bestimmte Metriken häufiger bei Seiten mit besserer Positionierung auftreten. Ob aber die Social Signals zu einer guten Platzierung führen oder ob vielmehr – was ebenfalls hoch-

gradig plausibel ist – Seiten mit guter Platzierung viele Social Signals erhalten, lässt sich somit nicht sagen. Damit ist erstens festzuhalten, dass über den Einfluss von Social Signals trotz mehrerer vorliegender Studien nach wie vor wenig bekannt ist. Zweitens wurde deutlich, dass die Klärung der Frage eine methodische Herangehensweise erfordert, mit der validere Ergebnisse erzielt werden können, als dies mit bisherigen Studien geleistet wurde.

6 Hypothesen und Forschungsfrage

Basierend auf den bisherigen Ausführungen lässt sich das zentrale Erkenntnissinteresse zum Einfluss der Social Shares auf das Ranking von Webseiten in den organischen Suchergebnissen von Suchmaschinen in folgende Hypothesen und Forschungsfrage zerlegen:

H1: Webseiten, die Social Shares von Usern mit hoher User Authority erhalten, erzielen eine bessere Platzierung im Google-Ranking als Webseiten, die Social Shares von Usern mit niedrigerer User Authority erhalten.

Social Shares entstehen vornehmlich in den drei großen sozialen Netzwerken Google+, Facebook und Twitter. Da Google+ das konzerneigene soziale Netzwerk der Suchmaschine Google ist, sind die nötigen Daten (Social Graph) für Google dort leichter zugänglich, deswegen lässt sich vermuten:

H2: Social Shares aus dem konzerneigenen Netzwerk Google+ wirken sich positiver auf eine Platzierung im Google-Ranking aus als Shares aus den Netzwerken Twitter und Facebook.

Sollten Social Shares als Rankingkriterium fungieren, stehen sie darüber hinaus in „Konkurrenz“ mit den anderen Rankingkriterien. Abschließend ist daher zu fragen:

FF1: Wie stark ist der Einfluss von Social Shares im Vergleich mit anderen Rankingkriterien?

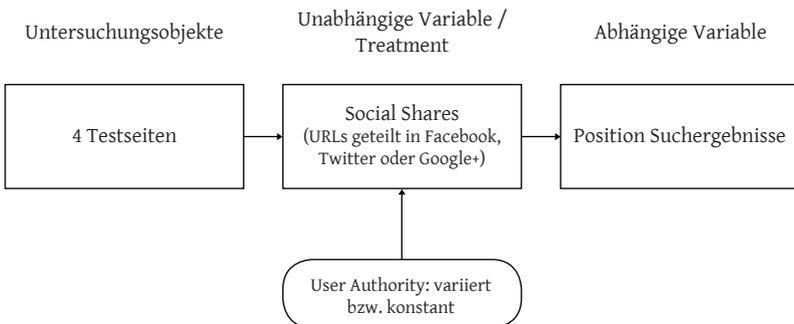
7 Methodik

Zur Überprüfung der Hypothesen und zur Beantwortung der Forschungsfrage wurden zwei Experimente durchgeführt. Experiment 1 prüft die Hypothesen 1 und 2. Hier stand die interne Validität der Ergebnisse im Vordergrund. Experiment 2 hingegen konzentriert sich auf die externe Validität und beantwortet die Forschungsfrage.

7.1 Untersuchungsanlage

Ausgangspunkt beider Experimente waren jeweils vier Testwebseiten, die eigens für die Experimente angelegt wurden. Deren URLs wurden in den drei Netzwerken Twitter, Google+ und Facebook als Social Shares geteilt. Jede Webseite erhielt dabei ausschließlich Social Shares aus jeweils einem Netzwerk. Diese Social Shares beschreiben das Treatment bzw. die unabhängige Variable im Experiment. Zusätzlich wurden diese Social Shares in ihrer User Authority variiert bzw. konstant gehalten. Die vierte Testseite fungierte als Kontrollgruppe, für die keine Social Shares generiert wurden. Die abhängige Variable bildet die jeweilige Position der einzelnen Testseiten in den Suchergebnissen von Google. Abbildung 1 zeigt den generellen Versuchsaufbau beider Experimente:

Abbildung 1: Überblick Versuchsaufbau beide Experimente



Die methodischen Herausforderungen der Experimente lagen zum einen darin, die Testseiten so zu generieren, dass sie technisch vergleichbar sind. Zum anderen musste das Treatment unter möglichst authentischen Bedingungen auf die Untersuchungsobjekte „verteilt“ und dabei die User Authority systematisch variiert bzw. auch gemäß realer Bedingungen konstant gehalten werden. Die folgenden Ausführungen beschreiben, wie dies in den beiden Experimenten realisiert wurde.

7.2 *Design der Testwebseiten*

Beim Design der Testseiten musste insbesondere auf die technische Vergleichbarkeit sowie eine ähnliche On-Page-Optimierung geachtet werden, damit Unterschiede in der Rankingplatzierung nicht durch Unterschiede auf den Seiten selbst hervorgerufen werden. Gleichzeitig durften die jeweiligen Testseiten aber auch nicht identisch sein, weil Google sogenannten Duplicate Content erkennt und im Ranking der Suchergebnisse herabstuft. Das bedeutet, dass die Inhalte einerseits zwar vergleichbar sein, sich andererseits aber so voneinander unterscheiden mussten, dass sie von Google nicht als Kopien identifiziert werden.

Die technische Umsetzung der Webseiten erfolgte in HTML und CSS. Jedes Testobjekt wurde bewusst einfach gehalten und bestand aus einer Startseite und – aufgrund der Impressumspflicht – einer Unterseite mit dem Impressum. Diese wurde allerdings mit dem Attribut „noindex, nofollow“ versehen, das den Crawlern anzeigt, die Seite nicht in den Index aufzunehmen.

Inhaltlich musste für die Seiten zunächst ein geeignetes Keyword ausgewählt werden, auf das die Seiten optimiert werden sollten. Keywords sind Schlüsselbegriffe, die Thema und Inhalt einer Seite möglichst präzise umschreiben und die erreichen sollen, dass die Seite im Index diesen Begriffen zugeordnet und bei einer entsprechenden Suchanfrage einem User angezeigt wird. Für die Sicherung der internen Validität wurde in Experiment 1 mit der „Savannenkatze“ ein *fiktives* Keyword ausgewählt. Der Vorteil eines fiktiven Keywords ist, dass die eigens angelegten Testwebseiten nicht in Konkurrenz zu anderen, bereits existierenden Webseiten stehen. In anderen Worten: Bei einer Google-suche nach dem Begriff „Savannenkatze“ wurden keine anderen Webseiten angezeigt. Auf diese Weise

lässt sich der Einfluss von Social Shares unter Ausschluss möglichst vieler Störfaktoren testen, da die Ergebnisse des Experiments nicht durch eventuelle Optimierungsmaßnahmen anderer Webseiten beeinflusst sind.

Die Optimierung der Testseiten in Experiment 2 hingegen erfolgte für ein bereits *existierendes* Keyword. Die Testseiten befinden sich daher in direkter Konkurrenz zu anderen Webseiten, wodurch sich der Einfluss der Social Shares unter realen Rahmenbedingungen testen lässt (externe Validität). Hierbei musste ein Keyword ausgewählt werden, das nicht zu stark umkämpft ist, so dass eine Positionierung der Testseiten in den vorderen Rankings von Google prinzipiell erreicht werden kann. Bei einer Testseite zum Thema „Versicherungen“ wäre dies beispielsweise nicht möglich gewesen, da die prominenten Plätze von sehr einflussreichen Seiten besetzt sind. Durch eine Recherche mit dem Google AdWords Keyword Tool, mit dem das monatliche Suchvolumen sowie der Wettbewerb für Suchbegriffe abgefragt werden kann, fiel die Wahl schließlich auf die Keyword-Kombination „Studentenwohnung Online“.

Die Nutzung der URLs für die jeweiligen Testseiten orientierte sich an den ausgewählten Keywords. Tabelle 1 zeigt die URLs sowie die Zuordnung zu Experimentalgruppen und Kontrollgruppe.

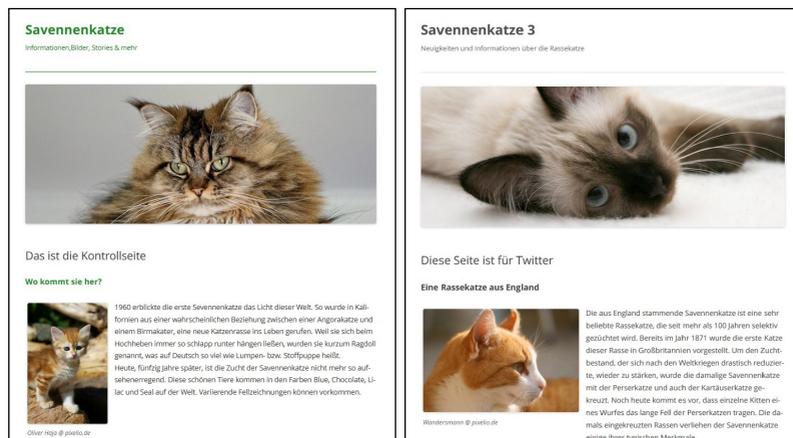
Tabelle 1: Verwendete Domains in den Experimenten und Zuordnung zu den Experimentalgruppen

Social Shares	Experiment 1	Experiment 2
Keine (Kontrollseite)	savennenkatze.de	studentenwohnung-online.de
Facebook	savennenkatze1.de	studentenwohnung-online3.de
Google+	savennenkatze2.de	studentenwohnung-online2.de
Twitter	savennenkatze3.de	studentenwohnung-online1.de

Die jeweiligen Seiten der beiden Experimente bestanden aus inhaltlich vergleichbaren Textteilen mit ähnlichem Umfang. Die Inhalte wurden selbst geschrieben und so angeordnet, dass sie nicht als Duplicate Content gewertet werden. Lizenzfreie Bilder der Bildagentur Pixelio ergänzten die Webseiten.

Des Weiteren war eine vergleichbare On-Page-Optimierung der Seiten zu gewährleisten. Dies geschah zum einen durch die Verwendung des jeweiligen Keywords im Titel und in h1-Überschriften der Webseiten. Zum anderen mussten die Keywords auf allen Seiten gleich häufig im Verhältnis zum Gesamttext (= Keyworddichte) und gleichmäßig (= Keywordlage) verwendet werden. Abbildung 2 zeigt Screenshots zweier Seiten aus Experiment 1 – die Kontrollseite (links) und die Seite, deren URL in Twitter geteilt wurde. Die Seiten zur Studentenwohnung hatte einen vergleichbaren Aufbau (Textanteil, Bilder).

Abbildung 2: Beispiel für die Testseiten aus Experiment 1



7.3 Treatment – Social Shares, User Authority und zeitlicher Ablauf

Für das Treatment mussten die URLs der erstellten Webseiten gezielt in den drei Netzwerken Facebook, Twitter und Google+ geteilt werden. Dieses Vorgehen wird auch als Linkaufbau bezeichnet. Grundvoraussetzung für den Linkaufbau war, dass die entsprechenden URLs der Testseiten immer nur in einem der Netzwerke geteilt werden, also die Shares für die Twitter-Seiten nur in Twitter, die Shares für die Facebook-Seiten nur in Facebook, usw.. Nur so

ließ sich ein möglicher Einfluss von Shares aus unterschiedlichen Netzwerken testen (Hypothese 2).

Beginn des Experiments war jeweils der Zeitpunkt, zu dem die Webseiten online gestellt wurden. In einer ersten Beobachtungsphase wurde zunächst geprüft, ob die Test- und Kontrollseiten nicht aufgrund anderer Faktoren (z.B. Links, die nicht zum Experiment gehörten) indexiert wurden. Dies war in beiden Experimenten nicht der Fall. Da Webseiten, die nicht verlinkt sind, aufgrund der oben beschriebenen Arbeitsweise der Crawler nicht in den Index einer Suchmaschine aufgenommen werden, musste vor dem Linkaufbau durch Social Shares dafür gesorgt werden, dass die Testseiten indexiert werden. Hierfür wurden auf die Seiten Links aus Blogs und Foren für die Kontrollseite (z.B. www.katzenblog.de) und Links aus eigens angelegten Fanpages (Google+ und Facebook) bzw. einer Twitterseite auf die entsprechenden Testseiten gesetzt, die schließlich zu einer Indexierung geführt haben (Woche 2).

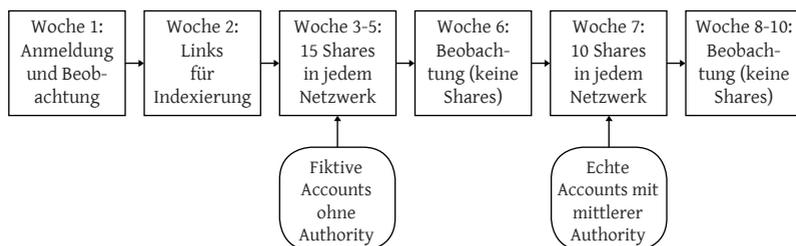
Danach begann der eigentliche Linkaufbau, der allerdings in beiden Experimenten unterschiedlich realisiert wurde:

Um zu überprüfen, ob die User-Authority einen Einfluss auf die Positionierung der einzelnen Testwebseiten hatte (Hypothese 1), wurden im Savannenkatzen-Experiment zunächst in jedem Netzwerk 15 fiktive Accounts angelegt, von denen in einem Zeitraum von drei Wochen während der Feldphase Social Shares für die jeweiligen Webseiten generiert wurden. Das bedeutet 15 Shares in Twitter für die Twitter-Testseite, 15 Shares in Facebook für die Facebook-Testseite und 15 Shares in Google+ für die Google+-Testseite. Um eine Identifikation der Accounts als „unecht“ zu verhindern, wurden bei der Erstellung unterschiedliche IP-Adressen und Namen verwendet. Diese Accounts besaßen allerdings weder Verbindungen zu anderen Usern, noch zeigten sie irgendwelche Aktivitäten im jeweiligen Netzwerk. Mit anderen Worten: Sie wiesen *keine User Authority* auf.

Für die Generierung von Social Shares durch Accounts *mit User Authority* wurden anschließend echte Accounts verwendet. Die Akquise der Teilnehmer erfolgte direkt über die Netzwerke, indem gezielt Personen angeschrieben und um ihre Mitarbeit gebeten wurden. Wichtig war es hierbei erstens, dass die User über ein öffentliches Profil verfügen, weil sie von Google ansonsten nicht gecrawlt werden können. Zweitens mussten die Accounts der Teilnehmer in allen Netzwerken eine vergleichbare Authority aufweisen. Wie in Kapitel 5 erläutert, bemisst sich die User Authority an unterschiedlichen Kriterien. Für die vorliegende Studie wur-

de die Höhe der Authority anhand der Anzahl an Freunden bzw. Followern eines Accounts sowie der Regelmäßigkeit von Postings und Statusupdates operationalisiert. Die Teilnehmer mit unterschiedlicher Authority wurden durch die Matched-Pairs-Methode (Schnell, Hill & Esser, 2008) als statistische Zwillinge (bzw. genauer gesagt Drillinge) ausgewählt. Da es äußerst schwierig und aufwendig ist, entsprechende Paare zu ermitteln, wurde aus forschungsökonomischen Gründen darauf verzichtet, weitere Kriterien zu berücksichtigen, wie zum Beispiel die Anzahl von Retweets und Reshares, die Postings und Tweets eines Accounts bekommen. Für das Savannenkatzen-Experiment war weniger entscheidend, welche Authority die User aufweisen, als vielmehr, dass sie *überhaupt* Authority aufweisen. Ausgewählt wurden Personen mit 50 bis 100 Freunden, die „mehrmals pro Woche“ im Netzwerk posteten. Pro Netzwerk wurden auf diese Weise jeweils zehn Personen ausgewählt, die in einem vorgegebenen Zeitraum von einer Woche die entsprechende URL entweder in Twitter, Facebook oder Google+ einmal geteilt haben (siehe Abbildung 3).

Abbildung 3: Zeitlicher Ablauf des Linkaufbaus in Experiment 1: „Savannenkatze“



In Experiment 2 zur Studentenwohnung stand die externe Validität im Vordergrund. Dafür mussten erstens die Webseiten für das bereits existierende Keyword optimiert werden. Zweitens war auch der Linkaufbau anders zu gestalten, da im Vergleich zu Experiment 1 mit einer höheren Anzahl an Social Shares gearbeitet werden musste. Insbesondere galt es, ein möglichst natürliches Wachstum der Social Shares zu simulieren, da zu plötzliche Erhöhungen der Verlinkungshäufigkeit von Google als Manipulationsversuche erkannt und sanktioniert werden (was auf-

grund der geringeren Anzahl an Social Shares in Experiment 1 nicht problematisch war). Dies bedeutet, es musste ein kontinuierlicher Linkaufbau durch Shares von Personen mit unterschiedlicher Authority erfolgen. Im Gegensatz zu Experiment 1 mussten daher über den gesamten Experimentalzeitraum Shares generiert werden. Zudem war eine größere Anzahl der Shares (insgesamt 60 in jedem Netzwerk) erforderlich, da zu erwarten war, dass in Konkurrenz zu anderen Webseiten mehr Signals über einen längeren Zeitraum notwendig waren, um Effekte zu erzeugen. Die besondere Herausforderung lag allerdings darin, die Shares von Personen unterschiedlicher Authority generieren zu lassen (= Simulation eines natürlichen Wachstums) und dabei gleichzeitig den möglichen Einfluss der Authority in den jeweiligen Netzwerken *identisch bzw. konstant* zu halten. Auch in Experiment 2 wurde daher wieder mit der Matching-Methode (statistischen Drillinge) gearbeitet. Erneut wurden dazu Personen in den Netzwerken direkt angeschrieben bzw. auch über einen Studierendenmailverteiler an der Technischen Universität Ilmenau um Mithilfe gebeten. Im Anschluss erfolgte die Einteilung der so rekrutierten Personen in Authoritygruppen. Kriterien hierfür waren die Anzahl der Freunde mit den drei Ausprägungen 0-50, 51-100 und > 100¹ sowie die Häufigkeit der Postings (in den vier Ausprägungen: täglich, mehrmals pro Woche, mehrmals im Monat und seltener). Auf diese Weise ließen sich insgesamt zwölf (3*4) Authoritygruppen definieren. Für die Validitätssicherung des Experiments mussten nun die rekrutierten Accountbesitzer den einzelnen Authoritygruppen so zugeordnet werden, dass in jeder Gruppe eine identische Anzahl an Personen mit identischer Authority pro Netzwerk enthalten waren (= statistische Drillinge). Das Ergebnis zeigt Tabelle 2.

Für die Durchführung von Experiment 2 bekam jeder dieser statistischen Drillinge auf den Tag genau vorgegeben, wann der jeweilige Social Share generiert werden sollte. Folgendes Beispiel verdeutlicht dies: An Tag X der Untersuchung sollte das erste der insgesamt sieben statistischen Trios der Authoritygruppe ‚0-50 Freunde‘ und ‚tägliches Posten‘ die entsprechende URL teilen, ebenso wie ein

1 Angesichts von Studien, die die durchschnittliche Anzahl an Freunden im sozialen Netzwerk Facebook mit 342 angeben (Weck, 2013), mögen die gewählten Werte der Freundesanzahl gering erscheinen. Der Grund für diese Vorgehensweise liegt in der vergleichsweise geringen Aktivität der User im Netzwerk Google+. In anderen Worten: Es wäre empirisch nur unter erheblichem Mehraufwand zu realisieren gewesen, Google+-User mit höherer als der angegebenen Authority zu rekrutieren.

Tabelle 2: Authoritygruppen mit statistischen Drillingen in Experiment 2

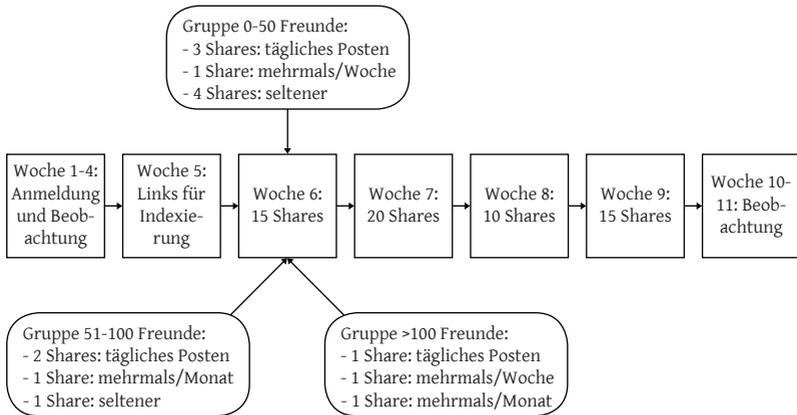
Häufigkeit Posting	Anzahl der Freunde		
	0-50	51-100	> 100
Täglich	7x Facebook	4x Facebook	4x Facebook
	7x Google+	4x Google+	4x Google+
	7x Twitter	4x Twitter	4x Facebook
Mehrere Male pro Woche	8x Facebook	4x Facebook	3x Facebook
	8x Google+	4x Google+	3x Google+
	8x Twitter	4x Twitter	3x Twitter
Mehrere Male im Monat	8x Facebook	5x Facebook	3x Facebook
	8x Google+	5x Google+	3x Google+
	8x Twitter	5x Twitter	3x Twitter
Seltener	6x Facebook	6x Facebook	2x Facebook
	6x Google+	6x Google+	2x Google+
	6x Twitter	6x Twitter	2x Twitter

statistisches Trio aus der Gruppe 51-100 Freunde mit mehrmaligem Posting pro Woche. An Tag Y wurden dann andere Trios gebeten, die entsprechenden URLs zu teilen. Auf diese Weise konnte über einen längeren Zeitraum hinweg ein „natürliches“ Linkwachstum von Personen mit unterschiedlicher Authority generiert werden, ohne dass unterschiedliche Authority-Niveaus Unterschiede zwischen den Testseiten hervorgerufen hätten. Abbildung 4 verdeutlicht diese Logik für den gesamten Untersuchungszeitraum. Am Beispiel der Woche 6 wird gezeigt, durch welche Drillinge die Shares in dieser Woche generiert worden sind.

7.4 Datenerhebung und Feldphase

Die Anmeldung der Testseiten erfolgte im ersten Experiment am 22.12.2012 und für das zweite Experiment am 19.01.2013. Nach der Beobachtungsphase und gleichzeitig mit der Indexierungsphase sowie parallel zum jeweiligen Linkaufbau erfolgte die Datenerhebung.

Abbildung 4: Zeitlicher Ablauf des Linkaufbaus in Experiment 2 mit Beispiel zum Linkaufbau in Woche 6



Der Untersuchungszeitraum für Experiment 1 war vom 07.01.2013 bis zum 10.03.2013 und für Experiment 2 vom 28.01.2013 bis zum 17.03.2013.

Für die Datenerhebung wurden mehrere Tools genutzt: Der „KeywordMonitor“ ist ein SEO-Tool, das täglich die Rankings der eigenen Webseiten für die angegebenen Keywords überwacht und dokumentiert. Der Vorteil gegenüber anderen Tools ist, dass auch fiktive Keywords analysiert werden können. Die Abfrage der Rankings ist allerdings auf die ersten 100 Ergebnisse bei Google beschränkt. Das Tool wurde vor Beginn der Durchführung implementiert. Als zweites Analyse Tool diente die „Google SEO Work Bench“. Dabei handelt es sich um ein Plug-in für den Browser Chrome mit dem es möglich ist, die Rankingposition von ausgewählten Webseiten für ein bestimmtes Keyword zu analysieren. Im Unterschied zum KeywordMonitor muss die Abfrage für die Suchbegriffe täglich aktiv selbst durchgeführt werden. Zusätzlich wurden die Ergebnisse der Tools mit dem Firefox-Browser im privaten Modus überprüft. Der private Modus ermöglicht die Nutzung von Webseiten, ohne dass hierbei Daten über die Webseitenbesuche auf dem Rechner gespeichert werden. Auf diese Weise werden die Ergebnisse in den Google-Suchergebnissen nicht durch vorherige Nutzungsgewohnheiten beeinflusst und zeigen bei jeder Anfrage das Ranking wie es sich bei einer erstmaligen Suche darstellen würde.

