

Digitale Öffentlichkeit

Wie algorithmische Prozesse
den gesellschaftlichen Diskurs beeinflussen
- Arbeitspapier -

Digitale Öffentlichkeit

Wie algorithmische Prozesse den gesellschaftlichen Diskurs beeinflussen - Arbeitspapier -

Konrad Lischka
Prof. Dr. Christian Stöcker
im Auftrag der Bertelsmann Stiftung

Impressum

© Juni 2017

Bertelsmann Stiftung

Carl-Bertelsmann-Straße 256

33311 Gütersloh

www.bertelsmann-stiftung.de

Verantwortlich

Konrad Lischka, Ralph Müller-Eiselt

Autoren

Konrad Lischka, Prof. Dr. Christian Stöcker

Lizenz

Dieses Arbeitspapier ist unter der Creative-Commons-Lizenz [CC BY-SA 3.0 DE](https://creativecommons.org/licenses/by-sa/3.0/de/) (Namensnennung – Weitergabe unter gleichen Bedingungen) lizenziert. Sie dürfen das Material vervielfältigen und weiterverbreiten, solange Sie angemessene Urheber- und Rechteangaben machen. Sie müssen angeben, ob Änderungen vorgenommen wurden. Wenn Sie das Material verändern, dürfen Sie Ihre Beiträge nur unter derselben Lizenz wie das Original verbreiten.

Titelbild: [Konrad Lischka](#) lizenziert unter [CC BY-SA 2.0](https://creativecommons.org/licenses/by-sa/2.0/)

DOI 10.11586/2017028 <https://doi.org/10.11586/2017028>

Inhalt

1	Vorwort	6
2	Einleitung	8
3	Worum es geht: Algorithmische Prozesse, die Öffentlichkeit beeinflussen	10
3.1	Intro: Captain Kirk und die Impfgegner.....	10
3.2	Für die Meinungsbildung besonders relevante Intermediäre	12
3.3	Begriffliche Grundlagen	15
4	Was sich ändert: Strukturwandel der Öffentlichkeit	18
4.1	Zentrale Aspekte des Strukturwandels.....	18
4.2	Kriterien, nach denen Redaktionen Relevanz bemessen	19
4.3	Kriterien, nach denen Intermediäre Aufmerksamkeit verteilen	20
4.4	Zwischenfazit: algorithmische Prozesse als Relevanzdeuter	25
5	Wozu das führt: aktuelle Ausprägung des Strukturwandels der Öffentlichkeit	28
5.1	Nur was man gerade sieht, zählt.....	28
5.2	Emotion bringt Reichweite.....	29
5.3	Verzerrte Kognitionen formen die Datenbasis sozialer Medien	31
5.4	Plattformdesign beeinflusst menschliches Verhalten	32
5.5	Gemessene Reichweite muss nicht gleich realer Reichweite sein	36
5.6	Polarisierung wächst, doch da wirkt viel mehr als nur das Internet	37
5.7	Zwischenfazit: Komplexe Wechselwirkung, verstärkt durch algorithmische Prozesse	40
6	Was sein soll: Leitwerte für Öffentlichkeit	43
6.1	Freiheit.....	44
6.2	Vielfalt	45
6.3	Wahrheit	45

6.4	Integration.....	46
6.5	Zwischenfazit: Anwendbarkeit der Leitwerte auf algorithmisch strukturierter Öffentlichkeit.....	47
7	Was man tun kann: Ansatzpunkte für Interventionen	51
7.1	Prozesse algorithmischer Entscheidungsfindung für Teilhabe gestalten.....	52
7.2	Systematischen Verzerrungen der menschlichen Wahrnehmung begegnen	57
8	Fazit.....	61
9	Literatur	66
10	Executive Summary (English)	81
11	Über die Autoren.....	85

1 Vorwort

Eine Nachricht macht die Runde auf Facebook. Der Friedensnobelpreisträger steht fest, die neuesten Wahlumfragen wurden veröffentlicht, eine Ministerin tritt zurück, etwas in dieser Art. Je mehr Menschen eine dieser Meldungen aufrufen oder mit einem Klick auf den „Gefällt mir“-Button loben, desto relevanter muss sie für ein breiteres Publikum sein. Und je mehr Menschen bei dieser Meldungen in ihrem Newsfeed auf den Schalter „Beitrag verbergen“ tippen, um weniger Meldungen wie diese zu sehen, desto weniger relevant ist das Thema wohl auch für die breite Masse. Oder?

So würde wahrscheinlicher jeder mit dem sogenannten *guten Menschenverstand* die Logik sozialer Netzwerke interpretieren: Je mehr positive Reaktionen ein Beitrag erhält, desto weiter oben sollte er im Nachrichtenstrom möglichst vieler Menschen stehen. Und umgekehrt: Je mehr negative Reaktionen kommen, desto schlechter sollte ein Thema platziert werden. So ähnlich hat Facebook die Reaktionen der Nutzer auch als Zeichen für Relevanz interpretiert. Bis jemand 2015 genauer hinsah und feststellte: Fünf Prozent der Facebook-Nutzer waren für 85 Prozent der über die Funktion „Beitrag verbergen“ ausgeblendeten Meldungen verantwortlich. Diese *Super Hider* getauften Nutzer waren ein Rätsel. Sie verstreckten fast alles, was in ihrem Nachrichtenstrom auftauchte. Sogar Beiträge, die sie kurz zuvor kommentiert hatten. Waren diese Menschen wirklich so unzufrieden mit den algorithmisch kuratierten Inhalten, die Ihnen Facebook präsentierte?

Unzufrieden waren die Super Hider ganz und gar nicht. Das ergaben qualitative Befragungen. Super Hider nutzen die Verbergen-Funktion nicht wie vorgesehen, um ihre Unzufriedenheit mit der Nachrichtenauswahl auszudrücken. Sie wollten gelesene Meldungen schlicht wegräumen, so wie einige Menschen ihren Posteingang durch fortwährendes Löschen von E-Mails sauber halten. Nach dieser Erkenntnis änderte Facebook sein Verfahren zur Einschätzung der Relevanz von Beiträgen und interpretierte das Verbergen von Beiträgen nun nicht mehr notwendigerweise als starkes Signal des Missfallens.

Diese Geschichte zeigt beispielhaft drei Entwicklungen, die das vorliegende Arbeitspapier untersucht:

- Algorithmische Prozesse beeinflussen den gesellschaftlichen Diskurs. Zum Beispiel, indem sie in sozialen Netzwerken und Suchmaschinen Mitteilungen priorisieren und so die Öffentlichkeit strukturieren.
- Diese algorithmischen Systeme interpretieren auch menschliche Reaktionen auf Beiträge als Kennzeichen der Relevanz dieser Inhalte.
- Die menschliche Wahrnehmung und Relevanzeinschätzung von Medieninhalten ist je nach Nutzungssituation kognitiven Verzerrungen unterworfen und deshalb schwierig auszuwerten.

Algorithmische Prozesse von sozialen Netzwerken und Suchmaschinen beeinflussen für mehr als die Hälfte aller Onliner in Deutschland, wie und welche Angebote redaktioneller Medien sie in der digitalen Sphäre wahrnehmen (vgl. Abschnitt [2.2](#)). Meinungsbildungsprozesse sind „ohne Intermediäre nicht mehr denkbar“, wie Schmidt et al. (2017: 98) bilanzieren.

Diese Entwicklung birgt Chancen und Risiken. Die Reaktionen und das Interesse des Publikums spielen dank algorithmischer Entscheidungsfindung eine größere Rolle. Das kann dazu führen, dass neue Themen und die Medienangebote neuer Akteure schneller Beachtung und Verbreitung finden. Hier beeinflussen Angebote wie soziale Netzwerke auch redaktionelle Medien, weil Redakteure sie als Recherchequelle und Anzeichen für das Publikumsinteresse nutzen. Ein Risiko dabei ist, dass fälschlicherweise Repräsentativität angenommen oder Popularität unhinterfragt mit Relevanz gleichgesetzt wird. Algorithmische Prozesse können zudem schon durch Fehler bei der Gestaltung Relevanz systematisch verzerrt einschätzen, wie das Beispiel der Super Hider veranschaulicht.

Dieses Arbeitspapier gibt einen Überblick bekannter Wechselwirkungen zwischen der Gestaltung von Intermediären wie Google oder Facebook und den psychologischen Mechanismen, die Wahrnehmung und

Verhalten der Nutzer bestimmen. Zunächst wird der Analysebereich mittels empirischer Daten zur Mediennutzung eingegrenzt. Und zwar auf für die Meinungsbildung besonders relevante soziale Netzwerke und Suchmaschinen (Abschnitt 2). Abschnitt 4 skizziert zentrale Aspekte des Strukturwandels der digitalen Öffentlichkeit: Flaschenhals ist nicht die Veröffentlichung, sondern die Verteilung von Aufmerksamkeit auf Veröffentlichtes. Darauf baut eine Analyse der Bedeutung solcher Prozesse für die öffentliche Wahrnehmung in Abschnitt 5 auf: Was ist anders bei Intermediären, die Prozesse algorithmischer Entscheidungsfindung nutzen? Abschnitt 6 untersucht, wie sich diese Folgen zu den Leitwerten für Öffentlichkeit verhalten, die bisher bei redaktionell kuratierten Medien gelten. Abschnitt 7 skizziert Ansatzpunkte für Interventionen, um die Auswirkungen des Strukturwandels der Öffentlichkeit in Einklang mit diesen Leitwerten zu bringen.

Dieses Arbeitspapier ist Teil einer Exploration zum Thema „Teilhabe in Zeiten von Algorithmen und Big Data“, in der sich die Bertelsmann Stiftung näher damit beschäftigt, wie sich Phänomene der digitalen Sphäre auf gesellschaftliche Teilhabe auswirken. Bislang erschienen ist eine Analyse internationaler Fallbeispiele von algorithmischer Prozessen aus Feldern wie Strafverfolgung und Bildung (Lischka und Klingel 2017) sowie ein Vorschlag für die Bestimmung des Wirkungspotenzials auf Teilhabe (Vieth Wagner 2017). Der gesellschaftliche Diskurs ist in unserem Verständnis teilhaberelevant. Denn Teilhabe im Sinne dieser Expertise umfasst die gleichberechtigte Einbeziehung von Individuen und Organisationen in politische Entscheidungs- und Willensbildung sowie die faire Partizipation aller an sozialer, kultureller und wirtschaftlicher Entwicklung. Und zur Willensbildung gehört medial vermittelte Öffentlichkeit. Wir widmen diesem Thema ein eigenes Arbeitspapier, weil die Wechselwirkungen komplex sind und diese algorithmischen Prozesse im Alltag vergleichsweise vieler Menschen in Deutschland eine Rolle spielen.

Die Analyse ist unser erster, tentativer Beitrag zu dieser Debatte. Wir veröffentlichen sie als Arbeitspapier unter einer freien Lizenz, um etwas zu einem sich schnell entwickelnden Feld zu leisten, worauf andere leicht aufbauen können.

Die Bauweise der Foren, in denen eine Gesellschaft diskutiert, formt auch ihre Debatten. Winston Churchill hat das im Oktober 1943 vor dem britischen Unterhaus einprägsam ausgedrückt, als er für eine Architektur plädierte, in der sich Regierungsfraktion und Opposition konfrontativ gegenüber sitzen: „We shape our buildings and afterwards our buildings shape us“ (Churchill 1943: 403). Heute prägen statt Parlamenten zunehmend Onlineplattformen und soziale Medien den gesellschaftlichen Diskurs. Deshalb gibt dieses Arbeitspapier einen Einblick in die Architektur der digitalen Sphäre.



Ralph Müller-Eiselt
Senior Expert
Taskforce Digitalisierung
Bertelsmann Stiftung



Konrad Lischka
Project Manager
Taskforce Digitalisierung
Bertelsmann Stiftung

2 Einleitung

Das Thema dieses Arbeitspapiers sind die Auswirkungen, die ADM-Systemen auf die künftige Konstitution von Öffentlichkeit haben könnten. [Kapitel 3 \(„Worum es geht“\)](#) zeigt, dass algorithmische Entscheidungsprozesse im Bezug auf Medieninhalte im Alltag einer großen Zahl von Menschen bereits heute eine wichtige Rolle spielen. Intermediäre wie Suchmaschinen und soziale Netzwerke liefern einem wachsenden Anteil der Bevölkerung heute als Vermittler Informationen, deren Überbringung noch bis vor wenigen Jahren traditionellen Medienanbietern vorbehalten gewesen wären. Über 57 Prozent der deutschen Internetnutzer informieren sich auch über Suchmaschinen und soziale Netzwerke, in den USA nutzen bereits 44 Prozent der Erwachsenen Facebook regelmäßig als Nachrichtenquelle.

[Kapitel 4 \(„Was sich ändert“\)](#) macht deutlich, dass sowohl Suchmaschinen als auch soziale Netzwerke bei der Sortierung und Zugänglichmachung von Inhalten teilweise andere Kriterien anlegen als etwa in Redaktionen arbeitende Journalisten das tun. Die Sortierung selbst wird von einer komplexen Wechselwirkung zwischen Nutzerverhalten und algorithmischen Systemen bestimmt: Nutzer interagieren mit einer bereits vom Algorithmus vorselektierten Auswahl von möglichen Inhalten, und zwar unter Umständen auf flüchtige, emotionale, jedenfalls nicht tiefer Verarbeitung basierender Weise. Dazu kommt, dass bestimmte Verhaltensweisen der Nutzer möglicherweise anderen Zwecken dienen als rein informierenden. So ist Verhalten in sozialen Netzwerken teilweise auf Identitätsmanagement ausgereicht. Nutzer reichen also Beiträge möglicherweise nicht primär deshalb weiter, weil sie den informierenden Inhalt gerne anderen Menschen zugänglich machen wollen, sondern weil der weitergereichte Inhalt zu ihrem Selbstbild passt oder sie damit die Zugehörigkeit zu einer Gruppe signalisieren wollen.

[Kapitel 5 \(„Wozu das führt“\)](#) präsentiert empirische Befunde und relevante Einflussfaktoren, die in der skizzierten Struktur und der gegenwärtigen Gestaltung der Intermediäre der Öffentlichkeit Wirkung entfalten. Die Faktoren befördern möglicherweise Polarisierung. Dabei geht es einerseits um psychologische Aspekte, die menschliche Wahrnehmung und menschliches Verhalten betreffen, etwa die Tendenz, Artikel auch dann weiterzureichen, wenn man sie selbst gar nicht gelesen hat, oder bestimmte kognitive Verzerrungen wie die Verfügbarkeitsheuristik. Andererseits spielen auch die Entscheidungen und Aktivitäten der Intermediäre selbst eine grundlegende Rolle, insbesondere was die Optimierung der eigenen Angebote auf maximale Interaktionswahrscheinlichkeit hin angeht. Die Plattformen der Intermediäre sind darauf ausgerichtet, einen bestimmten Kognitionsstil zu begünstigen, der schnelles aber unreflektiertes Verhalten wahrscheinlich macht. Die Kombination dieser Faktoren, verbunden mit einer Entmachtung traditioneller Gatekeeper und der ubiquitären Möglichkeit potentiell globalen Publizierens, kann zumindest bei manchen Menschen offenbar zu einer Verengung des medialen Weltbildes führen. Personen, die bereits extreme Ansichten vertreten, ziehen sich auf solchen Plattformen möglicherweise immer stärker in Gruppierungen Gleichgesinnter zurück, ignorieren Informationen, die ihre festgefügte Meinung in Frage stellen können beziehungsweise fühlen sich durch die aktive Abwehr solcher dissonanter Information sogar noch in ihren Vorstellungen bestärkt. Verstärkt wird diese Tendenz durch ADM-Systeme. Ein weiterer Faktor, dessen Einfluss sich derzeit zwar belegen aber kaum verlässlich quantifizieren lässt, ist die Manipulation von Relevanzsignalen durch verdeckte technische Eingriffe wie den Einsatz sogenannter Bots. Die Wechselwirkung von psychologischen Variablen, Design-Entscheidungen der Intermediäre und externer Verzerrung der von den Intermediären gemessenen Signale führt offenbar zumindest bei Teilen der Nutzerschaft zu einer verstärkten Polarisierung sowohl im Hinblick auf wahrgenommene Inhalte als auch im Hinblick auf gesellschaftliche und politische Einstellungen. Dieser Polarisierungsprozess ist augenscheinlich jedoch noch von einer Reihe weiterer, von ADM-Systemen zunächst unabhängigen Faktoren wie dem Lebensalter oder auch dem politischen System des jeweiligen Landes abhängig.

Die neue Öffentlichkeit, die durch die Vermittlung dieser Plattformen zustandekommt, folgt derzeit nicht den gleichen Kriterien wie die über das klassische Mediensystem vermittelte Öffentlichkeit, wie [Kapitel 6 \(„Was sein soll“\)](#) deutlich macht. Gesellschaftliche Leitwerte wie Achtung vor der Wahrheit oder Integration spielen weniger zentrale Rollen.

In [Kapitel 7 \(„Was man tun kann“\)](#) werden schließlich Lösungsansätze für die skizzierte Problemsituation beschrieben, die auf der Mikro-, der Meso- und der Makroebene ansetzen können. Einerseits erscheint es erfolgversprechend, auf diesen Ebenen direkt bei den Entwicklern und Betreibern von Plattformen anzusetzen, die ADM-Systeme einsetzen, etwa über eine Verbesserung der Beforschbarkeit solcher Systeme, Förderung externer Evaluation und die Etablierung von Leitwerten, die über die bislang bei den Intermediären vorherrschenden hinausgehen.

Auch mit Blick auf die Rolle der menschlichen Wahrnehmung in den skizzierten Prozessen sind Maßnahmen auf der Mikro-, Meso- und Makroebene denkbar. Zweifellos wünschenswert wäre es, die Gesellschaft in großer Breite zunächst für die Existenz von ADM-Systemen und deren Wirkungsweisen und möglichen Wechselwirkungen mit dem eigenen Verhalten zu sensibilisieren. Insbesondere gegen die Manipulation der öffentlichen Meinung durch gezielte Desinformationskampagnen könnte möglicherweise eine Immunisierung in Form von präziser Information über tatsächliche Sachverhalte sowie die Strategien und Methoden der Verbreiter von Desinformation zielführend sein. Auf der Makroebene könnte es auch hier sinnvoll sein, in unabhängige wissenschaftliche Arbeiten zu investieren, mit dem expliziten Ziel, der Optimierung auf kognitiv wenig anstrengende und damit für einen gesellschaftlichen Diskurs eher ungeeignete Kognitionsstile wissenschaftliche Studien entgegenzusetzen, die Benutzeroberflächen, Anreiz- und Belohnungssysteme für tiefe, gründliche und weniger fehleranfällige Kognitionsstile erforschen.

3 Worum es geht: Algorithmische Prozesse, die Öffentlichkeit beeinflussen

3.1 Intro: Captain Kirk und die Impfgegner

Der Mann, den die meisten nur als Captain Kirk kennen, hat im Frühjahr 2017 ein prägnantes Beispiel dafür geliefert, wie sich unsere Medien- und Kommunikationswelt im vorangegangenen Jahrzehnt verändert hat. Der Schauspieler William Shatner, der Kirk in der Fernsehserie *Star Trek* verkörperte und sich in dieser Rolle gern mit dem Bordarzt „Pille“ McCoy „kabelte“, geriet wieder einmal mit einem Mediziner in Streit. Diesmal aber im realen Leben und höchst öffentlich. Der Ablauf und das Ergebnis dieses öffentlichen Disputs wirft ein Schlaglicht auf eine ganze Reihe von Faktoren, die dafür sorgen, dass gesellschaftliche Teilhabe im Zeitalter der digitalen Kommunikation wohl neu ausgehandelt und geschützt werden muss.

Ausgangspunkt des Streits war eine Kurznachricht, die Shatner über den US-amerikanischen Kommunikationsdienst Twitter absetzte (Levinovitz 2017). Er warb darin für *Autism Speaks*, eine im Jahre 2005 in den USA gegründete Stiftung zur Enthinderungsselbsthilfe von Autisten für Autisten und deren Angehörige, die sich eigenem Bekunden zufolge für Menschen mit Autismus einsetzt, aber in der Vergangenheit mit Kritik und sogar Boykottaufrufen zu kämpfen hatte. Viele Menschen mit Autismus fühlen sich von der Organisation aus einer Reihe von Gründen nicht vertreten, zudem vertrat man bei Autism Speaks lange Zeit die wissenschaftlich unhaltbare Position, Autismus werde durch Impfungen verursacht.

Shatner hat auf Twitter 2,5 Millionen Follower, eine Reichweite also, die bis vor wenigen Jahren nur professionelle Medienorganisationen erreicht hätten. Was er dort sagt, hat Gewicht, und sei es nur, weil viele Menschen es lesen werden. Sein Autism-Speaks-Tweet bekam besonders viel Aufmerksamkeit – insbesondere von Personen, die der Organisation höchst kritisch gegenüberstehen.

Shatner, der sich schon seit einiger Zeit für Menschen mit Autismus einsetzt und keineswegs den Verschwörungstheorien der Impfgegner anhängt, wehrte sich nach Kräften gegen den Vorwurf, er unterstütze eine zweifelhafte Organisation. Diskussionen bei Twitter sind naturgemäß auf Wortmeldungen bis zu 140 Zeichen beschränkt, daher ist die Gefahr groß, dass es dort zu Missverständnissen, Verkürzungen und als Folge davon zu wütenden Debatten kommt. So auch in diesem Fall.

Als die Diskussion hitziger wurde, schaltete sich ein Mediziner ein, der sich mit den Verschwörungstheorien von Impfgegnern schon ausführlich beschäftigt hat. David Gorski, Onkologe an der Wayne State University in Detroit, betreibt unter anderem das Weblog *Science Based Medicine*, das sich für evidenzbasierte Medizin und gegen pseudowissenschaftliche Behauptungen einsetzt. Unter ideologisch motivierten Impfgegnern hat Gorski viele Feinde.

Gorski erklärte in einer ganzen Serie von Twitter-Nachrichten, gerichtet an Shatner, aber für jeden öffentlich sichtbar, detailliert und höflich, warum Autism Speaks eine umstrittene Organisation ist. Shatner reagierte so, wie viele das in dieser Situation tun würden: Er gab Gorskis Namen in das Suchfenster einer Suchmaschine ein.

Dann begann er das, was er dabei fand, unbesehen an seine 2,5 Millionen Twitter-Follower zu verteilen. Etwa einen Artikel über den Mediziner auf einer höchst dubiosen Webseite namens *TruthWiki*, betrieben von einem Impfgegner und Verschwörungstheoretiker. Gorski wird in dem Text als „bezahlter Lügner der Impfindustrie“ diffamiert, der eine „pseudowissenschaftliche Religion“ vertrete und zu „wahnsinnigen Wutausbrüchen“ neige. Ein weiterer ergoogelter Artikel über Gorski, den Shatner verbreitete, wirft dem Onkologen vor, er werbe für „medizinische Interventionen, die Krebs verursachen, weil er am Ende von Krebserkrankungen profitiert“. Weitere Links zu mutmaßlich justiziablen Falschbehauptungen über Gorski folgten.

Der Mediziner, der sich für die wissenschaftliche Methode und gegen Verschwörungstheorien, die gerade in diesem Bereich lebensgefährlich sein können, einsetzt, wurde einem Millionenpublikum als böartiger, raffgieriger Scharlatan vorgeführt.

Viele Nutzer¹ sprangen Gorski bei. Einer fragte Shatner, warum er nicht Gorskis Wikipedia-Eintrag, der die Faken korrekt darstellt, gelesen und verlinkt habe. Shatner antwortete, *TruthWiki* habe in den Google-Suchergebnissen weiter oben gestanden. Man könne „all das bei Google finden“, so der Schauspieler, als sei dies ein Merkmal für inhaltliche Qualität.

Die in Shatners Tweets verlinkten Seiten sind sämtlich Publikationsprojekte rechtsnationaler Verschwörungstheoretiker – mehrere wurden sogar von der gleichen Person gegründet. Die Kombination der Ablehnung evidenzbasierter medizinischer Verfahren bis hin zur Chemotherapie bei gleichzeitiger Begeisterung für rechtsradikale Thesen ist in gewissen Kreisen in den USA erstaunlich weit verbreitet.

Noch Tage später war Shatner damit beschäftigt, auf seinem Twitter-Account abfällige Bemerkungen über Gorski zu verbreiten, tatkräftig unterstützt und permanent bestärkt von aggressiven Impfgegnern. Dabei ist der Schauspieler selbst erklärtermaßen weder Impfgegner noch hängt er offenkundig Verschwörungstheorien an. Er fühlte sich wohl schlicht angegriffen und schlug zurück.

Der Fall zeigt exemplarisch eine ganze Reihe von Faktoren auf, die zu einem radikalen Wandel der Art und Weise geführt haben, wie Öffentlichkeit mittlerweile hergestellt wird.

- Auf den digitalen Kommunikationsplattformen der Gegenwart ist die traditionelle *Abgrenzung zwischen persönlicher und öffentlicher Kommunikation verschwunden*.
- Als Folge davon können einzelne Individuen, in diesem Fall William Shatner, plötzlich potenziell ein Millionenpublikum erreichen. In der Kommunikationswissenschaft wird in diesem Zusammenhang von der *Entmachtung der traditionellen Gatekeeper* – Journalisten, Pressestellen und so weiter – gesprochen.
- Dem Einzelnen stehen völlig neue, mächtige, aber keineswegs unfehlbare *Methoden der blitzschnellen Informationsgewinnung* zur Verfügung, im Beispiel die Suchmaschine Google.
- Diese Werkzeuge zur Informationsgewinnung folgen bestimmten *algorithmischen Regeln*.
- *Die Kriterien, nach denen die Algorithmen Relevanz bemessen*, decken sich vielfach *nicht mit den Qualitätskriterien*, die gute Journalisten oder Wissenschaftler anlegen würden.
- Die Algorithmen arbeiten vielfach *deskriptiv* – sie bilden etwa ab, auf welche Links in der Vergangenheit besonders häufig geklickt wurde. Die Ergebnisse werden von vielen Nutzern aber *als normativ wahrgenommen* („stand bei Google weiter oben“).
- Bei bestimmten, extrem umstrittenen Themen wie der falschen Hypothese, es bestehe ein Zusammenhang zwischen Autismus und Impfungen, können kleine, aber hoch motivierte Minderheiten durch ihr Nutzungsverhalten dafür sorgen, dass eine *massive Divergenz zwischen inhaltlicher Qualität und algorithmisch ermittelter „Relevanz“* entsteht.
- Die Art und Weise, wie über digitale Kanäle kommuniziert wird, kann zu *rascher Eskalation eigentlich trivialer Konflikte* führen, selbst dann, wenn es inhaltlich de facto gar keinen Dissens zwischen den Beteiligten gibt.

Nun hat sich das beschriebene Beispiel in den USA zugetragen. Einige der Faktoren und Mechanismen, die es möglich gemacht haben, spielen hierzulande bislang keine vergleichbar große Rolle. Twitter beispielsweise wird in Deutschland nur von gut zwei Prozent aller Internetnutzer regelmäßig zu Informationszwecken genutzt

¹ Aus Gründen der Einfachheit und besseren Lesbarkeit verwendet diese Publikation vorwiegend die männliche Sprachform. Es sind jedoch jeweils beide Geschlechter gemeint.

(Hasebrink, Schmidt und Merten 2016). Viele andere dagegen sind auch hierzulande schon daran beteiligt, eine neue Form der öffentlichen Debatte zu formen.

Dieses Arbeitspapier soll daher eine Reihe grundlegender Fragen hinsichtlich der gesellschaftlichen Relevanz von Prozessen algorithmischer Entscheidungsfindung (*algorithmic decision-making*, ADM) für den öffentlichen Diskurs klären.

Wie verändert sich Öffentlichkeit durch ADM-Prozesse? Wie soll, wie kann eine demokratische Öffentlichkeit im Zeitalter der Allgegenwart solcher Prozesse aussehen? Welche Ansatzpunkte für Interventionen sind erkennbar?

Zunächst aber soll in einem kurzen Überblick anhand der aktuell vorliegenden empirischen Daten geklärt werden, welche Relevanz Intermediäre und ADM-Prozesse bereits jetzt für Meinungsbildung und Öffentlichkeit haben.

3.2 Für die Meinungsbildung besonders relevante Intermediäre

Wenn man nach einem möglichst einprägsamen Bild davon sucht, wie die Digitalisierung unsere Öffentlichkeit in den vergangenen zehn bis fünfzehn Jahren verändert hat, lohnt ein Blick auf Fotos aus U-Bahnen irgendwo in der industrialisierten Welt. Rushhour-Bilder aus dem Jahr 2006 zeigen verlässlich einen komplizierten Tanz, um Kollisionen mit den Nachbarn zu vermeiden: Jeder zweite Fahrgast hatte damals eine Zeitung in der Hand. Die unhandlichen Papierstapel so zu halten, dass man dabei niemand anderem in die Quere kam, war eine täglich trainierte Kunst.