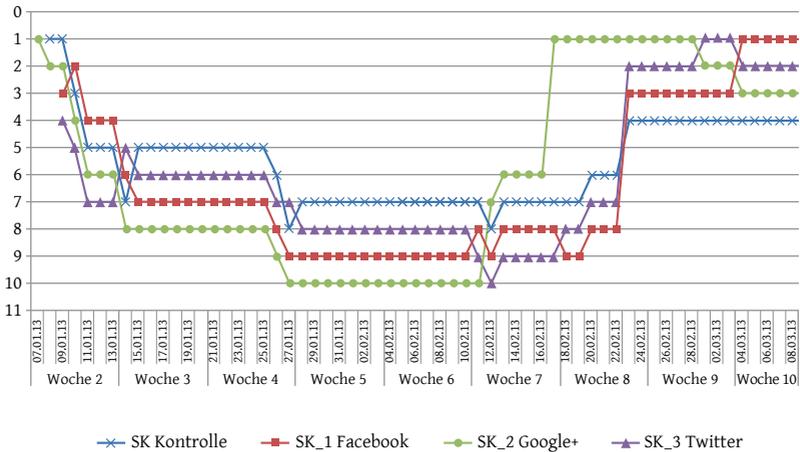
8 Ergebnisse

Da die Methodik in diesem Beitrag im Vordergrund steht, werden die Ergebnisse der Experimente nur kurz skizziert. Neben der übergeordneten Fragestellung, ob sich Social Shares überhaupt auf die Platzierung von Webseiten im Google-Ranking auswirken, wurde untersucht, ob auch die Authority sowie das Netzwerk, in dem ein Social Share generiert wird, einen Einfluss auf die Positionierung einer Webseite hat.

8.1 Experiment 1 – Savannenkatze

Abbildung 5 zeigt die Ergebnisse des ersten Experiments zum fiktiven Keyword „Savannenkatze“.

Abbildung 5: Ergebnisse des Experiments zur Savannenkatze



In der zweiten Woche wurden die Links für die Indexierung der Seiten gesetzt. Diese stammten aus Blogs und Foren für die Kontrollseite und aus eigens angeleg-

ten Fanpages für die Google+ und die Facebookseite bzw. einer Savennenkatzen-Twitterseite. Es zeigt sich zunächst, dass die Testseite für Google+ als erstes in den Index aufgenommen wurde, gefolgt von der Kontrollseite.

Mit Beginn des eigentlichen Experiments in Woche drei, in der mit den Shares von fiktiven Accounts ohne Authority begonnen wurde, sind zwei Entwicklungen zu beobachten: Zum einen steht die Kontrollseite während dieses Zeitraums weitestgehend konstant an oberster Stelle im Ranking. Gleichzeitig brechen alle Seiten in ihrer Rankingposition ein. Angesichts des fiktiven Keywords in diesem Experiment, durch das eigentlich keine Konkurrenzseiten existieren sollten, stellt sich natürlich die Frage, welche Seiten weiter oben gelistet waren. Dabei handelt es sich um die Facebook-Fanpage mit ihren insgesamt 5 Unterseiten, die sich auf den ersten Plätzen behauptete. Diese war im Experiment lediglich als „Linkgeber“ für die Indexierung der Facebook-Savennenkatze-Testseite eingeplant. Sie hat sich mit ihren Unterseiten auf das Experiment aber insofern ausgewirkt, als dass sie von Google gecrawlt, indiziert und für das Suchwort „Savennenkatze“ als relevant eingeschätzt wurde. Im Unterschied zur Facebook-Fanpage tauchen die Fanseiten in Google+ sowie die Savennenkatzen-Twitterseite jedoch nicht im Ranking auf.

In der siebten Woche, in der mit den Shares durch echte Accounts mit mittlerer Authority begonnen wurde, ändert sich das Bild. Zum einen konnten alle Seiten ihre Positionierung im Ranking verbessern, so dass die Facebook-Fanseite von den vorderen Positionen verdrängt wurde. Auch die Kontrollseite kann ihre vordere Positionierung nicht länger halten. Diese Ergebnisse zeigen zum einen, dass Social Signals ein Rankingkriterium sind, nach dem Google die Relevanz einer Seite beurteilt. Zum anderen weisen diese Ergebnisse darauf hin, dass die Authority der Nutzer einen Einfluss auf die Bewertung der Seiten hat. Erst zu dem Zeitpunkt, als Social Shares von echten Nutzern mit (mittlerer) Authority erzeugt wurden, verbesserte sich die Platzierung der Testseiten und sie schieben sich vor die Facebookseite. Dies bestätigt die Annahme, dass der Social Graph der Signalgeber als Rankingkriterium einen Einfluss hat.

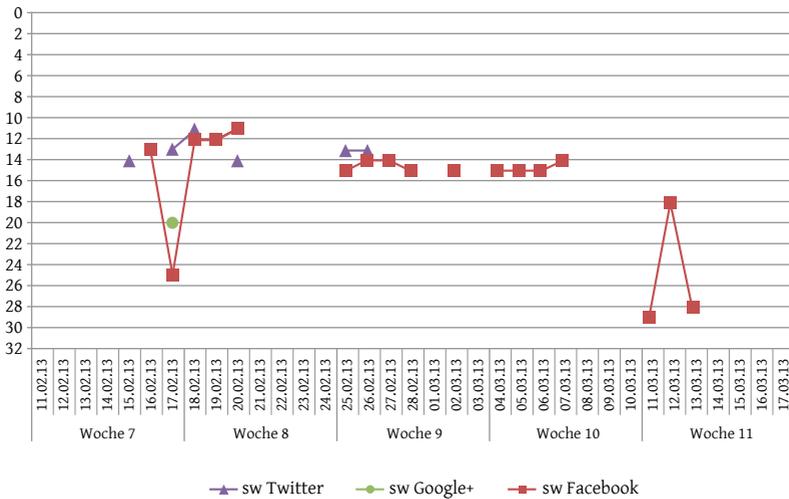
Am schnellsten und umfassendsten zeigen sich die Veränderungen bei der Google+-Testseite. Innerhalb einer Woche steigt sie im Ranking vom letzten auf den ersten Platz im Experiment. Bei den anderen Seiten dauert es länger, bis der Algorithmus auf die Social Shares reagiert. Damit scheinen die Shares von Google+ einen stärkeren Einfluss zu haben als die von Twitter und Facebook. Gleichzeitig deuten die Ergebnisse darauf hin, dass diese aber auch weniger nachhaltig sind, da ab

Woche 9 die Google+-Seite zunächst von der Twitter- und dann von der Facebook-Testseite „überholt“ wird.

8.2 Experiment 2 – Studentenwohnung Online

Abbildung 6 zeigt die Ergebnisse des zweiten Experiments, in dem der Einfluss der Social Shares auf die Positionierung der Testseiten im direkten Vergleich mit real existierenden Konkurrenzseiten getestet wurde.

Abbildung 6: Ergebnisse des Experiments zur Studentenwohnung



Es fällt auf, dass sich keine der Seiten im Experimentalzeitraum konstant unter den ersten 30 Platzierungen behaupten konnte. Erst in der siebten Woche, in der die meisten Shares produziert wurden, erreichen alle Testseiten mindestens einmal eine Positionierung in den Top 25 der Suchergebnisse. In Anbetracht des verhältnismäßig simplen Aufbaus und des geringen Umfangs sowie des jungen Domainalters der Experimentalseiten sind diese Positionierungen allerdings beachtlich.

Im Vergleich mit dem ersten Experiment überrascht allerdings, dass die Google+-Seite im gesamten Zeitraum nur einmal unter den ersten dreißig Rankingplätzen gelistet wird. Die Facebook-Testseite schafft die häufigsten Positionierungen und sie bleibt auch dann relativ weit vorne, wenn gar keine neuen Social Shares mehr von Facebook-Usern generiert wurden (Woche 10 und 11). Hier zeigt sich eine Parallele zum Savennenkatzen-Experiment. Dort hat es vergleichsweise lange gedauert, bis Google auf die Facebook-Shares reagiert hat. Sobald dies allerdings der Fall war, konnte mit den Facebook-Shares eine gute Platzierung erreicht werden. Vermutlich laufen die Crawler von Google seltener über die Facebook-Seiten, als dies bei Twitter und Google+ der Fall ist. Die Kontrollseite hingegen, für die nach der Indexierung keine weiteren Links generiert wurden, schaffte im gesamten Untersuchungszeitraum keine einzige Platzierung auf den ersten 30 Plätzen.

8.3 Zusammenfassung der Ergebnisse

Zusammenfassend ist zu den Ergebnissen beider Experimente festzuhalten: Die übergeordnete Forschungsfrage, ob Social Shares überhaupt einen Einfluss auf die Positionierung im organischen Ranking von Webseiten haben, ist eindeutig mit „Ja“ zu beantworten. In Experiment 1 zeigte sich darüber hinaus, dass auch die User Authority der Nutzer, die einen Social Share generieren, bei der Relevanzzuschreibung berücksichtigt wird. Hinsichtlich der Frage, ob die Shares aus unterschiedlichen Netzwerken unterschiedliche Gewichtung besitzen, konnten die Experimente keine eindeutigen Ergebnisse erzielen. Experiment 1 zeigte, dass die Google+-Testseite schneller indexiert wurde. Ebenso reagierte der Algorithmus schneller auf jene Shares, die im konzerneigenen Netzwerk generiert wurden. Die Ergebnisse von Experiment 2 hingegen legen keine generelle Bevorzugung von Google+ nahe. Im Gegenteil: Es scheinen eher die Facebook-Shares zu sein, die sich nachhaltiger auf die Positionierung im Ranking auswirken. Insgesamt betrachtet lassen die Ergebnisse aber keine großen Unterschiede im Einfluss von verschiedenen Netzwerken erkennen. Hinsichtlich der Frage, wie stark die Social Shares im Vergleich zu anderen Kriterien berücksichtigt werden, scheinen sie bislang noch keine herausragende Rolle zu spielen: Zwar zeigte Experiment 2, dass eine quantitative Zunahme der Shares (insbesondere) auf Facebook eine Po-

sitionierung unter den ersten 30 Platzierungen generierte. Dieser Befund ist aber mit Blick auf die verwendete Suchwortkombination „Studentenwohnung Online“ zu relativieren. In der Realität wird nach diesen beiden Suchworten vermutlich eher selten kombiniert gesucht. Eine Überprüfung der Rankings für das häufiger verwendete Suchwort „Studentenwohnung“ ergab, dass keine der Experimentalseiten im gesamten Zeitraum eine Platzierung unter den ersten 30 Treffern erzielen konnte. Damit ist festzuhalten, dass Google die Social Shares zwar als Rankingkriterium berücksichtigt, dass ihr Einfluss im Vergleich mit anderen Kriterien allerdings bisher noch gering zu sein scheint.

9 Methodische Reflexion

Auch wenn die vorliegende Untersuchung einige der kritisierten Mängel anderer Studien behebt, weist auch sie einige Limitationen auf: Zum einen wurden die Testwebseiten aus forschungsökonomischen Gründen verhältnismäßig einfach (Startseite und Impressum) sowie der Untersuchungszeitraum relativ kurz gehalten. Inwiefern sich insbesondere in Experiment 2 andere Ergebnisse gezeigt hätten, wären die Testseiten umfangreicher (und damit vergleichbar zu realen Konkurrenten) und der Untersuchungszeitraum länger gewesen, lässt sich nicht sagen. Zum vollen Verständnis der Frage, wie sich Social Signals auf die Positionierung einer Webseite auswirken, müssten neben den im Experiment untersuchten Social Shares auch Kommentierungen und Likes berücksichtigt werden. Wie bereits thematisiert, wurde die User Authority ausschließlich über die Anzahl der Freunde sowie die Häufigkeit der Postings bzw. Tweets operationalisiert, wodurch andere Kriterien (wie z.B. die durchschnittliche Anzahl von Retweets und Reshares) nicht berücksichtigt wurden. Abschließend wäre auch ein systematischer Vergleich zu klassischen Linkbuilding-Maßnahmen erforderlich, um die relative Bedeutung der Social Signals angemessen einschätzen zu können. Weitere Forschungen auf diesem Gebiet sollten daher a) mit umfangreicheren Testwebseiten arbeiten, b) die Entwicklung der Rankings über einen noch längeren Untersuchungszeitraum beobachten, c) Social Signals in allen ihren Ausprägungen mit einbeziehen, d) die User Authority noch umfassender operationalisieren und e) einen systematischen Vergleich mit „klassischen“ Linkbuildingmaßnahmen vornehmen.

Neben dieser unmittelbaren Reflexion der Methodik wurde eingangs auch die Frage gestellt, ob sich mit sozialwissenschaftlichen Methoden die Funktionsweise von Algorithmen untersuchen lässt. Die beiden Experimente haben verdeutlicht, dass die Kommunikationswissenschaft durchaus in der Lage ist, Erkenntnisse über solche Algorithmen zu gewinnen – und zwar ohne Zugang zu ihrem Code zu haben. Allerdings wurde auch deutlich, dass hierfür technisches Wissen und ausreichendes Verständnis der Funktionsweise der Algorithmen notwendig sind. Des Weiteren war insbesondere die „Produktion“ des Treatments mit einem hohen Aufwand verbunden. Dieser kann einer verhältnismäßig kurzen Halbwertszeit der Ergebnisse gegenüberstehen, da damit zu rechnen ist, dass auch im Bereich der Social Signals der Suchalgorithmus kontinuierlich angepasst wird. Diese Einschränkung gilt aber auch für „klassische“, sozialwissenschaftliche Experimente oder Befragungsstudien. Die Relevanz der Studie ergibt sich deswegen nicht so sehr aus den Detailbefunden, sondern zum einen aus den diskutierten methodischen Fragen und zum anderen aus dem emanzipatorischen Potenzial, wenn wir nicht nur Vermutungen darüber anstellen, sondern belastbare Ergebnisse dazu haben, wie soziale Kommunikation in technisch geprägten Umgebungen strukturiert und gesteuert wird.

Christin Hildebrandt, M.A. ist Online Marketing Manager bei Autoscout24

Dr. Christina Schumann ist wissenschaftliche Mitarbeiterin am Fachgebiet „Empirische Medienforschung und Politische Kommunikation“ der Technischen Universität Ilmenau

Prof. Dr. Jens Wolling ist Professor am Fachgebiet „Empirische Medienforschung und politische Kommunikation“ der Technischen Universität Ilmenau

Quellenverzeichnis

- Bishopnick, Y., & Ceyy, M. (2009). *Suchmaschinen-Marketing: Konzepte, Umsetzung und Controlling*. Berlin: Springer.
- Düweke, E., & Rabsch, S. (2012). *Erfolgreiche Websites: SEO, SEM, Online-Marketing, Usability*. Bonn: Galileo Computing.
- Enge, E., Spencer, S., Stricchiola, J., & Fishkin, R. (2012). *The Art of SEO. Mastering Search Engine Optimization (Theory in Practice, Bd. 2)*. Sebastopol, CA: O'Reilly Media.

- Erlhofer, S. (2011). *Suchmaschinen-Optimierung. Das umfassende Handbuch*. Bonn: Galileo Press.
- Fischer, M. (2012). Quo Vadis SEO. *Website Boosting*, 13(5-6), 40-50.
- Garner, R. (2012). *Search and social. The definitive guide to real-time content marketing*. Indianapolis: Wiley. Abgerufen von <http://site.ebrary.com/lib/alltitles/docDetail.action?docID=10615083>
- Glöggler, M. (2008). *Suchmaschinen im Internet: Funktionsweisen, Ranking Methoden, Top Positionen*. Berlin: Springer.
- Griesbaum, J., Bekavac, B., & Ritterberg, M. (2009). Typologie der Suchdienste im Internet. In D. Lewandowski (Hrsg.), *Handbuch Internet-Suchmaschinen. Nutzerorientierung in Wissenschaft und Praxis* (S. 18-52). Heidelberg: Akademische Verl.-Ges.
- Hamacher, T. (2010). *Suchmaschinenmarketing: Neue Marketing-Strategien für kleine und mittlere Wettbewerber*. Marburg: Tectum Verlag.
- Koch, D. (2007). *Suchmaschinen-Optimierung. Website-Marketing für Entwickler*. München: Addison-Wesley.
- Lewandowski, D. (2005). *Web information retrieval. Technologien zur Informationssuche im Internet*. Frankfurt am Main: Deutsche Gesellschaft für Informationswissenschaft und Informationspraxis.
- Maaß, C., Skusa, A., Heß, A., & Pietsch, G. (2009). Der Markt für Internet-Suchmaschinen. In D. Lewandowski (Hrsg.), *Handbuch Internet-Suchmaschinen. Nutzerorientierung in Wissenschaft und Praxis* (S. 3-17). Heidelberg: Akademische Verl.-Ges.
- Machill, M., Neuberger, C., & Schindler, F. (2002). *Transparenz im Netz*. Gütersloh: Bertelsmann Stiftung.
- Machill, M., Neuberger, C., Schweiger, W., & Wirth, W. (2003). Wegweiser im Netz: Qualität und Nutzung von Suchmaschinen. In M. Machill & C. Welp (Hrsg.), *Wegweiser im Netz: Qualität und Nutzung von Suchmaschinen* (S. 13-490). Gütersloh: Verlag Bertelsmann Stiftung.
- Noack, C. (2010). *Crossmedia Marketing. Suchmaschinen als Brücke zwischen Offline- und Online-Kommunikation*. Boizenburg: Hülsbusch.
- Schiff, D. (2013). *Social SEO: Unternehmer-Ratgeber zu Social Media, Google, SEO & SEM*. Heidelberg, München, Landsberg, Frechen, Hamburg: bhv.
- Schnell, R., Hill, P., & Esser, E. (2008). *Methoden der empirischen Sozialforschung*. München: Oldenbourg.

- Schulz, W., Held, T., & Laudien, A. (2005). *Suchmaschinen als Gatekeeper in der öffentlichen Kommunikation. Rechtliche Anforderungen an Zugänglichkeit und Transparenz bei Suchmaschinen im WWW*. Berlin: Vistas.
- Searchmetrics (2012). SEO Ranking Faktoren für Google Deutschland 2012. Abgerufen von <http://www.searchmetrics.com/de/services/whitepaper/seo-ranking-faktoren-deutschland/>
- SEOMoz (2011). 2011 Search Engine Ranking Factors. Abgerufen von <http://moz.com/article/search-ranking-factors/2011>
- Statista (2014). Suchmaschinenverteilung in Deutschland im Februar 2014. Abgerufen von <http://de.statista.com/statistik/daten/studie/167841/umfrage/marktanteile-ausgewaehlter-suchmaschinen-in-deutschland/>
- Sullivan, D. (2010, 1. Dezember). What Social Signals Do Google & Bing Really Count? *Search Engine Land*. Abgerufen von <http://searchengineland.com/what-social-signals-do-google-bing-really-count-55389>
- Tantau, B. (2011). „Social Linkbuilding“: Mehr Reichweite bei Google mit sozialen Netzwerken. *Website Boosting*, 5(6), 84-87.
- van Eimeren, B., & Frees, B. (2013). Rasanter Anstieg des Internetkonsums - Onliner fast drei Stunden täglich im Netz. *Media Perspektiven*, (7-8), 358-372. Abgerufen von http://www.ard-zdf-onlinestudie.de/fileadmin/Onlinestudie/PDF/Eimeren_Frees.pdf
- Weck, A. (26. April 2013). Die größte Facebook-Studie aller Zeiten. *t3n*. Abgerufen von <http://t3n.de/news/1-million-user-analysiert-grose-461058/>
- Wolling, J. (2002). Suchmaschinen – Gatekeeper im Internet. *Medienwissenschaft Schweiz*, (2), 15-23.

Empfohlene Zitierung: Jürgens, P., Stark, B., & Magin, M. (2015). Messung von Personalisierung in computervermittelter Kommunikation. In A. Maireder, J. Ausserhofer, C. Schumann, & M. Taddicken (Hrsg.), *Digitale Methoden in der Kommunikationswissenschaft* (S. 251-270). doi: 10.17174/dcr.v2.11

Zusammenfassung: Das Ziel personalisierter Online-Angebote ist, Rezipienten bei der Informationssuche zu unterstützen. Dabei greifen sie zwangsläufig in deren Auswahlentscheidungen ein und müssen deshalb als eigenständiger Einflussfaktor empirisch erschlossen werden. Die dahinterstehenden Algorithmen kommerzieller Anbieter sind für die Forschung größtenteils eine intransparente „Black Box“. Automatisierte Online-Experimente stellen eine selten eingesetzte Methode dar, die durch systematische Simulation von Nutzerverhalten die Funktionsweise von Personalisierungsalgorithmen ermitteln kann. Der Beitrag diskutiert zunächst Auswirkungen und Funktionsweise von Personalisierung und stellt daraufhin ein automatisiertes Online-Experiment am Beispiel der Google-Suche dar. Die vorgestellte Methode ermöglicht einen sozialwissenschaftlichen Zugriff auf die Funktionsweise und Inhalte von personalisierten Angeboten und fördert gleichzeitig Validität, Transparenz und Replizierbarkeit von Nutzungsstudien.

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

Pascal Jürgens, Birgit Stark & Melanie Magin

Messung von Personalisierung in computervermittelter Kommunikation

1 Das individualisierte Web

Mit der Angebotsexplosion im Internet ist die Aufmerksamkeit der Nutzer zu einem knappen Gut geworden. Neben Medienunternehmen suchen auch zahllose andere Webseiten, Rezipienten anzuziehen und zu binden, um Werbeerlöse zu generieren oder Verkäufe zu erzielen. Der wachsenden Konkurrenz begegnen insbesondere technologieaffine Firmen mit personalisierten Angeboten, die die Zufriedenheit der Nutzer erhöhen sollen (Teevan, Dumais, & Horvitz, 2010). Dabei werden Rezipienten bevorzugt solche Inhalte angeboten, die zu ihren individuellen Interessen passen – seien das Filme aus dem Angebot von T-Online, Artikel auf sueddeutsche.de oder Bücher bei Amazon.

Mit dem Siegeszug der digitalen Informationsübertragung geht dabei eine Verschiebung an Verantwortung einher: Wo Rezipienten zuvor eigenständig, geleitet von menschlichen Entscheidungs-, Bewertungs- und Handlungsstrategien Kommunikation bestimmten, können sie nun einen Teil dieser Entscheidungen auslagern. Die (vermeintlich) passendsten Bücher empfiehlt *Amazon*, die interessanteste Musik Spotify und die wichtigsten Statusnachrichten von Freunden kennt Facebook – und in jedem dieser Fälle ist die Empfehlung nicht das Produkt eines menschlichen Experten, sondern das eines Algorithmus. Das alltägliche Menü an Medieninhalten, Unterhaltung, Gütern, Informationen und sozialen

Kontakten wird also von neuen Intermediären mitbestimmt, die sich als Vermittler in die ursprünglich direkte Beziehung zwischen Rezipienten und den etablierten Anbietern bzw. deren Produkte schieben.

Diese Anbieter stehen dem Nutzer dabei meist nicht als Gatekeeper zur Seite, die Informationen weiterleiten bzw. vorenthalten, sondern vielmehr als Ratgeber, der aus einem großen Angebot einzelne Artikel hervorhebt. Die wichtigste Konsequenz solcher algorithmisch erzeugter Zusammenstellungen von Informationen ist, dass Inhalte von Rezipient zu Rezipient abweichen. Dadurch greifen Wirkungshypothesen zu kurz, die davon ausgehen, dass Webseiten zu einem gegebenen Zeitpunkt für jeden Besucher dieselbe Informationsauswahl darstellen. So ist beispielsweise denkbar, dass Nachrichtenangebote ihre Artikelauswahl an die individuellen politischen Einstellungen ihrer Nutzer anpassen.¹ Forschung zu politischer Selektivität (Selective Exposure) stünde dann vor der Herausforderung, zwischen dem Auswahlverhalten durch politisch motivierte Vermeidung bzw. Zuwendung und jenem durch die Personalisierung angebotener Artikel zu differenzieren bzw. ihre Verknüpfung zu beachten.

Die kommunikationswissenschaftliche Forschung sieht sich deshalb zunehmend mit dem Problem konfrontiert, bestehende theoretische und methodische Grundlagen der Online-Forschung um das Element personalisierter Inhalte zu erweitern. Ein prominentes Beispiel hierfür ist die Suchmaschinenforschung, die sich lange mit Verzerrungen durch Suchalgorithmen – wie der verstärkten Lenkung der Nutzer zu populären Webseiten (z.B. Hindman, 2009) – beschäftigt hat. Seit Suchergebnisse an das Verhalten der Suchenden angepasst werden, sind die bestehenden Befunde nicht länger übertragbar: Ob etwa Nutzer bevorzugt populäre („Mainstream“-)Seiten angezeigt bekommen, oder auch Nischenangebote auftauchen, kann von personalisierten Suchergebnissen und somit vom Nutzerverhalten abhängen (Jürgens, Stark, & Magin, 2014). Vor ähnlichen Herausforderungen stehen alle Forschungszweige, deren Gegenstand personalisierte Angebote beinhaltet. Hierzu zählen Nachrichtenseiten inzwischen ebenso wie E-Commerce-Anwendungen, Nachrichtenaggregatoren und soziale Netzwerkseiten.

1 In den USA setzt die Huffington Post bereits Personalisierung ein, um Nutzern eine interessensspezifische Artikelauswahl zu präsentieren (Reid, 2014).