Fotos von 2016, zehn Jahre später, zeigen ein völlig anderes Bild, auch wenn die meisten Passagiere ihre Aufmerksamkeit noch immer auf Medieninhalte richten: Egal ob in der U-Bahn von New York, Berlin oder Tokio – die Menschen starren auf ihre Smartphones. Das Jahr 2007, das Jahr, in dem Apple sein erstes iPhone vorstellte, markiert eine Veränderung von menschengeschichtlicher Bedeutung: Zwar gab es auch schon zuvor mobile Geräte, mit denen man E-Mails abrufen und auf speziell angepasste Webseiten zugreifen konnte. Erst mit dem iPhone aber war das neue Standardformat für die mobile Interaktion mit dem neuen, digitalen Medienuniversum gefunden: Ein rechteckiges Objekt, in der Regel hochkant gehalten, dessen Vorderseite vollständig aus einem berührungsempfindlichen Bildschirm besteht, mit mindestens einer Kamera und einer Reihe weiterer Sensoren ausgestattet, die etwa die Orientierung im Raum und diverse andere Nutzungsweisen ermöglichen. Das Smartphone ist der vorläufige Endpunkt einer Entwicklung, die der verblüffend weitblickende britische Kulturwissenschaftler Raymond Williams schon 1974 unter dem Schlagwort *mobile privatization* vorausahnte, einem Begriff, der mit *mobile Privatisierung* nur unzureichend übersetzt ist (Williams 1974). Williams bezeichnete mit dem Begriff eine in seinen Augen zu erwartende zunehmende Personalisierung sozialen Erlebens, die er auf technische Entwicklungen zurückführte. Die Markteinführung des Sony Walkman im Jahr 1979 gab Williams zum ersten Mal recht: Plötzlich konnten einzelne Personen im öffentlichen Raum sich ein eigenes akustisches Universum schaffen, ihren Alltag mit einem individuellen Soundtrack versehen.

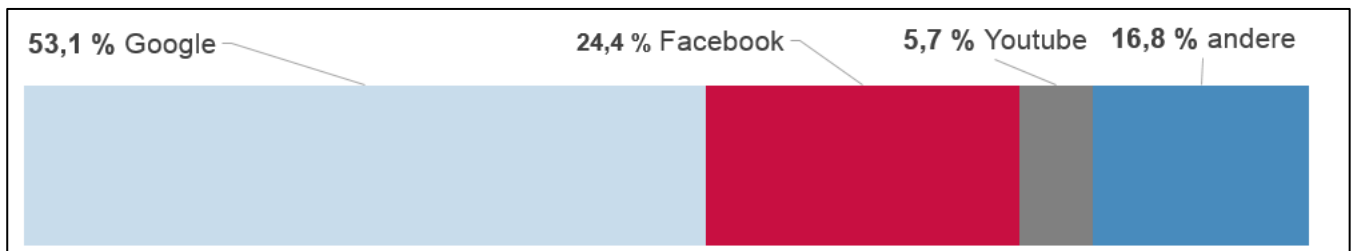
Mit dem Smartphone ist die mobile Totalindividualisierung der medialen und kommunikativen Umwelt nahezu jedes Einzelnen mittlerweile weit fortgeschritten. Zwischen den Zeitungslesern in den U-Bahnen von 2006 und den Smartphone-Nutzern in denen von 2016 besteht nämlich ein fundamentaler Unterschied: Bei den ersteren ist auf einen Blick ersichtlich, was sie gerade tun. Bei den letzteren ist es völlig unklar. Lesen Sie einen Artikel im Onlineangebot einer Zeitung oder Zeitschrift? Einen Fachartikel? Den Social Media Post eines Bekannten? Oder eines Prominenten wie William Shatner? Eine berufliche oder private E-Mail? Kaufen sie gerade ein? Suchen sie ein Hörbuch aus, ein Musikstück oder eine Playlist zur Beschallung der restlichen Bahnfahrt über Kopfhörer? Spielen sie ein Spiel? Diese Aufzählung ließe sich noch lange fortsetzen, und mit jedem weiteren Monat, der vergeht, kommen neue Nutzungsmöglichkeiten für die mobilen Touch-Computer hinzu. Personalisiert sind die Medien- und Kommunikationserfahrungen, die Smartphones ermöglichen, aber noch auf eine andere, schon aus dem stationären World Wide Web bekannte Weise: durch algorithmische Filterung.

In Deutschland hat sich für die Angebote, die solche Vermittlungs- und Filteraufgaben übernehmen, der Begriff *Intermediäre* durchgesetzt (Hasebrink, Schmidt und Merten 2016; Schmidt et al. 2017). Forscher des Hamburger Hans-Bredow-Instituts fassen darunter sowohl Suchmaschinen wie Google oder Bing als auch Videoportale wie Youtube, eher visuell orientierte Kommunikationsplattformen wie Snapchat und Instagram, aber auch herkömmliche Instant Messenger wie WhatsApp zusammen. Zahlen von TNS Infratest zufolge nutzen über 57 Prozent aller deutschen Internetnutzer regelmäßig solche Intermediäre, um sich zu informieren – also nicht nur, um sich mit Freunden auszutauschen oder unterhaltende Inhalte zu konsumieren. Suchmaschinen liegen dabei mit fast 40 Prozent Nutzung vorne, gefolgt von sozialen Netzwerken wie Facebook mit gut 30 Prozent und Videoportalen wie Youtube mit gut 9 Prozent. Instant Messenger nutzen demnach 8,5 Prozent in informierender Weise. Bei Letzteren spielen automatisierte Entscheidungsmechanismen derzeit in der Regel keine Rolle. Bei den Intermediären selbst gibt es klare Marktführer, was die informierende Nutzung betrifft: Google bei den Suchmaschinen (37,9 %) und Facebook bei den sozialen Netzwerken (24,1 %). Zum Vergleich: Microsofts Suchmaschine Bing wird nur von zwei Prozent der deutschen Onliner in informierender Weise genutzt, Twitter von 2,1 Prozent. Insgesamt kommen über 54 Prozent all jener, die sich im Internet informieren, regelmäßig über Suchmaschinen oder soziale Medien mit den Inhalten eher traditioneller Informationsanbieter, wie den Webseiten von Tageszeitungen, Zeitschriften oder TV-Sendern, in Kontakt.

Beschränkt man die Betrachtung nicht auf die informierende Nutzung, ist das Bild noch deutlich klarer: Über 95 Prozent aller deutschen Onliner nutzen täglich mindestens einen Intermediär (zu allen möglichen Zwecken), dabei liegt Google mit weitem Abstand (78,6 %) vorn, gefolgt von YouTube (42 %) und Facebook (41,8 %). Bei Instant-Messaging-Diensten, die für den Themenbereich dieses Arbeitspapier jedoch eine untergeordnete Rolle spielen, dominiert die Facebook-Tochter WhatsApp: Fast 75 Prozent aller deutschen Onliner nutzen den Dienst täglich.

Fragt man Nutzer, welche Angebote ihnen für ihre Informationssuche besonders wichtig sind, kommt man zu ähnlichen Ergebnissen (Ecke 2016), siehe auch Abbildung 1.

Abbildung 1: Die wichtigsten Intermediäre für Information aus Sicht der Nutzer



Angaben in Prozent; Tagesreichweite = Nutzung gestern. Basis: 23,252 Mio. Internetnutzer gestern (=Onliner) ab 14 Jahre in Deutschland, die gestern mindestens 1 Intermediär genutzt haben, n=669. Frage: „Sie haben gestern ja folgende Angebote als Quelle oder Kontakt zu Informationen zum Zeitgeschehen in Politik, Wirtschaft und Kultur genutzt. Welches dieser Angebote ist Ihnen da am wichtigsten?“.

Quelle: Kantar TNS. Berlin, in Ecke, 2016: 21).

Hasebrink und et al. (2016) kamen 2016 auf Basis dieser und anderer Daten zu dem Schluss, dass „Meinungsbildungsprozesse ohne Intermediäre nicht mehr denkbar“, diese gleichzeitig aber „nur ein Baustein im Prozess der Meinungsbildung“ seien. Auch das eigene soziale Umfeld „sowie die Berichterstattung publizistischer Medien, denen Vertrauen entgegengebracht wird“, seien „nach wie vor bedeutsam“ (a. a. O.).

Wie in nahezu jedem Bereich, den die Digitalisierung betrifft, gilt allerdings auch hier: Die genannten Zahlen bilden nur einen Zustand zu einem bestimmten Erhebungszeitpunkt im Jahr 2016 ab, Nutzungszahlen und -weisen sind in einem Zustand permanenter Veränderung begriffen. In den USA, in denen die Entwicklung in diesem Bereich stets ein wenig weiter fortgeschritten ist als hierzulande, nutzten beispielsweise schon Anfang 2016 44 Prozent aller Erwachsenen Facebook häufiger als „selten“ als Nachrichtenquelle (Gottfried und Shearer, 2016).

Dies- wie jenseits des Atlantiks gilt zudem: Je jünger die befragte Zielgruppe, desto höher der Anteil derer, die soziale Netzwerke, Videoplattformen und Suchmaschinen benutzen, um sich zu informieren – und desto geringer der Anteil derer, die dazu auf klassische Medien wie Tageszeitungen zurückgreifen (Mitchell et al. 2016).

Eine zweite wichtige Entwicklung ist die zunehmende Bedeutung sozialer Netzwerke als Einstiegspunkt und Kontaktvermittler für andere Inhaltenanbieter. Das US-Unternehmen Parse.ly, das Website-Betreibern Analysewerkzeuge anbietet und deshalb Zugriff auf die Traffic-Statistiken Tausender Anbieter hat, wertet jährlich auch aus, wie viele Besucher die erfassten Seiten woher bekommen. Ende 2012 lag Google als Traffic-Bringer für die erfassten Seiten mit einem Anteil von über 40 Prozent demnach noch klar vorne, Facebooks Anteil lag um die 10 Prozent. Mitte 2015 überholte der Anteil von Facebook erstmals den von Google, beide lagen nun knapp unter 40 Prozent.

Wer solche Intermediäre, ausgenommen Messenger, nutzt, um sich Informationen zu verschaffen, der begegnet algorithmischen Entscheidungen immer und grundsätzlich (Gillespie 2014y). Die in einem Facebook-Newsfeed angezeigten Beiträge, ob von Freunden des Nutzers oder aber von Medienanbietern, deren Facebook-Seiten der Betreffende folgt, werden nach bestimmten Kriterien automatisch sortiert. Bei Suchmaschinen gilt das Gleiche: Wer eine Google-Suche startet, bekommt eine maschinell aufgrund bestimmter Kriterien sortierte Ergebnisliste zurück. Erstaunlich ist, dass viele Nutzer das offenbar bis heute gar nicht wissen.

Als eine Forschergruppe aus Kalifornien 2015 Facebook-Nutzer befragte, wie die Reihenfolge der in ihrem eigenen Newsfeed angezeigten Beiträge wohl zustande komme, gingen über 62 Prozent der Befragten davon aus, dass ihnen alle Beiträge ihrer Freunde und aller Seiten, denen sie in dem Netzwerk folgen, angezeigt würden. Manche hatten zwar den Verdacht, dass da irgendetwas im Gange war, aber keine Vorstellung, was genau das sein könnte. Eine Teilnehmerin wird mit folgenden Worten zitiert: „Ich habe so 900 Freunde und glaube, ich sehe nur so 30 davon in meinem Newsfeed. Ich weiß also, dass da irgendetwas los ist, aber ich weiß nicht genau, was“ (Eslami et al. 2016).

Zusammenfassend lässt sich also sagen: Für einen wachsenden Teil der Bevölkerung in den Industrienationen sind Intermediäre, und hier vor allem Suchmaschinen und soziale Netzwerke, ein wesentlicher Faktor bei der Gewinnung und Rezeption teilhaberelevanter Information. Dies gilt umso mehr, je jünger die betrachtete Zielgruppe ist. Zwar spielen klassische Medienkanäle, insbesondere das Fernsehen, nach wie vor eine sehr wichtige Rolle. Der Trend der vergangenen Jahre zeigt jedoch klar, dass Informationsvermittlung durch Intermediäre und damit durch ADM-Prozesse eine immer wichtigere Rolle spielt. Gleichzeitig sind sich viele Nutzer keinesfalls der Tatsache bewusst, dass hier algorithmische Entscheidungsprozesse am Werk sind.

Einige Entwicklungen der jüngeren Zeit zeigen, dass auch die größten unter den neuen Aufmerksamkeitsverteilern, Facebook und Google, ihre wachsende Verantwortung durchaus reflektieren. Nach einer ausgiebigen Debatte über die Frage, welche Rolle über Facebook verbreitete Falschmeldungen (*Fake News*) möglicherweise bei der US-Präsidentschaftswahl 2016 gespielt haben könnten, äußerte sich Facebook-Chef Mark Zuckerberg, der entsprechende Einflussmöglichkeiten seines Unternehmens zuvor stets verneint hatte, so: „Wir wollen keine Falschmeldungen auf Facebook. Unser Ziel ist es, den Leuten die Inhalte zu zeigen, die ihnen am meisten bedeuten, und die Leute wollen korrekte Nachrichten. Wir haben bereits Arbeiten initiiert, die es unserer Community erlauben werden, Falschmeldungen und Fake News zu markieren, und wir können hier noch mehr tun“ (Zuckerberg 2016).

Am 31. Januar 2017 folgte dieser unspezifischen Ankündigung eine spezifische: Facebook wolle künftig „neue Signale berücksichtigen“, um „authentische Inhalte besser identifizieren und einstufen zu können“ (Lada, Li und Ding 2017). Bei Facebook ist der Gedanke, dass die eigenen Algorithmen zur Bildung der öffentlichen Meinung beitragen könnten, also durchaus angekommen, und dies schlägt sich auch in Form technischer Veränderungen nieder.

Bei Google gibt es eine entsprechende Qualitätskontrolle schon länger. Evaluatoren, sogenannte *Quality Rater*, tippen regelmäßig vorgegebene Suchanfragen in die Suchmaske des Unternehmens ein und beurteilen dann

nach einem Kriterienkatalog die Qualität der Ergebnisse. Das Handbuch für diese Evaluatoren umfasst auch einen Abschnitt über sogenannte *your money or your live (YMYL) pages* (Geld-oder-Leben-Seiten) (Google 2017). Diese spezielle Kategorie von Seiten umfasst dem Handbuch für Evaluatoren zufolge solche, die „sich auf die künftige Zufriedenheit, die Gesundheit oder finanzielle Stabilität der Nutzer auswirken könnten“. Neben Seiten mit Gesundheits- oder Finanzinformationen schließt diese Kategorie dem Handbuch zufolge explizit auch solche mit Nachrichteninhalten ein: „Webseiten, die wichtig sind, um eine informierte Bürgerschaft zu erhalten, einschließlich Informationen über Regierungsprozesse, Personen und Gesetze auf der lokalen, bundesstaatlichen oder der nationalen Ebene“, sowie „Nachrichten über wichtige Themen wie internationale Ereignisse, Wirtschaft, Politik, Wissenschaft und Technologie“. Die Evaluatoren werden aufgefordert, „Ihre Urteilskraft und Ihr Wissen über ihre Region“ zum Einsatz zu bringen (a. a. O.).

Die großen Intermediäre richten ihr Verhalten also bereits jetzt nach der Annahme aus, dass die Ergebnisse ihrer Systeme und Prozesse Einfluss auf die Information der Öffentlichkeit haben. Bevor wir uns der Frage zuwenden, wie die empirische Datenbasis für diese Annahme derzeit aussieht, wollen wir zunächst Klarheit über die hier verwendeten Begriffe und Konzepte schaffen.

3.3 Begriffliche Grundlagen

Dieses Arbeitspapier untersucht Angebote in der digitalen Sphäre, die sich auf die Entstehung von Öffentlichkeit und den gesellschaftlichen Diskurs auswirken. Gegenstand sind Angebote, die algorithmische Prozesse einsetzen. Für die Meinungsbildung besonders relevante Beispiele solcher Angebote kann man anhand des Nutzerverhaltens klassifizieren ([siehe Abschnitt 3.2](#)). Deutsche Internetnutzer nennen Google, Facebook, Youtube, Whatsapp und Twitter in dieser Reihenfolge am häufigsten als Quellen von oder Kontaktvermittler zu Informationen über das Zeitgeschehen in Politik, Wirtschaft und Kultur (Ecke 2016: 21). Deshalb betrachten wir diese Angebote und leiten übergreifende Gestaltungsprinzipien ab, die für unsere Analyse relevant sind.

3.3.1 Intermediäre und ihre Gestaltungsprinzipien

In der deutschen Debatte hat sich für die genannten Angebote der Oberbegriff *Intermediäre* etabliert. Bisweilen ist auch von *Informationsintermediären* die Rede, um nur Angebote von Bedeutung für den gesellschaftlichen Diskurs hervorzuheben. Die Begriff haben Schulz und Dankert (2016) geprägt, ohne dass „dieser Begriff klare Konturen oder eine theoretische Grundlage hätte“ (a. a. O.: 13). Für unsere tentative Analyse ist dieser Begriff ausreichend. Gemeint sind „digitale Dienste (...), die eine Vermittlungsrolle zwischen Nutzerinnen und Nutzern und Inhaltsangeboten vornehmen“ (a. a. O.: 19). Vermittlung als Gestaltungsprinzip ist den fünf aus Nutzersicht relevantesten Informationsintermediären in Deutschland gemeinsam, wie Tabelle 1 zeigt.

Tabelle 1: Gestaltungsprinzipien für die Meinungsbildung besonders relevanter Intermediäre

Merkmal	Google (Suche)	Facebook	YouTube	Whatsapp	Twitter
Vermittlung (zwischen Publikum und Angeboten Dritter; Funktionen z. B.: auffindbar machen, aggregieren)	ja	ja	ja	ja	ja
Auslieferung (komplette Inhalte Dritter werden über die Infrastruktur des Intermediärs ausgeliefert)	nein (Auszüge)	ja (teilweise)	ja	ja	ja (teilweise)
Strukturierung (Inhalte Dritter werden nach Relevanzeinschätzung des Intermediärs sortiert)	ja	ja	ja	nein	ja (optional)
Algorithmische Entscheidungsfindung (bei Relevanzeinschätzung und Auswahl angezeigter Inhalte)	ja	ja	ja	nein	ja (optional)

Soziales Netzwerk (Nutzer definieren untereinander bewusst Identität und Beziehungen)	nein	ja	ja	ja	Ja
--	------	----	----	----	----

Quelle: Eigene Darstellung.

Die anderen Gestaltungsprinzipien sind bei den fünf Angeboten unterschiedlich ausgeprägt. Es folgt eine kurze Diskussion der Unterschiede und dann als Ableitung eine positive Definition von Intermediären im Sinne dieser Expertise.

Auslieferung: Die meisten Anbieter setzen beide Modelle um: Sie verweisen einerseits lediglich auf andernorts veröffentlichte Inhalte Dritter, andererseits stellen sie selbst Dritten Infrastruktur zum Veröffentlichen und Verbreiten bereit. So erlaubt Facebook Verweise, aber auch die Veröffentlichung von Texten, Videos, Fotos über die Facebook-Infrastruktur, ebenso Twitter (allerdings nur für bestimmte Gattungen wie Fotos und Videos). Es existieren auch komplexe Mischformen wie Googles *Accelerated Mobile Pages*. In der deutschen Diskussion werden Angebote mit eigener Infrastruktur zum Speichern und Ausliefern oft als Plattformen bezeichnet (z. B. Schmidt et al. 2017: 9). Letztlich ist dieses Kriterium für diese Expertise zur Abgrenzung nicht entscheidend: Die Vermittlungsleistung wenden alle Anbieter auf alle Inhalte an – ganz gleich, auf welcher Infrastruktur diese bereitgestellt sind.

Strukturierung von Öffentlichkeit und der Einsatz von *Prozessen algorithmischer Entscheidungsfindung* sind ein gemeinsames Merkmal aller Angebote mit Ausnahme des Instant-Messaging-Dienstes Whatsapp. Whatsapp folgt einer Strukturierungslogik, die Facebook und Twitter vor Jahren aufgegeben haben: Wenn auf Whatsapp Menschen mit anderen einen Kontakt geknüpft haben, sehen sie alle Nachrichten dieser Person. In diesem Punkt ist Whatsapp ein Sonderfall. Das Angebot ähnelt eher für Dienste, die rein für interpersonale Kommunikation konstruiert sind, wie zum Beispiel E-Mail. In dieser gegenwärtigen Erscheinungsform sind Angebote wie Whatsapp als Informationsintermediäre für die Fragestellung dieser Expertise nicht relevant. Sie vermitteln zwischen Publikum und Dritten, strukturieren dabei aber Öffentlichkeit allein aufgrund bewusster Selektionssignale der Nutzer (Kontakt/kein Kontakt). Das unterscheidet sich kaum vom Gestaltungsprinzip einer Mailingsliste oder eines Forums. Wir fokussieren hingegen Anbieter, die algorithmische Entscheidungsfindung zur Strukturierung nutzen, das heißt zur „Ent- und Neubündelung von Informationen“ (a. a. O.: 20). Personalisierung ist in den allermeisten Fällen eine Erscheinungsform dieser Neubündelung. Allerdings sehen wir Personalisierung anders als etwa Schmidt (a. a. O.: 20) nicht als zwingendes Organisationsprinzip informationeller Intermediäre. Eine algorithmisch basierend auf Nutzerverhalten erstellte Liste der 20 wichtigsten Themen eines Tages ist relevant, auch wenn sie nicht personalisiert ist. Wenn alle Nutzer eines Angebots in einem Sprachraum auf der Startseite diese Themenauswahl sehen, beeinflusst auch das die öffentliche Meinungsbildung.

Soziales Netzwerk: Dieses Gestaltungsprinzip nutzen wir bei der Definition von Intermediären im Sinne dieser Expertise nicht. Die Mehrzahl der Angebote animiert Nutzer zwar zum Anlegen eines individuellen Profils und zum Knüpfen eines algorithmisch auswertbaren Beziehungsnetzes. Diese Signale sind hilfreich, aber keine Voraussetzung für die Strukturierung von Öffentlichkeit mittels algorithmischer Entscheidungsfindung. Das zeigen Angebote wie Twitter und die Google-Suche: Sie bieten auch nicht angemeldeten Nutzern eine strukturierte Auswahl, basierend auf einer Reihe verfügbarer Signale (Sprache, Standort usw.) und der Auswertung des Verhaltens anderer Nutzer.

Definition: Intermediäre der algorithmisch strukturierten Öffentlichkeit

Drei Merkmale zeichnen Intermediäre im Sinne dieser Expertise aus (in Anlehnung an Perset 2010: 9):

(1) Sie vermitteln zwischen Dritten, aus deren Interaktion Öffentlichkeit entsteht. Dabei handelt es sich unter anderem um Privatpersonen, journalistisch-redaktionelle Medien, Unternehmen, Politik und Verwaltung.

(2) Sie verbreiten und/oder erschließen von Dritten erstellte Inhalte. Dabei fügen Intermediäre diese Inhalte neu zusammen, basierend auf eigenen Prinzipien zum Einschätzen der Relevanz. Sie bestimmen die Bedingungen des Zugangs und die Mechanismen des Matchings.

(3) Für die Relevanzeinschätzung und Auswahl angezeigter Inhalte nutzen sie Prozesse algorithmischer Entscheidungsfindung.

3.3.2 Prozesse algorithmischer Entscheidungsfindung (ADM-Prozesse)

Wir verwenden den Begriff *Prozess algorithmischer Entscheidungsfindung* wie von Zweig (2017) für automatisierte Verfahren definiert, die

- durch einen Algorithmus eine Bewertung einer Situation oder eines Menschen vornehmen oder eine Vorhersage über die Wahrscheinlichkeit des Eintretens einer Situation treffen,
- darauf mithilfe der (auch indirekten) Aktivierung eines Elements reagieren, dessen Aktion unmittelbar das Leben eines Menschen betrifft.

Als Abkürzung für Prozess algorithmischer Entscheidungsfindung nutzen wir ADM-Prozess nach dem im Englischen üblichen Begriff *algorithmic decision-making*². Intermediäre, die Öffentlichkeit algorithmisch strukturieren, sagen die potenzielle Nachfrage nach einem Inhalt vorher und passen die Darbietung entsprechend der Prognose an (Napoli 2014b: 34). Diese Auswahl bestimmt, mit welchen Aspekten der Öffentlichkeit ein Mensch in einem bestimmten Angebot wie zum Beispiel Google oder Facebook konfrontiert wird.

De facto kommen die Auswahlentscheidungen auf noch ungleich komplexere Art und Weise zustande. In der Regel greifen mehrere Programme, die algorithmische Entscheidungsregeln abbilden, ineinander und beziehen zudem bestimmte variable Parameter, etwa sogenannte *Blacklists*, in die Entscheidungsfindung mit ein. Ein Beispiel wäre das sogenannte *BPjM-Modul* (BPjM o. J.), das die Bundesprüfstelle für jugendgefährdende Medien (BPjM) an Intermediäre wie Google weiterreicht, wo es in die Erstellung von Ergebnislisten einfließt, um zu verhindern, dass bestimmte jugendgefährdende oder indizierte Inhalte dort überhaupt auftauchen.

Für die vorliegende Expertise ist vor allem der algorithmische Anteil der Entscheidungsprozesse von Belang.

3.3.3 Teilhabe

Teilhabe im Sinne dieser Expertise umfasst die gleichberechtigte Einbeziehung von Individuen und Organisationen in politische Entscheidungs- und Willensbildung sowie die faire Partizipation aller an sozialer, kultureller und wirtschaftlicher Entwicklung. Es geht also erstens um Teilhabe an demokratischen Prozessen – und damit um politische Gleichberechtigung – und zweitens um Teilhabe an Errungenschaften eines sozialen Gemeinwesens, „angefangen von guten Lebens- und Wohnverhältnissen, Sozial- und Gesundheitsschutz, ausreichenden und allgemein zugänglichen Bildungschancen und der Integration in den Arbeitsmarkt bis hin zu vielfältigen Freizeit- und Selbstverwirklichungsmöglichkeiten“ (Beirat Integration 2013).

Teilhabe in diesem Sinne hat zur Voraussetzung, dass die materiellen Mittel aller über einem Mindestniveau liegen, das allen Menschen die Mitwirkung am gesellschaftlichen Leben möglich macht. Die Gewährleistung von sozialer und politischer Teilhabe setzt also eine „Sockelgleichheit der sozialen Grundgüter“ (Meyer 2016) voraus. Elemente dieses Sockels werden zum Beispiel in der Allgemeinen Erklärung der Menschenrechte und im Internationalen Pakt über wirtschaftliche, soziale und kulturelle Rechte beschrieben (Bundesgesetzblatt 1966). Um chancengerechte Teilhabe in diesem Sinne zu ermöglichen, sind gezielte Investition in die Entwicklung individueller Fähigkeiten nötig (Bertelsmann Stiftung 2011: 31). Es liegt in der Verantwortung des Staates und des Gemeinwesens, kontinuierlich jedes Individuum zu befähigen, Chancen tatsächlich zu nutzen.

² Verwendet zum Beispiel in: USACM 2017); (Ananny und Crawford 2016); Goodman und Flaxman 2016), Mittelstadt 2016a: Zarsky 2016.

4 Was sich ändert: Strukturwandel der Öffentlichkeit

4.1 Zentrale Aspekte des Strukturwandels

Die Mehrheit der in Deutschland für die Meinungsbildung besonders relevanten Intermediäre nutzt algorithmische Entscheidungsfindung zur personalisierten Auswahl und Strukturierung. Facebook personalisiert den News Feed seit September 2006 (Facebook 2006), Google die Suchergebnisse auch für ausgeloggte Nutzer weltweit seit 2009 (Horling und Bryant 2009), Twitter seit Anfang 2015 (Rosania 2015), wenn die Nutzer dem nicht explizit widersprechen (Twitter 2017). Die Gestaltungsprinzipien (siehe [Abschnitt 3.3.1](#)) dieser Intermediäre führen zu einem Strukturwandel der Öffentlichkeit. Wir fassen im Folgenden Chancen und Risiken von sechs für diese Expertise zentralen Aspekten der Entwicklung zusammen.