Sowohl die Funktionsweise als auch die Auswirkungen von Personalisierung sind in der Kommunikationswissenschaft noch weitgehend ungeklärt. Im Zuge der fortschreitenden Verbreitung der Technologie (Springer, 2014) wird es zunehmend wichtiger, zumindest empirisch überprüfen zu können, ob ein Angebot personalisiert ist und gegebenenfalls die Auswirkungen exemplarisch untersuchen zu können. Der vorliegende Aufsatz liefert hierzu einen ersten Beitrag: Zuerst stellt er die Verfahren der Personalisierung sowie damit verbundene Wirkungshypothesen kurz vor und entwickelt anschließend eine computergestützte Methode, anhand derer Personalisierungseffekte identifiziert und gemessen werden können. Abschließend werden potenzielle Weiterentwicklungen im Kontext technischer Innovationen diskutiert.

Personalisierung als Bias

Eine empirische Messung der Auswirkungen von Personalisierung erscheint auf den ersten Blick kaum umsetzbar: Welche Arten von Inhalten für welche Nutzer in welchem System bevorzugt angeboten werden, hängt nicht nur von einer Vielzahl an Kontextvariablen (Zeit und Ort des Abrufs), sondern auch von dem Verhalten des Nutzers ab und ist deshalb schwer vorhersagbar. Sowohl in der Informatik als auch in der Kommunikationswissenschaft stehen Bemühungen, die Auswirkungen dieser komplexen Interaktionen zu erfassen, erst am Anfang.² Erste theoretische und methodische Annäherungen können sich allerdings auf zwei bestehende Konzepte stützen, die zumindest eine erste Orientierung bieten:

- (1) In Bezug auf die *einheitliche Wirkung* von Softwaresystemen stellt sich zunächst die Frage nach einer von Personalisierung unabhängigen, systematischen Bevorzugung oder Benachteiligung spezifischer Inhalte. Diese Perspektive ist unter dem Stichwort des „Bias“ sowohl in der Kommunikationswissenschaft als auch in der Informatik bereits etabliert (Friedman & Kahn, 2008; Friedman & Nissenbaum, 1996). Im Kern von Bias-Hypothe-

2 Den größten Beitrag in diesem Bereich liefert wohl das interdisziplinäre Forschungsfeld der „Complexity Science“, in dem sich Informatik, Physik, Wirtschaftswissenschaften, Psychologie und weitere Disziplinen der Modellierung von Systemen mit nichtlinearen Verhaltensweisen annehmen (siehe Mitchell, 2009).

sen steht die Erkenntnis, dass Computersysteme trotz ihrer maschinellen Funktionsweise nicht automatisch neutrale Informationsvermittler sein müssen (Friedman & Nissenbaum, 1996; Rieder, 2004). Im Zuge der Verarbeitung können sie vielmehr Informationen gänzlich unterdrücken (als Gatekeeper, Filter oder Zensor) sowie in der Darstellung hervorheben oder herunterspielen. Eine Verzerrung oder „Bias“ liegt dieser Sichtweise folgend dann vor, wenn die Zugänglichkeit zu Informationen im Verarbeitungsprozess durch ein Computersystem verändert wird. Im Fall des Herausfilterns von Informationen ist der Bias vergleichsweise einfach zu definieren und empirisch nachzuweisen: Sobald Inhalte der Software als Input zur Verfügung stehen, aber nicht an den Nutzer weitergereicht werden, wirkt sie verzerrend. Dies ist beispielsweise der Fall, wenn Facebook-Nutzer nur ausgewählte Nachrichten von Freunden zu sehen bekommen (Bucher, 2012) oder wenn Suchmaschinen einzelne Seiten aufgrund von staatlichen Vorgaben nicht in den Ergebnislisten anzeigen (Taneja & Wu, 2013). Wenn es allerdings nicht um das Vorenthalten, sondern um die Hervorhebung von Informationen geht, wird die Definition von Verzerrungen schwieriger: Denn dann lässt sich ein Bias nur im Vergleich zu einem neutralen Maßstab konzipieren. So wäre ein verzerrendes System beispielsweise eine Hotelsuchmaschine, die Angebote der eigenen Firma grundsätzlich an oberster Stelle anzeigt, selbst wenn diese nicht die günstigsten sind (Vergleich der produzierten Reihenfolge am neutralen Maßstab Preis).

- (2) Der Zweck von personalisierten Angeboten besteht darüber hinaus darin, *unterschiedlichen Nutzern differierende Inhalte* vorzulegen. Hierdurch können solche Systeme Unterschiede in der Informationsauswahl verstärken und so Tendenzen zur Fragmentierung in der Nutzerschaft auslösen. Diese auch als „Filter Bubble“ (Pariser, 2011) bekannt gewordene und von zahlreichen Autoren vorgebrachte Hypothese (Sunstein, 2009) ist eine mögliche Folge der Überspezialisierung („Over-Specialization“). Aus dem Blickwinkel der Praxis erscheint dieses dystopische Szenario von vollständig isolierten Nutzern und einer komplett fragmentierten Öffentlichkeit allerdings nicht realistisch (hierzu auch Garrett, Carnahan, & Lynch, 2013). Eine vollständig personalisierte Informationsauswahl ist hier bereits deshalb nicht wün-

schenswert, weil vielfältige Angebote den Nutzern häufig einen größeren Mehrwert bieten (siehe Diaz, 2008; Adomavicius & Tuzhilin, 2005).

2 Personalisierung aus technischer Sicht

Im Kontext computervermittelter Kommunikation werden unter dem Begriff „Personalisierung“ Verfahren zusammengefasst, die digitale Inhalte in Abhängigkeit von bestimmten Merkmalen eines Nutzers anpassen (eine Übersicht über Definitionen unterschiedlicher Disziplinen liefern Fan & Poole, 2006; siehe auch Pitkow, Schütze, Cass, Cooley, Turnbull, Edmonds, et al., 2002). Für die weitere Argumentation werden unter Personalisierung (und in der Konsequenz unter personalisierten Angeboten) Systeme verstanden, die Inhalte einer Webseite gegenüber dem Nutzer durch Anwendung von Algorithmen anpassen.

Unter diese breite Definition fällt der regionale Wetterbericht einer Nachrichtenwebseite (Anpassung an den Ort des Nutzers) ebenso wie der Kontaktvorschlag einer Webseite für Partnervermittlung (Anpassung an Interessen und Charakter des Nutzers). Während im ersten Fall das dahinterstehende Computersystem vorhandene Daten³ ohne Interaktion mit dem Nutzer verwenden kann, sind die Merkmale im zweiten Fall nur explizit erfragbar (Sundar & Marathe, 2010). Der angestrebte Nutzen für den User ist in beiden Fällen gleich: „passende“ Informationen sollen den Nutzer bei der Selektion unter einer Vielzahl an potenziell uninteressanten Informationen unterstützen, ihn dadurch entlasten und schnell zu den „besten“ Inhalten führen, was in einer höheren Zufriedenheit mündet (Teevan et al., 2010).

Den Kern jedes personalisierten Angebots stellen die gesammelten Daten zu den vermuteten Interessen des Nutzers dar. In dem sogenannten „user model“ oder „user profile“ (Jannach, Zanker, Felfernig, & Friedrich, 2010, S. 1) werden Informationen gesammelt, die über den individuellen Geschmack und somit über

3 So kann beispielsweise der Ort eines Nutzers grob anhand der IP-Adresse des genutzten Geräts ermittelt werden (GeoIP). Eine frei verfügbare Datenbank mit entsprechenden Daten bietet beispielsweise MaxMind an. Die Open Source-Software „freegeoip“ kann als passende Rechercheplattform dazu verwendet werden.

die Gratifikationserwartungen des Nutzers Aufschluss geben. Ein solches Interessenprofil kann unter Einsatz von expliziten Fragen entstehen, zum Beispiel indem eine Dating-Webseite Selbsteinschätzungen des Charakters und persönliche Vorlieben abfragt. In vielen Fällen werden die Profile allerdings implizit aus den ohnehin bei der Benutzung des Angebots entstehenden Daten generiert. Dies ist etwa der Fall, wenn aus den in der Vergangenheit gekauften Produkten Vorschläge für zukünftige Kaufentscheidungen produziert werden.

Die Forschung zu *Recommender Systems* trennt weiterhin zwei Grundtypen von Algorithmen, die sich in Bezug auf die Empfehlungskriterien für „passende“ Inhalte unterscheiden (Ricci, Rokach, Shapira, & Kantor, 2010; Jannach et al., 2010):

- (1) Das sogenannte „*Collaborative Filtering*“ vergleicht lediglich unterschiedliche Profile miteinander und ermittelt, welche Rezipienten sich anhand des bisherigen Verhaltens ähnlich sind. Nach diesem Prinzip zeigt beispielsweise Google News personalisierte Nachrichten, die mit Hilfe der Profile ähnlicher Nutzer zusammengestellt werden (Das, Datar, Garg, & Rajaram, 2007). Ebenso bietet Amazon seinen Kunden Bücher an, die andere Leser mit ähnlichen Interessen bereits gekauft haben. Diese Gruppe an Verfahren ist relativ einfach und universell umsetzbar, da sie nicht auf die Eigenschaften der Produkte eingeht. Was für Objekte empfohlen werden ist egal: So lange genügend Datensätze vorliegen, lassen sich Personen mit ähnlichen Interessen identifizieren und dem Nutzer bislang unbekannte Einträge aus deren Profil anbieten.
- (2) Eine Weiterentwicklung des „*Collaborative Filtering*“ sind Empfehlungssysteme, die sich auch nach dem Inhalt empfohlener Texte richten („*Content-based recommendation*“, Jannach et al., 2010). Ein solches System kommt bei der US-Ausgabe der Huffington Post zum Einsatz. Dort ermittelt es aus den zuvor gelesenen Artikeln die Themenpräferenzen von Nutzern und passt die Startseite an, um die individuellen Präferenzen besser zu bedienen (Reid, 2014). Der Algorithmus ist dabei journalistischen Selektionsentscheidungen sehr ähnlich, da er erwartete Interessen auf Seiten des Rezipienten mit Merkmalen der empfohlenen Inhalte in Bezug setzt. Je deutlicher ein Objekt mit den Wünschen des Nutzers zusammenhängt, des-

to höher wird es eingestuft. Inhaltsbezogene Empfehlungssysteme haben den Vorteil, dass auch dann Empfehlungen abgegeben werden können, wenn bislang nur wenige Daten über andere Nutzer vorliegen. Die Kehrseite aus wirtschaftlicher Perspektive ist allerdings eine deutlich gesteigerte Komplexität für die Programmierer: Zunächst müssen die relevanten Merkmale der Artikel (z.B. Thema, Gegenstand, Textgattung usw.) erhoben werden. Außerdem ist nicht immer offensichtlich, aufgrund welches Merkmals ein Nutzer einen Artikel ausgewählt hat – ob er/sie also das Thema interessant findet, die Textgattung mag oder den Autor schätzt. Die genaue Beziehung muss deshalb erst durch statistische Verfahren ermittelt werden (Jannach et al., 2010). Ein in der Praxis erfolgreiches und häufig angeführtes Beispiel für ein solches Empfehlungssystem ist das Portal „Rotten Tomatoes“, auf dem Nutzer Filme finden können, die ihren eigenen Vorlieben (z.B. in Bezug auf Genre, auftretende Schauspieler usw.) entsprechen. Innerhalb dieser beiden Typen differenzieren sich einzelne System weiter aus. Je nach Anwendungsgebiet werden unterschiedliche Nutzer- und Inhaltsmerkmale gemessen und diese anhand verschiedener Algorithmen miteinander verglichen.

Von der verwendeten Definition von „Personalisierung“ ausgenommen sind hingegen Anwendungen, die der Nutzer selbst den eigenen Vorstellungen und Wünschen entsprechend anpassen kann, indem er beispielsweise das Aussehen einer Webseite oder die Auswahl der angezeigten Informationen ändert. Solche Anpassungen, die der Nutzer aktiv selbst vornimmt, sind in der Fachliteratur unter den Begriff „customization“ (also eine Modifikation) gefasst. Personalisierte Angebote hingegen funktionieren vollautomatisch und ohne Initiative des Rezipienten (siehe zu dieser Unterscheidung auch Sundar & Marathe, 2010).

3 Methodische Herausforderungen durch Personalisierung

Die Entwicklung von neuen Instrumenten in der Forschung zu computer-vermittelter Kommunikation ist mit dem Anspruch konfrontiert, Neu- und Weiterentwicklungen im Bereich der Personalisierung zu berücksichtigen. Insbesondere die Sicherung der Validität stellt hier eine zentrale Herausforderung dar.

Bei dieser Aufgabe steht die empirische Kommunikationswissenschaft von zwei Seiten aus unter Druck (siehe z.B. zu den Herausforderungen für die Suchmaschinenforschung auch Granka, 2010):

- (1) Die *Sicherung der internen Validität* stellt vorwiegend eine technische Herausforderung dar: Online-Angebote beziehen zunehmend Indikatoren in ihre Funktionsweise ein, die aus dem Verhalten der Nutzer abgeleitet sind. Webseiten versuchen für Leser attraktiver zu werden: Sie passen die Zusammenstellung der Startseite an die vermuteten Interessen des Nutzers an und gehen in einigen Fällen sogar auf besondere Details aus seinem Leben ein – was Google beispielsweise durch ein Glückwunschk-Doodle zum Geburtstag erreicht (Springer, 2014). Empirische Methoden stehen aufgrund der Vielzahl an personalisierten Angeboten unter einem anhaltenden Innovationsdruck, um nicht an interner Validität einzubüßen: So lange empirische Studien die genaue Zusammenstellung personalisierter Angebote nicht erheben können (siehe Emmer & Strippel in diesem Band) besteht die Gefahr, dass Personalisierung als Störfaktor auftritt. Untersucht beispielsweise ein Experiment innerhalb von Facebook die Reaktion von Nutzern auf Anfeindungen, dann können die Intensität der Beziehung zwischen „Angreifer“ und „Opfer“ sowie die daraus resultierende Personalisierung des Feeds als konfundierende Variable wirken: Sind die beiden nur weitläufig miteinander bekannt, wird dem „Opfer“ der Post des „Angreifers“ gegebenenfalls gar nicht angezeigt (siehe auch Bucher, 2012). Ein erster Schritt zur Untersuchung von personalisierten Angeboten muss demnach sein, die Funktionsweise der Personalisierung selbst zu messen. Da in den meisten Fällen unbekannt ist, welche Aspekte des Nutzerverhaltens (z.B. die angeklickten Artikel, das „Liken“ von Inhalten) und welche Kontextfaktoren (z.B. der verwendete Browser, der Zeitpunkt des Aufrufs, das verwendete Gerät) in die Personalisierung eingehen, ist die Simulation von Rezipientenverhalten ein besonders geeignetes Mittel. Dabei gibt sich der Forscher gegenüber einem Angebot als neuer Nutzer aus und interagiert systematisch mit der Webseite, wobei Kontextfaktoren wie Ort, Zeit und die verwendete Browser-Software kontrolliert werden. Zudem bleiben durch das automatisierte „Klickverhalten“ Faktoren konstant, deren Rolle für die Personalisierung nicht be-

kannt ist – wie etwa die Latenz zwischen Aufruf der Seite und dem Klick auf ein Ergebnis.

- (2) Ein wesentlich größeres Problem ist die *Sicherung der externen Validität*: In den meisten Fällen dürfte nicht bekannt sein, wie sich Rezipienten in ihrer Interaktion mit einer Seite verhalten. Da aber deren Inhalt selbst nun vom Verhalten abhängt, bleibt unklar, welchen Inhalten die Rezipienten de facto ausgesetzt sind (siehe Emmer & Strippel in diesem Band). Damit Erkenntnisse über das Rezeptionsverhalten auch unter den Vorzeichen einer Personalisierung generalisierbar sind, muss entweder bekanntes Nutzerverhalten nachgeahmt oder aber systematisch konzipiert werden. Die einfachste Form systematischen Verhaltens wäre dabei eine Monte-Carlo-Simulation (Mooney, 1997), die blind eine große Zahl zufälliger Stimuli anwendet und aus den Ergebnissen die Funktionsweise des Systems rekonstruiert. Ein solches Vorgehen ist jedoch nur dann pragmatisch durchführbar, wenn die Anzahl der möglichen Eingaben begrenzt ist. Für den programmatischen Zugang über eine API (Application Programming Interface) sind beispielsweise nur 100 Anfragen pro Tag umsonst – darüber hinaus sind \$5 pro 1.000 Anfragen zu bezahlen (Google, 2014).

Hinter der gestiegenen Anforderung an die Erhebungsmethodik stehen größtenteils technische Gründe. Personalisierung entsteht erst im Laufe der Zeit, da die entsprechenden Angebote zunächst Informationen über den Nutzer sammeln müssen, anhand derer die Webseite subsequent angepasst werden kann. Um überhaupt einen Eindruck der Personalisierung – etwa von Google News – zu erlangen, muss der Forscher zumindest die Webseite für zwei deutlich unterschiedliche Nutzungsstile im Zeitverlauf vergleichen. Am leichtesten fällt dies, wenn dazu Nutzerkonten mit Passwörtern erstellt werden. Denn dann ist jeweils sicher gestellt, dass nur die Webseiteninteraktion nach dem Login Einfluss auf das jeweilige Nutzerprofil hat. Sollen Inhalte systematisch erfasst werden, muss die dazu eingesetzte Software ebenfalls den passenden Benutzernamen und das Passwort verwenden, um Zugriff auf dieselbe Ansicht zu erhalten. Bereits an dieser Stelle können jedoch Komplikationen auftreten: Als Schutz vor Missbrauch (z.B. durch gezielte Manipulation mit kommerziellen Zielen wie im Fall von Search Engine Optimization) verhindern viele Betreiber von Webseiten den Zugriff durch

automatisierte Programme. Anhand der Hinweise von technischen Merkmalen wie zum Beispiel dem Namen des verwendeten Programms⁴ oder Verhaltensindikatoren (z.B. die Anzahl und Regelmäßigkeit von Anfragen) werden Zugriffe gegebenenfalls abgewiesen.

Der wissenschaftliche Zugriff wird weiterhin durch Programme erschwert, die auf dem Computer des Nutzers im Webbrowser ablaufen und die Seite erst dort in ihre letztlich sichtbare Form bringen. Die in der Programmiersprache *JavaScript* geschriebenen Anwendungen werden insbesondere eingesetzt, damit Webseiten schnell und einfach interaktiv bedient werden können. Klickt ein Nutzer eine entsprechende Schaltfläche an, wird keine komplett neue Seite geladen. Vielmehr fragt das Programm im Hintergrund neue Daten vom Server an, um diese dann in die bestehende Seite einzufügen und ggf. vorhandene Daten zu ersetzen (McCaw, 2011). Auf Basis dieser Technologien entwickelte und meist als „Web Apps“ bezeichnete Angebote, sind für den Großteil der Werkzeuge zur systematischen Speicherung und Archivierung von Webseiten nicht abbildbar. Wer beispielsweise eine Seite mit Suchergebnissen von Google mittels eines Computerprogramms herunterlädt⁵, erhält nur eine weitgehend leere „Hülle“. Die eigentliche Liste an Webseiten wird erst durch den Browser des Nutzers nachgeladen. Solche dynamischen Angebote stellen eine erhebliche Herausforderung für die reliable Erfassung der Inhalte von modernen Web-Angeboten dar (grundlegend aus der Perspektive digitaler Archive: Masanès, 2005, 2006; Lyman, 2002).

4 Das Öffnen der „Black Box“

Den Herausforderungen in Bezug auf die interne und externe Validität begegnet das hier vorgestellte Forschungsdesign durch eine automatisierte Simulation, die die bekannten Kontext-Drittvariablen (verwendete Soft- und Hardware, IP-Adresse, Zeitpunkt des Aufrufs) kontrollieren kann und eine authentische Simu-

4 Der sogenannte „User Agent“, den ein Programm bei HTTP-Anfragen mitsendet, kann allerdings frei manipuliert werden (RFC 2616, 14.43: <http://tools.ietf.org/html/rfc2616#section-14.43>)

5 Die zugehörige URL folgt schlicht dem Schema „<https://www.google.de/?q=test>“ mit „test“ als Suchbegriff.

lation von Nutzerverhalten erlaubt. Prinzipiell ermöglicht diese Methode den Zugriff auf personalisierte Angebote, durch den die angezeigten Inhalte passiv erhoben werden können und durch den auch aktiv mit ihnen interagiert werden kann. Eine erste nützliche Anwendung sind Studien mit quasi-experimentellem Design, die Licht auf die Funktionsweise von personalisierten Angeboten werfen.

Je mehr Online-Angebote Nutzer mit personalisierten Empfehlungen bei der Auswahl von Inhalten unterstützen, desto stärker nehmen sie dadurch Einfluss auf Rezeptionsentscheidungen (siehe zu Suchmaschinen etwa Pan, Hembrooke, Joachims, Lorigo, Gay, & Granka, 2007). Forschungsfelder wie die Mediennutzungsforschung, die Nutzerentscheidungen zuvor als Handeln des Rezipienten gegenüber einem fixen Angebot auffassen konnten, müssen ihren Fokus in Zukunft um Personalisierungseffekte erweitern.

Die größte Herausforderung ist dabei die Intransparenz der Funktionsweise: Wie genau ein Algorithmus Informationen auswählt und bewertet, kann mit abschließender Sicherheit nur an seinem Quellcode abgelesen werden. Ist ein Einblick in den Code nicht gestattet, dann muss die Funktionsweise der Software rekonstruiert werden. Hierzu setzt die verwendete Methode simuliertes Nutzerverhalten ein: Die grundlegende Arbeitsweise des Personalisierungsalgorithmus wird ermittelt, indem In- und Output unter systematisch veränderten Bedingungen verglichen werden. So könnte beispielsweise eine Nachrichtenwebseite von zwei simulierten Nutzern besucht werden, von denen einer nur Artikel über nationale Politik liest, während sich der zweite nur für internationales Geschehen „interessiert“. Wenn die Startseite anhand der Themenpräferenzen personalisiert wird, dann zeigt sich dies an einer im Zeitverlauf divergierenden Zusammenstellung der Themen.

Algorithmen sind komplexe Systeme, die viele verschiedene Variablen verwenden und oft schwierig nachzuvollziehenden Art und Weise verarbeiten (siehe z.B. McGeoch, 2002 zu Simulationsstudien von Algorithmen)⁶. Fälle, in denen ein großes

6 Einführende Darstellungen zu der Verwendung von Experimenten als Instrument in den Ingenieurwissenschaften und der Informatik finden sich bei Mason, Gunst und Hess (2003) und Feitelson (2006) mit Parallelen zur sozialwissenschaftlichen Experimentallogik. Kohavi, Longbotham, Sommerfield und Henne (2009) behandeln Web-Experimente mit menschlichen Probanden, wohingegen Paxson (2004) auf die Rolle von technischen Hürden bei Messungen von Internet-Anwendungen eingeht. Die Literaturstudie von Tichy, Lukowicz, Prechelt und Heinz (1995) betrachtet schließlich die Professionalisierung des Feldes in Bezug auf die Verwendung von Experimenten.

Spektrum aller In- und Outputs eines solchen Systems im Stil einer Monte-Carlo-Simulation ausprobiert werden können, dürften eine Seltenheit darstellen. Die meisten Web-Anwendungen erlauben vielmehr eine nahezu uneingeschränkte Bandbreite an Eingaben, die in Empfehlungen von Inhalten umgewandelt werden. Diese Konstellationen können selbstverständlich nicht alle getestet werden. Vielmehr müssen empirische Studien – wie auch bei klassischen Experimenten – zugespitzte Kontrastbedingungen finden, die mit einer zielführenden und theoriebezogenen Hypothese verknüpft sind. Sinnvolle Anknüpfungspunkte hierfür sind die eingangs angesprochenen Typen der Verzerrung: Wird beispielsweise die Zensur bestimmter politischer Positionen vermutet, dann muss verglichen werden, ob der Output des Algorithmus relevante Inhalte vermissen lässt, die eigentlich verfügbar wären.

Fallstudie Googleisierung der Informationssuche

Die Methode der automatisierten Simulation von Nutzerverhalten wurde im Rahmen des Forschungsprojekts „Googleisierung der Informationssuche“ das erste Mal eingesetzt, um Personalisierungseffekte bei Google-Suchanfragen zu messen (Jürgens et al., 2014). Das Experiment diente als empirischer Test der Hypothese, dass personalisierte Ergebnisse von Suchmaschinen zu einer Fragmentierung der von Nutzern wahrgenommenen Informationen führen („Filter Bubble“, Pariser, 2011). Die zu prüfende Nullhypothese war dementsprechend, dass die angezeigten Suchergebnisse für zwei Nutzer im Schnitt identisch sind, obwohl diese sich in ihrem Suchverhalten (das die Personalisierung bestimmt) unterscheiden. Im Folgenden werden die einzelnen Phasen der Methode exemplarisch dargestellt. Prinzipiell kann dasselbe Vorgehen mit nur geringfügigen Anpassungen auch auf andere Anbieter wie zum Beispiel Facebook mit dem EdgeRank-Algorithmus angewendet werden.