Tabelle 2: Zentrale Aspekte des Strukturwandels der Öffentlichkeit

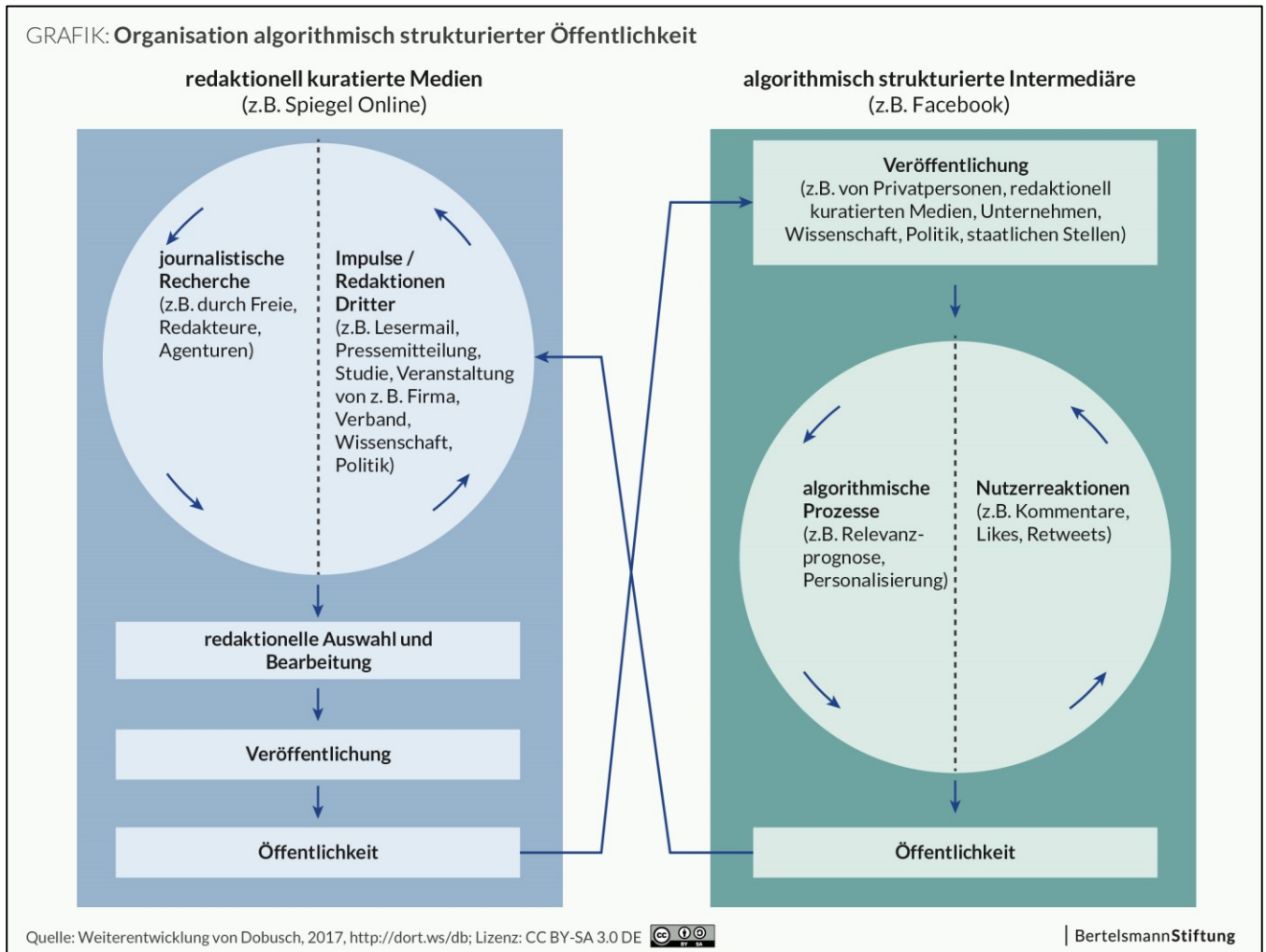
Aspekt	Beschreibung	Chancen	Risiken
Entkopplung von Veröffentlichung und Reichweite	Jeder kann veröffentlichen. Aber nicht jeder findet ein Publikum. Aufmerksamkeit entsteht erst durch das Zusammenwirken von Menschen und ADM-Prozessen (Dobusch 2017).	Größere, leichter zugängliche Medienvielfalt, weil Gatekeeper nicht über Veröffentlichung bestimmen.	Vielfaltsverengung bei Filterlogik durch Dominanz weniger Intermediäre und ADM-Prozesse.
Entbündelung von Publikationen	Reichweite wird auf Beitragsebene ausgehandelt, Platzierung innerhalb einer Publikation zieht Beiträge bei Intermediären weniger mit.	Größere Medienvielfalt. Chancen auf Aufmerksamkeit breiter gestreut. Medien können diverseres Publikum segmentieren.	Schwächung von Publikationen als Qualitätsmarker. Geringere Chancen für unpopuläre, aber relevante Themen.
Personalisierung	Nutzer erfahren mehr über ihre Interessengebiete.	Vertiefte Information, stärkere Mobilisierung.	Filterblasen Echokammern
Größerer Einfluss des Publikums auf Reichweiten	Nutzerreaktionen beeinflussen ADM-Prozess insgesamt und Reichweite jedes Beitrags.	Demokratisierung durch größeren Publikumseinfluss. Größere Dynamik der Agenda.	Verengung der Aufmerksamkeit auf in ADM-Prozessen chancenreiche Themen und Tonalitäten.
Zentralisierung der Auswahlinstanzen	Es gibt weniger Intermediäre als redaktionell kuratierte Medien.	Leichter zugängliche Medienvielfalt.	Vielfaltsverengung bei Filterlogik.
Wechselwirkung zwischen redaktioneller und maschineller Kuratierung	Redaktionell kuratierte Medien nutzen Reaktionen in Intermediären als Signal für Publikumsinteresse.	Größere Chancen auf Verbreitung für neue, nicht etablierte Themen und Akteure in redaktionell kuratierten Medien.	Irrige Annahmen über Repräsentativität der algorithmisch priorisierten Aussagen und Akteure. Reinwaschen durch Platzierung in redaktionellen Medien.

Ohne den oben skizzierten Strukturwandel der Öffentlichkeit sind Phänomene wie der Aufstieg und Fall der Piratenpartei schwer vorstellbar. Das birgt auch für die Organisation von Bewegungen um Themen herum ein neues Mobilisierungspotenzial, wie Dobusch schreibt:

„Dank Internet und digitalen Technologien werden heute Stimmen gehört und gesehen oder entstehen überhaupt erst, die vorher nicht öffentlich präsent waren. (...) Nie waren unsere Gesellschaften derart offen für Veröffentlichungen. Aber es ist nicht nur die Offenheit für Veröffentlichungen aller Art; es ist auch die algorithmisch vermittelte Verbreitung und Verstärkung über Plattformen wie Twitter, YouTube und Facebook, die diese Offenheit noch einmal verstärkt“ (Dobusch 2017).

Der Vergleich der beiden Öffentlichkeitsverläufe unten zeigt die neue, zentrale Rolle von Nutzerreaktionen und algorithmischen Prozessen. Beide bestimmen die Verteilung von Aufmerksamkeit über Intermediäre. Unsere Hypothese ist, dass Nutzerreaktionen und ADM-Prozesse sich dabei nicht eindeutig in eine lineare Kausalkette bringen lassen. Wie die Wechselwirkungen im Detail aussehen, analysiert der folgende Abschnitt.

Abbildung 2: Organisation algorithmisch strukturierter Öffentlichkeit (Weiterentwicklung von Dobusch, 2017)



4.2 Kriterien, nach denen Redaktionen Relevanz bemessen

Wie Journalisten entscheiden, was sie für berichtenswert halten, und wie Nachrichten und Informationen gewichtet werden, ist schon seit fast 100 Jahren Gegenstand einer intensiven Debatte und mittlerweile einer umfangreichen Forschungsliteratur. Der US-Journalist und Medienkritiker Walter Lippmann formulierte 1922 erstmals Kriterien für den *Nachrichtenwert* (*news value*) eines Ereignisses (Lippmann und Noelle-Neumann 1990). In den kommenden Jahrzehnten wurde eine Vielzahl solcher Kriterienkataloge veröffentlicht und intensiv diskutiert (z. B. Galtung und Ruge 1965; Schulz 1976). Siegfried Weischenberg führte 2013 als zentrale Nachrichtenfaktoren Bedeutung und Publikumsinteresse an, die er wiederum in eine Reihe weiterer von ihm identifizierter (Weischenberg 2001) Faktoren untergliederte: Ausmaß und Konsequenzen des Ereignisses als Unterpunkte des Faktors Bedeutung sowie Nähe, Prominenz der beteiligten Personen, Aktualität und Human Interest als Komponenten des Faktors Publikumsinteresse.

Wie die einzelnen Faktoren in Redaktionen gewichtet werden, hängt selbstverständlich auch von der Ausrichtung und dem Selbstverständnis des jeweiligen Mediums ab. Weischenberg zufolge steht bei sogenannten *harten*

Nachrichten die Bedeutung im Vordergrund, das Publikumsinteresse eher bei sogenannten *weichen Nachrichten*. Als harte Nachrichten werden etwa Meldungen aus Politik und Wirtschaft bezeichnet, weiche Nachrichten wären zum Beispiel Meldungen über Prominente oder andere Berichterstattung, die zumindest teilweise unterhaltenden Charakter hat.

4.3 Kriterien, nach denen Intermediäre Aufmerksamkeit verteilen

Facebook und Google gehen bei der Bestimmung dessen, was sie selbst als Relevanz definieren, anders vor als zum Beispiel Journalisten. Der folgende Abschnitt konkretisiert, was die Intermediäre dabei eigentlich tun. Das erscheint auf den ersten Blick trivial. Schließlich schaffen beide Intermediäre es, in jeweils einem Satz die wesentlichen Aufgaben ihrer algorithmischen Entscheidungsprozesse zu beschreiben:

- „Der Algorithmus bestimmt anhand von mehr als 200 Signalen, welche der unzähligen Seiten und Inhalte für Ihre jeweilige Suchanfrage am relevantesten sind“ (Google 2016).
- „Rather than showing people all possible content, News Feed is designed to show each person on Facebook the content that's most relevant to them“ (Boland 2014).

Nur: Was bedeutet relevant? Relevanz ist ein menschliches Konzept und somit nicht unmittelbar messbar. Wir analysieren im Folgenden exemplarisch, woran die beiden Intermediäre Google und Facebook Relevanz messen. Wir konzentrieren uns auf Beispiele für das Zusammenwirken menschlicher Wahrnehmung und maschineller Entscheidungen, aus denen sich Grundprinzipien ableiten lassen. Deshalb skizzieren wir im Folgenden lediglich beispielhaft einige der bekannten Faktoren, die die für die Meinungsbildung relevanten Intermediäre Google und Facebook nutzen. Wir nutzen dabei die Selbstdarstellung der Anbieter und die Ergebnisse externer Beforschung. Ziel ist es, einen Eindruck der für die Entscheidung relevanten und zugänglichen Faktoren zu erhalten – nicht Aussagen über ihre absolute Bedeutung zu treffen.

4.3.1 Google

Google spricht von „mehr als 200 Signalen“, anhand derer ein ADM-Prozess bestimmt, in welcher Reihenfolge die Treffer zu einer Suchanfrage auf der Ergebnisseite stehen (Google 2016). Die Analyse zeigt, dass Signale sich vier Kategorien zuordnen lassen:

- Inhalte der Suchergebnisse
- Eigenschaften des Suchenden
- Reaktionen anderer Nutzer auf die Ergebnisse
- menschliche Evaluation

Inhalte der Suchergebnisse

Die fürs Trefferranking ausgewerteten Eigenschaften der Inhalte haben mit dem Sinngehalt wenig oder nur mittelbar zu tun. Google nennt als Signale explizit beispielsweise die „Aktualität des Inhalts einer Website“, die „Wörter auf der Webseite“, „URL und Titel“ (Google 2016). Laut Google-Manager Andrey Lipattsev sind Signale aus dieser Kategorie die wichtigsten für die Bestimmung der Trefferrangfolge: „It's content, and links pointing to your site“ (Ratcliff 2016). Eine der wenigen wissenschaftlichen Untersuchungen zur Relevanz von Rankingfaktoren macht die Schlagwortdichte in URL, Titel und HTML-Header als ein einflussreiches Kriterium aus (Su et al. 2014). Es ist eine große, weitgehend unregulierte Dienstleistungsbranche entstanden, die sich der Optimierung von Webseiten für die Auffindbarkeit in Suchmaschinen widmet (Iredale und Heinze 2016). Die Anpassung an Veränderungen in der algorithmischen Entscheidungsfindung ist ein zentrales Merkmal der dort erbrachten Dienstleistungen.

Eigenschaften des Suchenden

Vom Suchenden selbst geht die geschriebene oder gesprochene Anfrage aus. Welche Intention sich hinter einer Anfrage verbirgt, hängt von weiteren Faktoren ab, zum Beispiel dem Wohnort des Nutzers oder der Sprache. Google selbst spricht abstrakt von „Personalisierung“ als einem Signal (Google 2016), welche genauen Merkmale des Suchenden sich dahinter verbergen, ist nicht öffentlich. Zur Prognose der Intention dient der laut mehreren Google-Vertretern (Levy 2016) dritteinflussreichste Ranking-Faktor: Der *Rankbrain* genannte ADM-Prozess sagt unter anderem die mögliche Bedeutung bisher nie gestellter Suchanfragen voraus, basierend auf Abstraktionen aus historischen Anfragen (Clark 2015). 2008 erklärte eine Google-Produktmanagerin in einem Blogbeitrag, für die Personalisierung würden der mutmaßliche Aufenthaltsort des Nutzers, kürzlich durchgeführte Suchanfragen sowie die Web History des genutzten Browsers herangezogen, vorausgesetzt der Nutzer hat der Erhebung dieser Daten zugestimmt (Garb 2008).

Reaktionen anderer Nutzer auf die Ergebnisse

Menschliche Reaktionen auf Suchergebnisse oder Inhalte bemisst der ADM-Prozess beispielsweise anhand der Anzahl anderer Websites, die auf Trefferseiten verlinken, sowie anhand der „Autorität dieser Links“ (Google 2016). Das Verhalten von Google-Nutzern auf den Trefferseiten wertet der ADM-Prozess ebenfalls zum Ranking aus, allerdings eher mittelbar. Wie viele Nutzer welche Treffer aufrufen und wie schnell sie wieder zur Trefferübersicht zurückkommen, wird als Signal der Qualität von Treffern gewertet. Aber die Informationen über dieses Nutzungsverhalten sind nicht so breit verfügbar und verlässlich wie die Verlinkungen (Ratcliff 2016) und werden daher anders ins Ranking der Suchergebnisse einbezogen. Signale aus sozialen Netzwerken fließen ebenfalls in die Entscheidung ein – „tweets and social mentions – which are basically more links to the page, more mentions of this context“ (a. a. O.).

Menschliche Evaluation

Google lässt auch Menschen gegen Bezahlung die Qualität von Suchergebnissen zu bestimmten Anfragen bewerten (siehe auch [Abschnitt 3.2](#)).

Evaluatoren bewerten die Qualität von Ergebnissen für spezifische Anfragen auf Basis eines 160-seitigen Handbuches (Google 2017). Sie beurteilen auf zwei zentralen Skalen: *Page Quality* (die Inhalte der Ergebnisse) und *Needs Met* (die Passung von Inhalten und Suchanfrage):

Page Quality bezieht sich auf die Ergebnisseiten allein. Evaluatoren sollen Expertise, Ansehen und Vertrauenswürdigkeit einer Seite bestimmen. Die Anleitungen im Evaluatoren-Handbuch zur Recherche lesen sich stellenweise wie ein Medienkompetenz-Ratgeber:

„Many other kinds of websites have reputations as well. For example, you might find that a newspaper website has won journalistic awards. Prestigious awards, such as the Pulitzer Prize award, are strong evidence of positive reputation. When a high level of authoritativeness or expertise is needed, the reputation of a website should be judged on what expert opinions have to say. (...) Look for information written by a person, not statistics or other machine compiled information“ (Google 2017: 16).

Die Skala Needs Met soll beziffern, wie gut die Treffer zur Anfrage und Intention des Nutzers passen. Extrembeispiel: Wer explizit nach *alternativen Sichtweisen* auf den Holocaust sucht, soll auch Verweise auf Holocaust-Leugnung sehen: „Pages which directly contradict well established historical facts should be rated Fails to Meet, unless the query clearly indicates the user is seeking an alternative viewpoint“ (a. a. O.: 119) Fragen wie „Did the holocaust happen?“ fallen dem Handbuch zufolge jedoch nicht in diese Kategorie. Google hatte nach einer Reihe von Medienberichten über Holocaustleugner-Webseiten in den Ergebnislisten Änderungen an den Richtlinien für die Evaluatoren vorgenommen (Sullivan 2017). Das Unternehmen führte ein zusätzliches Kriterium ein: Die Evaluatoren können Inhalte nun auch als *upsetting-offensive* markieren. Als Beispiel für diese Kategorie ist in der neuen Version des Manuals für die Evaluatoren explizit Holocaustleugnung aufgeführt.

Die Qualitätseinschätzungen der Such-Evaluatoren fließen laut Google als Signale nicht direkt ins Ranking einzelner Ergebnisse der Liveversion ein, sondern beispielsweise in die Bewertung von experimentellen Veränderungen des Ranking-Verfahrens (Underwood 2015). Es ist nicht öffentlich bekannt, wie viele Menschen als Evaluatoren arbeiten und in welchem Ausmaß Google diese Bewertungen nutzt. Der US-Suchmaschinenexperte Danny Sullivan schätzt die Zahl dieser Personen, die Google selbst Quality Raters nennt, auf weltweit über 10.000 (Sullivan 2017), in Deutschland sollen 600 Evaluatoren arbeiten (Spiegel Online 2017). Stellenanzeigen geben Hinweise darauf, dass Such-Evaluatoren zumindest zum Teil nicht direkt für Google arbeiten, sondern als freiberufliche Agenten von zu Hause aus für [Subunternehmen](#).

Tabelle 3: Exemplarische Übersicht öffentlich bekannter Signale für Googles Ranking von Suchtreffern

Kategorie	Signale
Eigenschaften der Suchenden	Vom Nutzer eingetippte oder gesprochene Suchanfrage.
	Vom ADM-Prozess Rankbrain prognostizierte Intention des Nutzers
	Persönliche Merkmale wie z. B. Wohnort
Inhalt der Suchergebnisse	Wie lange existiert die Seite?
	Welche Begriffe kommen wo wie oft vor?
Reaktionen anderer Nutzer auf die Ergebnisse	Wie viele Websites verweisen auf die Seite?
	Wer verweist auf das Angebot?
	Welche Begriffe verweisen?
	Wie schnell kehren Nutzer zur Suchübersicht zurück?
	Wer verweist in sozialen Netzwerken auf die Inhalte?
Menschliche Evaluation	Mittelbar: Wie bewerten Such-Evaluatoren Expertise, Ansehen und Vertrauenswürdigkeit einer Seite?
	Mittelbar: Wie bewerten Such-Evaluatoren die Passung der Ergebnisse zu einer spezifischen Anfrage?

Quellen: Google 2016; Google 2017; Christopher Ratcliff 2016; Mimi Underwood, 2015.

4.3.2 Facebook

Bei einem sozialen Netzwerk wie Facebook fehlt in den meisten Fällen ein eindeutiges Signal des Nutzers über sein Informationsinteresse. Denn in den meisten Fällen dürfte bei der Nutzung gar kein eindeutiges Informationsinteresse vorhanden sein: Die algorithmisch strukturierte Auswahl im News Feed ist ein Angebot an möglicherweise Interessantem – ein Strukturmerkmal, das der News Feed übrigens im Prinzip mit gewissen Boulevardmedien gemeinsam hat. Die Quellen der Inhalte aber sind bei Facebook andere: Es handelt sich um eine Mischung aus Meinungen, Unterhaltung, persönlicher oder medialer Kommunikation (Backstrom 2013). Wie sich diese Mischung zusammensetzt, stimmt Facebook mittels mehrerer Signalkategorien ab:

- Eigenschaften und Präferenzen des Nutzers
- Beziehungen des Nutzers zu Sendern
- Eigenschaften der Inhalte
- Reaktionen anderer Nutzer auf Inhalte
- menschliche Evaluation

Eigenschaften und Präferenzen des Nutzers

Werbekunden können bezahlte Beiträge basierend auf Profilinformatoren etwa über Alter, Wohnort, Geschlecht und Interessen in die News Feeds bestimmter Zielgruppen platzieren. Da solche Anzeigen dazu führen können,

dass Menschen beispielsweise bestimmte Facebook-Seiten abonnieren, kann diese Werbung auch Auswirkungen auf die Zusammensetzung des News Feed haben. Wer basierend auf einer auf seine Interessen zugeschnittenen Anzeige eine Seite abonniert, geht eine Beziehung mit diesem Angebot ein und sieht infolge dessen womöglich häufiger Inhalte dieses Angebots.

Facebook wertet für die Zusammenstellung der Feeds aus, welche Inhalte Nutzer priorisieren (über „als Erstes anzeigen“) oder verbergen (Facebook 2016b). Beide Signale beruhen auf vergleichsweise zielgerichteten Handlungen. Facebook wertet zur Strukturierung auch mit Sicherheit vom Nutzer unbewusst gesetzte Signale aus, zum Beispiel die Betrachtungsdauer bei einem bestimmten Element im News Feed. Wer an einer Stelle länger verharret, signalisiert Zustimmung ohne bewusste Interaktion (Oremus 2016).

Beziehung zwischen Nutzern und Sendern

Je häufiger ein Nutzer mit dem Facebook-Profil eines Freundes, einer Seite oder einer Person des öffentlichen Lebens interagiert hat, desto wahrscheinlicher ist eine Priorisierung weiterer Inhalte dieser Quelle (Backstrom 2013). Zu den Interaktionen können Likes, Klicks, Seitenbesuche, Kommentare, aber auch verlangsamtes Scrollen bei den Beiträgen dieser Absender zählen (Constine 2016).

Auch bei der Auswertung von Reaktionen anderer Nutzer auf bestimmte Inhalte zählen solche Profile mehr: Wenn sehr viele Facebook-Nutzer auf einen Beitrag anspringen, steigt die Wahrscheinlichkeit, dass dieser im News Feed einer Person höher priorisiert wird. Größer wird die Wahrscheinlichkeit, wenn viele Profile auf einen Beitrag anspringen, mit denen jemand häufig interagiert (Backstrom 2013).

Eigenschaften der Inhalte

Facebook priorisiert auch nach Beitragsart (Links, Videos, Fotos, Texte). Und zwar nach dieser Logik: Wer den Interaktionen zufolge ein Faible für Fotos hat, sieht mit höherer Wahrscheinlichkeit Fotos im persönlichen News Feed (Facebook 2016a). Welche darüber hinausgehenden Merkmale von Beiträgen als Signale ausgewertet werden, ist öffentlich kaum dokumentiert. Hier ein Hinweis darauf, welche weiteren Eigenschaften von Beiträgen der ADM-Prozess für das Ranking auswertet: „Videos played longer than three seconds, have sound activated, or are watched all the way through are prioritized in News Feed“ (a. a. O.).

Die offizielle Begründung für diese Bevorzugung lauter und von einigen Menschen intensiver genutzter Videos ist aufschlussreich: „We’ve found videos tend to be more engaging than any other content type“ (a. a. O.). Diese Aussage lässt drauf schließen, dass viele, wenn nicht sogar alle Signale letztlich der Prognose des „Engagements“ dienen, also der größtmöglichen Menge an Klicks, Likes, Shares oder Kommentaren. Was zu Reaktionen führt, ist relevant. Insofern müssen wir die Relevanz der Eigenschaften von Inhalten relativieren: Der öffentlichen Darstellung nach scheint Facebook bei der Priorisierung alle Eigenschaften des Inhalts nur in Relation zur Wirkung auf Nutzer zu betrachten. In dieser Logik gibt es keine von den persönlichen Reaktionen unabhängigen Merkmale von Inhalten, die eine Priorisierung rechtfertigen. Das entspricht den sogenannten *News Feed Values*, die Facebook darlegt: „Die Inhalte, die eine Person informativ oder interessant findet, können sich von den Inhalten unterscheiden, die eine andere Person für informativ oder interessant hält. (...) Unser Ziel besteht darin, die Meldungen zu liefern, von denen wir wissen, dass sie eine bestimmte Person am meisten interessieren“ (a. a. O.).

Facebook führt in den sogenannten *Gemeinschaftsstandards* Merkmale von Beiträgen auf, die unabhängig von den Publikumsreaktionen zu einer Entfernung aus dem News Feed führen, zum Beispiel bestimmte Formen von Nacktheit, Urheberrechtsverletzungen oder Aufrufe zu terroristischen Aktivitäten (Facebook o. J.).

4.3.2.1 Reaktionen anderer Nutzer

Beiträge mit vielen Likes, Kommentaren oder Shares haben eine höhere Wahrscheinlichkeit, weiter oben im News Feed von Menschen zu erscheinen – insbesondere, wenn der betreffende Nutzer häufiger mit den

Weiterverbreitern interagiert (Facebook 2016a). Zudem wertet der ADM-Prozess aus, wie schnell Menschen zu Facebook zurückkehren, nachdem sie einem Verweis im News Feed zu einer Quelle außerhalb gefolgt sind. Die Annahme dahinter ist: Wenn Menschen zu schnell wiederkommen, hat die Präsentation im News Feed mehr oder etwas anderes versprochen, als die beworbene Quelle einlöst (El-Arini und Tang 2014).

4.3.2.2 Menschliche Evaluation

Seit Mitte 2014 nutzt Facebook auch menschliche Evaluatoren. Sie bewerten die Qualität der ihnen angezeigten Zusammenstellung des News Feed, schreiben eine Einschätzung zur Platzierung jedes aufgeführten Beitrags. Anfangs wurden knapp 1.000 Mitarbeiter über einen Dienstleister in den Vereinigten Staaten für diese Arbeit bezahlt, seit Mitte 2015 arbeiten weltweit Evaluatoren (Oremus 2016). Wie viele es sind und welchen Einfluss ihre Bewertungen haben, ist unklar.

Ein bekanntes Beispiel für eine Veränderung des Ranking-ADM-Prozess wegen des Evaluatoren-Feedbacks: Menschen klicken schockierende oder beunruhigende Nachrichten im News Feed nicht an, obgleich sie deren Relevanz hoch einschätzen. Nachdem dieses Verhalten in Nutzerbefragungen (im sogenannten *Feed Quality Panel*) bestätigt wurde, operationalisierten Mitarbeiter ein neues Signal für Engagement – die Verweildauer bei einem Beitrag im Feed ohne weitere Interaktion (a. a. O.). Von Menschen zu speziellen Fragen bewertete Inhalte nutzt Facebook auch als Lernmaterial für Machine Learning. So beurteilen Evaluatoren unter anderem den Informationsgehalt von Beiträgen und die Vertrauenswürdigkeit der jeweiligen Quellen. Auf Basis dieser Daten wurde ein Algorithmus trainiert, der Merkmale informativer Inhalte erkennen soll (Kacholia 2013).

Außerdem befragt Facebook täglich Zehntausende normale, zufällig ausgewählte Nutzer weltweit zur Qualität ihres News Feed. Die Teilnehmer dieses „Feed Quality Program“ (Xu, Lada und Kant 2016) bewerten bewusst auf einer Skala von 0 bis 5, wie gern sie einzelne Beiträge in ihrem Feed sehen wollen (Zhang und Chen 2016).

Bei gemeldeten Verstößen gegen die Gemeinschaftsstandards prüfen Menschen die beanstandeten Beiträge, in Deutschland sind Mitarbeiter des Dienstleisters Arvato³ (Reinbold und Rosenbach 2016).

Tabelle 4: Exemplarische Übersicht öffentlich bekannter Auswahlkriterien von Facebook

Kategorie	Signal
Eigenschaften und Präferenzen der Nutzer	Bisherige Reaktion auf Formate (Likes/Verbergen)
	Betrachtungsdauer von Beiträgen
	soziodemographische Merkmale aus Profilingaben
Inhalt der Beiträge	Format (z. B. Foto, Video, Text)
	ähnliche Inhalte, die Tester gut bewerten
	Aktualität
Reaktionen anderer Nutzer auf die Beiträge	Anzahl Likes, Shares, Kommentare, Verbergen
	Betrachtungsdauer
	Rückkehrgeschwindigkeit
	Umfrageergebnisse zu Beiträgen
Beziehung zwischen Sender und Empfänger	Freundschaft, Tagging
	Klicks, Scrollverhalten, Seitenbesuche

³ Transparenzhinweis: Arvato gehört zum Konzernverbund der Bertelsmann SE & Co. KGaA, zu deren Gesellschaftern die Bertelsmann Stiftung zählt.

	Kommentare, Likes
Menschliche Evaluation	bezahlte Evaluation des News Feed
	Umfrageergebnisse zu Beiträgen
	menschliche Prüfung beanstandeter Beiträge
Quellen: Backstrom 2013; Facebook 2016a; Facebook 2016b; Oremus 2016; Zhang und Chen 2016.	

4.4 Zwischenfazit: algorithmische Prozesse als Relevanzdeuter

Die beiden untersuchten Intermediäre machen Relevanz an ähnlichen Kriterien messbar:

- Wie reagieren Nutzer auf Inhalte?
- Welche Eigenschaften des Inhalts korrelieren mit erwünschten Reaktionen der Nutzer?

Beide Kategorien von Signalen sind aus diversen Gründen zur Messung von wie auch immer definierter Relevanz mit Vorsicht zu betrachten.

4.4.1 Reaktionen von Nutzern

Vielfältige Nutzungsarten: Insbesondere bei Intermediären mit sozialen Netzwerken können Menschen aus sehr unterschiedlichen Bedürfnissen mit Beiträgen interagieren. Die Nutzungsweisen lassen sich in diese drei von Paus-Hasebrink, Schmidt und Hasebrink (2009) vorgeschlagenen Praktiken aufteilen: Beziehungs-, Identitäts-, Informationsmanagement (Überblick in: Schmidt et al. 2017: 21). Die Auseinandersetzung mit Inhalten und auch mit Informationen zu gesellschaftlichen Diskursen dient also nicht immer dem Informationsmanagement. Wie Menschen auf Inhalte reagieren, kann auch aus ganz anderen Motivlagen resultieren. Vielleicht favorisieren Menschen ja bestimmte Beiträge, weil viele gute Freunde diese bereits empfohlen haben. Die sichtbare Reaktion ist dann weniger ein Signal für die inhaltliche Qualität eines Beitrags. Die Reaktion eines Nutzers zeigt in diesem Fall eher an, dass jemand zu einer Gruppe gehört oder gehören will und dieser Beitrag ein gutes Mittel ist, um das zu zeigen. Der Einfluss unterschiedlicher Nutzungspraktiken dürfte bei sozialen Netzwerken größer als bei Suchmaschinen sein.

Kognitive Verzerrungen durch impulsive Reaktionen: Menschen geben die ausgewerteten Signale, wie Aufrufe von Verweisen, verlangsamtes Scrollen oder Favorisieren von Beiträgen, schnell und impulsiv ab. Es ist wahrscheinlich, dass hierbei der Denkmodus dominiert, den Kahneman als System 1 bezeichnet: Es „arbeitet automatisch und schnell, weitgehend mühelos und ohne willentliche Steuerung“ (Kahneman 2012a: 34). Die Stärken und Schwächen dieses Denkmodus führen zu systematischen kognitiven Verzerrungen der Wahrnehmung, zum Beispiel zu einer allgemeinen Bestätigungstendenz. Sie „begünstigt die unkritische Annahme von Vorschlägen und überzeichnet die Wahrscheinlichkeit extremer und unwahrscheinlicher Ereignisse“ (a. a. O.: 133). Die nicht hinreichend korrigierte Auswertung der auf System 1 zurückgehenden Signale als Relevanzmarker könnte dazu führen, dass zutiefst menschliche kognitive Verzerrungen Eingang in ADM-Prozesse finden. So könnte menschliche Fehlbarkeit durch maschinelles Entscheiden skalieren und gewissermaßen reingewaschen werden. Extrem subjektive und unbewusste Verarbeitungsprozesse drohen so zu scheinbar objektiven Relevanz- oder gar Qualitätsmaßen zu werden.

Reaktionen sind keine unabhängige Variable: Die Nutzer eines Intermediärs bestimmen nicht unbeeinflusst und frei darüber, welche Beiträge sie bevorzugen. Sie sehen immer nur eine Auswahl möglicher Beiträge. Sie reagieren auf eine algorithmisch erstellte Auswahl. Diese Reaktionen sind Grundlage der weiteren Personalisierung. Deshalb lässt sich keine Wirkungsrichtung isolieren: Weder bestimmt der Nutzer über den ADM-Prozess noch bestimmt der ADM-Prozess die Reaktionen der menschlichen Nutzer. Beides bedingt und beeinflusst einander fortwährend. Insofern ist das Versprechen von Facebook „Du kontrollierst deine Erfahrung“

(Facebook 2016b) nur bedingt gültig. Der einzelne Nutzer kontrolliert nur eingeschränkt die Strukturierung seines personalisierten News Feed. Die Zusammensetzung und Gewichtung aller Signale entzieht sich der Kontrolle des Einzelnen. Die Möglichkeit, einen lediglich chronologisch sortierten News Feed zu verwenden, ist keine wirkliche Alternative. Die schiere Menge an Beiträgen macht diese Lösung impraktikabel. Echte, praktikable Alternativen müssten anders aussehen: Zum Beispiel die Wahl zwischen unterschiedlichen Formen algorithmischer Strukturierung mit unterschiedlichen Freiheitsgraden und Anbietern. So könnte es möglich sein, den eigenen News Feed von einem algorithmischen System ordnen zu lassen, das nur Verweise zu Nachrichtenquellen mit transparent definierten Mindestanforderungen anzeigt (denkbare Kriterien: Einhaltung des Trennungsgebots von Redaktion und Verlag, z. B. Impressum, Mitwirkung in Selbstregulierung usw.). Oder ein alternatives System, das die ausgewerteten Reaktionen anderer Nutzer über einen längeren, vom Nutzer frei wählbaren Zeitraum auswertet (oben steht nur, was in den vergangenen 48 Stunden insgesamt von Freunden oft empfohlen wurde).

4.4.2 Eigenschaften der Inhalte in Relation zu Nutzerreaktionen

Die Analyse der Kriterien zeigt, dass beide Intermediäre Relevanz auf eine spezielle Art deuten: Relevanz ist personalisiert. Die optimale Passung von Inhalten an die Präferenzen Einzelner ist das oberste Ziel. Die aus diesem Relevanzverständnis abgeleitete Messbarmachung zielt also auf die Zufriedenheit der Nutzer ab. Zufriedenheit äußert sich in Nutzungsintensität und Häufigkeit.

Es ist aber lediglich eine Annahme, dass Zufriedenheit mit Relevanz gleichzusetzen ist. Dem lässt sich entgegenhalten, dass ein teilhabestärkender gesellschaftlicher Diskurs nicht allein durch die passgenaue Befriedigung möglichst vieler Einzelbedürfnisse entsteht, sondern beispielsweise durch einen Ausgleich möglichst vieler Bedürfnisse.

Relevanz als Befriedigung individueller Bedürfnisse äußert sich in Konzepten wie Googles „Needs Met (NM)“ oder der Orientierung von Facebook am Engagement der Nutzer. Google erfasst zwar einige Signale unabhängig von der Zuordnung zu Nutzerreaktionen, zum Beispiel die Verteilung von Schlagwörtern in Quellen und die von menschlichen Such-Evaluatoren bewertete Page Quality von Angeboten. Wie diese Faktoren in das Ranking von Suchergebnissen einfließen, ist allerdings nicht öffentlich.

Die Orientierung an persönlicher Passung führt dazu, dass sich die Qualität von Beiträgen immer durch die Zufriedenheit der Empfänger definiert. Eine mögliche Folge ist ein Wahrheitsrelativismus, wie ihn Facebook in den sogenannten *Werten des News Feed* formuliert:

„Unser Ziel besteht nicht darin, festzulegen, mit welchen Themen sich die Welt beschäftigen sollte. Unser Ziel besteht darin, Menschen mit Ideen zu verbinden – und Menschen die Geschichten zu zeigen, die für sie die größte Bedeutung haben. Es ist entscheidend für unsere Integrität, dass wir alle Perspektiven und Blickwinkel miteinbeziehen und Bewertungen verwenden, um den Menschen die Geschichten und Quellen zu bieten, die für sie am bedeutendsten und ansprechendsten sind“ (Facebook 2016b).

Hier wird implizit die an individueller Passung ausgerichtete algorithmische Entscheidungsfindung zum Superleitwert erhoben – Passung sticht Achtung vor der Wahrheit.

Dieses Vorgehen ist grundsätzlich zu hinterfragen, aber auch die Umsetzung verdient eine kritische Würdigung. Wenn Zufriedenheit sehr eingeschränkt an Seitenaufrufen, Nutzungsdauern, Favorisierungen und Ähnlichem festgemacht wird, bleiben qua Messweise denkbare Kriterien außen vor: Ausgeglichenheit, Wissensgewinn, konstruktiver Austausch mit Andersdenkenden, Konsens- und Wahrheitsfindung beispielsweise. Aus der Beobachtung und Analyse des Verhaltens Vieler lassen sich keine gültigen Aussagen für alle ableiten, wie Etzioni und Etzioni (2017) treffend für deskriptive Ansätze einer Ethik zusammenfassen: „Observing people will not teach machines what is ethical but what is common“ (a. a. O.: 5).

Die Unzulänglichkeiten dieses Ansatzes demonstrieren Extremfälle wie die hohen Reichweiten von Falschmeldungen auf Facebook im US-Präsidentschaftswahlkampf (Silverman 2016) oder die prominente Platzierung von Holocaust-Leugnung in Google-Suchergebnissen zu neutralen Suchanfragen (Mierau 2016).

Diese Beispiele zeigen: Auch Lügen provozieren starke Reaktionen, auch Lügen befriedigen menschliche Bedürfnisse und Interessen. Geht es dem Schöpfer der Lüge primär um Reichweite auf Intermediären mit algorithmisch sortierten Inhalten, kann die Lüge sogar deutlich erfolgreicher sein als die Wahrheit – schließlich kann ihr Verfasser sie auf maximale Interaktionswahrscheinlichkeit und maximales Engagement hin optimieren (siehe z. B. Silverman und Alexander 2016).

Im folgenden Abschnitt wollen wir einige Faktoren erörtern und analysieren, die in der skizzierten Struktur und der gegenwärtige Gestaltung der Intermediäre der Öffentlichkeit Wirkung entfalten.

5 Wozu das führt: aktuelle Ausprägung des Strukturwandels der Öffentlichkeit

Abschnitt 4 skizzierte den Strukturwandel der Öffentlichkeit. Der folgende Abschnitt 5 zeigt, welche Formen Öffentlichkeit in dieser neuen Struktur unter den aktuell geltenden Bedingungen annimmt.

5.1 Nur was man gerade sieht, zählt

In einem sozialen Netzwerk können als Grundlage der Meinungsbildung allein die Vorschau eines Artikels oder eines Videos sowie Kommentare und Empfehlungen anderer Nutzer dienen. Der referenzierte Inhalt selbst muss nicht in die Meinungsbildung einfließen. Wer Überschrift, Anreißertext, das Vorschaubild und 30 Kommentare sieht, wird den so präsentierten Inhalt selbst vielleicht gar nicht wahrnehmen.

Es gibt empirische Hinweise, dass Menschen Beiträge in sozialen Netzwerken empfehlen, die sie nicht gelesen haben. Die Untersuchungen dazu haben einige methodische Einschränkungen: Sie analysieren auf Beitragsebene, nicht auf Ebene einzelner Leser oder Zuschauer. Sie können sagen, ob es einen Zusammenhang gibt zwischen der Anzahl der Erwähnungen eines Artikels in sozialen Netzwerken und der Anzahl der Aufrufe oder der Lesetiefe dieses Beitrags im Onlineangebot des Mediums. Auf dieser Datenbasis kann man nicht mit Bestimmtheit sagen, aber begründet vermuten, wie Menschen sich verhalten. Die Erkenntnis: Es gibt keinen nennenswerten Zusammenhang zwischen der Anzahl der Erwähnungen eines Beitrags auf Twitter und der Anzahl der Aufrufe oder der Lesetiefe dieses Beitrags.

Der Statistikdienstleister Chartbeat hat 2013 untersucht, wie die Menge der Erwähnungen eines Artikels auf Twitter mit der Lesetiefe, der Scrolltiefe und dem Anteil komplett gelesener Artikel zusammenhängt. Es gibt nur eine schwache Korrelation: „Both at Slate and across the Web, articles that get a lot of tweets don't necessarily get read very deeply. Articles that get read deeply aren't necessarily generating a lot of tweets” (Manjoo 2013).

Zum selben Ergebnis kommt eine Studie zur Verbreitung von Artikeln der fünf Medien BBC, CNN, Fox News, New York Times und Huffington Post auf Twitter in einem Sommermonat 2015. Die Forscher analysieren, wann und wie oft Verweise auf Artikel dieser Medien auf Twitter herumgereicht wurden und die dabei verwendeten URLs des Verkürzungsdienstleisters bit.ly tatsächlich aufgerufen wurden (Gabelkov et al. 2016: 2). Die statistische Auswertung legt eine Interpretation nahe: Was geteilt wird, wurde nicht unbedingt gelesen. Die Autoren bilanzieren: „There seems to be vastly more niche content that users are willing to mention in Twitter than the content that they are actually willing to click on“ ([a. a. O.: 9](#)).

Diese Erkenntnisse deuten darauf hin, dass Intermediäre mehr als ein bloßer Vertriebsweg sind. Wenn die Masse an Hinweisen auf Artikel die Menge der tatsächlichen Aufrufe übertrifft, lassen sich zwei Effekte vermuten:

- *Kampf um Aufmerksamkeit:* Die Erfolgsfaktoren zur Reichweitenmaximierung in den jeweiligen Netzwerken beeinflussen, wie redaktionelle Medien ihre Inhalte dort präsentierten. Denn Überschrift, Anreißer und Vorschau für Beiträge dienen der Werbung mindestens ebenso sehr wie der Information. Sie sollen Aufmerksamkeit erregen und zum Aufruf der verlinkten Originalquelle animieren. Daher liegt es nahe, dass die Wirkweisen der Intermediäre antizipiert und beispielsweise Thesen überzeichnet, verkürzt und Sachverhalte zuspitzt werden (siehe auch Abschnitt 5.2).
- *Priming:* Die Einordnung und Kommentierung durch Dritte wirkt auf die Wahrnehmung der Beiträge. Wer eine Nachricht durch Empfehlung eines Dritten etwa mit dem Hinweis „Schlimm“ entdeckt, nimmt die Einordnung vor dem Eingearbeiteten wahr – wenn er die Originalquelle überhaupt noch zur Kenntnis nimmt.

In beiden Fällen nimmt man statt des kompletten Beitrags oft lediglich die Vorschau und die Kommentare anderer wahr. Diese kann zu voreiligen Schlussfolgerungen führen, dem WYSIATI-Effekt: „What you see is all there is (in etwa: ‚Nur was man gerade sieht, zählt‘)“ (Kahneman 2012a: 113).

5.2 Emotion bringt Reichweite

Es gibt einen Zusammenhang zwischen der in einem Beitrag geäußerten Stimmung, der Reichweite dieses Beitrags in sozialen Netzwerken und der im Folgenden geäußerten Stimmung in Beiträgen des Publikums. Darauf deutet eine Reihe von Untersuchungen hin.

Eine Studie von Stieglitz und Dang-Xuan (2012) der Beiträge und Kommentare auf den Facebook-Seiten von Parteien im Jahr 2011 zeigt: Emotional negativ aufgeladene Beiträge provozieren mehr Reaktionen als nicht emotional aufgeladene. Die Software-Textanalyse zählte die Häufigkeiten von Worten mit negativer (z. B. „enttäuscht“) und positiver (z. B. „Hurra!“) Konnotation und bewertete die Intensität der geäußerten Emotion auf einer Skala von 1 bis 5. Analysiert wurden 5.636 Beiträge und 36.397 Kommentare von 10.438 Nutzern auf den Facebook-Seiten von politischen Parteien zwischen März und September 2011 (CDU, CSU, SPD, FDP, B90/Die Grünen, Die Linke, Piratenpartei). Ein Kernergebnis: Negativität erzielt mehr Reaktionen. Je negativer Facebook-Beiträge (gemessen an Häufigkeit und Intensität genutzter Begriffe) formuliert sind, desto mehr Kommentare erhalten sie. Bei Beiträgen mit positiven Begriffen gibt es keinen solchen Zusammenhang.

Den Zusammenhang zwischen geäußelter Emotion und Reaktionen auf diese haben dieselben Forscher auch bei einem anderen sozialen Netzwerk (Twitter) untersucht und zum Teil bestätigt (Stieglitz und Dang-Xuan 2013). Der Unterschied bei dieser Analyse zwischen den Reaktionen auf positiv und negativ formulierte Beiträge ist nicht so groß wie bei der Facebook-Untersuchung. Als Reaktion werteten die Forscher auf Twitter allerdings nicht die Anzahl der Kommentare zu einem Beitrag, sondern die Anzahl der Retweets. Sie analysierten 165.000 Tweets der offiziellen Parteikonten von CDU, CSU, SPD, FDP, B90/Die Grünen, Die Linke und Piratenpartei im Umfeld der Landtagswahlen 2011 in Berlin, Baden-Württemberg und Rheinland-Pfalz mit derselben Methode wie beim Facebook-Datensatz (a. a. O.: 225). Die Erkenntnisse:

- Je stärker emotional aufgeladen ein Tweet formuliert ist, desto häufiger wird er per Retweet weiterverbreitet. Dieser Effekt ist bei geäußelter negativer Stimmung zum Teil (Sample zur Berliner Wahl, nicht Baden-Württemberg und Rheinland-Pfalz) stärker als bei geäußelter positiver.
- Je stärker emotional aufgeladen ein Tweet formuliert ist, desto weniger Zeit vergeht bis zur ersten Weiterverbreitung per Retweet. Hier ist kein Zusammenhang zwischen der Art der geäußerten Stimmung und der Zeitspanne feststellbar.

Die Facebook-Untersuchung von Stieglitz und Dang-Xuan (2012) zeigt einen weiteren Effekt: Emotionale Beiträge werden ähnlich emotional aufgeladen kommentiert. Je negativer ein Beitrag formuliert ist, desto negativer sind die Reaktionen. Und je positiver ein Beitrag geschrieben ist, desto positiver lesen sich die Kommentare dazu. Stieglitz und Dang-Xuan selbst leiten daraus ab: „Our results indicated that positive as well as negative emotions might diffuse in the following discussion“ (a. a. O.: 13).

Einen Zusammenhang zwischen geäußerten Emotionen in von Nutzern wahrgenommenen Beiträgen und dem Kommunikationsverhalten der Nutzer selbst hat ein weit größeres Experiment von Facebook Anfang 2012 experimentell belegt (Kramer, Guillory und Hancock 2014). Eine Woche lang zeigte Facebook 689.003 Nutzern Beiträge aus ihrem Newsfeed in Abhängigkeit von der darin geäußerten Stimmung an. Ein Teil der Nutzer sah weniger positive, ein Teil weniger negative und ein Teil komplett zufällig ausgewählte Beiträge. Die kleine, aber messbare Wirkung: Wer weniger Positives sieht, schreibt bei Facebook messbar negativer. Wer weniger Negatives sieht, schreibt messbar mehr Positives (a. a. O.: 8790).

Neben diesen zentralen Ergebnissen zeigt die Arbeit von Kramer et al. im Übrigen ein weiteres Mal die Macht von Facebook über seine Nutzer auf: Sobald der Intermediär in die Sortierung eingreift, hat das unmittelbare Auswirkungen, bis hin zur emotionalen Befindlichkeit der Nutzer (siehe Abschnitt 5.3).

Auf Basis der empirischen Befunde lässt sich nicht klar zuschreiben, für welche Wirkung die algorithmische Auswahl der Intermediäre und wofür menschliche Verhaltensmuster verantwortlich sind. Denn beides lässt sich in den algorithmischen Prozessen nicht mehr trennen: Wenn Menschen emotionale Beiträge häufiger weiterreichen und kommentieren, kann ein ADM-Prozess das als Signal für die Relevanz dieser Inhalte werten. Je mehr Retweets und Kommentare ein Beitrag erhält, als desto relevanter gilt er und desto relevanter könnte er für eine breitere Zielgruppe sein. Vielleicht wird auf Basis dieser Einschätzung ein emotionaler Beitrag in mehr Feeds angezeigt als andere, neutralere formulierte Beiträge. Nicht weil die geäußerte Stimmung analysiert wird, sondern weil der Prozess die Menge der Publikumsreaktionen in einer bestimmten Zeit als Maßstab für Relevanz heranzieht.

Unabhängig von der Wirkungsrichtung zeigen die empirischen Ergebnisse, dass menschliche Reaktionen auf Beiträge nur bedingt zur Relevanzeinschätzung taugen. Denn menschliche Reaktionen hängen offenkundig auch mit Faktoren zusammen, die niemand als Kriterien für die Qualität einer gesellschaftlichen Debatte heranziehen würde. Sehr negative Beiträge und viele heftige Reaktionen auf diese zeugen nicht zwangsläufig von Relevanz für den gesellschaftlichen Diskurs, in einigen Fällen gewiss nur von Lautstärke um der Freude an der Lautstärke willen (vergleiche zu kognitiven Verzerrungen [Abschnitt 3.4.1](#)).

Die Frage nach den angemessenen und tatsächlich von Intermediären genutzten Kriterien für Relevanz lässt sich nicht beantworten, ohne **Ziele** und **Konsequenzen** algorithmischer Prozesse zu analysieren. Bei beiden Themen gibt die Empirie zu emotionalen Beiträgen Denkstöße.

Zu den **Zielen**: Auf welches erwünschte Verhalten der Nutzer hin optimiert der Auswahlprozess eines Intermediärs das Angebot? Höheres Engagement? Klickraten? Kommentarhäufigkeit? Konstruktiven Diskurs? Zielkonflikte sind wahrscheinlich, einen möglichen zeigen die empirischen Befunde von Kramer et al. (2014). Sie haben bei dem Facebook-Experiment zur Gefühlsbeeinflussung auch entdeckt, dass Menschen, die weniger emotionale Beiträge in ihrem Feed sehen, danach selbst weniger aktiv sind. Kramer spricht von einem Rückzugseffekt: „We also observed a withdrawal effect: People who were exposed to fewer emotional posts (of either valence) in their News Feed were less expressive overall on the following days, addressing the question about how emotional expression affects social engagement online“ (a. a. O.: 8792). Will man das Engagement der Nutzer steigern, könnte man sich also für eine prominentere Platzierung stark emotionaler Beiträge entscheiden. Belohnen, was Menschen aufwühlt („to reward that which keeps us agitated“) – so formulieren Experten in einer Pew-Research-Befragung diese Sorge (Rainie, Anderson und Albright 2017: 12).

Weitere *Konsequenzen* emotionaler Beiträge in sozialen Netzwerken sind bislang wenig erforscht. Naheliegender erscheint die Vermutung, dass negativ gefärbte Beiträge destruktives Verhalten nach sich ziehen könnten. Eine Studie zu Troll-Kommentaren⁴ kommt basierend auf einem Experiment mit 667 Teilnehmern und einer Analyse von gut 16,5 Millionen Kommentaren auf CNN.com (Cheng et al. 2017) zu zwei Kernergebnissen:

- Menschen in schlechter Stimmung (im Experiment wurde die Stimmung durch einen Test beeinflusst und danach über einen Fragebogen erhoben) treten wahrscheinlicher als Trolle in Onlinediskussionen in Erscheinung.
- Menschen neigen eher dazu, in einer Diskussion als Trolle aufzutreten, wenn vorangegangene Kommentare ein solches Verhalten zeigen.

⁴ Definiert als „flaming, grieving, swearing, or personal attacks, including behaviour outside the acceptable bounds defined by several community guidelines for discussion forums“ (Cheng et al. 2017: 2)

Ein direkter Zusammenhang zwischen der Trollforschung und der Empirie über die in Beiträgen geäußerten Emotionen lässt sich nicht nachweisen. Wir wissen nicht, ob die in Facebook-Beiträgen geäußerte Stimmung nur auf die von Nutzern in Facebook-Beiträgen geäußerte Stimmung wirkt oder auch auf anderes Verhalten. Aber es erscheint zumindest plausibel, dass es eine solche Wirkung geben könnte. Wenn es so wäre, würde mit der Reichweite von emotional negativ aufgeladenen Beiträgen die Wahrscheinlichkeit von Beleidigungen und persönlichen Angriffen in Diskussionen steigen.

5.3 Verzerrte Kognitionen formen die Datenbasis sozialer Medien

Bestätigungsverzerrung (*confirmation bias*) bezeichnet ein psychologisches Phänomen, das aus der Kognitions- und Gedächtnisforschung stammt: Menschen füllen Wissenslücken bevorzugt mit dem aus, was sie ohnehin schon glauben, und sie sind eher bereit, Informationen Glauben zu schenken, die ihren bereits vorhandenen Überzeugungen entsprechen (siehe z. B. Klayman 1995). Kahneman formulierte das Konzept so:

„Contrary to the rules of philosophers of science, who advise testing hypotheses by trying to refute them, people (and scientists, quite often) seek data that are likely to be compatible with the beliefs they currently hold. The confirmatory bias of System 1 favors uncritical acceptance of suggestions and exaggeration of the likelihood of extreme and improbable events“ (Kahneman 2012b).

Mit System 1 meint Kahneman automatische, schnelle Verarbeitungsprozesse ohne deliberative Rückkopplung (siehe Abschnitt 4.4.1).

Es gibt neben dem *confirmation bias* noch eine Reihe weiterer kognitiver Verzerrungen, bei denen davon auszugehen ist, dass sie mit Blick auf Medien- und Kommunikationsinhalte in ständiger Wechselwirkung mit ADM-Prozessen stehen. Eine davon sei hier beispielhaft kurz skizziert:

Das Konzept der *Verfügbarkeitsheuristik* (siehe Tversky und Kahnemann 1974) beschreibt die Tendenz, die Häufigkeit von Ereignissen zu überschätzen, für die sich die betroffene Person leicht konkrete Beispiele in Erinnerung rufen kann (*ease of processing*). So wird beispielsweise die Wahrscheinlichkeit von Flugzeugabstürzen kurz nach in der Berichterstattung dominierenden realen Flugzeugunglücken vorübergehend überschätzt. Zudem wird die Wahrscheinlichkeit von Ereignissen überschätzt, mit denen die betreffende Person persönliche Erinnerungen oder besonders eindrückliche Beispiele verbindet (vgl. Kahneman 2012b).

Diese Verzerrung führt beispielsweise dazu, dass Ehepartner, die man nach ihrem prozentualen Beitrag zur Hausarbeit befragt, verlässlich Prozentsätze nennen, die addiert mehr als 100 Prozent ergeben. Beide Partner erinnern sich episodisch eher an ihre eigenen Beiträge als an die des Ehepartners und überschätzen daher ihren Anteil. Gleichmaßen wird die Wahrscheinlichkeit von gewaltsamen Todesarten wie Unfällen regelmäßig massiv gegenüber tatsächlich häufigen Todesarten wie Schlaganfällen überschätzt, weil erstere dank medialer Berichterstattung und persönlicher Anteilnahme vielen Menschen kognitiv leichter verfügbar sind. Innerhalb von Diskussionen etwa in sozialen Netzwerken, deren Inhalte von ADM-Systemen sortiert werden, kann sich eine derartige kognitive Verzerrung sehr einfach und effektiv ausbilden: Wer beispielsweise immer wieder Berichte über von muslimischen Einwanderern begangene Straftaten zu sehen bekommt, weil er in der Vergangenheit besonders intensiv mit derartigen Berichten interagiert hat oder sich in einer sozialen Gruppe bewegt, die solche Vorfälle besonders intensiv kommentiert, der wird über kurz oder lang die Wahrscheinlichkeit überschätzen, dass Muslime Straftaten begehen.

In diesem letzten Beispiel spielt eine weitere kognitive Verzerrung eine Rolle, nämlich der *base rate neglect*: Unter Umständen wird auch die Wahrscheinlichkeit, dass ein bestimmtes Verbrechen von einem Muslim oder einer Person mit Migrationshintergrund begangen wurde, überschätzt, weil die Grundgesamtheit der entsprechenden Bevölkerungsgruppen bei der impliziten Kalkulation dieser Wahrscheinlichkeit außer Acht gelassen wird. Die Tatsache, dass es in Deutschland beispielsweise sehr viel mehr Nichtmuslime als Muslime

gibt, wird bei der Einschätzung der Wahrscheinlichkeit nicht einbezogen, was zu einer Überschätzung des Prozentsatzes von Straftaten führen kann, die von Muslimen begangen werden.

Sowohl die Verfügbarkeitsheuristik als auch *base rate neglect* werden in der Literatur oft als Komponenten eines umfassenderen Konzeptes betrachtet, das der US-amerikanische Wissenschaftler Herbert Simon *bounded rationality* genannt hat (siehe Simon 1979). Im Kern geht es in der Debatte seit mittlerweile vier Jahrzehnten um die Frage, ob kognitive Heuristiken wie die von Tversky und Kahneman beschriebenen tatsächlich immer irrational und irreführend sein müssen oder ob sie unter bestimmten Bedingungen, in denen der jeweiligen Person keine umfassenden Informationen zur Verfügung stehen, womöglich im Rahmen der gegebenen Bedingungen sogar nützlich sind (Gigerenzer und Gaissmaier 2011). Eine umfassende Erörterung dieses Themenkomplexes würde den Rahmen dieser Expertise sprengen.

Festzuhalten bleibt in jedem Fall: Es gibt umfangreiche empirische Evidenz dafür, dass Menschen Informationen häufig nicht rational verarbeiten und dass kognitive Verzerrungen insbesondere dann auftreten, wenn Entscheidungen und Urteile schnell, tendenziell emotional und ohne Reflexion, ja „automatisch“ getroffen und gefällt werden. Eine vergleichsweise neue, auf klassisch lerntheoretischen Erkenntnissen aufbauende Wissenschaft der digital unterstützten Verhaltensmodifikation namens *Captology* (siehe Abschnitt 5.4) hat explizit zum Ziel, Verhaltensbeeinflussung unter anderem durch eine maximale Vereinfachung und Beschleunigung der notwendigen kognitiv-verhaltensmäßigen Prozesse zu erleichtern (Eyal 2014). Es liegt daher nahe, dass die Designentscheidungen der Intermediäre kognitive Stile priorisieren, die wiederum das Entstehen kognitiver Verzerrungen begünstigen.

5.4 Plattformdesign beeinflusst menschliches Verhalten

Wenn Intermediäre ihre eigenen Systeme funktional verändern, hat das Auswirkungen, und zwar einerseits auf die Erfahrung, die Nutzer machen, und andererseits auf das Verhalten der Nutzer. Mit anderen Worten: Die Intermediäre beeinflussen mit jeder Designänderung auch die Signale, die sie selbst messen, um Relevanz zu beurteilen. Dafür gibt es eine Reihe klarer empirischer Belege.

Ein Beispiel: Als Facebook im März 2008 eine neue Funktion freischaltete, einen Kontaktempfehlungsmechanismus namens „People You May Know“ (PYMK, „Leute, die Sie vielleicht kennen“), hatte das erhebliche Folgen für die Art und Form der Vernetzung unter den Nutzern des Angebots: die Anzahl der neuen Verknüpfungen zwischen Nutzern innerhalb eines gegebenen Zeitraums (Malik und Pfeffer 2016). Die Forscher kommen zu dem Schluss, dass „nicht nur der Prozess des Beitritts zu sozialen Netzwerkseiten beobachtbare Eigenschaften des Netzwerks hervorruft, sondern auch die Art und Weise, wie die Gestaltung des Angebots Nutzer beeinflusst“. Der „People You May Know“-Mechanismus fußt auf einer algorithmischen Auswertung der digital abgebildeten Bekanntenkreise von Facebook-Nutzern. Auf Basis dieser Auswertung werden andere Nutzer vorgeschlagen, die der Betroffene möglicherweise auch kennen könnte. Dies führte augenscheinlich zu einem großen Anstieg der über das Netzwerk abgebildeten Verknüpfungen zwischen Nutzern. Die in einem sozialen Netzwerk abgebildeten sozialen Verknüpfungen seien ein „nicht naturalistisches Maß sozialer Beziehungen“, folgern Malik und Pfeffer.