Zugang zu Angebot und Nutzerprofilen

Der wichtigste Vorteil der automatisierten Interaktion mit Web-Angeboten liegt in der Reliabilität, der Validität und Replizierbarkeit der Erhebungen (King, 1995). Voraussetzung dafür ist, dass ein Computerprogramm mit dem Angebot so interagieren kann, dass es dieselben Inhalte erhält, die auch ein menschlicher Nutzer sehen würde. Damit automatisierte Anfragen authentisch erscheinen, müssen sie einen Internet-Browser starten und darin Aktionen wie die Eingabe

von Text oder Mausklicks auslösen können. Für die Evaluation und Tests von Software wurden mehrere Programme entwickelt, die das erlauben: Projekte wie Selenium und PhantomJS erlauben dem Forscher, Aktionen innerhalb des Browsers zu automatisieren, ohne dass dieser auf dem Bildschirm sichtbar sein muss. Für die Messung der Google-Personalisierung wurde ein Programm erstellt, das PhantomJS nutzt, um zunächst einen neu angelegten Nutzer bei Google anzumelden. Dazu startet das Programm den Browser, öffnet die Google-Startseite, klickt auf den Link zur Anmeldung, trägt Benutzername und Passwort ein und landet schließlich erneut auf der Startseite. Bei jedem weiteren Start wird für den jeweiligen Nutzer der Login-Status über Cookies wieder hergestellt.⁷

Kontrolle von Kontextvariablen

Bei der automatisierten Verwendung der Google-Suchmaschine besteht stets die volle Kontrolle über alle technischen Merkmale des Browsers. Sofern keine Cookies geladen werden, startet das Programm bei jedem neuen Durchlauf in einem deterministischen, neutralen Zustand und übermittelt deshalb keine profilbezogenen Daten. So wird sichergestellt, dass die Webseite jeweils nur exakt die vorgesehene Nutzerinteraktion als Input erhält und nicht auf Spuren vergangener Nutzung (z.B. bestehende Cookies) zurückgreift. Sofern das Programm stets auf demselben Computer läuft, sind auch Änderungen der IP-Adresse bzw. des Ortes als Einfluss auf Personalisierung ausgeschlossen.

Stimulation von Personalisierung

Nachdem der Zugang zu dem jeweiligen Webangebot – hier Google – besteht, wird die Personalisierung durch gezielte Nutzung stimuliert. Dies setzt zumindest eine einfache Annahme voraus, welches Verhalten die Grundlage für die Anpassungen des Algorithmus bildet. Hierzu muss aus der Fachliteratur, aus firmeneigenen Publikationen und Experteninterviews zumindest ein plausibler Verhaltensindikator ermittelt werden, von dem angenommen wird, dass er sich auf die Personali-

7 Siehe hierzu die Dokumentation von PhantomJS: <http://phantomjs.org/api/phantom/property/cookies.html> bzw. Selenium http://docs.seleniumhq.org/docs/03_webdriver.jsp#cookies

sierung auswirkt. Die Google-Suche protokolliert nicht nur, welche Suchanfragen ein Rezipient sendet, sondern auch, auf welche Ergebnisse er klickt. Daher bestand die Manipulation aus regelmäßigen Suchen, bei denen jeweils zwei Nutzer nach Begriffen aus zwei unterschiedlichen Bereichen eines Themas suchten (zur Entwicklung des Stimulusmaterials siehe Jürgens et al., 2014). Nach jeder Suche wurde ein zufällig ausgewählter Link der ersten Ergebnisseite angeklickt (siehe Beispieldesign in Tabelle 1).

Für das Design der Manipulation ist die Anknüpfung an die eingangs beschriebenen Verzerrungstypen sinnvoll, um zum Beispiel wie im angeführten Beispiel auf simple Unterschiede zwischen zwei Profilen zu schließen. Prinzipiell ist es allerdings auch möglich, Personalisierung mit mehr als zwei Ausprägungen der Manipulation zu konzipieren, wenn etwa überprüft werden soll, ob die Anpassung an Nutzerinteressen bei verschiedenen politischen Präferenzen identisch ausfallen. Dann sind jeweils so viele simulierte Nutzer nötig wie Ausprägungen der UV vorhanden.

Tabelle 1: Beispiel eines Experimentaldesigns

Faktor: Bedeutungsebenen des Wortes „Stollen“	Treatment	Messung
<i>Experimentalnutzer 1</i>	<i>Sucht nach Begriffen rund um Weihnachten</i>	
<i>Experimentalnutzer 2</i>	<i>Sucht nach Begriffen rund um Bergbau</i>	<i>Suche nach „Stollen“</i>
<i>Kontrollsuche</i>	<i>Keine Suche</i>	

Messung der Personalisierung

Der eigentliche Test auf Personalisierung erfolgt in regelmäßigen Abständen durch den Vergleich eines einheitlichen Stimulus. Im Fall von Google sind dies Ergebnislisten, die beide Nutzer zu ein und demselben themenbezogenen Suchbegriff erhalten. Erhält ein Nutzer Treffer angezeigt, die dem anderen Profil feh-

len und gleichzeitig mit den Stimulus-Suchanfragen zusammen hängen, dann kann davon ausgegangen werden, dass dies eine Folge der Anpassung an das Nutzerprofil, also ein personalisiertes Ergebnis ist. Da Computersysteme deterministisch vorgehen ist für die Verallgemeinerung der Ergebnisse keine Stichprobe nötig: Wird die Personalisierung einmalig nachgewiesen, gilt das Ergebnis zumindest solange der Anbieter den Algorithmus nicht ändert. Für die Google-Suche konnten auch nach vierwöchiger Simulation keine nennenswerten, insbesondere keine mit dem Stimulus zusammenhängenden Personalisierungen nachgewiesen werden (eine ausführliche Darstellung findet sich in Jürgens et al., 2014).

In einer vereinfachten Darstellung läuft eine automatisierte Erhebung nach dem folgenden universellen Schema ab:

Tabelle 2: Ablauf eines Online-Experiments zur Messung von Personalisierungseffekten

1	Start des Browsers und laden des Skripts, das den Ablauf der Studie in Programmform beschreibt
	ggf. Laden von nutzerbezogenen Daten aus Cookies oder Login eines Nutzers
2	Öffnen der Initial-URL (z.B. Startseite des Angebots)
3	UV: Interaktion mit dem System mit dem Ziel, Personalisierung auszulösen. Dazu Durchführung von Handlungen, die aus der Vorrecherche mit der Personalisierung in Verbindung gebracht werden (z.B. „Liken“ von Artikeln, Suche nach Begriffen).
4	Protokollierung des Outputs durch Screenshots, Speichern des Quellcodes der Seite oder gezielte Extraktion von Daten
...	Beliebige Wiederholungen von Schritt 3 und 4
5	Messung von Personalisierungseffekten: Vergleich des System-Outputs bei identischen Nutzeranfragen (z.B. Vergleich der Zusammenstellung der Startseite)
6	Abmeldung, Beenden des Browsers, Speicherung der erhobenen Daten

5 Ausblick: Chancen und Risiken

Sozialwissenschaftler sind mit ihren Forschungsinteressen im Internet Leidtragende von rasch fortschreitender technischer Innovation, einem Wettrennen zwischen Anbietern und Manipulationsversuchen durch Dritte (siehe exemplarisch zu der Ko-Evolution von Suchmaschinen und deren Manipulation Metaxas, 2010; zu Spam siehe Brunton, 2013). In diesem „Arms Race“ versuchen Betreiber von Online-Webseiten der gezielten Manipulation ihrer Angebote – sei diese ökonomisch oder ideologisch motiviert – einen Riegel vorzuschieben (siehe etwa den Überblick bei Zerkowitz, 2011; und Jakobsson, 2012). Hierzu setzen sie verstärkt Lösungen ein, die automatisierte Anfragen verhindern, etwa indem Nutzer verzerrten Text lesen und in ein Formular eintippen müssen (sogenannte „Captchas“). Die Online-Forschung steht somit vor dem Problem, durch die veränderte Funktionsweise von Angeboten stärker auf automatisierte Verfahren angewiesen zu sein, während deren Einsatz gleichzeitig einen immer höheren technischen Aufwand mit sich bringt.

Sowohl automatisierte Erhebungen als auch die Simulation von Nutzerverhalten versprechen jedoch nicht nur validere, reliablere und replizierbare Erhebungen, sondern eröffnen der Kommunikationswissenschaft Zugang zu den Inhalten und Funktionsweisen von personalisierten Webangeboten. Bislang müssen sich Erhebungen zum Beispiel von der Nutzung von Nachrichtenwebseiten und sozialen Netzwerkseiten auf Daten verlassen, die vom Betreiber bereitgestellt werden (z.B. Protokolldaten, siehe Pass, Chowdhury, & Torgeson, 2006; und kritisch Zimmer, 2010). Die beschriebene Methode systematisiert dagegen die „normale“ Benutzung der Angebote und kann deshalb ohne privilegierten Zugriff durchgeführt werden, was den Forscher von einem Entgegenkommen des Anbieters unabhängig macht. Das „Verkörpern“ des Nutzers als Kunstpersona kann so helfen, intransparente, technisch determinierte Effekte aufzudecken. Die Vorgehensweise knüpft dabei an Erfahrungen aus der Experimentalforschung an und macht sich deren Stärke in der Formulierung von Fragestellungen und Hypothesen zu nutzen.

Eine Investition in die Methodenentwicklung kann langfristig die Beobachtbarkeit des digitalen Kommunikationsraums sichern (King, 2011). Die hier aufgezeigte Methodeninnovation wird – ebenso wie parallele Entwicklungen aus dem Bereich der *Digital Methods* (Rogers, 2013) – zusätzlich umso wichtiger, je stärker der durchschnittliche Forschungsaufwand von Studien in einer fortwährend

wachsenden digitalen Landschaft steigt (Lazer, Pentland, Adamic, Aral, Barabási, Brewer, et al., 2009). Der mit technischen Weiterentwicklungen steigende Aufwand setzt allerdings entweder interdisziplinäre Kooperationen mit der Informatik oder aber den Einkauf spezialisierter Expertise voraus.

Pascal Jürgens, M.A. ist wissenschaftlicher Mitarbeiter am Institut für Publizistik der Johannes Gutenberg-Universität Mainz

Prof. Dr. Birgit Stark ist Professorin für Kommunikationswissenschaft am Institut für Publizistik der Johannes Gutenberg-Universität Mainz

Dr. Melanie Magin ist Akademische Rätin am Institut für Publizistik der Johannes Gutenberg-Universität Mainz

Quellenverzeichnis

- Adomavicius, G., & Tuzhilin, A. (2005). Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions. *IEEE Transactions on Knowledge and Data Engineering*, 17(6), 734-749. doi: 10.1109/TKDE.2005.99
- Bucher, T. (2012). Want to be on the top? Algorithmic power and the threat of invisibility on Facebook. *New Media & Society*, 14(7), 1164-1180. doi: 10.1177/1461444812440159
- Brunton, F. (2013). *Spam*. MIT Press.
- Das, A. S., Datar, M., Garg, A., & Rajaram, S. (2007). Google News Personalization: Scalable Online Collaborative Filtering. *Proceedings of the 16th International Conference on World Wide Web* (S. 271-280), New York, NY, USA: ACM. doi: 10.1145/1242572.1242610
- Diaz, A. (2008). Through the Google Goggles: Sociopolitical Bias in Search Engine Design. In A. Spink & E. Zimmer (Hrsg.), *Web Search: Multidisciplinary Perspectives* (S. 11-34). Berlin: Springer. doi: 10.1007/978-3-540-75829-7_2
- Fan, H., & Poole, M. S. (2006). What Is Personalization? Perspectives on the Design and Implementation of Personalization in Information Systems. *Journal of Organizational Computing and Electronic Commerce*, 16(3-4), 179-202. doi: 10.1080/10919392.2006.9681199

- Feitelson, D. G. (2006). Experimental computer science: The need for a cultural change. Abgerufen von <http://www.cs.huji.ac.il/~feit/papers/exp05.pdf>
- Friedman, B., & Nissenbaum, H. (1996). Bias in computer systems. *Transactions on Information Systems*, 14(3), 330-347. doi: 10.1145/230538.230561
- Friedman, B., & Kahn, P. H., Jr. (2008). Human values, ethics, and design. In A. Sears & J. A. Jacko (Hrsg.), *The Human-Computer Interaction Handbook*. Hillsdale, NJ, USA: Lawrence Erlbaum Associates.
- Garrett, R. K., Carnahan, D., & Lynch, E. K. (2013). A Turn Toward Avoidance? Selective Exposure to Online Political Information, 2004-2008. *Political Behavior*, 35(1), 113-134. doi: 10.1007/s11109-011-9185-6
- Google (2014). Custom Search JSON/Atom API. Abgerufen von <https://developers.google.com/custom-search/json-api/v1/overview>
- Granka, L. A. (2010). The Politics of Search: A Decade Retrospective. *The Information Society*, 26(5), 364-374. doi: 10.1080/01972243.2010.511560
- Hindman, M. (2009). *The myth of digital democracy*. Princeton, NJ: Princeton UP.
- Jannach, D., Zanker, M., Felfernig, A., & Friedrich, G. (2010). *Recommender Systems*. Cambridge University Press.
- Jakobsson, M. (2012). *The Death of the Internet*. John Wiley & Sons.
- Jürgens, P., Stark, B., & Magin, M. (2014). Gefangen in der Filter Bubble? Search Engine Bias und Personalisierungsprozesse bei Suchmaschinen. In B. Stark, D. Dörr, & S. Aufenanger (Hrsg.), *Die Googleisierung der Informationssuche. Suchmaschinen zwischen Nutzung und Regulierung* (S. 89-135). Berlin, Boston: De Gruyter.
- King, G. (1995). Replication, Replication. *Political Science and Politics*, 28(3), 444-452. doi: 10.2307/420301
- King, G. (2011). Ensuring the Data-Rich Future of the Social Sciences. *Science*, 331(6018), 719-721. doi: 10.1126/science.1197872
- Kohavi, R., Longbotham, R., Sommerfield, D., & Henne, R. M. (2009). Controlled experiments on the web: survey and practical guide. *Data Mining and Knowledge Discovery*, 18(1), 140-181. doi: 10.1007/s10618-008-0114-1
- Lazer, D., Pentland, A., Adamic, L., Aral, S., Barabási, A.-L., et al. (2009). Computational Social Science. *Science*, 323(5915), 721-723. doi: 10.1126/science.1167742
- Lyman, P. (2002). Archiving the World Wide Web. Building a National Strategy for Preservation: Issues in Digital Media Archiving. Council on Library and Information Resources Washington, D.C. and Library of Congress, Page 38-51. Abgerufen von <http://www.clir.org/PUBS/reports/pub106/pub106.pdf>

- MacCaw, A. (2011). *JavaScript Web Applications*. O'Reilly Media, Inc.
- Masanès, J. (2005). Web Archiving Methods and Approaches: A Comparative Study. *Library Trends*, 54(1), 72-90. doi: 10.1353/lib.2006.0005
- Masanès, J. (2006). *Web archiving*. Berlin: Springer.
- Mason, R. L., Gunst, R. F., & Hess, J. L. (2003). *Statistical Design and Analysis of Experiments*. John Wiley & Sons.
- McGeoch, C. C. (2002). Experimental Analysis of Algorithms. In P. M. Pardalos & H. E. Romeijn (Hrsg.), *Handbook of Global Optimization, Volume 2* (S. 489-513). Boston, MA: Springer US. doi: 10.1007/978-1-4757-5362-2_14
- Metaxas, P. T. (2010). Web Spam, Social Propaganda and the Evolution of Search Engine Rankings. In J. Cordeiro & J. Filipe (Hrsg.), *Web Information Systems and Technologies. 5th International Conference, WEBIST 2009, Lisbon, Portugal, March 23-26, 2009, Revised Selected Papers* (S. 170-182). Berlin, Heidelberg: Springer. doi: 10.1007/978-3-642-12436-5_13
- Mitchell, M. (2009). *Complexity: A Guided Tour*. Oxford University Press.
- Mooney, C. (1997). *Monte Carlo Simulation*. Sage Publications.
- Pan, B., Hembrooke, H., Joachims, T., Lorigo, L., Gay, G., & Granka, L. (2007). In Google We Trust: Users' Decisions on Rank, Position, and Relevance. *Journal of Computer-Mediated Communication*, 12(3), 801-823. doi: 10.1111/j.1083-6101.2007.00351.x
- Pariser, E. (2011). *The Filter Bubble: What The Internet Is Hiding From You*. London: Penguin UK.
- Pass, G., Chowdhury, A., & Torgeson, C. (2006). A Picture of Search. *Proceedings of the 1st International Conference on Scalable Information Systems*, New York, NY, USA: ACM. doi: 10.1145/1146847.1146848
- Paxson, V. (2004). Strategies for sound internet measurement. *Proceedings of the 4th ACM SIGCOMM Conference*, 263-271. doi: 10.1145/1028788.1028824
- Pitkow, J., Schütze, H., Cass, T., Cooley, R., Turnbull, D., Edmonds, A., et al. (2002). Personalized search. *Communications of the ACM*, 45(9), 50-55. doi: 10.1145/567498.567526
- Reid, A. (2014, 20. Mai). Can personalisation save the news homepage? [journalism.co.uk](http://www.journalism.co.uk). Abgerufen von <http://www.journalism.co.uk/news/can-personalisation-save-the-news-homepage-/s2/a556786/>
- Rieder, B. (2004). Algorithmische Mediatoren: Software-Agenten als Gegenstand der Medienwissenschaften. *Medien Journal*, 28(1), 36-46.

- Ricci, F., Rokach, L., Shapira, B., & Kantor, P. B. (2010). *Recommender Systems Handbook*. New York: Springer.
- Rogers, R. (2013). *Digital Methods*. Cambridge, MA: MIT Press.
- Springer, R. (2014, 26. Mai). Personalizing Personalization: It's About More Than Just Knowing Your User's Name. *EContent*. Abgerufen von <http://www.econtentmag.com/Articles/News/News-Feature/Personalizing-Personalization-Its-About-More-Than-Just-Knowing-Your-Users-Name-96731.htm>
- Sundar, S. S., & Marathe, S. S. (2010). Personalization versus Customization: The Importance of Agency, Privacy, and Power Usage. *Human Communication Research*, 36(3), 298-322. doi: 10.1111/j.1468-2958.2010.01377.x
- Sunstein, C. R. (2009). *Republic.com 2.0*. Princeton, PA: Princeton University Press.
- Taneja, H., & Wu, A. X. (2013, 15. Mai). Does the Great Firewall really isolate the Chinese? Integrating access blockage with cultural factors to explain web user behavior. *arXiv.org*. Abgerufen von <http://arxiv.org/abs/1305.3311>
- Teevan, J., Dumais, S. T., & Horvitz, E. (2010). Potential for personalization. *ACM Transactions on Computer-Human Interaction*, 17(1), 1-31. doi: 10.1145/1721831.1721835
- Tichy, W. F., Lukowicz, P., Prechelt, L., & Heinz, E. A. (1995). Experimental evaluation in computer science: A quantitative study. *Journal of Systems and Software*, 28(1), 9-18. doi: 10.1016/0164-1212(94)00111-Y
- Zelkowitz, M. (2011). *Security on the Web*. London u.a.: Academic Press.
- Zimmer, M. (2010b). "But the data is already public": On the ethics of research in Facebook. *Ethics and Information Technology*, 12(4), 313-325.

IV TEXT- UND
INHALTSANALYSEN

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

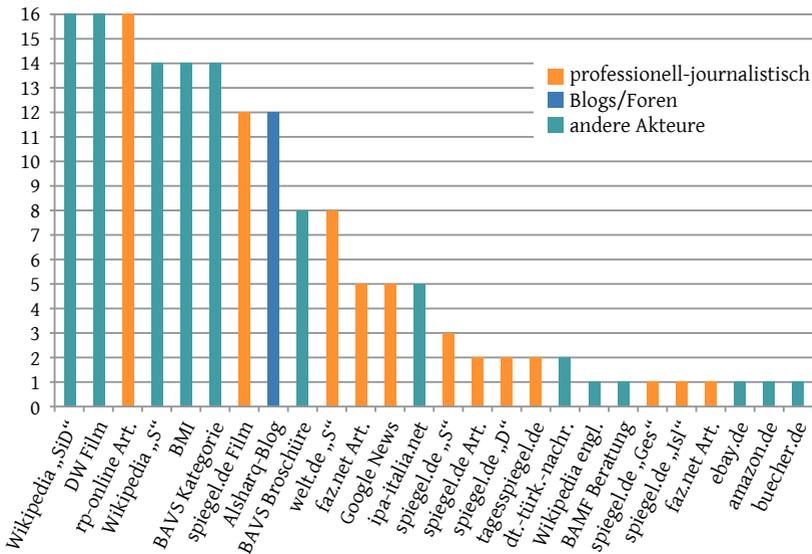
Informationen in erster Linie aus dem Internet beziehen, vermutlich deutlich weniger als früher von klassischen Medienangeboten definiert wird. Öffentlichkeit ist demzufolge heute weniger massenmedial-journalistisch geprägt, sondern besteht aus einem vielfältigeren Mix aus Quellen verschiedenster Herkunft 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“

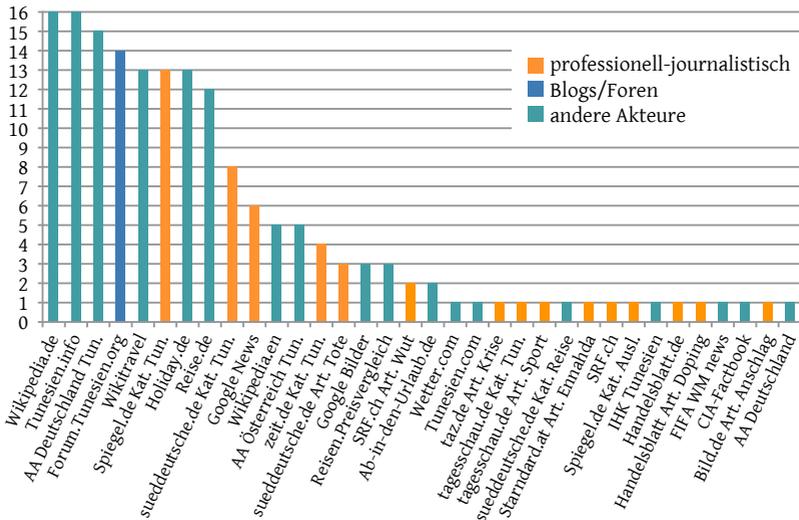


Bei den Angaben in Anführungszeichen handelt es sich nicht um Links auf konkrete Texte sondern auf thematische Kategorien innerhalb des jeweiligen Angebots. So haben zahlreiche Medien eigene Archive, die mit „Salafismus“ verschlagwortete Artikel versammeln. Die vollständigen Links können dem digitalen 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. Wunsch (Hrsg.), *Die Online-Inhaltsanalyse. Forschungsobjekt Internet* (S. 124-143). Köln: Herbert von Halem.
- Meier, S., Wunsch, C., Pentzold, C., & Welker, M. (2010). Auswahlverfahren für Online-Inhalte. In M. Welker & C. Wunsch (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.

Empfohlene Zitierung: Jungnickel, K., & Maireder, A. (2015). Der Multi-Step-Flow in Social Media: Ein Mehrmethodenansatz zur Analyse des Facebook-Newsfeed. In A. Maireder, J. Ausserhofer, C. Schumann, & M. Taddicken (Hrsg.), *Digitale Methoden in der Kommunikationswissenschaft* (S. 303-327). doi: 10.17174/dcr.v2.13

Zusammenfassung: Immer mehr Menschen erhalten Nachrichten und Informationen zu aktuellen Ereignissen über ihre Kontakte in sozialen Online-Netzwerken wie Facebook. Die Frage nach der Rolle von Facebook bei der Diffusion von Medieninhalten, und damit nach der Dynamik öffentlicher Kommunikationsprozesse, gewinnt so an Bedeutung. Vor dem Hintergrund des ‚Multi-Step-Flow of Communication‘ und Theorien sozialer Netzwerke ist besonders spannend, welche Inhalte Nutzer von welchen Kontakten (enge Freunde, entfernte Bekannte, Organisationen oder den Medien) auf Facebook erhalten. Um diese Frage zu beantworten, sind tradierte Methoden wie Inhaltsanalyse und Befragung für sich alleine jedoch unzureichend. Wir haben daher in einem Mehrmethoden-Design Befragung und Inhaltsanalyse kombiniert. 557 Facebook-Nutzer aus Deutschland und Österreich kopierten dafür bis zu fünf externe, auf Seiten außerhalb von Facebook führende Links in einen Online-Fragebogen, die sie zuletzt über ihren Facebook-Newsfeed erhalten hatten. Dabei zeigte sich beispielsweise, dass die Beziehungsqualität zwischen den Nutzern einen direkten Einfluss auf das Interesse an dem jeweiligen Inhalt hat, und zwar unabhängig von Thema und Urheber des verlinkten Inhalts. Bei der Untersuchung offenbarten sich jedoch auch einige methodische Probleme, für die der Beitrag sensibilisiert und Lösungsvorschläge anbietet.