Dieses Resultat macht einen weiteren wichtigen Punkt deutlich: Die Designentscheidungen der Intermediäre – hier die Einführung des PYMK-Features – wirken sich unter Umständen unmittelbar und massiv auf das Nutzerverhalten aus, im Beispiel die Anzahl der Verknüpfungen zu neuen „Freunden“, die hergestellt werden. Das Design der Intermediäre beeinflusst also stets auch die Signale, anhand derer sie selbst Relevanz messen.

Ein eindrucksvolles Beispiel für die Macht der Intermediäre lieferte eine Forschergruppe, an der auch Angestellte von Facebook beteiligt waren, im Jahr 2012. 61 Millionen Facebook-Nutzer wurden im Vorfeld der Wahlen zum US-Kongress im Jahr 2010 mit einer von Facebook selbst im eigenen Angebot verbreiteten Nachricht aufgefordert, sich zu informieren und zur Wahl zu gehen. Das Ergebnis in den Worten der Autoren:

„The results show that the messages directly influenced political self-expression, information seeking and real-world voting behaviour of millions of people. Furthermore, the messages not only influenced the users who received them but also the users’ friends, and friends of friends“ (Bond et al. 2012).

Konkret gehen die Autoren davon aus, dass 340.000 Personen oder 0,14 Prozent der Wahlberechtigten in den USA durch die Aktion zum Wählen animiert worden seien. Selbstverständlich ist dies ein in Demokratien wünschenswerter Effekt. Er zeigt aber auch, dass Intermediäre mit einer Reichweite in dieser Größenordnung unter Umständen wahlentscheidenden Einfluss ausüben können. Insbesondere, wenn man davon ausgeht, dass die durch solche Aktionen zusätzlich motivierten Wähler nicht gleichmäßig über das politische Spektrum verteilt sein dürften. Vor der potenziellen gesellschaftlich-politischen Wirkungsmacht anderer Intermediäre, nämlich von Suchmaschinenbetreibern, warnen Robert Epstein und Ronald Robertson. In einer Serie von Experimenten zeigten sie mit Wählern aus den USA und Indien, dass aktiv manipulierte Ergebnislisten das Wahlverhalten von Wahlberechtigten beeinflussen könnten (Epstein und Robertson 2015). Die Autoren selbst nennen ihren Befund *Search Engine Manipulation Effect* und geben den Anteil bis dahin unentschlossener Wähler, die mit derartigen Suchmaschinenmanipulationen in die eine oder andere Richtung beeinflusst werden könnten, mit 20 Prozent an. Diese Prozentangabe ist von anderer Seite aufgrund methodischer Kritikpunkte angezweifelt worden (Zweig 2017). Auch die Autorin dieser kritischen Betrachtung zur Arbeit von Epstein und Robertson, die Sozioinformatikerin Katharina Zweig, betrachtet die Ergebnisse aber als bedeutsam, insbesondere für Wahlen in Staaten mit Mehrheitswahlrecht. Der tatsächliche Effekt liege Zweig zufolge jedoch eher im Bereich von zwei bis vier Prozent.

Sowohl Google als auch Facebook können also unter Umständen gewaltigen Einfluss auf die von Nutzern wahrgenommenen Inhalte und mittelbar womöglich auch auf gesellschaftliche Prozesse ausüben. Auch weniger offenkundige Veränderungen eines Angebots können aber massive Auswirkungen auf Nutzerwahrnehmung und -verhalten haben.

Wie Kosinski et al. 2015 notierten:

„As users are more likely to interact with content and people suggested to them by Facebook, their behavior is driven not only by their intrinsic goals and motivations, but also (to some unknown extent) by the Facebook algorithms constantly adjusting their exposure to content and friends. For example, friends’ photos that appear on a given user’s Facebook news feed are clearly more likely to be liked. Essentially, largely unknown effects of personalization represent a general class of confounding variables characteristic for observational research and deserve further study“ (a. a. O.).

Auch für Google sind Fälle dokumentiert, in denen grundlegende Änderungen am Suchalgorithmus tief greifende Veränderungen in der Zusammenstellung von Ergebnislisten und damit auch der Erfolgswahrscheinlichkeit bestimmter Geschäftsmodelle zur Folge hatten, was sich wiederum auf das Verhalten von Seitenbetreibern und natürlich von Nutzern auswirkte. Google hatte eigenem Bekunden zufolge im Jahr 2011 angestrebt, „Inhalte von niedriger Qualität“ schlechter zu bewerten und damit in den Ergebnislisten weiter hinten anzuzeigen (Cutts 2011). Der Codename für die Änderung des Algorithmus war *Panda*. Ziel der Änderung war es, sogenannte *Content-Farmen*, also Webseiten, die keinerlei Inhalte von Wert, aber eine scheinbar hohe Suchmaschinenrelevanz aufwiesen, aus den Suchergebnissen zu verbannen. Solche Content-Farmen dienen ausschließlich monetären Zwecken, sie werden angelegt, um Besucher mithilfe suchmaschinenoptimierter Inhalte anzulocken und deren Aufmerksamkeit dann mittels auf den inhaltlich irrelevanten Seiten angezeigter Werbung zu monetarisieren.

Panda und auch weitere, in den darauffolgenden Jahren eingeführte umfassende Änderungen des Suchalgorithmus mit den Namen *Penguin* und *Hummingbird* führten in der Suchmaschinen-Optimierungsbranche zu teils wütenden Reaktionen (siehe z. B. Jenyns 2013). Ein ums andere Mal wurde deutlich, dass Google Geschäftsmodelle, die, wie in diesem Fall, auf einen Missbrauch des Suchalgorithmus setzen, von einem Tag auf den anderen zunichtemachen kann.

Wichtig ist im Zusammenhang dieser Expertise, dass jeder dieser grundlegenden Eingriffe in den Suchalgorithmus zwangsläufig massive – möglicherweise durchaus positive – Veränderungen für das Verhalten der Nutzer und damit für die von ihnen wahrgenommenen Medieninhalte mit sich bringt. Diversen Studien zufolge klickt eine überwältigende Mehrheit der Nutzer von Suchmaschinen ausschließlich auf Links, die auf der ersten Ergebnisseite erscheinen. In einer Untersuchung von van Deursen und van Dijk (2009) etwa zeigte sich, dass „91 Prozent [der Versuchspersonen] niemals über die erste Seite mit Suchergebnissen hinaus kamen“ (a. a. O.). Umgekehrt zeigen Analysen von Dienstleistern, die sich mit Suchmaschinenoptimierung und -marketing beschäftigen, dass schon die unterste, zehnte Position auf der ersten Ergebnisseite nur noch von einem sehr geringen Anteil aller Nutzer angeklickt wird. Unterschiedliche Erhebungen kommen für diese zehnte Position auf Werte zwischen 1 und 3,6 Prozent Durchklickrate (Dearinger 2011).

Eingriffe in die Systematik der Generierung solcher Ergebnisseiten wirken sich also aus Anbietersicht massiv auf die Wahrnehmbarkeit des eigenen Angebots und aus Nutzersicht massiv auf die wahrgenommenen Inhalte aus.

Der Einfluss der Intermediäre auf die Nutzererfahrung und damit auf die Gesellschaft als Ganzes, aber auch die Erfahrungen von Unternehmen und Organisationen, die sich auf Intermediäre als Reichweitenbringer verlassen, ist also groß. Bleibt die Frage: Sind die Entscheidungen und Veränderungen hier auch teilhaberelevant? Fest steht: Bei der Nutzung der neuen Intermediäre entstehen komplexe Wechselwirkungen zwischen Designentscheidungen der Betreiber und einer ganzen Reihe anderer Faktoren. Darunter sind psychologische Mechanismen, Merkmale von Inhalten und Versuche der Manipulation von außen. Sogar das politische System des Landes, in dem der jeweilige Nutzer lebt, spielt eine Rolle für die Dynamik, die innerhalb dieses komplexen Systems von Wechselwirkungen entsteht.

Die in den Abschnitten 0 und 5.2 aufgeführten Faktoren, die sich auf die Interaktionswahrscheinlichkeit und -intensität von Nutzern digitaler Angebote auswirken, sind den Intermediären selbstverständlich bekannt. Doch nicht nur das: Schon seit 15 Jahren beschäftigt sich ein neuer Zweig der Psychologie der Mensch-Maschine-Interaktion mit der Frage, wie digitale Technik menschliches Verhalten beeinflussen helfen kann. Ihr geistiger Vater, B.J. Fogg, hat diese Disziplin *Captology* getauft (als Kurzform von *Computers as Persuasive Technologies*). Im Vorwort zur ersten Ausgabe seines Buches „Persuasive Technology“ schrieb der berühmte Sozialpsychologe Philip Zimbardo, einer von Foggs Lehrern, es handele sich hier um den Ausgangspunkt einer neuen wissenschaftlichen Disziplin: „The very first book on a totally new field of intellectual inquiry, one that has important practical implications for a host of people across many different domains, all centered around persuasion“ (Fogg 2003).

Fogg zeigt in dem Buch, das heute als Meilenstein in der Geschichte der wissenschaftlichen Auseinandersetzung mit Mensch-Maschine-Interaktionen gilt, wie klassische lerntheoretische Konzepte und Verfahren aus Behaviorismus und Verhaltenstherapie mithilfe digitaler Systeme effizienter und effektiver auf das Verhalten von Menschen einwirken können. Als Foggs Buch erschien, existierten weder Facebook noch Instagram, und das Smartphone, wie wir es heute kennen, war noch nicht erfunden. Doch zahlreiche spätere Gründer und Mitarbeiter diverser Giganten des Silicon Valley durchliefen seine Kurse an der University of Stanford, darunter beispielsweise einer der Gründer von Instagram, Mike Krieger. Krieger entwickelte in Foggs Kurs gemeinsam mit einem Kommilitonen sogar eine auf Foggs Ideen basierende Anwendung zum Teilen von Fotos (Leslie 2016).

Die von Fogg entwickelten Methoden werden heute von nahezu allen Entwicklern von Endkunden-Softwareprodukten selbstverständlich angewendet. Der Autor des Buches „Hooked: How to build habit-forming products“, Nir Eyal, ist ein Alumnus von Foggs Labor in Stanford (Eyal 2014). Der Begriff *Habit* kann im Englischen sowohl für „Angewohnheit“ als auch für „Sucht“ stehen.

Gefeiert wird Eyal vor allem von Investoren, Start-up-Gründern und Vertretern der Werbe- und Marketingbranche, weniger von Wissenschaftlern. Seine angewandte Wissenschaft der digitalen Verhaltensmodifikation ist auf eine Art und Weise erfolgreich, die seinen Lehrer Fogg heute sehr beunruhigt. „I look at some of my former students

and I wonder if they're really trying to make the world better, or just make money“, wird er in einem Zeitschriftenartikel zitiert. „What I always wanted to do was un-enslave people from technology“ (Leslie 2016).

Captology betrachtet aus klassischen Lerntheorien bekannte Komponenten wie Auslösereize (*trigger*), Verhalten (*action*), Belohnung (*reward*) und Gewohnheit (*habit*), differenziert sie aus und überträgt sie in konkrete Anwendungsszenarien innerhalb digitaler Anwendungen. Ein kurzes Zitat aus *Hooked* soll hier der Illustration dienen:

„Facebook provides numerous examples of variable social rewards. Logging in reveals an endless stream of content friends have shared, comments from others, and running tallies of how many people have ‚liked‘ something. The uncertainty of what users will find each time they visit the site creates the intrigue needed to pull them back again. While variable content gets users to keep searching for interesting tidbits in their News Feeds, a click on the ‚Like‘ button provides a variable reward for the content’s creators“ (Eyal 2014: 78).

Ein zentrales Ziel der Optimierungsbemühungen der *Captologen* ist es, Hindernisse zwischen dem Nutzer und einem bestimmten Verhalten zu beseitigen. Auf Fogg's Theorien aufbauend, formuliert Eyal so:

„The pattern of innovation shows that making a given action easier to accomplish spurs each successive phase of the web, helping to turn the formerly niche behavior of content publishing into a mainstream habit. As recent history of the web demonstrates, the ease or difficulty of doing a particular action affects the likelihood that a behavior will occur. To successfully simplify a product, we must remove obstacles that stand in the user’s way“ (Eyal 2014: 51/52).

Vor dem Hintergrund der von Kahneman (2012) und anderen postulierten distinkten kognitiven Verarbeitungssysteme System 1 (schnell, intuitiv, automatisch, nicht anstrengend, anfällig für Fehler und Manipulation) und System 2 (langsam, deliberativ, geordnet, mit Anstrengung verbunden) ist dieses Gestaltungsziel erhellend: Ein zentrales Designprinzip der wissenschaftlichen Vordenker des Silicon Valley betrifft explizit die Erleichterung von System-1-Kognition und die aktive Umgehung von System-2-Kognition. Das hat vorteilhafte Folgen, wenn man als Erfolgskriterien Messwerte wie Klick-, Teil- oder Like-Raten, Publikationsfrequenz oder andere quantitative Maße des Nutzerverhaltens heranzieht. Legte man andere Kriterien, wie tiefe Verarbeitung wahrgenommener Inhalte, Verständnis oder die Qualität geteilter Inhalte an, käme man wohl zu anderen Designentscheidungen (siehe Abschnitt 7.2).

Hypothesen, die sich aus solchen lerntheoretischen Grundlagen ableiten lassen, werden von den Intermediären permanent untersucht. Sie unternehmen permanente experimentelle Manipulationen mit dem expliziten Ziel der Optimierung der Nutzerinteraktion: „At Facebook, we run over a thousand experiments each day“, so Facebook-Wissenschaftler Eytan Bakshy (2014) in einem Blogbeitrag. Manche dieser Experimente, so Bakshy, dienten der Optimierung kurzfristiger Ergebnisse, während andere die Basis für langfristige Designentscheidungen bildeten.

Auch von Google ist bekannt, dass schon kleinste Änderungen an der Benutzeroberfläche mit aufwendigen Experimenten im Livebetrieb und mit realen Nutzern vorbereitet werden – etwa, wenn es um die Frage geht, in welchem Blauton die Buchstaben von Textanzeigen gehalten werden sollen. Eine solche Farbänderung brachte dem Konzern 2014 Medienberichten zufolge ein Umsatzplus von 200 Millionen Dollar ein (Hern 2014).

Im Zusammenhang mit dem Thema dieses Arbeitspapiers ist diese Vorgehensweise vor allem aus einem Blickwinkel heraus von Bedeutung: Intermediäre verändern im Zuge der permanenten experimentellen Optimierung ihrer Angebote fortwährend Variablen, die sich auf die Messung ihrer eigenen Relevanzkriterien auswirken.

Wenn Interaktionshäufigkeit und -intensität gesteigert werden, wächst damit die Anzahl von Klicks auf den Like-Button bei Facebook oder die Klickraten von Links in Google-Treffern. Die Frage ist aber, welche Art von Auseinandersetzung mit Inhalten diese Variablen reflektieren. Eingestellt werden die Systeme in jedem Fall nicht

im Sinne einer optimal informierten Öffentlichkeit, sondern im Dienste der kommerziellen Interessen der Betreiber.

5.5 Gemessene Reichweite muss nicht gleich realer Reichweite sein

Ein Faktor, der viele der bislang geschilderten Messwerte und Ergebnisse in einem anderen Licht erscheinen lässt, wurde hier bislang ausgespart: die Möglichkeit, viele der beschriebenen Werte, die Intermediäre zur Messung vermeintlicher Relevanz heranziehen, mithilfe automatisierter Verfahren zu beeinflussen und zu verfälschen. Das derzeit vermutlich wichtigste Werkzeug dazu sind sogenannte *Bots*, automatisierte Softwarelösungen, die das Verhalten menschlicher Nutzer vortäuschen sollen. Manche dienen kommerziellen oder kriminellen Zwecken, sie „klicken“ etwa auf Werbeanzeigen, um Erlöse zu generieren oder das Werbebudget eines Konkurrenten mit wertlosem Traffic aufzubrauchen, andere durchsuchen das Netz nach Verwundbarkeiten, die Einfallstore für Hacker-Angriffe bieten könnten, wieder andere extrahieren beispielsweise Preisinformationen von den Seiten kommerzieller Konkurrenten, um so Wettbewerbsvorteile zu ermöglichen (*scraping*) (siehe z. B. Zeifman 2017).

Im Kontext dieser Expertise sind vor allem zwei Sorten von Bots relevant: sogenannte *Klickbots*, die realen Traffic auf Webseiten simulieren, und die jüngste Entwicklung in diesem Feld, sogenannte *Social Bots*⁵. Bei Letzteren handelt es sich um teilautonom handelnde Systeme, die sich auf sozialen Netzwerkplattformen ähnlich verhalten wie reale Nutzer, in Wirklichkeit aber nur Softwarelösungen sind, die Vorgaben folgen (Kollanyi, Howard und Woolley 2016). Solche Bots setzen beispielsweise automatisiert bestimmte Twitter-Nachrichten ab, oft angereichert mit ausgewählten Hashtags, oder retweeten Nachrichten anderer Accounts, um ihnen damit mehr Reichweite und vermeintliche Relevanz zu verschaffen. Bessi und Ferrara schätzen, dass etwa im Kontext der US-Präsidentenwahl 2016 allein auf Twitter rund 400.000 Bots aktiv waren, die insgesamt circa 3,8 Millionen Tweets abgesetzt hätten, „was etwa einem Fünftel der gesamten Konversation [über die Präsidentenwahl auf Twitter, Anm. d. Verf.] entspricht“. Kollanyi, Howard und Woolley (2016) schätzen sogar, dass im Vorfeld der US-Wahl „etwa ein Drittel der Pro-Trump-Aktivität auf Twitter von Bots und anderen hochautomatisierten Accounts getrieben wurde, verglichen mit etwa einem Fünftel der Pro-Clinton-Aktivität“.

Howard und Kollanyi (2016) fanden auch für die Debatte über den Austritt Großbritanniens aus der Europäischen Union (*Brexit*) einen substanziellen Anteil von Twitter-Nachrichten, die sie auf automatisierte Aktivitäten zurückführen. Weniger als ein Prozent aller Twitter-Accounts, die relevante Hashtags wie *#Brexit* oder *#StrongerIn* benutzt hätten, zeichneten für beinahe ein Drittel aller zu diesen Themen veröffentlichten Tweets verantwortlich. Die beobachteten Bots seien während der Phase vor der Abstimmung vor allem genutzt worden, um „Botschaften zu verstärken“, so die Autoren, etwa indem sie dafür sorgten, dass „die Familie der Hashtags, die mit Argumenten für das Verlassen der EU verknüpft waren, dominierte“ (a. a. O.).

Hegelich und Janetzko (2016) untersuchten ein Bot-Netz, das sich thematisch auf den Ukraine Konflikt und Russland konzentriert und augenscheinlich dazu dient, pro-russische Botschaften abzusetzen. Die Autoren kommen zu dem Schluss, dass das Verhalten der untersuchten Bots „nicht von einer einfachen, deterministischen Struktur von Befehl und Gehorsam zwischen einem menschlichen Botmaster und einer Armee von Bots“ geprägt sei. Es handele sich vielmehr um das Ergebnis „komplexer Algorithmen, die zu einem hohen Grad an Autonomie bei den Bots führen“. Die Programme verhielten sich so, dass sie auf den ersten Blick wie menschliche Nutzer wirkten, und folgten gleichzeitig abstrakten Regeln wie „nimm einen populären Tweet und hänge folgende Hashtags an“. Die Tarnstrategien dieser Systeme machten es „extrem schwer, die Bots zu identifizieren und ihr politisches Ziel zu erkennen“. Derartige Bot-Netze, so die Schlussfolgerung der Autoren, seien „eine neue Entwicklung von hoher politischer Relevanz“ (a. a. O.).

⁵ Transparenzhinweis: Einer der Autoren dieser Expertise, Christian Stöcker, ist aktuell an einem vom BMBF geförderten Forschungsprojekt zur Entdeckung und Abwehr automatisierter Propagandaangriffe („PropStop“) beteiligt (siehe <https://www.wi.uni-muenster.de/de/news/1975-bmbf-projekt-zur-propaganda-erkennung-online-medien-gestartet>).

Die zitierten Studien beziehen sich allesamt auf Twitter, vor allem, weil sich Bots dort vergleichsweise leicht implementieren, aber auch erkennen lassen. Doch auch auf Facebook und anderen Intermediären gibt es derartige, als normale Nutzer getarnte, in Wahrheit automatisierte Accounts, die bestimmten politischen oder kommerziellen Zwecken dienen (siehe z. B. Lill et al. 2012).

Für das Thema dieser Expertise ist das Phänomen *Social Bots* vor allem aus einem Grund relevant: Bots sind geeignet, die Signale zu verzerren, anhand derer die Intermediäre Relevanz messen. Beispiel *Fake News*: Im US-Wahlkampf erzeugten diverse frei erfundene Meldungen hohe Interaktionszahlen bei Facebook, weit höhere als korrekte Berichterstattung in traditionellen Medien. Das bekannteste Beispiel ist die frei erfundene Behauptung, Papst Franziskus habe sich dafür ausgesprochen, für Donald Trump zu stimmen (vgl. Silverman 2016). Womöglich wurden die gemessenen Reichweite- und Interaktionswerte jedoch massiv von Bots verfälscht. Wie viele Likes und Shares von echten Nutzern stammen, ist bislang nicht zu klären.

Wie die obige kurze Zusammenstellung von Studien bereits anschaulich zeigt, sind solche automatisierten Propagandasysteme gerade bei politisch umkämpften Themen massiv im Einsatz. Es erscheint daher plausibel, dass Bots, die sich beispielsweise gemäß der von Hegelich und Janetzko (2016) beschriebenen Maxime „Nimm einen populären Tweet und hänge folgende Hashtags an“ verhalten, Desinformationsartikeln über Intermediäre massive zusätzliche Verbreitung verschaffen könnten. Das wiederum könnte sich auf die Einstufung der entsprechenden Links oder Posts durch die ADM-Systeme der Intermediäre auswirken. Die Aktivitäten der Sortieralgorithmen und der algorithmisch betriebenen Manipulations-Bots könnten auf diese Weise also in eine Spirale gegenseitiger Verstärkung geraten. Das so entstehende Bild der digitalen Öffentlichkeit hätte dann mit einem Abbild der tatsächlichen Verhältnisse nur noch wenig gemein. Gemessene Reichweite muss also nicht immer gleich realer Reichweite sein.

5.6 Polarisierung wächst, doch da wirkt viel mehr als nur das Internet

Der Begriff *Filterblase* geht auf den US-amerikanischen Autor Eli Pariser (2011) zurück. Die Filterblase definiert Pariser selbst so: „Das persönliche Informationsuniversum, das Sie online bewohnen – einzigartig und nur für Sie aufgebaut von den personalisierten Filtern, die das Web jetzt antreiben“ (a. a. O.) Mit personalisierten Filtern meint Pariser in etwa das, was wir in dieser Expertise als ADM-Prozesse bezeichnen: individuell auf einzelne Nutzer zugeschnittene Filterung und Sortierung von Inhalten auf Basis algorithmischer Verfahren.

Die Idee, dass digitale Kommunikationsangebote eine Verengung des medialen Weltbildes befördern könnten, stammt nicht von Pariser. Sie wurde bereits 2001 von dem US-amerikanischen Rechtswissenschaftler Cass Sunstein (2001; 2008) thematisiert. Sunstein spricht von „Informationskokons“ oder „Kommunikationsuniversen, in denen wir nur hören, was wir auswählen, was uns angenehm ist, was uns gefällt“ (Sunstein 2008). Ihm geht es dabei allerdings vor allem um Personalisierung auf Basis persönlicher Auswahlentscheidungen, zunächst noch weniger um die Wirkung algorithmischer Selektionssysteme. Sunstein formulierte eine Sorge, die auch in neueren Arbeiten zu diesem Themenkomplex immer wieder durchklingt: dass personalisierte Informationsuniversen zu politischer Radikalisierung führen könnten und dass der Prozess der demokratischen Willensbildung beeinträchtigt werden könnte, wenn Nutzer von Intermediären keine ausreichende Vielfalt konkurrierender Standpunkte mehr wahrnehmen. Der Gesellschaft könne auf diese Weise der „soziale Kitt“ abhanden kommen, so Sunstein, in Übereinstimmung etwa mit Jürgen Habermas, der 2008 warnte, das Publikum zerfalle „im virtuellen Raum in eine riesige Anzahl von zersplitterten, durch Spezialinteressen zusammengehaltene Zufallsgruppen. Auf diese Weise scheinen die bestehenden nationalen Öffentlichkeit eher unterminiert zu werden“ (Habermas 2008).

Bei Facebook war man insbesondere mit der Vision Parisers von einer durch weitreichende Personalisierung zersplitterten Gesellschaft augenscheinlich unzufrieden. 2015 erschien bei *Science* eine Studie, die anhand eines gewaltigen Facebook-Datensatzes prüfen sollte, ob Nutzer eines Intermediärs wirklich vor allem ideologisch zu ihren eigenen Einstellungen passende Inhalte zu sehen bekommen (Bakshy, Messing und Adamic 2015). Die

drei Autoren sind Angestellte von Facebook, was für sich genommen in der Wissenschaftswelt bereits als Politikum galt. *Science* stellte dem Forschungsbericht deshalb einen weiteren Artikel eines unabhängigen Wissenschaftlers zur Seite. David Lazer von der Northeastern University lobte darin die Bereitschaft von Facebook, sich der Diskussion zu stellen, warnte aber auch vor der Gefahr, dass „die einzigen, die Facebook erforschen können, Forscher bei Facebook sind“ (Lazer 2015).

Die Autoren selbst kamen in ihrer Arbeit zu dem Schluss, dass das Filterblasen-Problem eher ein imaginäres sei: „Unsere Arbeit legt nahe, dass Personen über soziale Medien mehr gegensätzlichen [politischen] Diskursen ausgesetzt werden, als dies in der digitalen Realität der Fall wäre, die sich manche ausmalen“ (Bakshy, Messing und Adamic 2015). An dieser Textstelle findet sich ein Literaturverweis auf Parisers Buch.

Tatsächlich aber zeigen die Daten der Facebook-Wissenschaftler, dass der Algorithmus zu einer gewissen Verzerrung hinsichtlich der politischen Ausrichtung wahrgenommener Inhalte führte: „Konservative sehen etwa fünf Prozent weniger Inhalte aus dem anderen politischen Lager im Vergleich zu dem, was ihre Freunde tatsächlich teilen, während Liberale etwa acht Prozent weniger ideologisch anders gefärbte Inhalte sehen“ (Bakshy, Messing und Adamic 2015). Der Einfluss der individuellen Entscheidungen der Nutzer sei aber weit größer, so die Autoren. Ja, solche individuellen Entscheidungen schränkten „die Konfrontation mit Inhalten, die der eigenen Haltung zuwiderlaufen, stärker ein als Algorithmen“ (a. a. O.).

Sowohl diese Schlussfolgerung als auch die methodische Herangehensweise der Facebook-Forschergruppe zog unmittelbar nach Erscheinen des Artikels heftige Kritik einiger Sozialwissenschaftler nach sich (zusammenfassend siehe Lumb 2015). Insbesondere wurde kritisiert, dass Bakshy, Messing und Adamic den signifikanten Effekt der algorithmischen Sortierung kleinzureden versuchten, indem sie ihn mit dem Effekt der Entscheidungen der Nutzer selbst verglichen. Das eine müsse unabhängig vom anderen betrachtet werden, so etwa die Soziologin Zeynep Tufekci (2015). Neu, bemerkenswert und signifikant sei schließlich der Effekt der algorithmischen Sortierung auf die Zusammensetzung der wahrgenommenen Medieninhalte.

Letztlich spielt vermutlich die **Wechselwirkung individueller Entscheidungen mit ADM-Systemen** eine zentrale Rolle bei der Frage, ob sich das Weltbild eines Facebook-Nutzers durch den Intermediär verengt oder nicht.

Eine Forschergruppe aus Wissenschaftlern vom italienischen Institute for Advanced Study in Lucca und Kollegen von anderen Einrichtungen haben die Ausbreitung von Gerüchten, insbesondere Verschwörungstheorien, in sozialen Netzwerken untersucht. Das Team kommt über die Auswertung von Facebook-Daten zu dem Schluss, Nutzer hätten die Tendenz, sich dort in Interessengemeinschaften zu sammeln, sodass sie vor allem entsprechende Inhalte zu sehen bekämen. Das führe zu „Bestätigungsverzerrung, Spaltung und Polarisierung“. Im Kontext sozialer Medien führe der *confirmation bias* (siehe Abschnitt 5.3) „zur Verbreitung von verzerrten Narrativen, angefacht von unbestätigten Gerüchten, Misstrauen und Paranoia“, so Michela Del Vicario et al. (2016).