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

Katrin Jungnickel & Axel Maireder

Der Multi-Step-Flow in Social Media

Ein Mehrmethodenansatz zur Analyse
des Facebook-Newsfeed

1 Einleitung

Mit der zunehmenden Relevanz von Social Media für die Verbreitung von Medieninhalten gerät auch die Nutzung dieser Dienste zur Beobachtung von und Beteiligung an öffentlicher Kommunikation verstärkt in den Fokus. Vor allem Facebook hat sich in den letzten Jahren auch im deutschsprachigen Raum zu einer viel genutzten Plattform für die Kommunikation von Nachrichten entwickelt, auf die Menschen (unter anderem) zugreifen, um sich über das aktuelle Tagesgeschehen zu informieren (BITKOM, 2011). Die zentrale Darstellungsform von Informationen auf Facebook ist der *Newsfeed*, eine rückwärts-chronologisch sortierte und laufend aktualisierte Liste jener Mitteilungen, die von den jeweiligen Kontakten eines Nutzers entweder selbst veröffentlicht oder aber ‚geteilt‘ – also zur Weiterverbreitung ausgewählt – wurden. Neben Privatpersonen, Organisationen und Institutionen nutzen auch Medienunternehmen das soziale Netzwerk zunehmend als Kanal zur Verbreitung öffentlich relevanter Inhalte. Dabei vermitteln sie diese einerseits an ihre eigenen Kontakte und können andererseits durch das ‚Teilen‘ oder ‚Liken‘ ihrer Mitteilungen durch die Nutzer auf eine weitere Diffusion hoffen. 2012 haben die 23 Millionen deutschen Facebooknutzer (Busemann, 2013) insge-

samt 21,3 Millionen mal Artikel von den Top 15 der deutschen Online-Medien weitergeleitet (Schiller, Heimbach, Strufe, & Hinz, 2013), die in den individuellen Newsfeeds der Nutzer mit Urlaubsfotos, Musikvideos und Blogposts um die Aufmerksamkeit der Nutzer konkurrieren.

Bislang fand die wissenschaftliche Untersuchung von Facebook hauptsächlich innerhalb von sechs Forschungsrichtungen statt (Weissensteiner & Leiner, 2011): Untersucht wurden (1) Zusammenhänge zwischen Personeneigenschaften und Nutzung, (2) Nutzungsmotive, (3) Selbstdarstellung, (4) die Wirkung von Benutzerprofilen auf die Betrachter, (5) der Konflikt zwischen Präsenz, Privatsphäre und Datenschutz, (6) sowie die Möglichkeiten zur Beschaffung von Sozialkapital. Daneben spielte auch die Bedeutung von Facebook für die strategische Kommunikation (Gesundheitskommunikation, Unternehmenskommunikation, politische Kommunikation) eine Rolle. Methodisch griff man bei der Untersuchung insbesondere auf Befragungen der Nutzer, teils auch auf Inhaltsanalysen einzelner Facebook-Profile zurück.

Welche konkreten Online-Inhalte sich über Facebook verbreiten und was Nutzer an öffentlich relevanten Informationen auf ihrem Newsfeed erhalten, blieb dabei jedoch bislang weitgehend unerforscht. Für die Beantwortung dieser Fragen sind sowohl die Inhaltsanalyse als auch die Befragung als Methoden *für sich alleine* jedoch unzureichend:

- Mit *Inhaltsanalysen* kann man zwar (öffentliche) Profile von Facebook-Nutzern analysieren, nicht aber ihren jeweiligen Newsfeed, auf dem sie von anderen geteilte Inhalte erhalten. Dazu würde man die Erlaubnis der Nutzer benötigen, sich unter ihren Namen bei Facebook online oder über eine App einzuloggen. Dies ist sowohl aus forschungspraktischer und forschungsethischer Perspektive als auch aus Datenschutzgründen problematisch. Zudem kann die Bestimmung des Beziehungsverhältnisses zwischen Nutzern (z.B. *weak ties*, *strong ties*) nur durch die Befragung der Nutzer selbst vorgenommen werden.

- Bei *Befragungen* stellen sich Probleme bei der Bestimmung der Inhalte des Facebook-Newsfeed. Offene Fragen würden dabei wohl wenig spezifische Antworten liefern. Bei geschlossenen Fragen nach konkreten Inhaltsformen (z.B. politische Inhalte, Nachrichten etc.) müssten die Einschätzungen zur Klassifikation den Befragten überlassen werden – eine überaus problematische Vorgehensweise.

Um diese Schwierigkeiten zu umgehen, haben wir in einem Mehrmethoden-Design Befragung und Inhaltsanalyse kombiniert. Wir baten die Nutzer, einen Online-Fragebogen auszufüllen und sich parallel dazu bei Facebook einzuloggen. Neben einigen soziodemografischen Angaben (Alter, Geschlecht, Bildung, Beruf, Facebook-Nutzung und Anzahl der Freunde) bestand die Hauptaufgabe darin, die letzten fünf Links, welche die Befragten in ihrem Newsfeed erhalten hatten, in den Fragebogen zu kopieren. Zu jedem Link mussten sie zudem angeben, wie stark sie dieser interessiert, wer ihn gesendet hat und ob derjenige noch Zusatzinformationen zum Link dazu geschrieben hat.

Dieses Vorgehen brachte methodisch einige Vorteile mit sich: Zum einen hatten wir die Möglichkeit, die verlinkten Inhalte selbst mit Hilfe eines standardisierten Codebuchs inhaltsanalytisch im Hinblick auf Produzent, Thema und Datentyp (Bild, Video, Text) zu klassifizieren – was aufgrund der Heterogenität der Angebote eine ganz eigene Herausforderung darstellte. Zum anderen konnten wir die Nutzer selbst nach der Form ihrer Beziehung zu denjenigen befragen, die die Medieninhalte geteilt hatten, und diese Angaben mit der Inhaltsanalyse verknüpfen.

Gleichzeitig stießen wir mit unserer Methode jedoch auch auf vielfältige Herausforderungen. Nach einer kurzen Vorstellung der zentralen Forschungsfragen, der Methode und der Ergebnisse dieser Studie, liegt der Fokus dieses Beitrags auf den methodischen Vor- und Nachteilen, denen wir im Zuge unserer Arbeit begegnet sind. Der Text versteht sich daher als Reflexion unseres Forschungsprozesses. Er soll für potenzielle Stolpersteine bei Studien zu öffentlich relevanten Kommunikationsprozessen in sozialen Online-Netzwerken sensibilisieren und für zukünftige Forschungsprojekte Anregungen zur Bewältigung dieser Probleme geben.

2 Studie „Links auf Facebook“

Seit der People's Choice-Studie von Lazarsfeld, Berelson und Gaudet (1944) zur Informationsdiffusion bei den amerikanischen Präsidentschaftswahlen 1940 haben sich Kommunikationswissenschaftler vielfach mit der Verbreitung politischer Informationen und Meinungen von den Massenmedien zu Bürgern beschäftigt und dabei gezeigt, dass die Diffusion meist als ‚Multi-step flow‘ über die

Vermittlung vielfältiger Intermediäre erfolgt (Weimann, 1994). Massenmediale Inhalte werden insbesondere von so genannten Meinungsführern an diejenigen in der Gesellschaft verbreitet, die weniger Medien nutzen und somit dort seltener mit diesen Inhalten in Kontakt kommen.

Während diese Diffusionsprozesse bis ins späte 20. Jahrhundert fast ausschließlich an interpersonelle Kommunikation (persönliches Gespräch, Telefongespräch) gebunden waren, stehen heute vielfältige internetbasierte Kommunikationsmittel zur Verfügung. Insbesondere soziale Online-Netzwerke geben ihren Nutzern außerordentlich niederschwellige Möglichkeiten, um Mitteilungen inklusive Links zu Web-Dokumenten als ‚Statusmeldung‘ zu publizieren und Beiträge anderer mit der Funktion ‚Teilen‘ weiter zu verbreiten. Die Gruppen, die dabei erreicht werden können, sind um ein Vielfaches größer als in persönlichen Gesprächen: Einerseits können Facebook-Nutzer durchschnittlich 190 Kontakte (Ugander, Karrer, Backstrom, & Marlow, 2011) direkt erreichen, zum anderen werden Mitteilungen vielfach erneut mit anderen Kontakten geteilt und erscheinen dann auch auf deren Newsfeeds.

Dabei ist zu beachten, dass Facebook-Newsfeeds nicht einfach die Mitteilungen aller Kontakte wiedergeben, sondern zusätzlich gefiltert werden. Dabei ermittelt Facebook's ‚Edge Rank‘-Algorithmus die Relevanz einzelner Beiträge im Wesentlichen an Hand der Interaktionsdichte zwischen den jeweiligen Nutzern, der Interaktion anderer Nutzer mit dem Beitrag (z.B. Anzahl der Kommentare und ‚Gefällt mir‘-Angaben), sowie dem Alter des Beitrags (Kincaid, 2010). Facebook errechnet so bei jedem Aufruf des Newsfeed erneut, welche Elemente auf Grund dieser Relevanzkriterien angezeigt und welche unterdrückt werden.

Bucher (2012) diskutiert diese Mechanismen als Disziplinierung der NutzerInnen, die die ständig drohende Unsichtbarkeit der eigenen Kommunikationselemente zu befürchten haben, der sie mit der Generierung von Interaktion entgegenwirken können. Aus einer Rezeptionsperspektive liegt für die Nutzer stets ein große Unsicherheit darüber vor, welche Beiträge ihrer Kontakte sie nun bekommen und welche sie gegebenenfalls ‚verpassen‘. In der Makroperspektive auf den ‚Multi-Step-Flow‘ auf Facebook bedeutet dies, dass Informationsdiffusion nicht nur von den konkreten kommunikativen Handlungen der Beteiligten abhängt, sondern auch von der Verarbeitung der digitalen Manifestation dieser Handlungen durch Programmoperationen.

2.1 Forschungsfragen

Auf Facebook sind nicht nur Individuen aktiv, sondern auch Unternehmen, Organisationen und Institutionen. Nutzer können die Mitteilungen dieser kollektiven Akteure abonnieren und geben damit sowohl politischen Parteien, Initiativen und Interessensverbänden als auch Medienunternehmen die Möglichkeit, sie direkt zu erreichen. So integriert Facebook potenziell alle Stufen des Mehrstufenflusses in ein und dasselbe sozio-technische System.

Dass Facebook die Möglichkeiten liefert, Inhalte im Sinne eines Mehrstufenflusses zu verbreiten, bedeutet jedoch nicht unbedingt, dass Nutzer auch von ihnen Gebrauch machen und sie tatsächlich nutzen, um Meinungen und Informationen zu öffentlich relevanten Themen zu verbreiten. Studien in anderen Online-Netzwerken, insbesondere Twitter, deuten jedoch darauf hin. So stellten Wu, Hofman, Mason und Watts (2011) fest, dass es zwei Gruppen von Nutzern gibt: *Intermediäre*, die ihre Informationen hauptsächlich direkt von den Massenmedien erhalten, also vielen Medien und Journalisten auf Twitter folgen und *andere Nutzer*, die ihre Informationen zum Großteil von den Intermediären erhalten. Diese Intermediäre sind generell aktiver, haben mehr Follower als der Rest der Twitter-Community und fungieren damit als Meinungsführer. Ebenso konnten An, Cha, Gummadi und Crowcroft (2011) zeigen, dass Massenmedien ihr Publikum auf Twitter durch Retweets durchschnittlich um 28 Prozent vergrößern können und Twitter-Nutzer durch Retweets mit deutlich mehr massenmedialen Inhalten in Kontakt kommen, als nur mit denen, denen sie direkt folgen.

Aus diesen Vorüberlegungen ergab sich für uns zunächst die folgende Forschungsfrage:

F1: Gibt es den Mehrstufenfluss der Kommunikation zu Themen von öffentlicher Relevanz auch bei Facebook und wenn ja, wie sieht dieser aus?

Wenn Links zu gesellschaftlich relevanten Themen über den Mehrstufenfluss in den Facebook-Newsfeed eines Nutzers gelangen, stellt sich im zweiten Schritt die Frage, ob sich der Nutzer, der sie erhält, auch mit den Inhalten beschäftigt. Anders als im persönlichen Gespräch, bei dem es in der Regel als unhöflich gilt, wenn man seinem Gesprächspartner nicht zuhört, ist es bei Facebook nicht ungewöhnlich, dass Nutzer nicht allen Beiträgen ihres Newsfeeds Aufmerksamkeit

schenken. Wahrscheinlich nehmen sie nur einen Teil der Beiträge tatsächlich wahr und verfolgen noch weniger auch wirklich weiter (also klicken sie an, um den Inhalt anzusehen, zu dem der Link führt). Das Interesse an einem Link können potentiell verschiedene Faktoren beeinflussen: das Thema des Inhalts, die Sprache (deutsch, englisch), der Produzent (z.B. eine bestimmte Medienmarke oder eine Organisation), der Datentyp (Text, Video, Bild, etc.) oder derjenige, der den Link verbreitet hat und im Folgenden als „Sender“ der Botschaft bezeichnet wird. Daraus ergibt sich die zweite Forschungsfrage:

F2: Welche Eigenschaften eines auf Facebook verbreiteten Links (Inhalt, Produzent des Inhalts, Datentyp, Sprache) und der Beziehung zum Sender bestimmen, wie interessant er für Nutzer ist?

Der Sender der Botschaft ist dabei nicht mit dem Produzenten gleichzusetzen, denn er hat in der Regel den von ihm verlinkten Inhalt nicht selbst erstellt, sondern nur weiterverbreitet. Dennoch hat ein Sender die Möglichkeit, bei den Empfängern ein Interesse am Link zu wecken, weil jene aus Erfahrung dessen Meinungen und Tipps schätzen. Wie stark dieser Faktor wirkt, hängt dabei vermutlich auch von der Beziehungsqualität zwischen Sender und Empfänger ab. Granovetter (1973, S. 1361) definiert die Beziehungsqualität zwischen zwei Personen als eine Kombination aus der miteinander verbrachten Zeit, der emotionalen Intensität, der Intimität und der gegenseitigen Hilfe. Enge Beziehungen werden demnach als *strong ties* bezeichnet, entfernte Bekannte hingegen als *weak ties*.

Ein Blick auf die bisherige Forschung zeigt, dass interpersonaler Einfluss eher in engeren Beziehungskonstellationen stattfindet. So werden politische Themen beispielsweise häufiger mit engen Freunden oder in der Familie besprochen (Straits, 1991; Morey, Eveland, & Hutchens, 2012). Schwache Bindungen haben demgegenüber den Vorteil, dass sie als Brückenverbindungen zwischen Mitgliedern verschiedener Freundeskreise fungieren und so dafür sorgen, dass Informationen und neue Ideen auch außerhalb ihres eigenen sozialen Umfelds verbreitet werden (Weimann, 1983). In sozialen Online-Netzwerken wie Facebook haben Nutzer zahlreiche Kontakte, wovon einige zum engeren Freundeskreis zählen, viele aber auch nur Teil des erweiterten Bekanntenkreises sind (Haythornthwaite, 2002; de Zuniga & Valenzuela, 2011). Es besteht also die Hoffnung, dass die

persönlichen Netzwerke der Nutzer bei Facebook heterogener und diverser sind und sie so durch Kontakte, zu denen eher schwache Bindungen bestehen, mit vielen neuen Informationen und Ideen in Berührung kommen – anstatt in sogenannten Echo Chambers nur in ihren bereits bestehenden Ansichten bestärkt zu werden (Wojcieszak & Mutz, 2009). Die Voraussetzung dafür ist jedoch, dass sich die Nutzer auch für die Botschaften interessieren, die ihre entfernten Bekannten auf Facebook verbreiten.

2.2 Methode

Um die Forschungsfragen zu beantworten, entschieden wir uns für ein Mehrmethodendesign, dessen Stärken und Schwächen in Kapitel 3 genauer erläutert werden. Dazu baten wir im Februar 2012 Facebook-Nutzer aus Deutschland und Österreich, bis zu fünf externe, auf Seiten außerhalb von Facebook führende Links, die sie zuletzt über ihren Facebook-Newsfeed erhalten haben, in einen Online-Fragebogen zu kopieren. Die Teilnehmer wurden im Schneeballverfahren über Posts auf Facebook und Twitter ausgehend von studentischen Accounts an der Universität Wien sowie über eine universitätsinterne Mailingliste der TU Ilmenau rekrutiert. Zusätzlich zur Linkadresse gaben sie jeweils an, welche Beziehung sie zu der Person haben, die den Link geteilt hat, und wie sehr sie an dem Inhalt interessiert sind. Von 745 Befragungsteilnehmern kopierten 557 Nutzer insgesamt 2.635 Links in den Fragebogen, von denen wir 2.070¹ mitsamt ihrer Zielseiten anschließend inhaltsanalytisch untersuchten.²

Unser Sample war dabei nicht repräsentativ und bestand vorwiegend aus jüngeren und höher gebildeten Personen (77% sind Studierende, 58% unter 25 Jahren). Daher können wir auch keine generalisierenden Aussagen zu den auf Facebook verbreiteten Inhalten treffen, sondern lediglich grundlegende Tenden-

- 1 Die restlichen Links erfüllten entweder die Bedingung nicht, auf eine Facebook-externe Webseite zu verweisen, waren in einer anderen Sprache als Deutsch oder Englisch verfasst, sodass der Inhalt nicht codiert werden konnte, oder führten auf eine mittlerweile nicht mehr existierende Webseite (siehe Abschnitt 3.5).
- 2 Die Inhalte wurden von zwei Codiererinnen erfasst. Die Intercoderreliabilität (Holsti-Koeffizient) lag in allen untersuchten Variablen bei $>.70$.

zen ablesen bzw. existierende Kommunikationsflüsse aufdecken. Die verlinkten Inhalte beschäftigten sich zu 14 Prozent mit einem politischen Thema und zu 22 Prozent mit Produkt- oder Unternehmensinformationen. 37 Prozent der Inhalte stuften wir als öffentlich relevant³ ein. Diese drei Themenkategorien schließen sich jedoch gegenseitig nicht aus, sondern traten häufig in Kombination miteinander auf. Die Links mit einem politischen Thema wurden beispielsweise zu 100 Prozent auch als öffentlich relevant eingestuft; Links mit Produkt- oder Unternehmensinformationen hingegen nur zu 20 Prozent. Da wir uns insbesondere für den Informationsfluss von Themen mit öffentlicher Relevanz interessierten, werden zur Beantwortung der Forschungsfragen im Folgenden nur die 783 Links berücksichtigt (kopiert von n=392 Befragungsteilnehmern), die zu einer Webseite mit einem solchen Thema führten.

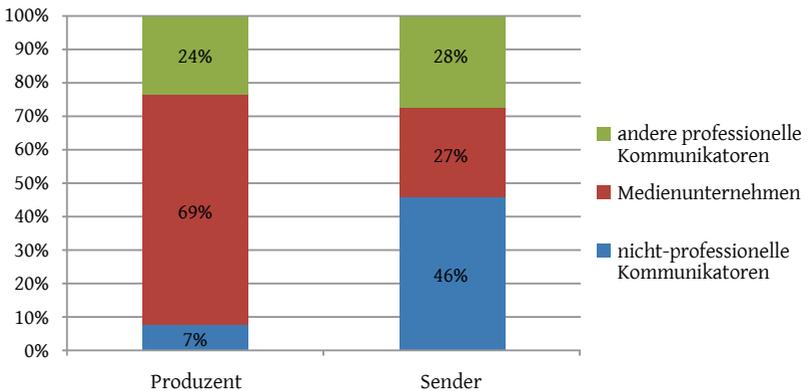
2.3 Ergebnisse

Um Forschungsfrage 1 zu beantworten, haben wir nach Indizien für einen Mehrstufenfluss auf Facebook gesucht und dazu die Angaben der Befragten zum Sender des Links mit dem durch die Inhaltsanalyse ermittelten Produzenten des Linkinhalts verglichen. Dazu teilten wir die Sender und Produzenten in drei Hauptkategorien ein: (1) nicht-professionelle Kommunikatoren, (2) Medienunternehmen (professionelle, in Kollektiven organisierte Akteure im Rahmen einer Organisation, deren Hauptzweck in der Verarbeitung und Veröffentlichung von Informationen zur Bereitstellung für die Öffentlichkeit besteht) und (3) andere professionelle Kommunikatoren (z.B. Unternehmen, Organisationen, Institutionen, Künstler etc.).

- 3 Dabei definierten wir einen Inhalt dann als öffentlich relevant, wenn er irgendeine Nachricht, einen Bericht oder einen Meinungsbeitrag enthielt, der öffentlich diskutiert wird und Themen wie Politik (Innen- und Außenpolitik, politischer Aktivismus), Wirtschaft (Arbeitsmarkt, Finanzen, Immobilien), Justiz (Polizeieinsätze, Kriminalität, Rechtsprechung), Soziales (Verwaltung, aktuelle soziale Probleme, Bildung), Umwelt und Klima (Klima, Naturkatastrophen, Naturereignisse, Umweltschutz), aber auch Kunst und Kultur (Rezension und Bewertung), Medien, Sport, Stars, Prominente, Mode, Lifestyle, Reisen, Ratgeber (allgemeine Tipps und Tricks, populärwissenschaftlich aufgearbeitete Forschung), Tiere oder Regionales aufgriff.

Abbildung 1 zeigt, dass die große Mehrheit der Links zu öffentlich relevanten Themen von professionellen Kommunikatoren (Medien und anderen) produziert wurde, ein hoher Anteil der Sender jedoch aus nicht-professionellen Kommunikatoren besteht, was bereits auf einen Mehrstufenfluss bei Facebook hindeutet.

Abbildung 1: Link-Sender und Link-Produzenten (n=738)



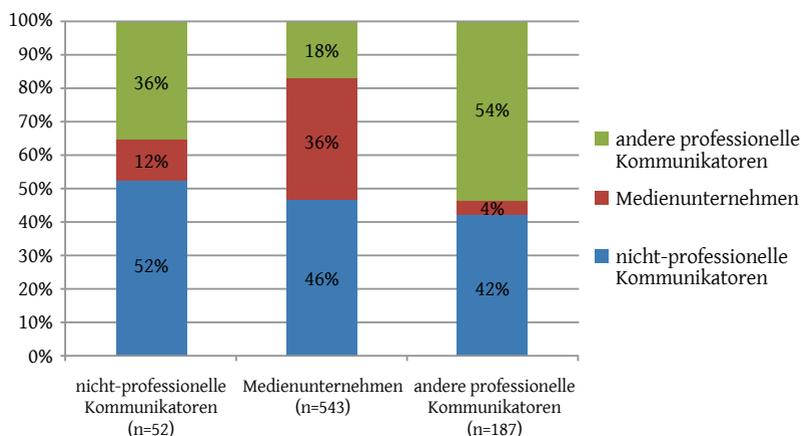
Nicht-professionelle Kommunikatoren als Sender lassen sich noch einmal nach der Stärke der Beziehung zum Empfänger unterteilen. Dabei bestand der überwiegende Teil der nicht-professionellen Sender aus entfernten Bekannten (55%), gefolgt von guten Bekannten (30%) und nahestehenden Freunden oder Verwandten (15%). Dieses Ergebnis deckt sich mit anderen Studien, die in sozialen Onlinenetzwerken ebenfalls einen großen Anteil an weak ties entdeckten (u.a. de Zuniga & Valenzuela, 2011).

Interessanterweise beschränken sich die Sender nicht darauf, Links einfach zu teilen. Sowohl Medien (85%) als auch andere professionelle Kommunikatoren (85%) und nicht-professionelle Kommunikatoren (67%) schreiben mehrheitlich noch einen kurzen Text zu dem Link dazu, bevor sie ihn an ihre Kontakte weiterleiten.

Dass der Mehrstufenfluss im Social Web nicht nur in eine Richtung verläuft, verdeutlicht Abbildung 2 (siehe nächste Seite). Der am häufigsten vorkommende

Informationsfluss ist tatsächlich der von einem professionellen Kommunikator über einen nicht-professionellen Kommunikator hin zum Empfänger. So wurden zum Beispiel 46 Prozent der von Medienunternehmen produzierten Inhalte von nicht-professionellen Kommunikatoren auf Facebook geteilt. Aber auch Inhalte von nicht-professionellen Kommunikatoren haben Medien (12%) und andere professionelle Kommunikatoren (36%) auf Facebook verbreitet. Der Kommunikationsfluss bei Facebook verläuft so wenig hierarchisch ab.