In einer weiteren Studie zeigten einige an der eben zitierten Studie beteiligte Forscher zudem, dass gerade Verschwörungstheoretiker auf scheinbar paradoxe Weise reagieren, wenn sie mit Informationen konfrontiert werden, die ihre Ansichten infrage stellen könnten: Sie ignorieren Fakten, die der Verschwörungstheorie widersprechen – oder sie wenden sich sogar noch stärker der Echokammer Gleichgesinnter zu (Zollo et al. 2015). Die Arbeiten der italienischen Forscher legen nicht zuletzt eins nahe: Die Inhalte, die Menschen in sozialen Netzwerken teilen, spielen eine wichtige Rolle für das Identitätsmanagement (siehe z. B. Schmidt 2016). Man teilt, was zum eigenen Weltbild passt (siehe auch An, Quercia und Crowcroft 2013).

Eine wichtige Frage in diesem Zusammenhang betrifft die Konfrontation mit den eigenen Einstellungen entsprechenden oder widersprechenden Informationen, insbesondere bei bereits radikalisierten Personen. Magdalena Wojcieszak (2010) kam in einer Studie mit Teilnehmern US-amerikanischer Onlineforen für Neonazis zu dem Schluss, dass „die Nutzer solcher Plattformen die Angaben, die sie dort vorfänden, verwenden, um „Gegenargumente zu entkräften und Erklärungen zu generieren, die ihre eigenen Vorannahmen stärken“

a. a. O.). Die Kommunikationswissenschaftlerin zitiert ein anonymes Mitglied eines der untersuchten rechtsextremen Foren mit dem Satz: „Wir existieren in einer Welt voller Einflüsse, aber wir sind überwiegend immun dagegen, weil wir uns selbst fortgebildet haben.“ Unter *Fortbildung* wird dort etwa das Sammeln von Argumenten für die Leugnung des Holocaust oder die Überlegenheit der „weißen Rasse“ verstanden. Zumindest bei bereits teilweise radikalisierten Personen kann die Verfügbarkeit extremistisch geprägter Kommunikationsumgebungen – seien es Internetforen oder Facebook-Gruppen – also offenbar zu einer weiteren Verstärkung der Radikalisierung führen (a. a. O.).

Empirische Arbeiten, die klare Belege für oder gegen eine weitreichende Wirkung digitaler Kommunikationsplattformen oder von ADM-Systemen auf die demokratische Willensbildung nachweisen, sind allerdings nach wie vor dünn gesät. Es gibt sowohl Hinweise auf eine entsprechende Verengung der rezipierten Medieninhalteintermediäre (z. B. An, Quercia und Crowcroft 2013) als auch Autoren, die zur Gelassenheit mahnen, da es „bislang wenig empirische Evidenz“ gebe, „die Sorgen über Filterblasen rechtfertigen würden“ (Borgesius et al. 2016).

Eine der wenigen groß angelegten Studien zum Thema kommt zu einem ambivalenten Schluss. Flaxman, Goel und Rao (2016) untersuchten den Medienkonsum einer Stichprobe von 50.000 in den USA lebenden Internetnutzern. Die zentrale Schlussfolgerung im Wortlaut:

„We showed that articles found via social media or web-search engines are indeed associated with higher ideological segregation than those an individual reads by directly visiting news sites. However, we also found, somewhat counterintuitively, that these channels are associated with greater exposure to opposing perspectives. Finally, we showed that the vast majority of online news consumption mimicked traditional offline reading habits, with individuals directly visiting the home pages of their favorite, typically mainstream, news outlets. We thus uncovered evidence for both sides of the debate, while also finding that the magnitude of the effects is relatively modest“ (a. a. O.).

Anders als Pariser (2011) und Sunstein (2001;2008) sind die Autoren der Meinung, dass „soziale Medien und Suchmaschinen zwar zur Segregation beizutragen scheinen, der Mangel an Binnenvariation bei einzelnen Nutzern jedoch primär auf direkte Seitenaufrufe [von Online-Nachrichtenangeboten; Anm. d. Verf.] rückführbar scheint“ (a. a. O.). Die „parteiische“ Berichterstattung, die Nutzer wahrnehmen, stamme zudem weiterhin von Angeboten, die die Autoren eher dem „Mainstream“ zuordnen, „von der ‚New York Times‘ im linken Spektrum bis zu Fox News im rechten“. Ideologisch extremere Seiten wie *Breitbart News* hätten „anscheinend keine qualitative Auswirkungen auf die Dynamik des Nachrichtenkonsums“ (a. a. O.). Dabei ist allerdings zu berücksichtigen, dass die Daten der zitierten Studie in einem Zeitraum zwischen März und Mai 2013 erhoben wurden, also lange vor der Endphase des jüngsten US-Präsidentenwahlkampfes.

Eine 2017 erschienene Studie kommt hinsichtlich der Bedeutung von *Breitbart News* und ähnlichen Publikationen zu einem deutlich anderen Schluss: Die politisch extremen Nachrichtenangebote hätten moderateren Angeboten wie Fox News mit ihrer „hyper-parteiischen“ Berichterstattung die Themen diktiert und sogar die Berichterstattung liberalerer Medien massiv beeinflusst, „insbesondere mit Blick auf die Berichterstattung über Hillary Clinton“. Dabei hätten die sozialen Medien als Verbreitungskanal eine zentrale Rolle gespielt. Daten über die Social-Media-Verbreitung von 1,25 Millionen Artikeln, die zwischen dem 1. April 2015 und der Wahl im November 2016 online veröffentlicht wurden, zeigten, dass „sich um ‚Breitbart‘ herum ein rechtes Mediennetzwerk als abgegrenztes und isoliertes Mediensystem entwickelt hat, das soziale Medien als Rückgrat benutzt, um einen hyper-parteiischen Blickwinkel in die Welt zu senden“ (Benkler et al. 2017).

Weitgehender Konsens herrscht unter den in diesem Bereich arbeitenden Wissenschaftlern darüber, dass Mediennutzer insbesondere in politisch stark polarisierten Gesellschaften wie den USA immer stärker ideologisch einschränken, wo sie sich informieren, zeigten zum Beispiel Iyengar und Hahn (2009), dass sich US-Republikaner Fox News zu- und von eher liberal eingestellten Medien wie CNN und NPR abwenden, während sich unter Demokraten das umgekehrte Muster ergab. Dieser Mechanismus könne angesichts der durch das

digitale Medienangebot gewachsenen Wahlmöglichkeiten „zu einer weiteren Polarisierung des Nachrichtenpublikums beitragen“, so die Autoren. Gerade politisch extreme Publikationen wie Breitbart News sind erst durch digitale Distributionsmöglichkeiten einem breiteren Publikum zugänglich. Diese Polarisierung sei dennoch keine reine Folge des Internet, so Iyengar und Westwood (2015) in einer weiteren Studie: Die „affektive Polarisierung hinsichtlich der eigenen und der anderen Partei“ sei bereits „in den vergangenen vier Jahrzehnten substanziell angewachsen“.

Die Autoren ließen beispielsweise bekennende Anhänger der Republikaner und der Demokraten angebliche Stipendien an in Wahrheit fiktive Highschool-Abgänger vergeben. Die Noten der fiktiven Bewerber erwiesen sich dabei als zweitrangig, die angebliche Parteizugehörigkeit dagegen spielte bei den Entscheidungen eine zentrale Rolle: Republikaner vergaben Stipendien an Republikaner, Demokraten an Demokraten. Die in den USA oft sehr aggressive Rhetorik politischer Akteure und die Bereitschaft zur Herabsetzung des politischen Gegners habe dazu geführt, dass Parteigänger beider großen Parteien das Gefühl hätten, „Animositäten gegenüber den Parteigängern der anderen Seite freimütig ausdrücken und diese diskriminieren zu können“, so die Schlussfolgerung der Autoren. Ein auf Mehrheitswahlrecht basierendes Zweiparteiensystem wie in den USA begünstigt möglicherweise diese ausgeprägte Form der Polarisierung (siehe z. B. Trilling, van Klingeren und Tsfati 2016).

Auch Boxell (2017) weisen darauf hin, dass diese wachsende Polarisierung zumindest nicht allein dem Internet angelastet werden kann:

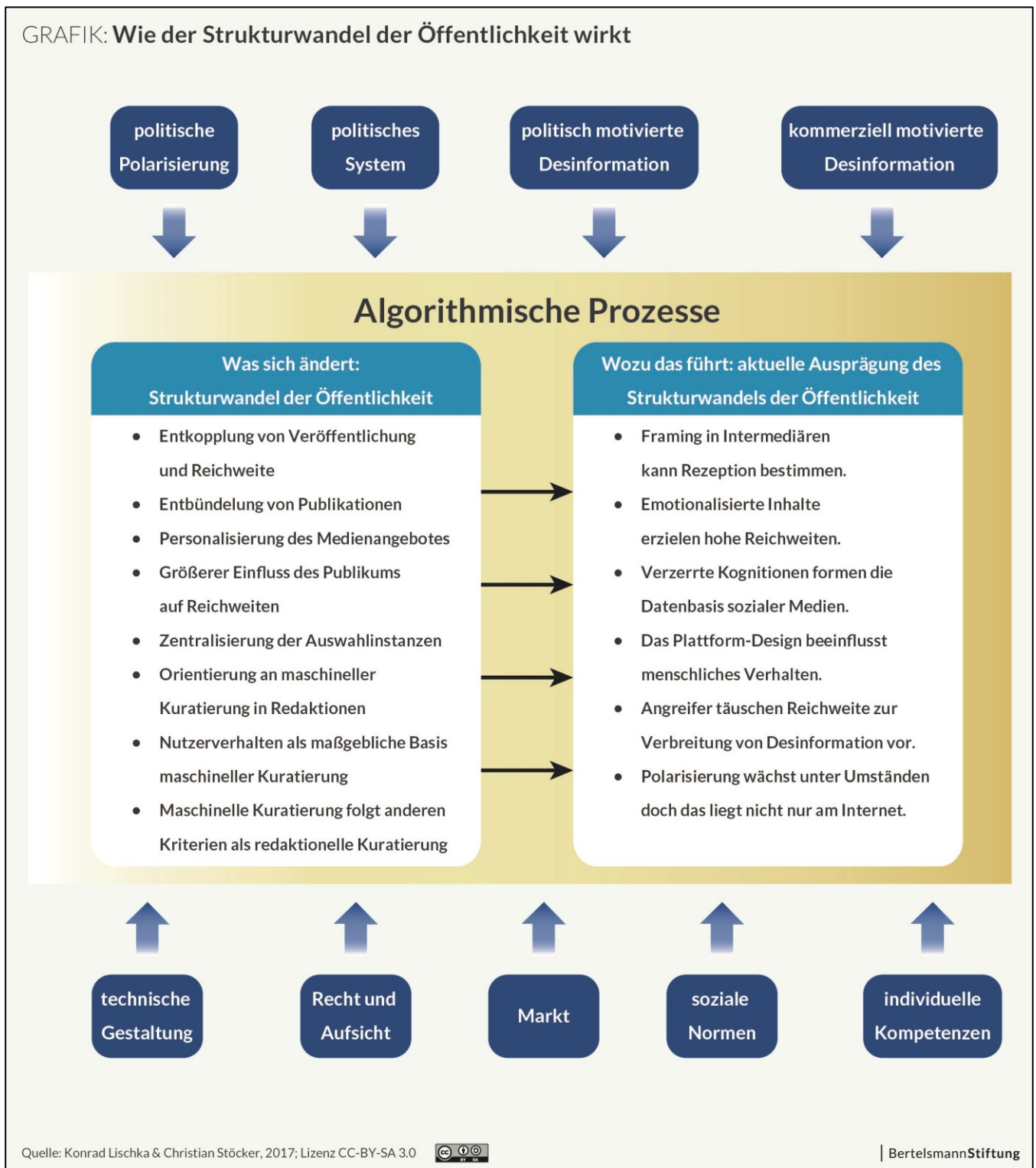
„We find that the increase in polarization is largest among the groups least likely to use the internet and social media. A normalized index of our nine polarization measures increases by 0.18 index points overall between 1996 and 2012. Among respondents aged 75 and older, the increase is 0.38 index points, whereas for adults under age 40, the increase is 0.05 index points. Across intermediate age groups, the growth in polarization is consistently higher among older respondents. Polarization increases more for the old than the young in eight of the nine individual measures. A similar pattern emerges for groups of respondents divided by our broader index of predicted internet use“ (a. a. O.).

Insgesamt zeigen sich klare Anhaltspunkte für eine komplexe Wechselwirkung zwischen Intermediären, ADM-Systemen, gewachsenem und diversifiziertem Medienangebot, politischer Polarisierung und dem politischen System eines Landes, wie etwa Borgesius et al. (2016) betonen: „The effect of personalised news on polarisation is conditional on the political system.“ So bilanziert eine aktuelle Studie aus den Niederlanden, in denen es kein Zweiparteiensystem gibt, dass sich dort keine eindeutigen Bezüge zwischen zur eigenen politischen Einstellung kongruenter Information und Polarisierung herstellen lassen. Mit anderen Worten: Ob jemand extremere politische Positionen vertritt, hängt in den Niederlanden nicht davon ab, ob er selektiv Medieninhalten mit einer bestimmten politischen Ausrichtung ausgesetzt wird (Trilling, van Klingeren und Tsfati 2016).

5.7 Zwischenfazit: Komplexe Wechselwirkung, verstärkt durch algorithmische Prozesse

Wie die vorangegangenen Abschnitte zeigen, entsteht Öffentlichkeit vermittelt durch digitale Intermediäre als Ergebnis einer komplexen Wechselwirkung unterschiedlicher Faktoren. Auf der individuellen Ebene spielen bestimmte Nutzungsweisen eine Rolle, die wiederum durch die Gestaltung der jeweiligen Angebote begünstigt werden. So werden Medieninhalte etwa häufig nur auf Basis ihrer Überschriften und Anreißertexte weitergereicht, ohne dass der Weiterreichende sie tatsächlich vollständig rezipiert hat. Zudem werden stark emotionalisierende Inhalte besonders häufig geteilt und intensiv diskutiert.

Abbildung 3: digitale Öffentlichkeit – Strukturwandel, Einflussfaktoren und aktuelle Ausprägung



Diese Signale gehen in die algorithmische Sortierung von Inhalten ein und erhöhen so die Wahrscheinlichkeit, dass besonders emotionalisierende Inhalte ein noch breiteres Publikum erreichen. Aus der Psychologie bekannte kognitive Verzerrungen wie die Verfügbarkeitsheuristik treten vermutlich in Wechselwirkung mit solchen Mechanismen: Im Zweifel wird das Weltbild einzelner Nutzer maßgeblich von Inhalten bestimmt, die weder sie selbst noch derjenige, dessen Weiterleitungsaktion ihnen diese Inhalte verfügbar gemacht hat, tatsächlich vollständig rezipiert haben. Solche Verzerrungen und ohne großen kognitiven Aufwand durchgeführten Aktionen werden begünstigt von den Gestaltungszielen und -prinzipien der Intermediäre: Als Leitwert gilt eine möglichst

ausgeprägte Einfachheit für alle angestrebten Aktivitäten. Hindernisse, die zu kognitiver Verlangsamung führen könnten, werden identifiziert und nach Möglichkeit beseitigt.

Externe Akteure schalten sich mit automatisierten Manipulationsmechanismen, zum Beispiel Bots, in diese Prozesse ein und sorgen aus kommerzieller oder propagandistischer Motivation heraus dafür, dass bestimmte Inhalte und Konzepte durch ADM-Prozesse mehr Aufmerksamkeit bekommen, als das ohne sie der Fall wäre. Schließlich führt das Zusammenspiel individueller Nutzerentscheidungen, algorithmischer Sortiersysteme und psychologischer Faktoren zumindest bei Teilen der Nutzerschaft zu einer verstärkten Polarisierung sowohl im Hinblick auf wahrgenommene Inhalte als auch im Hinblick auf gesellschaftliche und politische Einstellungen. ADM-Prozesse sind in diesem komplexen Gefüge nur ein Faktor, aber einer, der in Wechselwirkung mit allen übrigen Faktoren steht und somit als Verstärker sowohl menschlich-kognitiver Unzulänglichkeiten als auch gezielter technischer Manipulationen gelten kann.

6 Was sein soll: Leitwerte für Öffentlichkeit

Die Analyse zentraler Wirkfaktoren des Strukturwandels der Öffentlichkeit legt nahe, dass Intermediäre zunehmend Einfluss auf Meinungsbildung und gesellschaftlichen Diskurs nehmen. Wir haben an den zwei Intermediären Facebook und Google-Suche exemplarisch herausgearbeitet, dass die bekannten Rankingsignale und die Methode der Auswertung individuelle Relevanz für jeden Nutzer als eine Art übergeordneten Leitwert der Strukturierung offenbaren. Nun wollen wir diesen Leitwert der digitalen Öffentlichkeit mit jenen Prinzipien vergleichen, die in Deutschland Grundlage der von redaktionellen Medien geschaffenen Öffentlichkeit sind. Diese arbeiten wir im folgenden Abschnitt heraus. Wir fokussieren hier die Prinzipien, auf denen Normen basieren, wie sie in Professionsethik, Institutionen und Regulierung operationalisiert sind.

Der Vergleich beobachteter Leitwerte besonders relevanter Intermediäre mit denen redaktioneller Medien ist aus unserer Sicht nicht nur zulässig, sondern notwendig. Intermediäre erfüllen für viele Nutzer mit redaktionellen Medien vergleichbare Funktionen. Zum Beispiel die für unsere Untersuchung relevanten Funktionen Orientierung über das politische, wirtschaftliche und kulturelle Zeitgeschehen und Teilhabe am gesellschaftlichen Diskurs. Daran ändert sich nichts, weil für einige Nutzer andere Funktionen im Mittelpunkt stehen.

Für die Frage nach Leitwerten der Öffentlichkeit erscheint es hilfreicher, die Beschaffenheit der Kommunikation statt die des Übertragungsweges zu analysieren. Und hier sind keine grundsätzlichen Unterschiede zwischen Intermediären und redaktionellen Medien auszumachen, die einen Vergleich der Leitwerte verbieten würden. Eine dafür geeignete Systematisierung schlägt Jan-Hinrik Schmidt (2016: 286) vor: Er unterscheidet diese drei Kommunikationsmodi:

- Publizieren (*one to many*): z.B. redaktionelle Medien, die sich online an ein breites Publikum richten (Spiegel Online)
- Konversation (*one to one/one to few*): z. B. Chat-Gruppen
- Mensch-Maschine-Interaktion: z. B. Datenbankabfragen, Suchmaschinen

Betrachten wir die Kommunikationsmodi der in dieser Expertise bisher näher analysierten Angebote, zeigt sich eine zunehmende Konvergenz: Bei redaktionell kuratierten Medien dominiert zwar die *One-to-many-Kommunikation*, aber in den Foren der Angebote und den Social-Media-Präsenzen pflegen die Medienmarken zum Teil auch *One-to-few-Kommunikation* (zum Beispiel bei Antworten auf Kritik oder Leserhinweise in Kommentaren).

Auf den ersten Blick dominiert in sozialen Netzwerken *One-to-few-Kommunikation*. Menschen reichen persönliche Erfahrungen weiter oder geben Hinweise auf One-to-many-Quellen, versehen mit einem persönlichen Kommentar, einer Empfehlung oder Kritik für ihr soziales Umfeld in dem Netzwerk. Aber auch redaktionell kuratierte Medien kommunizieren über ihre Profile in sozialen Netzwerken one to many. Wenn ein Angebot wie *Buzzfeed* oder die *Huffington Post* Videos und Beiträge direkt bei Facebook einstellt, ist dieser Kommunikationsmodus kaum von der Publikation in der eigenen App- oder Onlinepräsenz zu unterscheiden. Beide Kommunikationsmodi sind bei einem Intermediär wie Facebook zudem untrennbar mit Mensch-Maschine-Interaktion verwoben: Den Kommentar eines Bekannten zu einem Artikel auf Spiegel Online sehe ich nur, weil der ADM-Prozess zur Priorisierung so gestaltet ist.

Da die Art der Kommunikation der redaktionell und algorithmisch kuratierten Medien und Intermediäre einander ähnlicher wird, spricht vieles dafür, dass auch die Leitwerte einander ähnlicher werden sollten. Wir plädieren nicht für identische Leitwerte oder gar eine Übernahme von einem System ins andere. Doch es gibt Anlass genug für einen Vergleich und eine Prüfung. Redaktionell kuratierte Medien und algorithmisch kuratierte Intermediäre unterscheiden sich in vielen Punkten. Doch wo Intermediäre und redaktionelle Medien Menschen zur Orientierung über das politische, wirtschaftliche und kulturelle Zeitgeschehen und zur Teilhabe am gesellschaftlichen Diskurs dienen, spricht vieles für ähnliche Leitwerte. Im Folgenden skizzieren wir anhand ausgewählter Prinzipien, welche Ansprüche die Gesellschaft in Deutschland an Öffentlichkeit stellt.

Gesellschaftlicher Diskurs ist nach diesen Prinzipien mehr als die Aggregation individueller Entscheidungen, bei denen Qualität durch individuelle Passgenauigkeit definiert ist. Es ist ein Prozess, der über die Summe der Passungen im Einzelfall hinausgeht. Diese aus dem Grundgesetz abgeleiteten Leitwerte konkretisieren, wie eine positive Ordnung der Öffentlichkeit aussehen kann. Die Leitwerte konkretisieren, was Teilhabe und Teilhabeorientierung bezogen auf Öffentlichkeit und gesellschaftlichen Diskurs bedeuten und wie eine positive Ordnung der Öffentlichkeit aussehen kann, die für die breite Partizipation aller gestaltet ist.

6.1 Freiheit

Das Grundgesetz garantiert Meinungs- und Informationsfreiheit als Grundrechte. Aus diesen folgen Grundrechtspositionen wie der Anspruch auf eine positive Ordnung. Eine positive Ordnung zeichnet sich durch Qualitäten wie Achtung der Wahrheit und Integration aus statt lediglich durch die negative Abwesenheit staatlicher Eingriffe in die Meinungsfreiheit (siehe Abschnitt 6.3.2). Beide Grundrechtspositionen können einander widersprechen: Das Recht auf freie Meinungsäußerung der Programmgestalter im öffentlich-rechtlichen Rundfunk ist zum Beispiel durch die Programmgrundsätze als eine Operationalisierung der positiven Ordnung beschränkt, insbesondere durch den vom Verfassungsgericht aus dem Grundgesetz abgeleiteten Anspruch auf umfassende und wahrheitsgemäße Information. „Es ist Sache des Gesetzgebers, solche Kollisionen zum Ausgleich zu bringen“ hält das Bundesverfassungsgericht im dritten Rundfunkurteil fest (Bundesverfassungsgericht 1981: 321).

Anders gesagt: Die Freiheit der Meinungsäußerung ist in bestimmten Fällen beschränkt, um etwa die Informationsfreiheit zu sichern. Die Abwesenheit von Eingriffen sichert nicht eine positive Ordnung. Diese Ausprägung des Leitwerts Freiheit rührt aus einem spezifischen Verständnis der Grundrechte her: „Bestimmend ist hier die Ansicht, dass Individuen vor allem deshalb frei sind, weil sie in einer Gesellschaft leben, die aufgrund bestimmter Vorgaben und Beschränkungen partikularen Handelns ein Höchstmaß an potenzieller Freiheit für alle ihre Mitglieder ermöglicht“ (Heesen 2016: 52). Das Verfassungsgericht formuliert explizit, dass bei einer bestimmten Bedeutung eines Mediums für die Meinungsbildung ein bloßes Abwehrrecht gegen staatlichen Einfluss auf das freie Spiel der Marktkräfte nicht genügt: „Bei dieser Sachlage würde es dem verfassungsrechtlichen Gebot, die Freiheit des Rundfunks zu gewährleisten, nicht gerecht werden, wenn nur staatliche Eingriffe ausgeschlossen würden und der Rundfunk dem freien Spiel der Kräfte überlassen würde; ...“ (Bundesverfassungsgericht 1981: 323).

Darin unterscheidet sich das vom Grundgesetz und Bundesverfassungsgericht bestimmte Verständnis der Meinungs- und Informationsfreiheit grundlegend von dem des US-Mediensystems. In den USA reicht das Prinzip der Meinungsfreiheit sehr weit, ausgeweitet durch die Entscheidung des Supreme Court (US Supreme Court 1969: Brandenburg v. Ohio 395 U.S. 444, dass die allgemeine, öffentliche Befürwortung von Gesetzesverstößen und Gewalt (in dem Fall von einem Ku-Klux-Klan-Vertreter gegenüber Schwarzen) zulässig ist, sofern sie nicht unmittelbar Menschen dazu aufruft, selbst sofort so zu handeln. Öffentlichkeit ist in diesem Verständnis ein Marktplatz der Ideen, der weitgehend dem Wettbewerb der Akteure überlassen sein sollte. In dieser Logik ist Gegenrede das erste Mittel der Wahl gegen Lügen, rassistische Äußerungen oder Hetze. Wer angegriffen wird, muss sich wehren. Nicht weniger, sondern mehr Meinungsäußerung schafft Abhilfe - diesen Kerngedanken hat der Richter Louis Brandeis 1927 in einem Sondervotum zu einer Entscheidung des Obersten Gerichtshofs (in dem Fall gegen die Redefreiheit eines Gründers der Communist Labor Party) formuliert:

„If there be time to expose through discussion the falsehood and fallacies, to avert the evil by the processes of education, the remedy to be applied is more speech, not enforced silence. Only an emergency can justify repression. Such must be the rule if authority is to be reconciled with freedom“
(*Justitia US Supreme Court 1927: 377*).

Facebook beruft sich zum Beispiel auch in Deutschland auf Möglichkeiten und Strategien der Gegenrede (Herbold 2016).

6.2 Vielfalt

Eine freie und umfassende Meinungsbildung ist nur möglich, wenn im medialen Gesamtangebot die „für die freiheitliche Demokratie konstitutive Meinungsvielfalt zur Darstellung gelangt“ (Bundesverfassungsgericht 1981: 321). Dieses Pluralismusgebot sehen beispielsweise Dörr und Deicke als eine der bedeutendsten Strukturvorgaben aus dem Demokratieprinzip des Grundgesetzes abgeleitet:

„Vielfalt setzt das Bestehen und Kommunizieren von verschiedenen Meinungen und meinungsrelevanten Informationen voraus, um so zur Grundlage eines Kommunikationsprozesses in der Bevölkerung werden zu können. Es bildet sogar das Fundament für die Entfaltung der von Art. 1 Abs. 1 GG vorausgesetzten individuellen und politischen Autonomie, da das Individuum dafür auf die Kenntnis von Fakten sowie Meinungen und Haltungen angewiesen ist“ (Dörr und Deicke 2015: 13).

Zwei Pluralismus-Modelle prägen die Umsetzung dieses Leitwerts in Deutschland:

- *Binnenpluralität*: Ein Angebot für sich genommen muss ausgewogen vielfältig gestaltet sein (z. B. öffentlich-rechtliche Rundfunkangebote).
- *Außenpluralität*: Das gesamte Spektrum unterschiedlicher Angebote zusammen ist vielfältig (z. B. Presse).

Auf diesem Leitwert beruhen in Deutschland Rundfunkregelungen, wie reichweitenabhängige Verpflichtungen zu regionalen Fensterprogrammen oder Sendefenstern für unabhängige Dritte.

6.3 Wahrheit

6.3.1 Presse: Wahrheitsfindung als Prozess und berufsethischer Leitwert

In fast allen europäischen Journalismus-Kodizes ist eine Wahrheitsnorm „zentraler und wichtiger Bestandteil“ (Bentele 2016: 63), etwa Ziffer 1 der berufsethischen Grundsätze des Deutschen Presserats: „Die Achtung vor der Wahrheit, die Wahrung der Menschenwürde und die wahrhaftige Unterrichtung der Öffentlichkeit sind oberste Gebote der Presse. Jede in der Presse tätige Person wahrt auf dieser Grundlage das Ansehen und die Glaubwürdigkeit der Medien“ (Deutscher Presserat 2017 4).

Aus diesem Grundsatz folgen einige Konkretisierungen und Operationalisierungen, auf die wir im Folgenden eingehen. Wichtig ist das Grundverständnis dieser Norm: Der Kodex postuliert nicht, dass die Wahrheit in jedem Fall berichtbar ist. Die Formulierung „Achtung vor der Wahrheit“ gibt vielmehr ein Ziel vor, auf das hin beispielsweise Prozesse zu optimieren sind.

Der Pressekodex enthält eine Reihe von Konkretisierungen dieses Leitwerts, zum Beispiel:

- „Unbestätigte Meldungen, Gerüchte und Vermutungen sind als solche erkennbar zu machen.“ (a. a. O.: 4).
- „Bei Veröffentlichungen, die ein Eigeninteresse des Verlages betreffen, muss dieses erkennbar sein.“ (ebd.).
- „Die Presse darf eine Person als Täter bezeichnen [bezieht sich auf den Zeitraum vor Urteilen; Anm. d. Verf.], wenn sie ein Geständnis abgelegt hat und zudem Beweise gegen sie vorliegen oder wenn sie die Tat unter den Augen der Öffentlichkeit begangen hat“ (a. a. O.: 16).

Diese Konkretisierungen stammen aus Entscheidungen des Presserats über Beschwerden wegen möglicher Verstöße gegen den Kodex. Der Presserat entwickelt durch die Spruchpraxis den Kodex fortwährend weiter, konkretisiert und aktualisiert so die Berufsethik. Getragen wird der Presserat von Verleger- und Journalistenorganisationen; gegründet wurde er 1956 als System freiwilliger Selbstkontrolle, um ein geplantes

Bundespressegesetz abzuwehren. Die Mitglieder der Beschwerdeausschüsse des Presserats werden (über das Plenum) von den Trägern des Presserats bestimmt – das entspricht der Logik eines berufsethischen Selbstkontrollorgans.

Der Ansatz von Pressekodex und Presserat geht davon aus, dass Wahrheitsfindung ein Prozess ist und wahrhaftige Unterrichtung entscheidend darauf basiert, bestimmte erprobte Methoden der Erkenntnisgewinnung anzuwenden. Das erinnert an das Prinzip wissenschaftlichen Arbeitens: Aussagen müssen nachprüfbar, nachvollziehbar und falsifizierbar sein, um als wahr gelten zu können, bis sie widerlegt sind.

6.3.2 Rundfunk: Ein Grundrecht auf wahrheitsgemäße Information

Das Bundesverfassungsgericht stellt in seinem 3. Rundfunkurteil klar, dass aus dem Grundrecht auf Informationsfreiheit im Grundgesetz (Art. 5 Abs. 1 Satz 1 Halbs. 2) ein „Anspruch auf umfassende und wahrheitsgemäße Information“ folgt (Bundesverfassungsgericht 1981: 321). Das Verfassungsgericht urteilt, dass die Rundfunkfreiheit mehr als nur abwehrende Bedeutung wie die klassischen Freiheitsrechte hat. Bloße Staatsfreiheit bedeute noch nicht, dass „freie und umfassende Meinungsbildung“ (a. a. O.: 320) möglich werde. Dazu bedürfe es vielmehr einer „positiven Ordnung“.

Diese positive Ordnung muss der Gesetzgeber gestalten. Bei der Rundfunkorganisation kommt es nach dem Urteils des Verfassungsgerichts allein darauf an, dass „freie, umfassende und wahrheitsgemäße Meinung im dargelegten Sinne gewährleistet ist“ (a. a. O.: 321). In dessen Umsetzung verpflichten die gesetzlichen Programmgrundsätze den öffentlich-rechtlichen Rundfunk entsprechend, beispielsweise legt das WDR-Gesetz in § 5 Abs. 4 fest, dass der WDR „der Wahrheit verpflichtet sein“ soll.

Die Einhaltung solcher Programmgrundsätze überwachen für den öffentlich-rechtlichen Rundfunk die mit Vertretern unterschiedlicher gesellschaftlicher Gruppen (von Land zu Land unterschiedlich) besetzten Rundfunkräte. Diese Gremien sind nach einer anderen Logik als der Presserat zusammengesetzt: Sie sollen nicht eine Berufsethik entwickeln, sondern die Kontrolle der Allgemeinheit operationalisieren. Entsprechend sollten Rundfunkräte durch ihre Besetzung die Allgemeinheit abbilden. Dass dieses Ideal nicht erreicht wird, fasst Donges pointiert zusammen: „Betrachtet man beispielsweise die Gremien der öffentlich-rechtlichen Rundfunkanbieter, so zeigt sich, dass dort nicht die Gesellschaft insgesamt repräsentiert wird, sondern eher der ältere, männliche und statushohe Teil davon“ (Donges: 97).

Auch dieses System ist offen für Beschwerden Dritter. Anders als beim Presserat können Rundfunkräte aus den Programmbeschwerden aber keine Konkretisierungen oder Aktualisierungen der Programmgrundsätze ableiten, sie behandeln nur Einzelfälle.

6.4 Integration

Das Ideal der deliberativen Öffentlichkeit sieht die gesellschaftliche Integration als zentrales Merkmal einer positiven Ordnung der Öffentlichkeit. Das Beratschlagen, der Austausch von Argumenten im gesellschaftlichen Diskurs soll zu einem gewissen Ausgleich und Konsens führen. Das Idealbild zeichnet eine Deliberation, „in der Menschen die mögliche Vielfalt von Positionen und Meinungen anerkennen und sich auf dieser Basis verständigungsorientiert miteinander austauschen“ (Schmidt et al. 2017: 26).⁶

Das Ideal der deliberativen Öffentlichkeit sieht den Widerstreit der Meinungen nicht als fortwährenden Kreislauf, sondern als Prozess mit Entwicklungsstufen. Voraussetzungen für das Gelingen dieses Kreislaufs sind Austausch, Rationalität, Argumentation:

⁶ „Diese Charakterisierung beruht auf den Gedanken von Jürgen Habermas (1981), ohne der Komplexität seiner Öffentlichkeitstheorie hier gerecht werden zu können“ (Schmidt et al.: 26).

„Die Idee öffentlicher Deliberation ist ja gerade, Konsens im Durchgang durch Dissens anzustreben. Kritik und Problematisierung gehören ebenso zum Diskurs wie der Versuch einer rationalen Auflösung von Dissens. Und von einer seriösen Diskussion können wir kaum reden, wenn die Beteiligten nicht in irgendeiner Form darauf abzielen, sich gegenseitig durch Argumente zu überzeugen und sich gegebenenfalls überzeugen zu lassen“ (Peters 2002: 31 f.)

Das Ideal gesellschaftlicher Integration durch deliberative Öffentlichkeit klingt im Auftrag der öffentlich-rechtlichen Rundfunkanstalten an, den gesellschaftlichen Zusammenhalt in Bund und Ländern zu fördern. Dies sollen sie leisten, indem sie „einen umfassenden Überblick über das internationale, europäische, nationale und regionale Geschehen in allen wesentlichen Lebensbereichen“ geben (die medienanstalten 2016: 19).

Dieses Idealbild prägt die Diskussion über Funktion und Wirkung von Intermediären in Deutschland, zum Beispiel bei der Diskussion über Filterblasen und Echokammern. Bei diesen Phänomenen stellt sich auch die Frage nach der Vielfalt neu: Intermediäre bieten zwar Zugang zu einer großen Vielfalt von Inhalten und Quellen, priorisieren aber je nach Personalisierungsart und -grad nur einen kleinen und homogenen Ausschnitt dieser potenziellen Vielfalt. Die Frage in der Logik der hier herausgearbeiteten Leitwerte muss lauten: Welcher Binnenpluralismus bei der Priorisierung durch Intermediäre ist nötig, damit gesellschaftliche Integration durch deliberative Öffentlichkeit gelingen kann?

Gesellschaftliche Integration ist auch auf europäischer Ebene ein Leitwert der Öffentlichkeit: So warnte die High Level Expert Group on Media Diversity and Pluralism, dass Menschen aufgrund personalisierter Medieninhalte weniger unterschiedliche Meinungen wahrnehmen könnten, was die Öffentlichkeit und den demokratischen Willensbildungsprozess beeinträchtigen könne (Vike-Freiberga et al. 2013). Auch der Europäische Rat formulierte Warnungen hinsichtlich der Auswirkungen von algorithmischer Sortierung in Suchmaschinen (Council of Europe 2012a) und sozialen Netzwerken (Council of Europe 2012b) auf die Meinungsbildung.

6.5 Zwischenfazit: Anwendbarkeit der Leitwerte auf algorithmisch strukturierter Öffentlichkeit

Der folgende Abschnitt diskutiert, inwieweit die oben als Leitwerte skizzierten, redaktionellen Medien zugesprochenen Ansprüche an Öffentlichkeit für algorithmisch strukturierte Angebote gelten können. Wir analysieren, ob diese Leitwerte angelegt und wie sie interpretiert werden können, wo Intermediäre Menschen zur Orientierung über das politische, wirtschaftliche und kulturelle Zeitgeschehen und zur Teilhabe am gesellschaftlichen Diskurs dienen.

Unsere Überlegungen basieren auf Erkenntnissen zum Strukturwandel der Öffentlichkeit ([siehe Abschnitt 4](#)) und der Wirkung algorithmischer Prozesse sowie deren Wechselwirkung mit einer Reihe unterschiedlicher Faktoren ([siehe Abschnitt 5](#)): Die Verfügbarkeit einer größeren Menge teils ideologisch extremerer Medieninhalte über solche Angebote, das Auswahl- und Empfehlungsverhalten einzelner Nutzer, die algorithmische Sortierung von Inhalten auf Basis für Außenstehende teils intransparenter Signale, die Designentscheidungen und teils verdeckt durchgeführten experimentellen Variationen der Intermediäre, die Einwirkung auf Manipulation digitaler Öffentlichkeiten ausgerichteteter automatisierter Systeme sowie weitere, äußere Faktoren, wie etwa das politische System eines Staates.

6.5.1 Freiheit als Freiheit der Relevanzdeutung

Intermediäre wie die in beispielhaft untersuchten Angebote von Facebook und Google strukturieren Öffentlichkeit nach ihren eigenen Prinzipien und Werten. Diese Gestaltungsentscheidungen ([detailliert in Abschnitt 4.3](#)) eines Intermediärs sind unserer Auffassung nach letztlich publizistische:

- Kriterien, nach denen ein Intermediär Relevanz gewichtet
- Signale, anhand derer er diese Konzepte operationalisiert und misst

- Wie stark er die einzelnen Faktoren gewichtet

Die Gewichtung und Auswahl von Meinungen, Informationen und Äußerungen über das Zeitgeschehen ist eine Meinungsäußerung. Diese Ansicht fasst Volokh in einem vielbeachteten Auftragsgutachten für Google zusammen:

„For example, a newspaper also includes the materials that its editors have selected and arranged, while the speech of DrudgeReport.com or a search engine consists almost entirely of the selected and arranged links to others' material. But the judgments are all, at their core, editorial judgments about what users are likely to find interesting and valuable“ (Volokh und Falk 2012: 4 f.).

Intermediäre sind in diesem Sinne Gatekeeper der Informationsvielfalt (*gatekeepers of information diversity*), wie Morganti et al. (2015) es formulieren. Philip Napoli (2014a) geht noch weiter und spricht von *automated media*.

Bei Intermediären ist der Einfluss der Nutzer auf die Strukturierung größer als bei einem redaktionell kuratierten Medium. Aber die bewusst und unbewusst vom Nutzer gesetzten Signale sind abhängig von den Gestaltungsentscheidungen des Intermediärs darüber, welche Signale möglich sind und welche gefördert werden. Deshalb sind Nutzerentscheidungen keine unabhängige Variable ([siehe Abschnitt 5.4](#)), sondern stehen in einer fortwährenden Wechselwirkung mit der Gestaltung der Intermediäre (siehe Abschnitt 4.4).

Angesichts der Relevanz von Intermediären für die Meinungsbildung (siehe Abschnitt 3.2) scheint ein Ausräumen zwischen der Freiheit der Meinungsäußerung von Intermediären und dem Recht auf umfassende und wahrheitsgemäße Information der Allgemeinheit geboten (siehe Abschnitt 6.1). Wie sich Informationsintermediäre in die vom Bundesverfassungsgericht entwickelte Systematik der Informationsfreiheit und der dienenden Rundfunkfreiheit fügen, ist bisher nicht abschließend geklärt. Allerdings lassen sich der jüngeren Rechtsprechung doch Hinweise entnehmen, die beispielsweise Fetzer zu dem Fazit zusammenfasst:

„Inhalte und Dienste, die über das Internet verbreitet werden und sich an die Allgemeinheit richten und die von Relevanz für die individuelle und kollektive Meinungsbildung sind, werden grundsätzlich vom sachlichen Schutzbereich des Art. 5 Abs. 1 Satz 2 Alt. 2 GG erfasst“ [Pressefreiheit und Rundfunkfreiheit, Alt. freie Entfaltung der Persönlichkeit; Anm. d. Verf.] (Fetzer 2015: 8).

6.5.2 Vielfalt als Vielfalt algorithmischer Prozesse

Die Übertragung des Leitwerts Vielfalt auf Intermediäre offenbart eine implizite Annahme: Angebotsvielfalt korreliert mit Nutzungsvielfalt. Das ist die Ratio hinter dem Pluralismusgebot (vgl. Burri 2013). Bestimmte Auflagen der Rundfunkregulierung in Deutschland orientierten sich zwar an Schwellenwerten tatsächlicher Nutzung, doch darüber hinaus gibt es keine nennenswerte Konzeptualisierung des Verhältnisses von Angebots- und Nutzungsvielfalt. Auf dem gegenwärtigen Stand der Empirie lässt sich nicht abschließend bewerten, wie sich die Nutzungsvielfalt im Zuge der gewachsenen Angebotsvielfalt in der digitalen Sphäre entwickelt hat. Es fehlt Empirie zu tatsächlicher Nutzungsvielfalt. Wir wissen nicht, wie groß und wie vielfältig zum Beispiel die personalisierten Newsfeeds sind. Das ist wenig untersucht, auch weil Informationsintermediäre von außen unabhängig nur schwer beforschbar sind.

Diese Fragen beziehen sich allesamt auf eine Ebene der informationellen Nutzung: redaktionell erstellte Inhalte. Wie viele existieren, wie plural sind sie? Davon ist eine zweite Ebene der Vielfalt zu trennen, die für Intermediäre äußerst relevant erscheint: die Verfahrensvielfalt bei Selektion und Priorisierung. Jeder Intermediär konzeptualisiert Relevanz und gewichtet Signale entsprechend. Diese publizistische Leistung unterliegt nicht der Kontrolle der Nutzer. Wenn ich eine algorithmische Strukturierung meines Facebook-News-Feeds wünsche, kann ich nicht unter verschiedenen Verfahren wählen. Die Diskussion über Vielfalt bei Intermediären ist also auf drei Ebenen zu führen:

- algorithmische Priorisierungsverfahren (Prozess),
- Angebot (Output) und

- Nutzung (Outcome).

Wobei Prozess und Angebot die Nutzung beeinflussten und die Nutzung auch auf Priorisierung und Angebot wirken kann. Algorithmische Prozesse ermöglichen diese Wechselwirkung und vermitteln zwischen alle Ebenen.

Die Verengung der Verfahrensvielfalt ist nicht mit Phänomenen aus redaktionell kuratierten Medien zu vergleichen. Es geht hier nicht – oder zumindest derzeit nicht – um die Priorisierung basierend auf weltanschaulichen Linien (z. B. konservativer vs. liberaler Intermediäre). Es geht vielmehr darum, ob die Vielfalt der von Intermediären eingesetzten Relevanzkonzepte verengt ist. Dass zum Beispiel bei einigen Intermediären negativ emotionalisierte Beiträge messbar größere Reichweiten erzielen als neutrale (vgl. Abschnitt 5.4), resultiert aus einer bestimmten Gestaltung der Plattformen: Die Nutzeroberflächen sind auf leichtgängige Reaktionen optimiert, die eine bestimmte Form der Interaktion bevorzugen, das algorithmische System nutzt diese Interaktionen als Merkmale von Relevanz. Andere Verfahren zur Relevanzdeutung könnten zu anderen Ergebnissen kommen und würden in Summe nicht negativ emotionalisierte Beiträge im selben Grad fördern. Möglich wäre dies aber nur bei größerer Vielfalt der nutzbaren algorithmischen Prozesse.

Es geht zweifelsohne ein gewisses Maß an Vielfaltsverengung mit der Tatsache einher, dass sehr wenige Intermediäre extrem hohe Reichweiten und Nutzungszeiten erzielen und jeder von diesen Anbietern ausschließlich die eigenen Verfahren zur Relevanzdeutung zulässt und einsetzt.

Daraus folgt: Die Entwicklung vielfaltsfördernder Interventionen bei Intermediären muss auf einer anderen und breiteren analytischen Grundlage stehen als Vielfaltssicherung im klassischen Rundfunk und in der Presse. Vielfalt ist auch als Vielfalt der algorithmischen Priorisierungsverfahren (Prozessebene) zu untersuchen und zu adressieren. Zudem braucht es bei der Analyse der tatsächliche Nutzung und des tatsächlich von Intermediären ausgespielten inhaltlichen Angebots an Inhalten einen anderen analytischen Zugang als bei redaktionellen Medien. Um die Wirkung algorithmisch gesteuerter Personalisierung auf die Vielfalt von Angebot und Nutzung untersuchen zu können, benötigt unabhängige Forschung Zugang zu entsprechenden Interaktionsdaten.

6.5.3 Achtung vor der Wahrheit

Wenn algorithmische Relevanzdeutung Öffentlichkeit strukturiert, dann muss sie sich auch am Leitwert der Achtung vor der Wahrheit messen lassen. Es gibt keine stichhaltigen Argumente, die Bezüge im Grundgesetz und in der Rechtsprechung nicht auf Intermediäre anzuwenden, die zur Meinungsbildung beitragen. Die entscheidende Frage ist: In welchem Maß müssen sich diese Angebote der Achtung vor der Wahrheit verpflichten? Offenkundig weiter als beispielsweise Facebook in seinen *Werten des News Feed*, die keinen Bezug auf diesen Leitwert nehmen (Facebook 2016b). Aber sehr wahrscheinlich nicht so weit, wie es Regulierung und Selbstregulierung von redaktionell kuratierten Medien verlangen.

Die Diskussion über gezielt platzierte virale Falschmeldungen (sogenannte *Fake News*) in Deutschland zeigt, dass Intermediäre auch an diesem Leitwert gemessen werden. Den Begriff *Fake News* verwenden wir nicht. Wir sprechen von Falschinformationen, die „absichtlich produziert und gestreut und dabei so komponiert werden, dass sie die Logiken der sozialen Medien ausnutzen“ (Reinbold 2017).

Zur Achtung vor der Wahrheit gehört bei Intermediären auch die Frage, inwiefern die ausgewerteten Signale das Vortäuschen größerer Publikumsreaktionen ermöglichen (siehe Abschnitt 5.5). Nicht nur die priorisierten Inhalte, auch die von Intermediären vermittelten Reaktionen des (vermeintlichen) Publikums müssen sich am Leitwert der Achtung vor der Wahrheit messen lassen.

6.5.4 Gesellschaftliche Integration

Dieser Leitwert ist am schwersten von redaktionell kuratierten Medien auf Intermediäre zu übertragen. Die Personalisierung der algorithmischen Entscheidungsprozesse steht per se in Spannung zu diesem Leitwert, auch wenn die Sorgen vor Filterblasen und Echokammern angesichts der vorliegenden empirischen Befunde nur teilweise begründet erscheinen (siehe Abschnitt 5.6). Bei diesen Phänomenen fokussiert der Leitwert der

gesellschaftlichen Integration die Outcomeebene stärker als der Leitwert der Vielfalt: Integration bezieht sich auf die tatsächliche Nutzung und die Qualität der ausgelösten Interaktionen. Wenn das Angebot eines Intermediärs zum Beispiel dazu führt, dass Nutzer mit vielfältigen Inhalten konfrontiert sind, diese aber bestehende Polarisation in Gruppen verstärken, fördert der spezifische algorithmische Prozess zwar Vielfalt, nicht aber Integration.

Die integrierende Wirkung einer deliberativen Öffentlichkeit bezieht sich aber nicht allein auf Personalisierung, Auswahl und Priorisierung, sondern auf den Diskurs, der sich auf dieser Basis vollzieht. Diese Ebene fehlt oft in Debatten über Filterblasen-Effekte: Alle Informationsintermediäre, die über eine soziale Komponente ihrer Plattform den Diskurs ermöglichen, prägen durch die Gestaltung und Moderation die Qualität des gesellschaftlichen Diskurses unmittelbar. Zahlreiche Befunde deuten darauf hin, dass dies große Wirkung auf das Nutzerverhalten hat (für eine Übersicht siehe Diakopoulos 2016b).

Ein das Ideal der deliberativen Öffentlichkeit berücksichtigendes Design von Kommentaren und Interaktionsmöglichkeiten würde vielleicht andere Signale auswerten und als Erfolgsmarker nutzen. Ein grundlegender Unterschied im Design ist der zwischen auf einen Konsens hin optimierten Systemen, wie *Mediawiki* (der Betriebssoftware der Wikipedia) oder *LiquidFeedback*, und fortlaufend unbeschränktem Kommentieren. Einen Einblick, wie Intermediärsdesign öffentliche Meinungsbildung beeinflusst, geben Lobo und Lauer in ihrer Analyse der Piratenpartei:

„Für die Piratenbasis hat Twitter ab etwa 2010 als emotionales Ventil funktioniert, als Schiffsschaukel der Empörungsmaximierung. (...) Die dominierenden sozialen Medienplattformen eignen sich hervorragend zur Mobilisierung, zur Gruppenbildung und zur blitzartigen Informationsverbreitung – aber nicht zur produktiven, politischen Diskussion; deshalb müssen diese Debattenräume völlig anders strukturiert werden als Twitter oder Facebook, es braucht eigene, funktionierende Plattformen für die digitale Demokratie“ (Lobo und Lauer 2015: loc. 2436 ff.).

Inzwischen erreicht beispielsweise Facebook in dieser Funktion gut die Hälfte der 23,25 Millionen Deutschen, die sich an einem Durchschnittstag auf algorithmisch kuratierten Angeboten über das Zeitgeschehen informieren: Sie gewinnen hier nicht nur einen Einblick in das Tagesgeschehen, sondern auch in die vermeintliche gesellschaftliche Diskussion darüber. Die aktuelle Debatte über Themen wie Hassrede und Desinformation in sozialen Netzwerken legt nahe, dass die dort wahrgenommene gesellschaftliche Diskussion häufig jedoch nicht als produktiv und konstruktiv, sondern im Gegenteil als aggressiv und oft aversiv wahrgenommen wird. Mit Blick auf gesellschaftliche Teilhabe erscheint dies derzeit wie eine vertane Chance: Nie zuvor in der Geschichte haben sich so viele Menschen in quasi-öffentlichen Räumen an Debatten auch über politische und gesellschaftliche Themen beteiligt. Das Ergebnis aber wird nicht als Bereicherung für die deliberative Demokratie, sondern derzeit primär als Belastung, ja als Gefahr wahrgenommen und diskutiert.

7 Was man tun kann: Ansatzpunkte für Interventionen

Die bisherige Untersuchung zeigt: Die Leitwerte von Freiheit, Wahrheit, Vielfalt und gesellschaftlicher Integration sind grundsätzlich zur Analyse von Intermediären geeignet, die auch zur Orientierung über das politische, wirtschaftliche und kulturelle Zeitgeschehen dienen und Teilhabe am gesellschaftlichen Diskurs ermöglichen. Die Anwendung dieser Leitwerte auf digitale Öffentlichkeit und ihren vermittelnden algorithmischen Prozessen braucht teils einen anderen analytischen Zugang, teils eine andere Operationalisierung.

Handlungsoptionen, um den Leitwerten Gültigkeit zu verschaffen, sind im Wesentlichen in zwei Bereichen zu suchen:

- algorithmische Entscheidungsfindung
- menschliche Wahrnehmung

Beide Bereiche haben einen großen Einfluss auf die öffentliche Meinungsbildung in der digitalen Sphäre ([siehe Abschnitt 4.1](#)). Sie stehen in enger Wechselwirkung, Kausalität lässt sich nach heutigem Erkenntnisstand selten eindeutig beschreiben: Kognitive Verzerrungen beeinflussen, welche Signale Menschen bei Intermediären abgeben. Prozesse algorithmischer Entscheidungsfindungen werten diese Signale als Proxys für Relevanz und skalieren dabei unter Umständen die in den Daten abgebildeten Verzerrungen, die dann wiederum Grundlage für die Priorisierung bestimmter Medienangebote und im nächsten Schritt weiterer menschliche Urteile sind. Zusätzlich werden die Relevanzsignale womöglich von außen technisch manipuliert. Für diese beiden inhaltlichen Bereiche unterscheiden wir drei Handlungsebenen:

- Makroebene: gesellschaftlicher Rahmen, Leitwerte, Regulierung, staatliche Akteure
- Mesoebene: Unternehmen, öffentlich-rechtliche Institutionen, Organe der Selbstkontrolle
- Mikroebene: Mediennutzer, Entwickler, Journalisten

Analysieren wir in dieser Matrix vermeintlich klar benannte Phänomene wie *Fake News*, zeigt sich, dass dahinter vielfältige Veränderungen auf unterschiedlichen Ebenen stehen: Es ist über Intermediäre einfacher, ohne etablierte Medienmarke ein breites Publikum zu erreichen. Es ist einfacher, durch gesteuerte Konten eine kritische Masse an Interesse für bestimmte Inhalte vorzutäuschen. Und es ist einfacher, mit optimierten Inhalten bestimmte kognitive Verzerrungen zu nutzen, um eine kritische Masse an Nutzerreaktionen anzuregen. Diese Vielfalt an zusammenwirkenden Faktoren spricht dafür, dass eine Handlungsoption allein keine Abhilfe schaffen kann und daher eine systematische Untersuchung notwendig ist (vgl. Tabelle 5).

Tabelle 5: Faktoren für die Verbreitung bewusst lancierter Falschmeldungen über Intermediäre

Handlungsebene	Algorithmische Entscheidungsfindung	Menschliche Wahrnehmung
Makro	Mangelnde Vielfalt der ADM-Prozesse Unzureichende Evaluation	Starke Breitenwirkung von Intermediären mit Konstruktionsprinzipien, die bestimmte kognitive Verzerrung fördern
Meso	Auswahl und Gewichtung von Signalen Fehlende Gültigkeit des Leitwerts Achtung vor der Wahrheit bei Intermediären Fehlen unabhängiger Gremien (analog z. B. zum Presserat)	Fehlen von Maßnahmen zum Eindämmen der Wirkung kognitiver Verzerrungen
Mikro	Einschätzung einzelner Falschmeldungen als hoch relevant Entscheidungen von Entwicklern und Evaluatoren	Weitergabe von Falschmeldungen; mögliche Ursachen: Fehlwahrnehmung der ADM-Prozesse, Dominanz anderer als der Informationsbedürfnisse (z. B. Stärkung von Gruppenzugehörigkeit, Identität u. Ä.)

Quelle: Eigene Darstellung.

Im Folgenden stellen wir Ansatzpunkte für mögliche Verbesserungen auf den einzelnen Feldern dieser Matrix vor. Als Verbesserungen verstehen wir dabei, den exemplarisch herausgearbeiteten Leitwerten von Öffentlichkeit mehr Geltung bei der algorithmischen Strukturierung von Öffentlichkeit zu verleihen. Wir entwickeln und empfehlen aber nicht einzelne Lösungskonzepte, sondern identifizieren vielmehr lohnende Ansätze für weitere Überlegungen (vgl. Tabelle 6).

Tabelle 6: Ansatzpunkte und Handlungsoptionen

Handlungsebene	Algorithmische Entscheidungsfindung	Menschliche Wahrnehmung
Makro	Gesellschaftliche Verständigung über Leitwerte Sicherung der Vielfalt der Relevanzprognosen Herstellung externer Beforschbarkeit	Forschung und externe Evaluation
Meso	Förderung externer Evaluation Entwicklung alternativer Relevanzprognosen Institutionelle Verankerung und dynamische Entwicklung von Leitwerten (z. B. Professionsethik, Algorithmenrat als Selbstkontrollorgan, Impact Review bei ADM-Änderungen)	Sicherung der Vielfalt der Foren für menschliche Interaktion Immunisierung
Mikro	Individuelle Verankerung von Leitwerten (z. B. in Ausbildung) Förderung individueller Kompetenzen im Umgang mit algorithmischen Systemen	Vielfältiger individueller Input Sensibilisierung für die Wirkung kognitiver Verzerrungen

Quelle: Eigene Darstellung.

Voraussetzungen einer gesellschaftlichen Debatte über die Öffentlichkeit, die wir als Gesellschaft wollen, sind: transparente, erklärbare, überprüfbare und korrigierbare ADM-Prozesse sowie Verantwortlichkeit und Anwendungskompetenz bei allen Akteuren. Wenn wir nicht wissen, woraufhin und mit welchem Ziel wir ADM-Prozesse optimieren, ist es unmöglich, darüber zu diskutieren, ob diese Prozesse gut oder schlecht sind im Sinne der Teilhabe und der Leitwerte, die eine Teilhabeorientierung für die Gestaltung von Öffentlichkeit als positive Ordnung konkretisieren.

7.1 Prozesse algorithmischer Entscheidungsfindung für Teilhabe gestalten

7.1.1 Gesellschaftliche Verständigung über Leitwerte digitaler Öffentlichkeit (Makroebene)

Eine gesellschaftliche Übereinkunft über die Leitwerte von Öffentlichkeit in der digitalen Sphäre ist Voraussetzung für eine Gestaltung fürs Gemeinwohl. In Deutschland allerdings ist beispielsweise die Vielfalt von Intermediären und algorithmischen Entscheidungsprozessen nicht als Aufgabe der Vielfaltssicherung verankert (siehe Abschnitt 7.1.2).

Es gibt keine Entscheidung des Bundesverfassungsgerichts zu der Frage, in welchem Maß die *positive Ordnung* der Rundfunkfreiheit in die digitale Sphäre zu übertragen ist. Die Zuständigkeiten zwischen Bund und Ländern, zwischen Telekommunikations- und Medienrecht sind nicht abschließend geklärt. Eine Klärung könnten Bund oder Länder durch aktive Gestaltung dieses Schwebezustands herbeiführen. Um auf der Makroebene Intermediärsvielfalt zu fördern, braucht es Sensibilisierung für die Folgen von Vielfaltsverengungen auf diesem Gebiet und überzeugende Konzepte zur Gestaltung der Daseinsvorsorge in der digitalen Sphäre. Und dann gegebenenfalls gesetzliche Umsetzung.