Abbildung 2: Link-Sender in Abhängigkeit vom Link-Produzenten



Anmerkung: χ^2 signifikant, $p < .001$

Damit ein Link die Nutzer aber nicht nur potenziell sondern auch faktisch erreicht, muss er ihr Interesse wecken. Auf der Suche nach Einflüssen auf das Interesse am Link haben wir eine mehrfaktorielle Varianzanalyse mit dem Interesse am Link als abhängige Variable gerechnet. Das Interesse wurde im Fragebogen anhand einer 5-stufigen Likert-Skala von 1='interessiert mich gar nicht' bis 5='interessiert mich sehr' abgefragt. Als unabhängige Variablen integrierten wir die Themenvariable ‚politischer Inhalt‘, den Produzenten des Inhalts (nicht-professioneller Kommunikator, Medienunternehmen, anderer professioneller Kommunikator),

den Sender des Links (nahestehender Freund oder Verwandter, guter Bekannter, entfernter Bekannter, Medienunternehmen, anderer professioneller Kommunikator), den Datentyp (Schrift, Bilder, Video/Audio/Anwendungen, Startseiten) sowie die Sprache (Deutsch, Englisch) und eine Information darüber, ob der Sender den Link noch kommentiert hat oder nicht. Tabelle 1 zeigt die Ergebnisse der Varianzanalyse, wobei hier nur die Haupteffekte und nicht die Interaktionseffekte zwischen den einzelnen Variablen berücksichtigt wurden.

Tabelle 1: Einflussfaktoren auf das Interesse am Link

	F-Wert	Signifikanz
Politischer Inhalt	30,262	,000
Sender	12,136	,000
Datentyp	1,971	,117
Produzent	,233	,793
Zusatztext vom Sender	,138	,710
Sprache	,117	,732

Anmerkung: MANOVA, korrigiertes $R^2 = .087$

Einen signifikanten Einfluss auf das Interesse am Link hatten demnach nur der Sender und das Thema. Das Interesse an politischen Inhalten (MW=3,50; SD=1,28) war generell höher als an nicht-politischen Inhalten (MW=3,07; SD=1,18). Dieser Befund könnte jedoch der demografischen Zusammensetzung unseres Samples geschuldet sein, da wir hauptsächlich höher gebildete Menschen befragt haben, bei denen das Interesse an Politik vermutlich höher ausgeprägt ist.

Interessanter sind die Ergebnisse zu den Produzenten und Sendern der Links. Der Produzent des Inhalts hatte keinen signifikanten Einfluss auf das Interesse am Link. Es war also weniger wichtig, ob der Inhalt ursprünglich von Medien, anderen professionellen Kommunikatoren oder als User Generated Content (UGC) erstellt wurde. Bedeutsamer war stattdessen der Sender des Links, also derjenige, der ihn bei Facebook geteilt hat. Tabelle 2 zeigt, dass das Interesse an Links, die

von nahestehenden Freunden oder Verwandten (*strong ties*) verbreitet wurden, am höchsten war. Demgegenüber hatten die Befragten an Links von guten Bekannten, Medienunternehmen oder anderen professionellen Kommunikatoren weniger Interesse. Am wenigsten interessierten sie sich für Links von entfernten Bekannten (*weak ties*).

In einem zweiten Schritt haben wir nur mit den signifikanten Variablen ‚Sender‘ und ‚politischer Inhalt‘ ein gesättigtes Modell varianzanalytisch berechnet. Dabei deckten wir zusätzlich einen signifikanten Interaktionseffekt ($p < .05$) zwischen den beiden unabhängigen Variablen auf: Während es bei Medien und anderen professionellen Kommunikatoren kaum einen Unterschied machte, ob der Inhalt politisch war oder nicht, sah das bei nicht-professionellen Kommunikatoren ganz anders aus: Sowohl bei Links von nahestehenden Freunden als auch bei denen von guten oder entfernten Bekannten war das Interesse an politischen Inhalten deutlich höher als an nicht-politischen Inhalten (siehe Tabelle 2).

Tabelle 2: Interesse am Link in Abhängigkeit vom Link-Sender

Sender	N	Interesse am Link (MW)		
		Gesamt (n=780)	Politischer Inhalt (n=291)	Nicht- politischer Inhalt (n=489)
Medienunternehmen ^b	209	3,38	3,52	3,28
Andere professionelle Kommunikatoren ^b	216	3,37	3,49	3,34
Nahestehende Freunde oder Verwandte ^a	51	3,78	4,05	3,61
Gute Bekannte ^b	107	3,31	3,83	2,90
Entfernte Bekannte ^c	197	2,73	3,16	2,42

Anmerkung: ANOVA, Post-Hoc-Test Tamhane, unterschiedliche Buchstaben markieren signifikante ($p < .05$) Gruppenunterschiede (für das Gesamtinteresse); Interesse auf einer Skala von 1= ‚interessiert mich gar nicht‘ bis 5= ‚interessiert mich sehr‘

Zusammenfassend zeigt sich, dass Menschen in sozialen Online-Netzwerken viele Kontakte haben, zu denen häufig eine nur schwache Bindung besteht. Allerdings scheinen sie wenig an den Informationen interessiert zu sein, die ihnen diese entfernten Bekannten senden. Stattdessen verlassen sie sich auf enge Bindungen oder gleich auf professionelle Kommunikatoren, da sie deren Seiten vorher bereits abonniert haben und somit schon ein Grundinteresse an den dort bereit gestellten Inhalten vorhanden ist.

Bei der Interpretation der Ergebnisse ist jedoch zu beachten, dass die erklärte Varianz des Modells relativ gering ist und somit noch andere Variablen das Interesse am Link bestimmen. Weiterhin darf ein generelles Problem nicht vergessen werden, auf das auch Aral (2011, S. 217-218) bereits hinwies: Das Interesse am Link muss nicht zwangsläufig dadurch gesteigert worden sein, dass ein nahestehender Freund ihn verbreitet hat. Stattdessen kann ein erhöhtes Interesse auch ein Resultat bereits bestehender ähnlicher Interessen zwischen engen Freunden sein. Woran ein höheres Interesse an Links von *strong ties* nun genau liegt, ist jedoch zweitrangig, wenn es um die Frage geht, ob *weak ties* neue Ideen in homogene Netzwerke tragen können. Die Antwort darauf muss differenziert ausfallen: Das Potential ist vorhanden, da die meisten Links zu öffentlich relevanten Themen tatsächlich von entfernten Bekannten kommen. Allerdings ist das Interesse an diesen Inhalten eher gering, sodass fraglich ist, inwieweit sie im Alltag von den Empfängern wirklich wahrgenommen und rezipiert werden. Weiterhin ist unklar, ob eine hohe Anzahl von *weak ties* innerhalb der Kontakte zwangsläufig dazu führt, dass das persönliche Netzwerk tatsächlich heterogener wird. Zumindest in Bezug auf das Alter konnten wir eine große Homogenität zwischen Sendern und Empfängern feststellen: Bei 90 Prozent der *strong ties* und auch bei 75 Prozent der *weak ties* betrug der Altersunterschied höchstens fünf Jahre. Allerdings hatte die Altersdifferenz insgesamt keinen Einfluss auf das Interesse an Links. Ob sich Sender und Empfänger auch in Bezug auf generelle Interessen oder politische Einstellungen gleichen, haben wir nicht erhoben, wäre jedoch eine interessante Fragestellung für weitere Untersuchungen.

3 Methodische Herausforderungen

Das von uns verwendete Mehrmethodendesign stellt einen neuen Ansatz dar, um Kommunikationsprozesse im Social Web zu untersuchen. Grundsätz-

lich sind die Anwendungsmöglichkeiten der vorgestellten Methodenkombination vielfältig. So könnte man die Befragungsteilnehmer auch bitten, anstelle von Links andere Inhalte für die spätere Analyse zu kopieren. Dabei sind jedoch einige Schwierigkeiten in der Umsetzung zu beachten. Die folgende Diskussion greift diese Hürden auf und stellt Lösungsansätze und Verbesserungsmöglichkeiten vor, um zukünftigen Studien den Einsatz des Designs zu erleichtern.

3.1 *Stichprobe und Abbruchquote*

Wie bei vielen anderen Facebook-Studien, konnten wir auch in unserer Untersuchung keine für Facebook-Nutzer repräsentative Stichprobe ziehen. Die Befragten wurden im Schneeballverfahren über Facebook und studentische E-Mail-Verteiler rekrutiert. Dadurch besteht unsere Stichprobe hauptsächlich aus jungen (Durchschnittsalter: 25), höher gebildeten (98% Hochschulreife) Personen mit relativ vielen Freunden auf Facebook (im Durchschnitt 237). Durch Rückgriff auf ein repräsentatives Online-Panel könnte man hier eine bessere Verallgemeinerbarkeit der Ergebnisse erreichen.

Jedoch lässt sich auch dann nicht verhindern, dass die Abbruchquote bei diesem Design vermutlich höher ausfallen wird als bei anderen Untersuchungen. In unserer Studie waren 25% der Teilnehmer nicht bereit, die Links aus ihrem Facebook-Newsfeed in den Fragebogen zu kopieren. Sie brachen, obwohl sie die Befragung bereits gestartet und Fragen zur Soziodemografie beantwortet hatten, bei der Aufforderung zum Linkkopieren die Befragung ab. Unter den „Abbrechern“ waren mehr Männer und weniger Menschen mit einem Medienbezug in Studium und/oder Beruf. Außerdem ist diese Gruppe etwas weniger aktiv auf Facebook und hat dort auch weniger Freunde (Tabelle 3). Es zeigt sich also, dass ein etwas komplexeres Design wie dieses verstärkt die weniger medienaffinen Personen verschreckt, was zu einer weiteren Verzerrung der Stichprobe führen kann.

Im Pretest mit 45 Teilnehmern haben wir drei Hauptgründe für den Abbruch der Befragung identifizieren können:

- (1) *Großer Aufwand:* Einige Pretest-Teilnehmer gaben an, dass der Wechsel zwischen dem Fragebogen und Facebook in Verbindung mit dem Kopieren der Links und dem Klassifizieren der Sender nervig und aufwändig sei.

- (2) *Langeweile*: Der Umstand, dass bei fünf Links immer die gleichen Fragen beantwortet werden mussten, hat einige Pretest-Teilnehmer gelangweilt.
- (3) *Datenschutzbedenken*: Obwohl Links eigentlich wenig Persönliches über die Befragten verraten, war es einigen Teilnehmern dennoch unheimlich, Informationen aus ihrem Newsfeed in den Fragebogen zu kopieren. Manche Befragte hatten auch Bedenken, dass Facebook die Daten des Fragebogens speichern kann, da man während der Befragung gleichzeitig bei Facebook eingeloggt sein musste. Interessant war, dass einige Teilnehmer nur deshalb weniger Datenschutzbedenken hatten, weil sie von einer befreundeten Person gebeten wurden, den Fragebogen auszufüllen. Die Empfehlung von Freunden stellte somit einen Vorteil unseres Schneeballsystems dar, der bei einem anonymen Online-Panel nicht gegeben wäre.

Die Punkte 1 und 2 konnten wir relativieren, indem wir den Fragebogen im Vergleich zum Pretest für die Hauptuntersuchung noch einmal deutlich verkürzt haben. Ursprünglich sollten die Befragten noch Datum und Uhrzeit sowie die An-

Tabelle 3: Vergleich zwischen Teilnehmern und Abbrechern

	Teilnehmer (n= 557)	Abbrecher (n= 188)
Facebook wird täglich genutzt*	88%	80%
männlich**	48%	61%
Herkunft: Deutschland	82%	82%
Hochschulreife	98%	96%
Nicht berufstätig	63%	55%
Studenten	78%	73%
Medienbezug*	42%	32%
Unter 25 Jahre	59%	54%
Mehr als 200 Facebook-Freunde**	51%	35%

*Anmerkung: Chi²-Tests, Signifikanz: * p<.05; ** p<.01*

zahl der Gefällt-mir-Angaben und Teilungen des Links angeben. Um die Befragten nicht zu langweilen oder zu lang zu beschäftigen, haben wir uns schließlich auf die wichtigsten Angaben zu Interesse und Sender beschränkt. Punkt 3 ließ sich als Kernelement unseres Designs jedoch nicht ändern und ist vermutlich auch hauptverantwortlich für die Abbruchquote von 25 Prozent.

3.2 *Auswahl der zu kopierenden Inhalte*

Eine zweite Herausforderung stellte die Auswahl der zu kopierenden Inhalte dar. Die erste Möglichkeit besteht darin, die Befragten zu bitten, eine bestimmte Anzahl *aller* zuletzt erhaltenen Beiträge in den Fragebogen zu kopieren. Dies erschien jedoch aus zwei Gründen nicht praktikabel: Zum einen vermuteten wir, dass darunter einige sehr persönliche Beiträge von Freunden zu finden sind, bei denen die Befragten noch mehr Bedenken haben, diese weiter zu geben. Zum anderen sind diese persönlichen Beiträge für die Erforschung des Mehrstufenflusses von öffentlich relevanten Themen auch weniger relevant. Aus diesem Grund entschieden wir uns, die Teilnehmer nur *Links zu externen Seiten* kopieren zu lassen, da der Mehrstufenfluss einerseits im Kern das Weiterleiten von Beiträgen anderer Produzenten beinhaltet und andererseits bei Links die Wahrscheinlichkeit sehr hoch ist, auf Themen mit öffentlicher Relevanz zu stoßen.

Weiterhin galt es zu beachten, dass die Inhalte des Facebook-Newsfeeds dynamisch sind und sich stetig verändern. Nutzer können bei den Einstellungen wählen, ob sie die „neuesten Meldungen“ oder die „Hauptmeldungen“ angezeigt bekommen. Bei den Hauptmeldungen bestimmt der oben diskutierte ‚Edge Rank‘-Algorithmus, welche Meldungen für den Nutzer wahrscheinlich am interessantesten sind. Selbst wenn ein Nutzer diese Einstellungen nie verändert, sieht der Newsfeed doch jedes Mal, wenn er sich wieder bei Facebook einloggt, anders aus, weil neue Beiträge hinzugekommen sind.

Uns stellte sich somit die Frage, wie wir mit diesem sich ständig verändernden Newsfeed umgehen und ob wir den Befragten eine bestimmte Grundeinstellung sowie einen bestimmten Zeitrahmen für die zu kopierenden Links vorschreiben. Wir entschieden uns aus zwei Gründen dagegen: Einerseits bedeutet es für die Befragten zusätzlichen Aufwand, ihre Grundeinstellungen zu ändern bzw. statt einfach die letzten fünf Links zu kopieren, in ihrem Newsfeed scrollen zu müs-

sen um zum von uns vorgegebenen Zeitpunkt zu gelangen. Zweitens schränkt ein solches Vorgehen die externe Validität der Ergebnisse ein. Wir interessierten uns dafür, was der Einzelne in einer möglichst natürlichen Nutzungssituation mit seinen eigenen Grundeinstellungen zu sehen bekommt. Mit der Anweisung, die *letzten* fünf erhaltenen Links zu kopieren, erhielten wir eine große Variation an Links, die in einem Zeitraum von mehreren Tagen oder auch nur mehreren Stunden (je nachdem wie häufig jemand Links auf Facebook erhält) verbreitet wurden. Auch dies entspricht am ehesten einer natürlichen Nutzungssituation, in der sich Nutzer in der Regel das anschauen, was sie zuletzt erhalten haben – unabhängig davon, wann genau die Inhalte verbreitet wurden.

3.3 Medienkompetenz der Befragten

Befragungen haben generell den Nachteil der Reaktivität. Neben einer Tendenz zur sozialen Erwünschtheit spielt oft auch die Medienkompetenz der Befragten eine Rolle, die bestimmt, wie wahrheitsgetreu sie den Fragebogen ausfüllen (können). In unserer Befragung gab es in diesem Zusammenhang zwei besondere Hürden:

- (1) *Auswahl der Links:* Wir baten die Teilnehmer, die letzten fünf Links zu externen Websites zu kopieren. Obwohl wir Links zu Facebook-internen Applikationen oder Fotoalben ausschlossen, kam es doch hin und wieder vor, dass die Befragten diese Links ebenfalls kopierten. Dieses Problem ließ sich für uns jedoch im Nachhinein leicht beheben, indem wir solche Links von der Analyse ausschlossen.
- (2) *Klassifikation der Sender:* Die Nutzer mussten angeben, wer ihnen die jeweiligen Links gesendet hat. Handelte es sich dabei um einen nicht-professionellen Kommunikator, nahmen wir an, dass es für die Befragten recht unkompliziert war, die Beziehungsqualität (nahestehender Freund, guter Bekannter, entfernter Bekannter) einzuschätzen. Bei professionellen Kommunikatoren war die Klassifikation schon schwieriger. Aus diesem Grund bildeten wir später aus den Antworten nur noch zwei Hauptkategorien von professionellen Kommunikatoren: Medienunternehmen und andere

professionelle Kommunikatoren – in der Hoffnung, dass die Befragten den Begriff „Medien“ noch einigermaßen gut einschätzen können. Allerdings war auch hier zu befürchten, dass es den Befragten bei Grenzfällen wie beispielsweise journalistischen Blogs schwerfiel, den Sender des Links richtig zu klassifizieren. Ähnliche Probleme kamen später auch in der Inhaltsanalyse (siehe Kapitel 3.4) auf uns zu.

3.4 *Inhaltsanalytische Kategorien*

Auf Facebook und generell im Social Web ist ein Kommunikationsraum entstanden, in dem die Dichotomien zwischen persönlichem Gespräch und massenmedialer Kommunikation (Krotz, 2000) sowie zwischen Rezipienten und Produzenten (Guenther & Schmidt, 2007) verschwimmen. Es zeigen sich Kommunikationszusammenhänge mit höchst unterschiedlichen und fluiden Akteurskonstellationen, in denen sich klassische Rollenbilder ebenso auflösen, wie tradierte Inhalts- und Strukturkategorien für Medieninhalte. So stoßen relevante Teile des zentralen Begriffs- und Konzeptrepertoirs, das die Kommunikationswissenschaft in ihrer jahrzehntelangen Auseinandersetzung mit massenmedial vermittelter Kommunikation entwickelt hat, an ihre Grenzen (Couldry, 2009).

Bis in die jüngste Vergangenheit wurden mit kommunikationswissenschaftlichen Inhaltsanalysen primär massenmediale und insbesondere journalistische Texte analysiert, da das Interesse der Kommunikationswissenschaft vielfach auf der Vermittlungsrolle der Massenmedien für gesellschaftliche Kommunikationsprozesse lag. Teilweise an die journalistische Praxis selbst angelehnt, sind aus diesen Analysen vielfältige Kategoriensysteme hervorgegangen, die Texte thematisch (z.B. Politik, Chronik), stilistisch (z.B. Bericht, Kommentar) oder inhaltlich (z.B. ‚soft news‘, ‚hard news‘) klassifizieren. Weiterhin lösen journalistische Texte zumeist bestimmte Erwartungen an Struktur (z.B. Titel, Teaser), Stil (z.B. Qualitäts- vs. Boulevardjournalismus) und Inhalt (z.B. W-Fragen, Nachrichtenwerte) aus, da sie in einem professionellen, institutionalisierten Umfeld entstanden sind.

Auf Facebook werden unzählige Inhalte verlinkt, deren Entstehungskontexte der Grundlage für tradierte kommunikationswissenschaftliche Kategorien und Analysesysteme entbehren. In unserer Stichprobe fanden sich u.a. eine Seite

mit Team-Ranglisten von MMOGs⁴, die Übersicht über aktuelle Landungen am Frankfurter Flughafen und eine Lebensmittelwarnung für Laugen-Brezeln der Verbraucherzentrale, um drei besonders unterschiedliche Beispiele zu nennen. Dabei entzogen sich diese Inhalte nicht nur bestehenden inhaltlichen Klassifizierungssystemen – ohne Information zum konkreten ‚Sharing‘-Kontext (dem Text der Statusmeldung) war in vielen Fällen nicht einmal klar, was eigentlich der Inhalt ist. So wurden Seiten geteilt, auf denen unzählige verschiedene ‚Inhalte‘ neben- und untereinander standen: Videos, kürzere und längere Texte, Listen, einzelne Bilder und komplette Fotogalerien, die zum Teil nicht einmal einen eindeutigen thematischen Zusammenhang erkennen ließen.

Entsprechend musste unser ursprüngliches Ansinnen, alle gefundenen Inhalte in einem einheitlichen Kategoriensystem zu kategorisieren, früh im Forschungsprozess fallen gelassen werden. Aber auch die für die Beantwortung der Forschungsfragen notwendige Minimal-Kategorisierung in ‚Inhalte von öffentlicher Relevanz‘ und ‚politische Inhalte‘ hatte ihre Tücken: Während bei professionell erstellten Inhalten zumeist gut erkennbar ist, ob sie das Themenfeld ‚Politik‘ betreffen, zum Beispiel weil politische Akteure genannt werden, war dies bei vielen der Webseiten unserer Stichprobe nicht so. Gerade in den als Internet-Meme weit verbreiteten humorvollen Bild/Text-Kombinationen (zumeist ein Foto mit wenigen, in breiter Sans-Serif-Schrift gesetzten Wörtern) wurde vielfach auf tagespolitisches Ereignisse angespielt, ohne jedoch Akteure oder politische Konzepte zu nennen. Erst nach der Anlage eines sehr breiten Politikverständnisses und einer mehrfachen Nachschulung der Codiererinnen konnte hier eine zufriedenstellende Interdecoderreliabilität erreicht werden.

Bei der Bestimmung der ‚öffentlichen Relevanz‘ eines Inhalts stießen wir an die Grenzen der begrifflichen und konzeptionellen Bedeutung von ‚Öffentlichkeit‘ als Sphäre gesamtgesellschaftlicher Vermittlung und Selbstbeobachtung (Gerhards, 1994): Ist der Spielbericht von einem Amateurfußballturnier, die Nachricht über die Einführung neuer virtueller Waffen in einem Multiplayer-Spiel oder die Petition gegen die Errichtung einer lokalen Umfahrungsstraße für ‚die Öffentlichkeit‘ relevant, beziehungsweise *für welche*? Mit der Entscheidung, all jene Themen als

4 Massively Multiplayer Online Games, bezeichnet Computerspiele, die über das Internet gleichzeitig von mehreren Tausend Spielern gespielt werden können.

„öffentlich relevant“ zu kategorisieren die „auch in einer Tageszeitung (General-Interest) behandelt werden könnten“, so unser Codebuch (siehe Fußnote 3 weiter oben), haben wir zwar einen pragmatischen und in Anbetracht der Forschungsfragen wohl durchaus sinnvollen Weg eingeschlagen, das Problem an sich aber nicht gelöst.

Als zusätzliche und etwas anders gelagerte Herausforderung erwies sich die Bestimmung der Quelle eines bestimmten Medieninhalts. Angelehnt an den Mehrstufenfluss fragen wir in der Studie insbesondere nach dem Diffusionsprozess zwischen dem Produzenten eines Medieninhaltes und dem Rezipienten, und so ist die Bestimmung des Akteurs, der den jeweiligen Medieninhalt erstellt hat, von besonderer Bedeutung. Dabei war uns die Unterscheidung zwischen Medienunternehmen, anderen professionellen Kommunikatoren (insbesondere politische wie Parteien, NGOs, etc.) und nicht-professionellen Kommunikatoren wichtig, da es uns unter anderem um die Relevanz des Journalismus für Facebook-Newsfeeds ging.

Abgesehen von der klassischen Inhaltsanalyse-Problematik der Zuordnungsregeln zu diesen Kategorien, hatten wir mit dem Umstand zu kämpfen, dass sich bei vielen Inhalten mehrere mögliche ‚Produzenten‘ anboten, die wir vielfach auch unterschiedlichen Kategorien zugeordnet hätten: Ist die Quelle eines von einem anonymen Nutzer auf YouTube hochgeladenen Ausschnitts einer ZDF-Sendung dieser Nutzer oder das ZDF? Wer ist der Produzent eines Spielmitschnitts aus einem MMOG auf einer einschlägigen Plattform: die Spieler, die Plattform, oder gar der Spielhersteller? Und wie verhält sich dies bei einem Eintrag in einem Wiki? Die auf Facebook geteilten digitalen Objekte verunmöglichen vielfach die eindeutige Bestimmung einer Quelle, weil in den Prozess der Produktion, Distribution und zuweilen Rekonfiguration von Medieninhalten unterschiedliche Akteure eingebunden sind und diese Einbindung auch sichtbar ist. Dieses Problem stellt sich insbesondere, aber nicht nur, im Kontext von User-Generated-Content-Plattformen wie YouTube, Flickr oder Pinterest.

In unserer Studie haben wir schließlich diejenigen als Produzenten identifiziert, die offensichtlich für die Erstellung des Inhalts verantwortlich sind und dessen finale Version im Rahmen einer kreativen Eigenleistung erstellt haben. So ist der Produzent eines MMOG-Mitschnitts der Spieler (User Generated Content), da er durch seine Spielaktionen den Inhalt maßgeblich mitgestaltet hat. Demgegenüber ist der Produzent des ZDF-Videoausschnitts bei YouTube das ZDF selbst, da der

Nutzer den Beitrag lediglich hochgeladen aber nicht weiter kreativ bearbeitet hat. Dieses Vorgehen ist angelehnt an Neuberger (2008), der bei den Rollen öffentlicher Kommunikation zwischen Anbietern (=Betreiber eines Angebots) und Leistungserbringern (=verantwortlich für die Inhalteproduktion) unterscheidet.

3.5 Zugänglichkeit der Inhalte

Die oben skizzierten Probleme bei der Kategorisierung von Medieninhalten lassen sich pragmatisch lösen; für das Problem der teilweisen Unzugänglichkeit von digitalen Objekten gilt das nicht. Im Zuge der Codierung haben sich unterschiedliche Problemlagen gezeigt, durch die der Zugriff auf die jeweiligen Inhalte für die Forscher erschwert oder sogar verunmöglicht wurde:

- Die *zeitliche Verfügbarkeit* ist bei einigen Medieninhalten eingeschränkt. Zwischen den Zeitpunkten der Datenerhebung und der Codierung lagen mehrere Wochen und einige Objekte waren nach diesem Zeitraum nicht mehr verfügbar (308 von insgesamt 2.635 Links, entspricht 12%). Wurde auf der jeweiligen Website eine entsprechende Fehlermeldung angezeigt, konnte der Umstand der Unzugänglichkeit von den Codiererinnen entsprechend notiert werden. Manche Webseiten leiten die Nutzer bei nicht (mehr) verfügbaren Inhalten jedoch ohne Hinweis auf die Startseite ihres Angebots. In diesen Fällen war nur durch den Unterschied zwischen ursprünglich geteilter URL und der im Moment der Codierung im Browser angezeigten URL erkennbar, dass es sich nicht um den ursprünglich verlinkten Inhalt handelt.
- Auch die *geographische Verfügbarkeit* ist bei bestimmten Objekten eingeschränkt. So sind beispielsweise Videos, die von einer österreichischen IP-Adresse problemlos rezipiert werden können, von einem Standort in Deutschland auf Grund des Lizenzstreits zwischen Google und der Verwertungsgesellschaft GEMA nicht verfügbar. Wir haben uns dafür entschieden, dass alle entsprechenden Videos von der österreichischen Codiererin analysiert werden.
- Einige Inhalte waren in *anderen Sprachen* als Deutsch und Englisch verfasst, und damit von den Codiererinnen nicht lesbar (71 von 2.635 Links, entspricht 2%).

- Bestimmte Objekte scheinen nur sichtbar, wenn der jeweils betrachtende Nutzer in jenem (semi-öffentlichen) System *eingeloggt* ist, in dem der Inhalt abgelegt wurde. Diese Links sind nicht codierbar (186 von 2.635 Links, entspricht 7%).
- Schließlich führen zahlreiche Links auf Seiten, deren *Inhalte* sich unverkennbar häufig *verändern*. Dies betrifft vor allem Startseiten von Webangeboten. Hier besteht zwar kein grundsätzliches Zugänglichkeitsproblem, es ist jedoch im Rahmen der vorgestellten Studienanlage im Moment der Codierung klar, dass zum Zeitpunkt der Erhebung ein anderer Inhalt eingeblendet und damit möglicherweise ein anderer Inhalt gemeint war. Diese Links (297 von 2.635 Links, entspricht 11%) haben wir jedoch trotz etwas eingeschränkter Aussagekraft codiert und in die Analyse einbezogen.

Insgesamt konnten auf Grund dieser Probleme nur 79 Prozent oder 2.070 der 2.635 von den Befragten kopierten Links auch analysiert werden, wobei die größte Hürde die begrenzte zeitliche Verfügbarkeit der Inhalte darstellte. Zukünftige Studien sollten aus diesem Grund darauf achten, die Inhalte der Websites möglichst bald nach Ende der Befragung zu codieren oder die Websites sofort nach der Befragung offline zu speichern, sodass die Inhalte unverändert zur Verfügung stehen.

4 Zusammenfassung

Ausgehend von Fragen zum Mehrstufenfluss öffentlicher Kommunikation auf Facebook haben wir eine Studie vorgestellt, die mit einer Kombination von Befragung und Inhaltsanalyse versucht, Einblick in die individuell strukturierten Newsfeeds von Facebook-Nutzern zu erhalten. Die direkte Ansprache der Nutzer durch eine Befragung war notwendig, da die Objekte des Forschungsinteresses (Links in Statusmeldungen) nur für die individuellen Nutzer selbst zugänglich sind. Die Befragung gab uns außerdem die Möglichkeit, das Interesse der Nutzer an den Links und ihre individuelle Beziehung zu den jeweiligen Intermediären zu bestimmen, und dadurch Aussagen über die Zusammenhänge zwischen Diffusionsmoment (Wer teilte den Link?) und Medieninhalt zu machen. Durch die Inhaltsanalyse der kopierten Links konnten wir zusätzlich vermeiden, die Kategorisierung der Objekte der Einschätzung der Befragten zu überlassen.

Das Vorgehen ist gewinnbringend, wie die Ergebnisse zeigen, aber auch herausfordernd. Das liegt zum einen an den komplexen und ungewöhnlichen Anforderungen an die Befragungsteilnehmer, die nicht nur Fragen beantworten, sondern auch konkrete Inhalte (Links) bereitstellen sollten. Ein Viertel der Befragten lehnte diese Form der Beteiligung ab, oder zeigte sich nicht konzentriert oder kompetent genug, die Anforderungen gemäß der Vorgabe zu erfüllen – obwohl das Erhebungsinstrument im Vorfeld mehrfach getestet und entsprechend vereinfacht wurde. Zukünftige Studien sollten die Bereitschaft der Nutzer für die Teilnahme an solchen Befragungen jedoch nicht überschätzen und sich darauf konzentrieren, die Schwelle für die Teilnahmebereitschaft möglichst niedrig anzusetzen. Grundsätzlich sind auch technisch unterstützte Lösungen vorstellbar: So könnten die Nutzer einer Anwendung kurzfristig erlauben, Daten aus ihrem Facebook-Profil zu speichern. Dies würde den Aufwand verringern, könnte aber andererseits zu einer verstärkten Ablehnung der Teilnahme aufgrund von Datenschutz-Bedenken führen, und würde gleichzeitig vielfältige forschungsethische Probleme aufwerfen.

Ein anderer Herausforderungskomplex betrifft die Inhaltsanalyse der verlinkten Dokumente. Hier zeigt sich, wie vielfältig sich das Feld öffentlicher (politischer) Kommunikation heute präsentiert, und wie schwierig die hochgradig diversen Inhalte und Akteure mit tradierten (massen-)kommunikationswissenschaftlichen Klassifikationssystemen zu bestimmen sind. Kurzfristig können wir für ähnlich gelagerte Inhaltsanalysen nur empfehlen, sich nicht an althergebrachten Kategorien zu orientieren, sondern diese in qualitativen Vorstudien auf die konkrete Fragestellung hin zugeschnitten selbst zu entwickeln. Mittel- und langfristig sollte die Kommunikationswissenschaft ein neues empirisch fundiertes Angebot für die Einordnung der diversen Kommunikationsobjekte, abseits medientechnisch definierter Begriffe wie ‚Weblog‘ oder ‚Infoportal‘ machen, die dem weitreichenden Wandel öffentlicher Kommunikationsprozesse, abseits der ‚Vorlage‘ Massenkommunikation, gerecht wird.

Katrin Jungnickel, M.A. ist wissenschaftliche Mitarbeiterin am Institut für Medien- und Kommunikationswissenschaft der Technischen Universität Ilmenau

Dr. Axel Maireder ist wissenschaftlicher Leiter des Social Media Intelligence Center der GfK und war bis Ende 2014 Mitarbeiter am Institut für Publizistik und Kommunikationswissenschaft der Universität Wien

Quellenverzeichnis

- An, J., Cha, M., Gummadi, K. P., & Crowcroft, J. (2011). Media landscape in Twitter: A world of new conventions and political diversity. *Proceedings of the Fifth International Conference on Weblogs and Social Media* (July 2011). Abgerufen von <http://www.cl.cam.ac.uk/~jac22/out/twitter-diverse.pdf>
- Aral, S. (2011). Commentary-Identifying Social Influence: A Comment on Opinion Leadership and Social Contagion in New Product Diffusion. *Marketing Science*, 30(2), 2172-23. doi: 10.1287/mksc.1100.0596
- BITKOM (2011): Soziale Netzwerk werden zum Informationskanal. Abgerufen von http://www.bitkom.org/70397_70419.aspx
- Bucher, T. (2012). Want to be on the top? Algorithmic power and the threat of invisibility on Facebook. *New Media & Society*, 14(7), 1164-1180. doi: 10.1177/1461444812440159
- Busemann, K. (2013). Wer nutzt was im Social Web? *Media Perspektiven*, (7-8), 391-399.
- Couldry, N. (2009). Does “the Media” Have a Future? *European Journal of Communication*, 24(4), S. 437-449. doi: 10.1177/0267323109345604
- de Zuniga, H. G., & Valenzuela, S. (2011). The Mediating Path to a Stronger Citizenship: Online and Offline Networks, Weak Ties, and Civic Engagement. *Communication Research*, 38(3), 397-421. doi: 10.1177/0093650210384984
- Gerhards, J. (1994). Politische Öffentlichkeit. Ein system- und akteurstheoretischer Bestimmungsversuch. In F. Neidhardt (Hrsg.), *Öffentlichkeit, Öffentliche Meinung, Soziale Bewegungen. Sonderheft 34 der Kölner Zeitschrift für Soziologie und Sozialpsychologie* (S. 77-105). Opladen: Westdeutscher Verlag.
- Granovetter, M. S. (1973). The Strength of Weak Ties. *American Journal of Sociology*, 78(6), 1360-1380. Abgerufen von <http://www.jstor.org/stable/2776392>
- Guenther, T., & Schmidt, J. (2008). Wissenstypen im “Web 2.0” - eine wissenssoziologische Deutung von Prodnutzung im Internet. In H. Willems (Hrsg.), *Weltweite Welten. Internet-Figurationen aus wissenssoziologischer Perspektive* (S. 167-188). Wiesbaden: VS Verlag.
- Haythornthwaite, C. (2002). Strong, Weak, and Latent Ties and the Impact of New Media. *The Information Society*, 18, 385-401. doi: 10.1080/01972240290108195
- Kincaid, J. (2010, 22. April). EdgeRank: The Secret Sauce That Makes Facebook’s News Feed Tick. *TechCrunch*. Abgerufen von <http://techcrunch.com/2010/04/22/facebook-edgerank/>

- Krotz, F. (2000). Öffentlichkeit und medialer Wandel. Sozialwissenschaftliche Überlegungen zu der Verwandlung von Öffentlichkeit durch das Internet. In W. Faulstich, & K. Hickethier (Hrsg.), *Öffentlichkeit im Wandel: neue Beiträge zur Begriffsklärung* (S. 210-223). Bardowick: Wissenschaftler-Verlag.
- Lazarsfeld, P., Berelson, B., & Gaudet, H. (1944). *The People's Choice. How the voter makes up his mind in a presidential campaign*. New York: Columbia University Press.
- Morey, A. C., Eveland, W. P., & Hutchens, M. J. (2012). The "Who" Matters: Types of Interpersonal Relationships and Avoidance of Political Disagreement. *Political Communication*, 29(1), 86-103. doi: 10.1080/10584609.2011.641070
- Neuberger, C. (2008). Neue Medien als Herausforderung für die Journalismustheorie: Paradigmenwechsel in der Vermittlung öffentlicher Kommunikation. In C. Winter, A. Hepp, & F. Krotz (Hrsg.), *Theorien der Kommunikations- und Medienwissenschaft. Grundlegende Diskussionen, Forschungsfelder und Theorieentwicklungen* (S. 251-268). Wiesbaden: VS Verlag.
- Straits, B. (1991). Bringing strong ties back in: Interpersonal Gateways to political information and influence. *Public Opinion Quarterly*, 55(3), 432-448. doi: 10.1086/269272
- Schiller, B., Heimbach, I., Strufe, T., & Hinz, O. (2013). Development of the Social Network Usage in Germany in 2012. Abgerufen von <http://www.emarkets.tu-darmstadt.de/forschung/social-network-usage-in-germany-study/>
- Ugander, J., Karrer, B., Backstrom, L., & Marlow, C. (2011). The Anatomy of the Facebook Social Graph. Abgerufen von <http://arxiv.org/abs/1111.4503>
- Weimann, G. (1983). The Strength of Weak Conversational Ties in the Flow of Information and Influence. *Social Networks*, 5, 245-267. doi: 10.1016/0378-8733(83)90027-8
- Weimann (1994). *The influentials. People who influence people*. New York: State University of New York Press.
- Weissensteiner, E., & Leiner, D. (2011). Facebook in der Wissenschaft: Forschung zu sozialen Onlinenetzwerken. *Medien & Kommunikationswissenschaft*, 59(4), 526-544.
- Wojcieszak, M.E., & Mutz, D. (2009). Online Groups and Political Discourse: Do Online Discussion Spaces Facilitate Exposure to Political Disagreement? *Journal of Communication*, 59(1), 40-56. doi: 10.1111/j.1460-2466.2008.01403.x
- Wu, S., Hofman, J. M., Mason, W. A., & Watts, D. J. (2011). Who Says What to Whom on Twitter. WWW 11, Hyderabad, India. doi: 10.1145/1963405.1963504

Suggested Citation: Eble, M., & Stein, D. (2015). Utilisation of Audio Mining Technologies for Researching Public Communication on Multimedia Platforms. In A. Maireder, J. Ausserhofer, C. Schumann, & M. Taddicken (Eds.), *Digitale Methoden in der Kommunikationswissenschaft* (S. 329-345). doi: 10.17174/dcr.v2.14

Abstract: The number and volume of spoken language corpora which are generally available for research purposes increase significantly. That is due to the wide adoption of audio-visual communication on news websites and social web platforms. The respective messages that are published by professional and individual communicators are subject to online content analysis. To date, such analyses strongly rely on manually operated processes which come along with a huge effort for transcribing spoken language corpora into textual content. Hence, challenges like the ever increasing volume, velocity and variety of multimedia content need to be faced. Audio Mining technologies are capable of reducing the effort for turning speech into text significantly. Using these technologies via application programming interfaces (APIs), it is demonstrated how a hybrid approach enables researchers to reduce the time that is needed for analysing news content by an order of magnitude.

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

Michael Eble & Daniel Stein

Utilisation of Audio Mining Technologies for Researching Public Communication on Multimedia Platforms

1 Public communication on online multimedia platforms as research subject

Everyday public communication fulfils several functions regarding transparency, validation and orientation for modern societies (Neidhardt, 1994; Donges & Imhof, 2005). Audio-visual media content has become an essential part of public communication, and it takes place on Internet platforms for two reasons: First, such content is constantly being produced by professional publicists, editors, journalists etc. for multimedia platforms like BBC.co.uk. Secondly, it is also being used in professionally motivated public communication as well as individual communication, e.g., on social web platforms like YouTube. The production and consumption of audio-visual media content enable online-based publics for certain topics like elections or sport events (van Eimeren & Frees 2012, p. 371). In addition to that, social web platforms have been established as infrastructures for networked publics and follow-up communication through strongly interlinked audio-visual media content (Schmidt, 2011 & 2013). A current example for this is the uptake of applications and formats in the area of “Social TV” and “Second Screen” (Eble, 2013b).

Overall, the number and volume of spoken language data which are generally available for research purposes in the field of public communication increase significantly. The *growing importance of audio-visual messages* for public communication can be tracked down to three reasons: First, editorial staff at online media organisations increase their production and publication of audio-visual content. Secondly, media organisations digitize and open up their archived audio-visual content in order to attract more users to their content and brand. Thirdly, user-generated content is no longer restricted to text but includes more and more video content instead.

Public communication that takes place on Internet platforms and makes use of audio-visual messages is subject to *online content analysis*. However, several challenges need to be addressed in order to successfully carry out such analyses of multimedia content. Given that context, the paper at hand sets out to contrast methodical challenges and approaches from the field of automatic speech analysis (section 2). Furthermore, it aims for strengthening the bridge between computer science and communication studies by giving an overview on the utilisation of Audio Mining technologies for analysing public communication on multimedia platforms (section 3). A hybrid approach is emphasised that takes into the account the advantages arising from collaborative human and computational analysis (section 4). The authors are convinced that such an approach is inevitable in order to cope with future challenges in online communication research.

2 Methodical challenges in the analysis of audio-visual media content

Content analysis is an essential and well-known research method within media and communication science (Wirth, 2001). According to Früh (2007, p. 27) it is understood as an empirical, non-reactive process for the “systematic, inter-subjectively comprehensible description of content and formal characteristics of messages” that enables the quantitative investigation of messages.¹

1 The quote has been translated by the authors from German as given in the source to the English version at hand.

Thus, quantitative content analysis is about generalised statements that are based on the systematic study of amounts of text. Therefore, content analysis sets out to identify certain patterns using a code book that is based on theoretical frameworks and individual research questions (Früh, 2007, p. 39). In addition to a formal descriptive analysis, a diagnostic approach is possible: One can infer from parameter values of the content to the communicator, the recipient and the communication situation, for example (Merten, 1995, p. 23; Früh, 2007, p. 44). In general, content analysis is characterized by a high degree of flexibility, as it can be applied on text, image, audio and video at any time, arbitrarily often and with varying research questions (Früh, 2007, p. 39). The method is mainly applied to content that is produced and published by public media institutions (Rössler, 2010, p. 34), which is due to the discipline's research tradition as well as its concentration on media services that are most relevant for society (Zeller & Wolling, 2010, p. 143).

Meanwhile, online content analysis has been well-established in order to study the structures and messages of content published by institutional online media (Welker & Wünsch, 2010). Rössler and Wirth (2001, p. 284) propose a typology for online content analyses that includes 1) a content-centred approach and 2) a user-centred approach. The first approach consists of area or category analyses (e.g. genre, format, groups of organisations) as well as focus analyses (e.g. events, issues, persons). In the second approach, they distinguish between publicity analyses (e.g. on the basis of traffic or other usage metrics) and selectivity analyses (e.g. according to user behaviour). A combination of the content-centred and the user-centred approach can increase the insights gained through this method (Merten, 1995, p. 119; Zeller & Wolling, 2010).

Online content analysis needs to deal with heterogeneous media services and interlinked multimedia messages: Due to current and upcoming technical and editorial patterns of news production, distribution and usage, the respective data is distributed among various sources. That is, research data are characterised not only by increasing volume, but also by velocity and variety of online content. With respect to such challenges, Schweitzer (2010, p. 45) sees

2 The quote has been translated by the authors from German as given in the source to the English version at hand.

the need for intensive discussions on specific recommendations for future research practice. This also seems to be strongly required according to Rössler (2010, p. 41), who argues that expertise in the field of online content analysis could become a “key qualification for empirical research in all social science disciplines”.²

With these considerations in mind and with respect to audio-visual content, challenges of online content analysis can be described in more detail as follows (Eble, 2013a; Eble, Ziegele, & Jürgens, 2014):

- The *definition of research samples* can be challenging or almost impossible in terms of the sample’s determinability and definition of the selection unit.
- The subsequent *collection of audio-visual data* is challenging especially in terms of its (from time to time hardly predictable) dynamic and constant flow (e.g. live streams) as well as its volume. That requires appropriate automatic software agents (crawler).
- Having collected the data that is needed for one’s research question leads to the need for *archival and transfer of data* which comes along with challenges like storing raw (i.e. unstructured) and analysed (i.e. structured) data in storage systems most suitable for the respective purpose (e.g. HDFS, NoSQL or SQL databases)³. Storing the initial raw data apart from the results of its processing puts the researcher in the position to re-analyse the messages based on future research questions.
- Finally, for the *analysis of data*, issues like avoiding live coding, encoding and re-coding, ensuring intra-coder and inter-coder reliability as well intersubjective confirmability, etc. have to be tackled (McMillan, 2000, p. 93; Rössler & Wirth, 2001, p. 296; Weare & Lin, 2000, p. 287). Furthermore, researchers have to deal with legal and ethical issues in terms of data privacy and personal rights (Eble et al., 2014).

3 HDFS: Hadoop Distributed File System; NoSQL: Not only SQL; SQL: Structured Query Language.

Given that context, the following section focuses on the automatic transformation of spoken content into textual content via audio mining technologies in order to enable researchers to carry out the actual content analysis much faster. Thus, the content within audio-visual media, e.g., the transcript of the interview with a politician derived from speech, becomes conveniently accessible for manual coding by human annotators which use tools like SPSS or QDA Miner and apply a specific coding scheme. That is, the paper at hand primarily addresses challenges occurring with data analysis. In addition, some insights concerning challenges of storing and transferring data are provided.

3 Supporting content analysis through Audio Mining technologies

In order to illustrate challenges that arise in communication science research when working with audio-visual media content, and to elaborate on approaches that computer science could offer, we make use of the following (conceptual) use case scenario:

Jessica Müller, PhD student, is investigating the (presumed) change of journalistic reporting on renewable energies and their positive and negative effects as a focus analysis on the topic “renewable energy”. This content analysis is focusing on four major broadcast news shows in the time frame of January till December 2014, working with ARD, ZDF, RTL and SAT1 which have the highest media coverage (selected in terms of publicity aspects). The creation of appropriate categories has already been conducted, and the code book is present. In order to move on, Jessica now needs access to transcribed text of the spoken content of all the video files.

Based on the author’s experience a manual transcription of these broadcast news shows seems impractical. Assuming a real-time factor of seven for the manual transcription of words (in a very basic manner such as: words, interjections and speaker changes; more elaborate schemes such as, e.g., GAT2 (Selting et al., 2009) would take considerably more time) in an audio signal means that a daily broadcast of a quarter of an hour over the course of a year would result in roughly eighty working days. It thus seems reasonable to employ automatic speech analysis (ASA) as a pre-processing step, at the very least to identify the videos relevant for further, manual post-processing. In direct comparison to other media processing techniques, ASA is comparatively old and well-established. Many new

systems have adopted a deep neural network (DNN) paradigm (Hinton, Deng, Yu, Dahl, Mohamed, Jaitly, & others, 2012), and it seems that these machine learning architectures that consist of multiple non-linear transformations are quite capable in handling speech input. With modern computer power, they are also real-time capable and outperform former approaches with more rigid architectures quite consistently.⁴

Jessica processed the videos with automatic speech analysis. On a computer with eight processor cores, 365 hours of video (assuming a proper audio/video compression and a resolution of 512x288, this equals to roughly 150 GB) in less than two days. She is searching for the first-best video that matches her search query “renewable energy” and checks the transcript for quality. Especially those proportions of the video where the anchor is speaking are quite intelligible – once she adjusts to the fact that the transcripts feature no punctuation marks. In an interview with a local resident that has quite a thick accent, the quality drops significantly. Jessica also notices that some words she is looking for are transcribed rather odd or even wrong, e.g., “null Energie Haus” rather than “Nullenergiehaus”; and “Off Shoah Wind Parks” rather than “Offshore-Windparks”.

Especially for broadcast news, ASR is considered to be well-established. For planned speech with clean audio quality (i.e., sufficient sample rate with no or little background noise) and a known speech domain (news, sport, tourism) it shows very good results (Jurafsky & Martin, 2000). However, it is also true that deviations⁵ from this ideal scenario are still in the focus of current research (Goldwater et al., 2010). These deviations include: spontaneous speech, together with hesitations, interjections and incomplete sentences; background noise (and especially music, which has strong similarities to speech characteristics because of its harmonic nature); multiple, simultaneous speakers; and pronunciation variants due to speech disorders, accentuated speech or loan words (German: “getwittert”, “unfriended”, “bromance”). The underlying reason for some of these errors is the internal process of the ASR system, which is typically built around recognizing phonemes before mapping phoneme chains to actual words. Unknown

4 For the systems used for our own research projects, we witness an error reduction of about one-third.

5 The term “deviation” is meant to be mathematical here and indicates a statistical deviation with respect to the models derived from ideal data.

words (i.e., never seen before during the ASR training phase) and neologisms such as “Nullenergiehaus” are thus as problematic as words that might have been encountered in previous texts but are pronounced in an uncommon way (e.g., loan words such as “Offshore-Windparks” or local pronunciation variants).

In order to quantify the quality of the transcripts, and to decide whether they are appropriate for her studies in spite of the errors, Jessica corrects the transcripts of a bunch of videos by hand and computes the word error rate. For the selected videos, she obtains an error rate of 15 percent.

The common quality criterion for automatically derived transcripts is the word error rate (WER), which is the number of transcription errors weighted over the total number of reference words. Transcription errors include insertions, deletions and substitutions (also called “Levenshtein distance” or “edit distance”; Levenshtein, 1966). In an ideal scenario, WERs that are below ten percent are often witnessed. Studies suggest that a WER of up to 25 percent is still perceived to be intelligible (or at least understandable) by humans, and that WERs of up to 40 percent still render the transcripts to be usable by further text mining algorithms (Munteanu et al., 2006). ASR systems also typically mark words with a confidence measure, which indicated how sure the system was during the recognition. They can thus be employed to discard those words that are considered to be most error-prone.

Jessica feels that the WER of 15 percent is reasonable to obtain a first impression on the data, but for reliable conclusions for her analysis she deems this error rate to be too high. She decides to go for a hybrid approach, by using data mining techniques in order to identify those parts of the data that are important for her research, and to manually correct the transcripts of these parts in a second run.

In general, technologies from the field of data mining have the task to derive new, i.e., not already explicitly stated, patterns and information from (largely arbitrary) data. The algorithms and techniques employed draw their theoretical background from the fields of artificial intelligence (machine learning) (Han & Kamber, 2006) and are typically statistically-based.⁶ Depending on the source and

6 I.e., they do not consist of manually defined rules but are based on statistical data-observation.

modality of the data, the techniques are also grouped into “text mining”, “web mining” (here, this is primarily the analysis of web-based information such as interlinking of websites, server statistics, and user behaviour) and “social mining” (activity within social web platforms). Following an established definition (Fayyad, Piatetsky-Shapiro, Smyth, 1996), aspects of data mining are (a) grouping of similar content, (b) classification, (c) regression analysis, (d) association analysis, (e) outlier detection and (f) summary.

Jessica wants to employ key word extraction on her documents in order to find those excerpts that are the most relevant to her term “renewable energy”. She computes the 40-best (i.e., the 40 most probable candidates based on the underlying statistical model) key words automatically, but multiplies these values by the confidence of the ASR transcript output, in order not to overestimate possibly erroneously recognized words.⁷

A keyword is part of the document itself and highlights words that seem to disambiguate the document in comparison to other documents; thus, a keyword is not a stand-alone description of a document’s content but rather a highlighted feature in comparison to a set of other given documents. A well-established technique is the term frequency inverse document frequency (TF-IDF)⁸ measure, i.e., a word is considered the most relevant if it appears considerably often in the given document but on the same time does not appear in many other documents. TF-IDFs also show good performance for ASR transcripts (Schneider & Tschöpel, 2010). For Jessica’s query, the key word actually is composed of two words (“renewable energy”), which is typically coped by modern algorithms. Especially in morphologically rich languages such as German, the words are often reduced to the stem form by automatic means and compound words can be split so that the

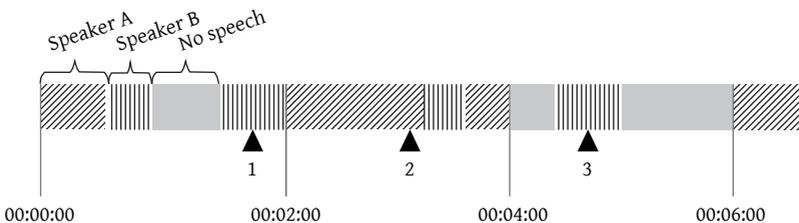
- 7 The TF-IDF algorithm used for keywords will be very keen on words that seem to be unique for a specific document/transcript and thus give it high rankings; thus, especially for words where the speech recognized did not assign high probabilities, i.e., it was unsure if this word was recognized correctly, the TF-IDF score needs to be adjusted in order to not over-emphasize wrong transcript parts.
- 8 The term “frequency” appears twice in this name since there are two frequencies compared against each other: the frequency of the words within the document versus the frequency within all reference documents.

algorithms do not treat similar words such as “Energie”, “Energien” and “Solar-energie” as completely different entries.⁹

Employing the key word technique, Jessica now samples a smaller corpus of 200 videos which contain the key word “renewable energy” with a reasonable ASR confidence. This is quite a step ahead from the originally 1.400 videos in her collection. However, those videos are still quite long and poly-thematic, and large proportions are not relevant for her analysis. Since she knows that each topic is introduced by the anchor, she decides to segment the videos based on speaker changes. After the segmentation, Jessica marks a few proportions where an anchor speaking (there are roughly 5-10 anchors per broadcast show), which results in the new segments to be mostly mono-thematic and much smaller than before.

Segmenting audio data based on speaker changes is also called “diarization”. This does not require known speakers; instead, the segments are assigned to alpha-numerical values such as speakerA-speakerB-speakerA-speakerC. This is called speaker identification (SID) if the individuals are known in advance. One state-of-the-art approach usable for both tasks is the i-vector paradigm (Dehak et

Figure 1: Exemplary result of a segmentation and analysis process performed by Audio Mining



Example for segment no. 1

Speaker B: “Renewable energy is one of the most important topics in the bread and pretzel industry.”

⁹ E.g., by Snowball stemming as part of the Lucene project (lucene.apache.org)

¹⁰ C.f. <http://ivectorchallenge.nist.gov>

al., 2011). Here, every segment is projected into a sparse dimensional vector space where different speakers have a large distance to each other, while different segments of the same speaker are located more closely. In international benchmark activities where different research groups compete against each other on the same data, i-vectors show very good performance;¹⁰ moreover, they have the benefit that the sparse feature space makes it quite convenient to store after the first processing, thus building up characteristics for each segment that can be used fast for later retrieval, e.g., when the importance of a speaker becomes apparent only after the first processing.

The segments that Jessica marked as relevant now have a total duration of 300 minutes. Jessica has them corrected manually based on the existing transcripts, which she is relieved to find to be way faster than transcribing from scratch, i.e., about double real-time. For a manually confirmed transcription of the data she is interested in, she now needs roughly 2 days instead of 80. Additionally, her computer was busy for 2 days whilst the ASR was running, and she spent another working day for the speaker diarization and recognition. With the weekend approaching soon, she decides to further improve her selection before she carries on with the entries in her code book. One thing that she is interested in is an estimation of how much segments are similar to the ones she has identified so far, i.e., documents that discuss renewable energy without mentioning this particular term.

Deriving similarity measures between different documents is used for recommendation (“users looking at this document are also interested in...”) or classification of text (e.g., to identify content not suitable for minors based on the language; Bardeli et al., 2013). Similarity measures can be based on quite a large selection of features, but often also employ a bag-of-words approach, which takes the vocabulary of the document into account.¹¹ The techniques working with these similarity measures thus can identify documents that belong to the same topic but use different vocabulary, which can be extremely helpful when the meaning of specific terms change or if they are actually invented and established over the course of time.

11 A bag-of-words is a vector of integer numbers which counts word appearances given an overall vocabulary but ignores the position of the words within the document (so, typically, many entries will consist of zeros whenever those particular words do not appear within the document). It is an essential feature in many text mining applications.

The techniques depicted here are only a small excerpt of the tools that data mining can offer. For example, rather than looking at text transcripts only, investigating phenomena on paralinguistic aspects such as prosody, emotion, the structure of the speaker changes within heated discussions or the linking strategy of a user can be of interest. Audio fingerprinting approaches, which identify known data that has been used in other contexts, can generate valuable information, e.g., by identifying when snippets of an interview are referenced in subsequent shows (Bardeli et al., 2012).

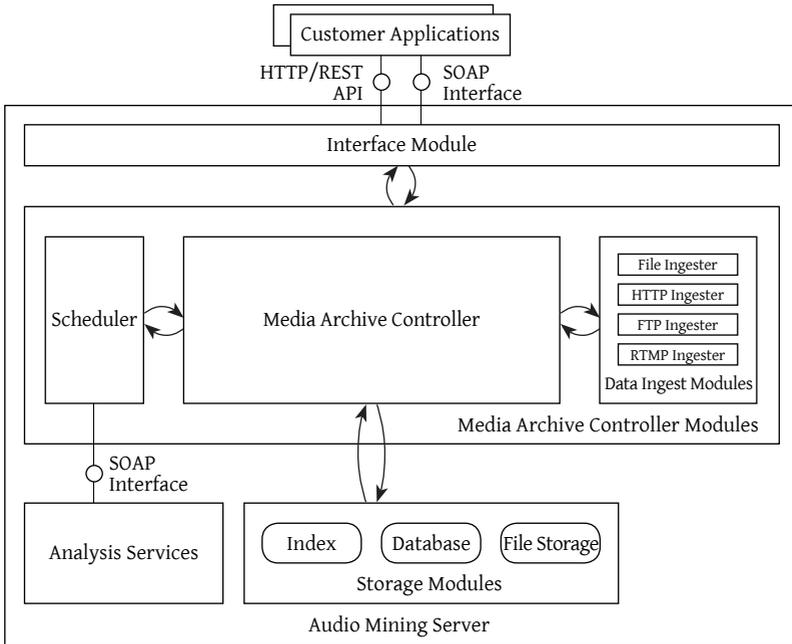
For most of these techniques, open-source solutions exist (e.g. for transcribing and maintaining: ELAN, EXMARaLDA; for natural language processing: Mallet, NLTK, Weka). However, while the algorithms itself come usually for free, they have to be trained in advance with appropriate training data (e.g. for German-language news content) before they become fully usable.

4 Transfer: System architecture for Audio Mining technologies

With respect to the use case scenarios and technologies as described above, a system architecture for researching public communication on multimedia platform can be described as follows (see FIcontent wiki, 2013): *Storage Modules* take care of capturing all the video data to file storage (e.g. as MPEG-7) and result from processing in indices and databases (e.g. XML, Solr), while *Media Archive Controller Modules* deal with all processes that need to be carried out for ingesting audio-visual media content coming from several sources and scheduling its analysis (e.g. downloading video data, processing data, exporting transcripts). A component called “*Analysis Services*“ performs the actual automatic speech analysis that is relevant for research purposes and has been described above. Finally, the *Interface Module* is designed to enable application development on top of the Audio Mining system. That is, a browser-based web interface enables its users to view transcripts of spoken video content. Figure 2 (next page) comprises the modules described beforehand.

Since systems are often specifically designed for certain research purposes, there is a need for the flexible implementation of tools for automatic speech analysis via open and documented application programming interfaces (APIs). For the

Figure 2: Architecture of Audio Mining system



Source: Fcontent wiki, 2013

system described above, the API documentation¹² and a developers' guide can be found online. Therefore, using an API to automatic speech analysis for German-language content in a specific study design could enable researchers to implement scenarios as the two following (Eble, 2013a, p. 238.): The *combination of focus analyses and publicity analyses on multimedia platforms* sets out to investigate certain topics of public interest (e.g. “renewable energies”) both regarding their initial messages on

12 API documentation: <http://wiki.mediafi.org/doku.php/fcontent.socialtv.enabler.audiomining>

13 Developers' guide: <http://wiki.mediafi.org/doku.php/fcontent.socialtv.enabler.audiomining.developerguide>

news sites like interviews with politicians on BBC.co.uk and their follow-up communication on social web platforms like YouTube and Facebook. Due to recent developments like Social TV, the linkages between public communication and interpersonal communication become more important. The *(near) real-time capturing of audio-visual news streams* puts researchers in the position to perform content analysis of unforeseeable news events. Due to increasing, dynamic streams of data relevant for news events and its diffusion and follow-up communication, the need for continuous data collection and according capturing systems increases, too. Otherwise, the short periods of diffusion and follow-up communication tend to prevent a reasonable retrospective study (see De Fleur, 1987, p. 125). Thus, a continuous and automatic data collection for a standardized set of parameters of a certain number of wide-reaching news sites would be a promising approach, i.e. for (comparative) re-analysis. Note that issues regarding data privacy and protection need to be addressed in such online research. Current algorithms are capable of ensuring data protection during as well as after the collection and processing of multimedia data.

5 Conclusion and outlook

The paper contrasts methodical challenges and approaches from the field of automatic speech analysis. Furthermore, it provides information about Audio Mining technologies and their use for analysing public communication on multimedia platforms in order to strengthen the bridge between computer science and communication studies. It is demonstrated how the specific use case scenario of a focus analysis on the topic “renewable energies” can be supported by automatic approaches: Speaker segmentation and identification enable researchers to efficiently split audio-visual corpora into individual parts that can be skimmed quickly. Following that, speech analysis allows for turning spoken words into text so that the content becomes accessible for content analyses. It has also been shown how a concrete system implementation of Audio Mining can look like. The FContent API enables developing applications for research purposes in order to carry out content analysis on multimedia platforms and to automatically handle heterogeneous multimedia data.

In further research, at least three topics need attention: First, the combination of automatic processing and human analysis needs to be improved in terms

of tools that are easy to use. To date, several data sets are available for training, like those provided by the *Linguistic Data Consortium* (LDC; www ldc.upenn.edu) or the *European Language Resources Association* (ELRA; www.elra.info). Secondly, algorithms and their implementations need further improvement in order to handle the ever increasing amounts of audio-visual news content in (near) real-time. Thirdly, collaboration between communication studies and computer science needs to be strengthened in order to deeply understand each other's concepts, research questions and approaches.

Dr. Michael Eble is Consultant at mm1 Consulting & Management PartG in Stuttgart

Dr. Daniel Stein is Senior Research Scientist at the Fraunhofer Institute for Intelligent Analysis and Information Systems IAIS in Sankt Augustin

References

- Bardeli, R., Becker, S., Bergholz, A., Kolb, I., Korte, H., Maaskant, W., Paaß, G., Schneider, D. Stein, D., & Tschöpel, P. (2013). *Studie zum technischen Jugendmedienschutz: Möglichkeiten und Grenzen von Verfahren zur Detektion jugendschutzrelevanter Web-Inhalte*. Sankt Augustin: Fraunhofer IAIS.
- Bardeli, R., Schwenninger, J., & Stein, D. (2012). Audio Fingerprinting for Media Synchronisation and Duplicate Detection. Media Synchronisation Workshop, Berlin, Germany, October 2012.
- De Fleur, M. L. (1987). The Growth and Decline of Research on the Diffusion of the News, 1945-1985. *Communication Research*, 14(1), 109-130. doi: 10.1177/009365087014001006
- Dehak, N., Kenny, P., Dehak, R., Dumouchel, P., & Ouellet, P. (2011). Front-End Factor Analysis for Speaker Verification. *IEEE Transactions on Audio, Speech, and Language Processing*, 19(4), 788-798.
- Donges, P., & Imhof, K. (2005). Öffentlichkeit im Wandel. In H. Bonfadelli, O. Jarren, & G. Siegert (Eds.), *Einführung in die Publizistikwissenschaft* (p. 147-175). Bern, Stuttgart, Vienna: Haupt.
- Eble, M. (2013a). *Medienmarken im Social Web: Wettbewerbsstrategien und Leistungsindikatoren von Online-Medien aus medienökonomischer Perspektive*. Berlin: LIT.

- Eble, M. (2013b). Social TV, Second Screen und vernetzte Öffentlichkeiten: Kommunikationswissenschaftliche Perspektiven auf Schnittstellen zwischen Fernsehen und Social Web. In U. Breitenborn, G. Frey-Vor, & C. Schurig (Eds.), *Medienumbrüche im Rundfunk seit 1950 – Jahrbuch Medien und Geschichte 2013* (p. 73-89). Cologne: Herbert von Halem.
- Eble, M., Ziegele, M., & Jürgens, P. (2014). Forschung in geschlossenen Plattformen des Social Web. In: M. Welker, M. Taddicken, J. Schmidt, & N. Jakob (Eds.), *Handbuch Online-Forschung. Sozialwissenschaftliche Datengewinnung und -auswertung in digitalen Netzen* (p. 128-154). Cologne: Herbert von Halem.
- Fayyad, U., Piatesky-Shapiro, G., & Smyth, P. (1996). From data mining to knowledge discovery in databases. *AI magazine*, 17(3), 37. doi: 10.1609/aimag.v17i3.1230
- Fcontent wiki (2013). Audio Mining Specific Enabler for Social connected TV platform. Retrieved from <http://wiki.mediafi.org/doku.php/fcontent.socialtv.enabler.audiomining>
- Früh, W. (2007). *Inhaltsanalyse: Theorie und Praxis*. Konstanz: UVK.
- Goldwater, S., Jurafsky, D., & Manning, C. D. (2010). Which words are hard to recognize? Prosodic, lexical, and disfluency factors that increase speech recognition error rates. *Speech Communication*, 52(3), 181-200. doi: 10.1016/j.specom.2009.10.001
- Han, J., & Kamber, M. (2006). *Data mining: concepts and techniques*. San Francisco, USA: Morgan Kaufmann.
- Hinton, G., Deng, L., Yu, D., Dahl, G. E., Mohamed, A. R., Jaitly, N., et al. (2012). Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups. *IEEE Signal Processing Magazine*, 29(6), 82-97.
- Jurafsky, D., & Martin, J. H. (2000). *Speech and language processing an introduction to natural language processing, computational linguistics, and speech*. New Jersey: Pearson.
- Levenshtein, V. I. (1966). Binary codes capable of correcting deletions, insertions, and reversals. *Soviet Physics Doklady*, 10, 707-710.
- McMillan, S. J. (2000). The Microscope and the moving Target: The Challenge of applying Content Analysis to the World Wide Web. *Journalism and Mass Communication Quarterly*, 77(1), 80-98. doi: 10.1177/107769900007700107
- Merten, K. (1995). *Inhaltsanalyse. Einführung in Theorie, Methode und Praxis* (2. Auflage). Opladen: Westdeutscher Verlag.

- Munteanu, C., Baecker, R., Penn, G., Toms, E., & James, D. (2006). The effect of speech recognition accuracy rates on the usefulness and usability of webcast archives. *Proceedings of the SIGCHI conference on Human Factors in computing systems*, 493-502. doi: 10.1145/1124772.1124848
- Neidhardt, F. (1994). Öffentlichkeit, öffentliche Meinung, soziale Bewegungen. In F. Neidhardt (Eds.), *Öffentlichkeit, öffentliche Meinung, soziale Bewegungen* (p. 7-41). Wiesbaden: Westdeutscher Verlag.
- Rössler, P. (2010). Das Medium ist nicht die Botschaft. In M. Welker & C. Wünsch (Hrsg.), *Die Online-Inhaltsanalyse. Forschungsobjekt Internet* (S. 31-43). Köln: Herbert von Halem.
- Schmidt, J. (2011). *Das neue Netz :Merkmale, Praktiken und Folgen des Web 2.0* (2. Auflage). Konstanz: UVK.
- Schmidt, J. (2013). Onlinebasierte Öffentlichkeiten: Praktiken, Arenen und Strukturen. In C. Fraas, P. Meier, & C. Pentzold (Eds.), *Online-Diskurse. Theorien und Methoden transmedialer Online-Diskursforschung* (p. 35-56). Köln: Herbert von Halem.
- Schneider, D., & Tschöpel, S. (2010). A lightweight keyword and tag-cloud retrieval algorithm for automatic speech recognition transcripts. *11th Annual Conference of the International Speech Communication Association*, Imakuhari, Japan, September 2010, 1277-1280. Abgerufen von <http://publica.fraunhofer.de/dokumente/N-198426.html>
- Selting, M., et al. (2009): Gesprächsanalytisches Transkriptionssystem 2 (GAT 2). *Gesprächsforschung – Online-Zeitschrift zur verbalen Interaktion*, 10, 353-402.
- Van Eimeren, B., & Frees, B. (2012). Ergebnisse der ARD/ZDF-Onlinestudie 2012. *Media Perspektiven*, (7-8), 362-379.
- 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
- Welker, M., & Wünsch, C. (2010). *Die Online-Inhaltsanalyse. Forschungsobjekt Internet*. Köln: Herbert von Halem.
- Wirth, W. (2001). Zum Stellenwert der Inhaltsanalyse in der kommunikations- und medienwissenschaftlichen Methodenausbildung. In W. Wirth & E. Lauf (Hrsg.), *Inhaltsanalyse. Perspektiven, Probleme, Potentiale* (S. 353-361). Köln: Herbert von Halem.

Zeller, F., & Wolling, J. (2010). Struktur- und Qualitätsanalyse publizistischer Onlineangebote: Überlegungen zur Konzeption der Online-Inhaltsanalyse. *Media Perspektiven*, (3), 143-153.

Acknowledgments

This work is partially supported by the Collaborative Project FI-CONTENT 2 (www.mediafi.org) funded by the European Commission through the 7th Framework Programme (FP7-603662, Future Internet Public Private Partnership FI-PPP).

Herausgeber/innen und Reihenherausgeber/innen

Mag. (FH) Julian Ausserhofer ist wissenschaftlicher Mitarbeiter am Institut für Journalismus und PR der Fachhochschule Joanneum in Graz und promoviert am Institut für Publizistik- und Kommunikationswissenschaft der Universität Wien. Er forscht zu digitalen Methoden in den Sozialwissenschaften und im Journalismus sowie zu den Dynamiken sozialer Medien in Nachrichten und Politik. Besonders interessieren ihn die Interaktionen zwischen Nutzern, Datenbanken, Algorithmen und anderen Akteuren in der Entstehung vernetzter Öffentlichkeiten. In den vergangenen Jahren war er jeweils für mehrere Monate zu Gast am Berliner Humboldt Institut für Internet und Gesellschaft sowie am Center for Science, Technology, Medicine and Society der Universität von Kalifornien, Berkeley.

Prof. Dr. Martin Emmer ist Professor für Publizistik- und Kommunikationswissenschaft mit dem Schwerpunkt Mediennutzung an der Freien Universität Berlin. Er hat in München und Berlin studiert und an der Technischen Universität Ilmenau 2004 mit einer Arbeit zur politischen Mobilisierung durch das Internet promoviert. Seine Forschungsschwerpunkte sind Mediennutzung, computervermittelte Kommunikation, politische Kommunikation und empirische Methoden.

Dr. Axel Maireder ist wissenschaftlicher Leiter des GfK Global Social Media Intelligence Center in Wien, in dem alle Web- und Social-Media-Forschungsprojekte des GfK Konzerns abgewickelt werden. Zuvor war er bis Ende 2014 wissenschaftlicher Mitarbeiter am Institut für Publizistik und Kommunikationswissenschaft an der Universität Wien und dort zuletzt Leiter eines Forschungsprojektes zur Strukturanalyse öffentlicher Kommunikation im Internet. Maireder unterrichtet an der FH Joanneum Graz und der Universität Wien.

Dr. Christina Schumann ist wissenschaftliche Mitarbeiterin am Fachgebiet „Empirische Medienforschung und politische Kommunikation“ am Institut für Medien und Kommunikationswissenschaft der Technischen Universität Ilmenau. Ihre For-

schungsschwerpunkte liegen im Bereich der Computervermittelten Kommunikation und der Rezeptions- und Wirkungsforschung. Insbesondere befasst sie sich mit Qualitätsforschung zu Computerspielen und Themenverdrossenheit als Vermeidungsphänomen der Nachrichtenrezeption. Seit November 2014 ist sie stellvertretende Sprecherin der Fachgruppe „Computervermittelte Kommunikation“ der Deutschen Gesellschaft für Publizistik und Kommunikationswissenschaft.

Dr. Monika Taddicken leitet die Abteilung für Kommunikations- und Medienwissenschaften an der Technischen Universität Braunschweig. Sie hat an der Georg-August-Universität Göttingen studiert und an der Universität Hohenheim zum Thema Methodeneffekte bei Web-Befragungen promoviert. Ihre Forschungsgebiete sind Onlinekommunikation, Medienwirkungen und Mediennutzung, Methoden und Wissenschaftskommunikation. Seit November 2012 ist sie Sprecherin der Fachgruppe „Computervermittelte Kommunikation“ der Deutschen Gesellschaft für Publizistik und Kommunikationswissenschaft.

Prof. Dr. Martin Welker ist Professor für Journalismus und Unternehmenskommunikation an der Hochschule für Medien, Kommunikation und Wirtschaft (HMKW) in Köln. Er hat an der Universität Mannheim Politikwissenschaft, Anglistik, Philosophie und Volkswirtschaftslehre studiert und dort im Jahr 2001 promoviert. Dann Vertretungsprofessor für Journalistik an der Universität Leipzig und Habilitation, anschließend Vertretungsprofessor an der Technischen Universität Braunschweig. Welker ist Herausgeber der „Neuen Schriften für Online-forschung“ im Herbert von Halem Verlag.

Prof. Dr. Jens Wolling ist Professor für Empirische Medienforschung und politische Kommunikation an der Technischen Universität Ilmenau. Er hat an der Freien Universität Berlin studiert und an der Technischen Universität Dresden zum Thema „Politikverdrossenheit durch Massenmedien?“ promoviert. Von 2003 bis 2006 war er Professor für Online-Forschung an der Ludwig-Maximilians-Universität München. Seine Forschungsgebiete sind Medienwirkungen und Mediennutzung, politische Kommunikation, Online-Kommunikation sowie Energie- und Umweltkommunikation.