7.1.2 Vielfalt der Relevanzprognosen sichern (Makroebene)

Wenn wenige Auffassungen von Relevanz und wenige entsprechende algorithmische Verfahren im gesellschaftlichen Diskurs in der digitalen Sphäre dominieren, schmälert das potenziell Vielfalt. Nötig ist ein erweitertes Verständnis von Vielfaltssicherung, nämlich auch als Vielfalt der algorithmischen Verfahren, der Betreibermodelle und der Ziele zur Strukturierung von Öffentlichkeit in der digitalen Sphäre.

Um hier Ansätze für mehr Vielfalt auf der Makroebene zu entdecken, ist eine Erkenntnis entscheidend: Es ist weder technisch noch funktional zwingend, dass Menschen eine digitale Infrastruktur (z. B. ein soziales Netzwerk) ausschließlich mit einem einzigen algorithmischen Verfahren (des Infrastrukturbetreibers) zur Relevanzeinschätzung nutzen können. Diese Art der Gestaltung dominiert derzeit bei Intermediären in der digitalen Sphäre, ist aber mit Blick auf die gesamte Geschichte des World Wide Web und des Internet noch immer eine vergleichsweise neue Entwicklung. Mitte der Nullerjahre ermöglichten offene Standards wie *Rich Site Summary* (RSS) oder *Outline Processor Markup Language* (OPML) und so einfache Hilfsmittel wie CSV-Adressbuchdateien die Übertragbarkeit gesammelter Quellen und sozialer Verbindungen. Menschen konnten so ihre eigenen strukturierten Signale (Welche Quellen schätze ich und wen kenne ich?) bei unterschiedlichen Anbietern algorithmischer Strukturierung auswerten. Was an Portabilität heute noch möglich ist, zeigen Angebote wie *Nuzzel*, die basierend auf den *Application Programming Interfaces* (API) von Twitter und Facebook deren abrufbare soziale Infrastruktur trotz der stark eingeschränkten Nutzbarkeit für eine spürbar alternative Informationsauswahl nutzen.

Alternativen sind machbar: Theoretisch lässt sich ein Social Graph oder ein Suchindex als Infrastruktur betreiben und frei auch mit anderen Verfahren und Zielen auswerten. Daraus folgt: Auch bei Intermediären sind Modelle der Außen- und Binnenpluralität denkbar. Man muss sich von der Vorstellung lösen, dass die Beziehungen, welche die Infrastruktur eines Intermediärs erfasst (etwa zwischen Profilen oder zwischen Angeboten im indexierten Web), untrennbar mit einem einzigen Verfahren zu Auswertung verbunden sind.

Vielfaltssicherung bei Intermediären muss neben der Angebotsvielfalt an redaktionellen Medien weitere Aspekte in den Blick nehmen, die bislang zu wenig beachtet werden: zum einen die Vielfalt der Verfahren zur Strukturierung von Öffentlichkeit, zum anderen die tatsächliche Vielfalt der Nutzung. Die Nutzungsvielfalt muss in einer höheren Auflösung erfasst werden. Erkenntnisse auf Beitragsebene sind dabei mindestens so relevant wie Gesamtreichweiten ausgewählter Medienangebote, so wie sie beispielsweise der Medienvielfaltsmonitor erfasst. Denn nur auf Ebene einzelner Beiträge lassen sich Erkenntnisse über potenziell vielfaltsverengende Effekte wie etwa Filterblasen (siehe Abschnitt 5.6) gewinnen. Für die Wirkung algorithmischer Strukturierung der Öffentlichkeit auf Vielfalt ist die Medienaufsicht mangels Kompetenz und Forschung blind.

Mittelstadt skizziert hier Handlungsbedarf- und Handlungsmöglichkeiten (zur externen Beforschbarkeit und Evaluation siehe auch Abschnitt 7.1.2):

„One possibility for managing algorithmic auditing would be a regulatory body to oversee service providers whose work has a foreseeable impact on political discourse by detecting biased outcomes as indicated by the distribution of content types across political groups ...” (Mittelstadt 2016b: 4998).

Mit welchen Instrumenten auf dieser Ebene Vielfalt gesichert werden kann, muss eine gesellschaftliche Debatte bestimmen, die bis heute nicht begonnen hat. Die Bandbreite der möglichen, aber bislang nicht genutzten Instrumente ist groß: Es gibt keine Konzepte für einen öffentlich-rechtlichen Rahmen für Intermediäre, die nicht auf Refinanzierung durch Reichweitenmodelle und Werbung optimiert sind. Es fehlen Modelle, wie unternehmerisch verfasste alternative Intermediärsmodelle gefördert werden könnten, um zugleich Vielfalt zu sichern. In den Niederlanden hingegen hat der vom Kulturministerium finanzierte *Stimuleringsfonds voor de Journalistiek* etwa durch die frühe Unterstützung von Angeboten wie *Blendle* gezeigt, dass Potenziale existieren. Betreibermodelle des dritten Sektors für Intermediäre sind in Deutschland unerprobt, in den Vereinigten Staaten zeigen die *Wikimedia Foundation* und die *Mozilla Foundation* beispielhaft, dass solche alternativen Betreibermodelle zu mehr Angebotsvielfalt führen können.

7.1.3 Externe Beforschbarkeit herstellen (Makroebene)

Wissenschaftler, die nicht für die jeweiligen Intermediäre arbeiten, stehen bei der Erforschung und Evaluation der Vorgänge auf algorithmisch kuratierten Angeboten vor einer Reihe von Problemen. *Application Programming Interfaces* (API) etwa dienen ursprünglich externen Entwicklern, beispielsweise Anwendungen für die jeweilige Plattform zu entwickeln, beispielsweise Spiele oder andere Apps, die innerhalb von Facebook laufen und dazu unter Umständen auf Teile der Facebook-Datenbank zugreifen, etwa die Freundeslisten der jeweiligen Nutzer. API können theoretisch auch von Forschern genutzt werden, um auf Basis der auf solchen Plattformen in großer Menge anfallenden Daten bestimmte Auswertungen vorzunehmen oder Hypothesen zu testen. Praktisch aber ist der Zugriff auf diese Schnittstellen für solche Zwecke derzeit oft eingeschränkt.

„In contrast to the Twitter API, using the API of Facebook means that researchers have to request permission to collect nonpublic data from the participants through a Facebook app“ (Lomborg und Bechmann 2014).

Eine Erklärung für die Einschränkungen, mit denen sich Forscher hier konfrontiert sehen, liefern Puschmann und Ausserhofer:

„Facebook has greatly restricted access to user data through the API out of privacy concerns, as have other platforms. When dubious actors acquire large amounts of data that are clearly not used for the API’s intended purpose, this often leads to a tightening of policies by the API’s operators, if only because providing and sustaining the performance of an API is not trivial computationally“ (Puschmann und Ausserhofer 2017 in: Schäfer und van Es 2017).

Auch Twitter hat die Möglichkeiten, auf seine API zuzugreifen, im Laufe der Jahre eingeschränkt, wie Puschmann und Ausserhofer notieren:

„Initially offering broad access to data in the first years of its operation in order to encourage development of derivate services, such as software clients for unsupported platforms, the company reasserted its control by making access to data more restrictive in several successive steps over recent years“ (a. a. O.).

Puschmann und Burgess fassen die derzeitige Gemengelage hinsichtlich der Verfügbarkeit von Nutzer- und Nutzungsdaten derartiger Plattformen am Beispiel Twitter wie folgt zusammen:

„Platform providers and users are in a constant state of negotiation regarding access to and control over information. Both on Twitter and on other platforms, this negotiation is conducted with contractual and technical instruments by the provider, and with ad hoc activism by some users. The complex relationships among platform providers, end users, and a variety of third parties (e.g., marketers, governments, researchers) further complicate the picture. These nascent conflicts are likely to deepen in the coming years, as the value of data increases while privacy concerns mount and those without access feel increasingly marginalised“ (Puschmann und Burgess 2013, in: Weller 2013).

Eine weitere Problematik entsteht aus den Geschäftsbedingungen der Intermediäre: Facebook zum Beispiel untersagt – in der lautereren Absicht, Manipulationen etwa durch Bots zu verhindern – das Anlegen von Profilen, die nicht mit einer realen Person korrespondieren, sogenannten *Sockenpuppen*. Das aber schränkt die Möglichkeit von Wissenschaftlern, die Wirkungsweise des Sortieralgorithmus mithilfe isolierender Variation zu beobachten und zu prüfen, massiv ein.

Für einen evidenzbasierten Diskurs über die Wirkung algorithmisch strukturierender Intermediäre braucht es mehr Empirie. Die unabhängige Beforschbarkeit und Evaluation von Intermediärsleistungen ist eingeschränkt. Nötig sind:

- **Transparenz:** Offenlegung der genutzten Daten, ihrer Gewichtung sowie der Konsequenzen

- Erklärbarkeit: Die Nachvollziehbarkeit von Entscheidungen (und Kriterien) ist Voraussetzung für Diskurs
- Überprüfbarkeit: Kontrolle und Bewertung algorithmischer Entscheidungen durch unabhängige Dritte

Zu diesen Themen lässt sich vieles aus der Debatte über die gesellschaftliche Evaluierbarkeit von ADM-Prozessen allgemein lernen und übertragen, die in den Vereinigten Staaten etwas weiter ist als hierzulande. Welche Anforderungen hier an reichweitenstarke Intermediäre zu stellen sind, muss konkretisiert werden. Nur auf Basis hinreichend klar formulierter Gütekriterien lässt sich ermitteln, inwieweit Intermediäre im Detail Leitwerte von Öffentlichkeit erfüllen.

Konkrete Modelle für Transparenzanforderungen formuliert beispielsweise Diakopoulos (2016a: 60), der für die externe Evaluation relevante Informationen über die Funktionsweise von ADM-Prozesse in diese fünf Kategorien einteilt:

- Wo sind Menschen involviert?
- Welche Daten nutzt der Prozess?
- Wie sieht das Modell aus?
- Welche Schlussfolgerungen leitet das Verfahren ab?
- Wo und wann ist das Verfahren im Einsatz?

7.1.4 Externe Evaluation fördern (Mesoebene)

Für die Umsetzung externer Evaluation existiert eine Reihe von Vorschlägen, deren Umsetzbarkeit auszuarbeiten wären. Zum Beispiel: abgestufte Transparenz gegenüber Institutionen und der Öffentlichkeit: Wo öffentliche Transparenz heikel ist, weil sie zum Beispiel missbraucht werden könnte, schlägt Tutt als Lösung abgestufte Transparenz vor.

„On the lighter end, an agency could require that certain aspects of certain machine-learning algorithms (their code or training data) be certified by third-party organizations, helping to preserve the trade secrecy of those algorithms and their training data. Intermediately, an agency could require that companies using certain machine-learning algorithms provide qualitative disclosures (analogous to SEC disclosures) that do not reveal trade secrets or other technical details about how their algorithms work but nonetheless provide meaningful notice about how the algorithm functions, how effective it is, and what errors it is most likely to make“ (Tutt 2016: 18).

Algorithmische Prozesse, die sich laufend fortentwickeln, müssen anders evaluiert werden als statische Verfahren. Sandvig et al. (2014) unterscheiden fünf Ansätze für die Erforschung solcher Systeme von außen:

1. Quellcode-Prüfung
2. nicht invasive Methoden wie die Befragung von Nutzern
3. Extrahieren von Daten aus zugänglichen Quellen (Scrapen) zur Analyse
4. systematische Tests mithilfe von Sockenpuppen-Konten (zu den Schwierigkeiten dieses Ansatzes s. o.)
5. kollaborative Erforschung mit freiwilligen Mitgliedern oder bezahlten Mitgliedern Erforschung per Crowdsourcing

Sandvigs Crowdsourcing-Idee erinnert an das repräsentative Fernsehpanel, das in Deutschland zur Analyse von TV-Quoten genutzt wird. Die Parallele zeigt: Aus der Fachdebatte über die Erforschung und Evaluation von ADM-Prozessen allgemein ließen sich viele konkrete Konzepte für die Untersuchung algorithmischer Prozesse bei Intermediären transferieren.

Diese Beispiele stellen beileibe keine erschöpfende Abbildung der mittlerweile umfangreichen Debatte über die wissenschaftliche Nutzung von Daten aus den Beständen der Betreiber dar, werfen aber ein Schlaglicht auf mögliche regulatorische Ansatzpunkte. Vielleicht sind bestimmte Transparenzaufgaben nötig, damit die dazu

notwendige Datenbasis verfügbar wird. Eine forschungsfreundliche Regelung der Zugriffsrechte und -möglichkeiten auf die Datenbanken der Betreiber könnte einiges zur Transparenz von öffentlichkeitsrelevanten ADM-Systemen beitragen, ohne eine Offenlegung der Algorithmen selbst erforderlich zu machen, gegen die in der Tat gewichtige Gründe sprechen (siehe z. B. Diakopoulos 2016a).

Die Entwicklung und die Erprobung von Prototypen solcher Instrumente können der erste und der dritte Sektor durch gezielte, lösungsorientierte Förderung vorantreiben. Wie Sandvig schreibt: Eine Gestaltung algorithmischer Systeme und Prozesse für das Gemeinwohl im Sinne breiter Teilhabe braucht öffentliche und zivilgesellschaftliche Aktivität: „... regulating for auditability also implies an important third-party role for government, researchers, concerned users, and/or public interest advocates to hold Internet platforms accountable by routinely auditing them. This would require financial and institutional resources that would support such an intervention: a kind of algorithm observatory acting for the public interest“ (Sandvig et al. 2014: 18).

7.1.5 Alternative Relevanzprognosen entwickeln (Mesoebene)

Abhängig von den Erkenntnissen solcher Evaluationen wären Anstrengungen für die Verbesserung der Signale zum Beispiel für Relevanz im Sinne der Leitwerte nötig. Es gibt eine Reihe von Projekten zu Lösungen für einzelne Herausforderungen, die auf Relevanz für und Übertragbarkeit nach Deutschland untersucht werden könnten. Im Rahmen des *The Trust Project* an der University of Santa Clara entwickeln Forscher zum Beispiel einen Katalog von Signalen für redaktionelle Qualität, die sich aus dem HTML-Quelltext von Websites ziehen lassen (Filloux 2017). Wenn das konsistent funktioniert und valide Ergebnisse liefert, wäre dies eine wichtige Ergänzung zur Relevanzeinschätzung von Intermediären wie Facebook.

7.1.6 Leitwerte institutionell verankern und dynamisch entwickeln (Mesoebene)

Im Bereich der redaktionell kuratierten Medien in Deutschland sichern und entwickeln zahlreiche Formen der Regulierung, Koregulierung und reinen Selbstregulierung die Leitwerte von Öffentlichkeit. Einrichtungen wie der Presserat und die Rundfunkräte sind exemplarische Instrumente, die dazu dienen, diese Anforderungen umzusetzen:

- **Angemessenheit:** Vor dem Einsatz eines neuen ADM-Prozesses braucht es Verständigung über Zielsetzung, gesellschaftliche Wirkung und grundsätzliche Zulässigkeit in einem breiten Dialog.
- **Verantwortlichkeit:** Zuständigkeiten für guten Einsatz algorithmischer Prozesse müssen auf jeder Stufe des Prozess klar verortet sein und wahrgenommen werden.

Es ist zu prüfen, wie sich Einrichtungen wie der Presserat auf Intermediäre übertragen lassen, die Öffentlichkeit strukturieren. In den Vereinigten Staaten sind erste Versuche professionsethischer Grundsätze von der *Association for Computing Machinery (ACM)* (USACM 2017), *Fairness, Accountability, and Transparency in Machine Learning (FAT/ML)* (2016) und den *Asilomar-KI-Prinzipien* gelegt. Diese Dokumente behandeln zwar ADM-Prozesse im Allgemeinen, sind jedoch eine gute Basis für den Transfer auf ADM-Prozesse, die Öffentlichkeit strukturieren. Die Entwicklung des Pressekodex und Presserats in Deutschland zeigt, dass eine dynamische Spruchpraxis und Konkretisierung am Einzelfall auch zur Wirkung auf individueller Ebene führt: Die Entscheidungen des Presserats in ähnlichen Fällen dienen im journalistischen Alltag oft als Maßstab für das eigene Handeln. Vergleichbare Effekte lassen sich mittelfristig von einer Professionsethik für ADM-Entwickler erhoffen.

Als konkrete Instrumente und Träger einer Professionsethik sind verpflichtende soziale Folgeabschätzungen denkbar, wie sie die Gruppe FAT/ML für ADM-Prozesse im Allgemeinen vorschlägt: „When the system is launched, the statement should be made public as a form of transparency so that the public has expectations for social impact of the system“ (FAT/ML 2016).

7.1.7 Leitwerte individuell verankern durch Professionsethik (Mikroebene)

Eine in der oben beschriebenen Art entwickelte und institutionalisierte Professionsethik muss auf individueller Ebene bei allen Personen wirken, die algorithmische Systeme gestalten. Um das zu erreichen, könnten berufsethische Aspekte Teil der Ausbildung werden (Zweig 2017).

7.1.8 Individuelle Kompetenzen im Umgang mit algorithmischen Systemen fördern (Mikroebene)

Entwickler, professionelle Anwender und Bürger brauchen Wissen über Grundlagen, Fehlerquellen und Folgen algorithmischer Prozesse allgemein sowie die Annahmen und Nebenwirkungen einzelner Prozesse, mit denen sie konfrontiert sind. Hier braucht es Sensibilisierung, Bildungs- und Beratungsangebote (etwa durch Verbraucherschutzstellen), aber auch Instrumente wie standardisierte, öffentliche, allgemeinverständliche Selbsteinschätzungen der Entwickler eines algorithmischen Prozesses über Grundannahmen, Nebenwirkungen und Ähnliches (Zweig 2017).

7.2 Systematischen Verzerrungen der menschlichen Wahrnehmung begegnen

7.2.1 Forschung und externe Evaluation (Makroebene)

Die anwendungsorientierte Forschung zur Optimierung von Benutzeroberflächen für die digitalen Plattformen der Intermediäre konzentriert sich derzeit primär auf das Absenken von Hürden und Schwellen. Das Ziel ist es in der Regel, Benutzerinteraktion so einfach und reibungslos wie möglich zu gestalten, um eine möglichst hohe Interaktionshäufigkeit zu erzielen (siehe Abschnitt 5.4). Eyal listet eine Reihe beispielhafter Fragen auf, die der Designer gemäß den Ansätzen von B. J. Fogg zufolge beantworten müsse, um „die Wahrscheinlichkeit zu erhöhen, dass ein bestimmtes Verhalten auftritt“:

„Hat der Nutzer es eilig? Ist das Verhalten mit zu hohen Kosten verbunden? Ist der Nutzer nach einem langen Arbeitstag erschöpft? Ist das Produkt zu schwierig zu verstehen? Befindet sich der Nutzer in einem sozialen Kontext, in dem das Verhalten als unangemessen wahrgenommen werden könnte? Liegt das Verhalten so weit von der normalen Routine des Nutzers entfernt, dass seine Fremdheit abschreckt?“ (Eyal 2014).

Eyal empfiehlt sogar explizit, kognitive Heuristiken, die eigentlich Verzerrungen darstellen, wie Anker- oder Framingeffekte (siehe auch Kahneman 2012a; Tversky und Kahnemann 1974) gezielt auszunutzen, um bestimmte Verhaltensweisen besonders einfach und effektiv auszulösen.

Mit anderen Worten: Captology, die Wissenschaft der digitalen Verhaltensmodifikation, basiert maßgeblich auf der Optimierung von Angeboten für System-1-Verarbeitung (Kahneman 2012b), also schnelle, wenig anstrengende, tendenziell emotionale, fehleranfällige Formen von Kognition, die für viele komplexere Aufgaben, wie die Auseinandersetzung mit gesellschaftlichen Zusammenhängen, ungeeignet sind.

Zu begrüßen wäre deshalb ein Forschungsprogramm, das in die andere Richtung strebt: Anwendungsforschung, die anstrengende, aber tiefe und gründliche Auseinandersetzung mit Inhalten begünstigt (System-2-Verarbeitung). Zu beantworten wären nicht Fragen wie „Wie kann ich ein extrem einfaches Verhalten möglichst häufig und verlässlich auslösen?“ – etwa mittels eines Klicks auf den „Like“-Button –, sondern Fragen wie „Wie kann ich dafür sorgen, dass jemand einen Artikel nur dann weiterreicht oder kommentiert, wenn er ihn tatsächlich rezipiert und im Idealfall auch verstanden hat, wenn er also kognitiven Aufwand betrieben hat, bevor er oder sie zur Tat schreitet?“

7.2.2 Vielfalt der Foren für menschliche Interaktion sichern (Mesoebene)

Die Gestaltung der digitalen Angebote beeinflusst, welche Art menschlicher Interaktion dominiert. Maßgeblich auf schnelle, wenig anstrengende, tendenziell emotionale Formen von Kognition optimierte digitale Angebote (vgl. Abschnitt 7.2.1) fördern bestimmte Formen des Umgangs und der Diskussionskultur. Wie Lobo und Lauer (2015)

feststellen: „Die dominierenden sozialen Medienplattformen eignen sich hervorragend zur Mobilisierung, zur Gruppenbildung und zur blitzartigen Informationsverbreitung – aber nicht zur produktiven, politischen Diskussion“ (a. a. O.: loc. 2436 ff.).

Ein Ansatzpunkt für Interventionen ist hier die Gestaltung alternativer Plattformen für begrenzte Anwendungsbereiche durch gemeinwohlorientierte Akteure. Es geht nicht um neue soziale Netzwerke, sondern darum, mittels klar einzugrenzender Zielgruppen und Einsatzfelder zu ermitteln, wo Bedarf an digitalen Angeboten, deren algorithmische Prozesse andere kognitive Verarbeitungsprozesse fördern, besteht. Angebote also, die auf andere Ziele hin optimiert sind, als einfaches Verhalten endlos möglichst häufig und verlässlich auszulösen. Alternatives Ziel der Optimierung könnte unter anderem sein: Konsensfindung in einem strukturierten Prozess mit definiertem Ende. Beispiele für Projekte mit solchen Zielen: das von der Bundeszentrale für politische Bildung unterstützte Aula-Projekt, bei dem Schüler eigene Ideen zur Verbesserung ihrer Schule mittels einer auf Liquid Democracy basierenden Plattform entwickeln, Mehrheiten finden, Kompromisse aushandeln und abstimmen (Dobusch 2015). Das *Coral Project* von Mozilla Foundation, Knight Foundation, New York Times und Washington Post entwickelt Software und Handreichungen für eine bessere Diskussionskultur in Foren (The Coral Project 2016).

7.2.3 Immunisierung durch Angebotsdesign (Mesoebene)

Es gibt Ansätze, mit denen nach Meinung einiger Forscher gezielter Desinformation bzw. deren Wirkung prophylaktisch entgegengewirkt werden kann. Van der Linden et al. (2017) etwa zeigten experimentell mit einer großen Stichprobe, dass insbesondere detaillierte Warnungen vor bevorstehenden Desinformationsversuchen mit Blick auf Thesen zum menschengemachten Klimawandel die Wirkung derartiger Desinformation tatsächlich neutralisieren können. Die Autoren sprechen von einer „Inokulation“ gegen Desinformation, wie sie von Leugnern des menschengemachten Klimawandels insbesondere in den USA mit großer Vehemenz betrieben wird. Allerdings bleibt unklar, inwieweit sich dieses Ergebnis auf andere Themenbereiche und Szenarien generalisieren lässt. Zudem wird die gleiche „Inokulations“-Methode offenbar auch von den Verbreitern von Desinformation durchaus erfolgreich angewendet: Sowohl die fortgesetzte Beschimpfung der kritisch über seine Amtsführung berichtenden Medien als „dishonest media“ durch US-Präsident Donald Trump als auch der in Deutschland in neurechten Kreisen beliebte Kampfbegriff „Lügenpresse“ zur pauschalen Diskreditierung der deutschen Medienlandschaft können als Versuche gedeutet werden, kritische Berichterstattung durch eine „Inokulation“ mit umgekehrten Vorzeichen als zweifelhaft zu brandmarken.

Ein in der einschlägigen Forschungsliteratur in jüngerer Zeit populäres Konzept zur Erklärung bestimmter Wahrnehmungs- und Verarbeitungsverzerrungen insbesondere im Hinblick auf politisch konnotierte Information ist der des *politically motivated reasoning* (siehe z. B. Taber, Cann und Kucsova 2009). In den Worten von Taber und Kollegen: „Citizens’ prior attitudes toward the people, groups, or issues implicated in political arguments strongly bias how they process those arguments, through selective exposure or selective judgment processes“ (a. a. O.).

Dieses Konzept fußt auf älteren sozialpsychologischen Konzepten wie dem der kognitiven Dissonanzreduktion (Festinger 2001). Eine Forschergruppe hat kürzlich jedoch auf Ergebnisse hingewiesen, die angesichts der Probleme, die *politically motivated reasoning* verursacht, Hoffnung macht: Kahan et al. (2016) wollen gezeigt haben, dass „wissenschaftliche Neugier“ (*scientific curiosity*) ein schützender Faktor sein könnte, der in der Lage ist, die verzerrenden Auswirkungen politisch motivierter Kognition auszugleichen. Den Autoren zufolge existiert ein – in der Fachwelt lang hoch umstrittener – Persönlichkeitszug, der sich charakterisieren lässt als Freude an der Überraschung über Erkenntnisse, die dem eigenen Weltbild zuwiderlaufen. Diesen Zug nennen die Autoren *wissenschaftliche Neugier*. Er sei bei allen Menschen mehr oder minder ausgeprägt vorhanden und könne bei Personen, bei denen diese Ausprägung besonders stark ist, den Verzerrungen politisch motivierter Kognition entgegenwirken. Die Forschung in diesem Bereich steht allerdings noch an ihrem Anfang – und es ist unklar bis zweifelhaft, ob wissenschaftliche Neugier etwa durch Bildungsanstrengungen auch bei Personen geweckt werden könnte, bei denen dieser Zug nicht schon naturgemäß ausgeprägt ist.

Ein weiterer Ansatzpunkt, um einigen der in Abschnitt 0 erläuterten Wirkfaktoren entgegenzuwirken, könnten Designveränderungen aufseiten der Intermediäre sein, die der langsamen, geordneten Verarbeitung den Vorzug vor der derzeit favorisierten schnellen System-1-Verarbeitung einräumen (siehe Abschnitt 5.4). Möglichkeiten wären beispielsweise die gezielte Aktivierung sogenannter *Metakognitionen*, also des Nachdenkens über das eigene Denken (Alter et al. 2007), über die aktive Verzögerung bestimmter Reaktionen auf Inhalte oder über Methoden, die eine emotionale Distanzierung vom jeweiligen Gegenstand begünstigen (siehe z. B. Costa et al. 2014; Geipel, Hadjichristidis und Surian 2015).

Ein System, das augenscheinlich einen derartigen Effekt erzielen soll, erprobt derzeit das Web-Angebot des norwegischen öffentlich-rechtlichen Fernsehens: Bei *NRKbeta* müssen Nutzer, die einen Kommentar unter bestimmten Artikeln hinterlassen wollen, zunächst drei Multiple-Choice-Fragen beantworten, die den Inhalt des Artikels betreffen (Lichterman 2017). Ohne auf psychologische Erkenntnisse oder Theorien zu rekurrieren, begründete einer der Redakteure des betreffenden Angebots die Maßnahme unter anderem so: „Wenn Sie dafür 15 Sekunden aufwenden, sind das vielleicht die 15 Sekunden, die dem Wutmodus, in dem Leute sonst oft kommentieren, die Wucht nehmen“ (a.a.O.). Zudem stelle das System sicher, dass Menschen, die den Artikel kommentieren, ihn tatsächlich gelesen hätten.

7.2.4 Vielfältiger individueller Input (Mikroebene)

Ein erster Schritt der Immunisierung gegen mögliche teilhaberelevante Verzerrungen der individuellen medialen Umwelt ist zweifellos ein Hinarbeiten auf eine gewisse Vielfalt im medialen Input des Einzelnen. Verzerrte Abbilder der Öffentlichkeit oder gar Irreführung durch gezielte Desinformation über Intermediäre fallen naturgemäß weniger ins Gewicht, wenn der Einzelne neben diesen Informationsquellen noch weitere nutzt. Vor diesem Hintergrund kann beispielsweise im schulischen Kontext die Entwicklung einer Grundsensibilität für den Wert medialer Vielfalt und medialer Qualität eine vermutlich immunisierende Wirkung haben.

7.2.5 Sensibilisierung für die Wirkung kognitiver Verzerrungen (Mikroebene)

Nicht nur, aber auch weil Entwicklungen, wie die in den vorangegangenen Abschnitten vorgeschlagenen, einige Zeit in Anspruch nehmen werden, sollten Maßnahmen zur Kontrolle und Überwachung der Auswirkungen algorithmischer Entscheidungsfindung von solchen zur individuellen Sensibilisierung für diesen Themenbereich flankiert werden. Dies könnte zum Beispiel in Form spezieller Unterrichtseinheiten oder Fortbildungsveranstaltungen stattfinden. Unabhängig von der konkreten wissenschaftlichen Entwicklung in diesen Bereichen bleibt festzuhalten, dass eine an Fakten, Objektivierbarkeit und gründlicher Quellenprüfung orientierte Bildung und Ausbildung unerlässlich bleiben, ja noch deutlich wichtiger werden wird, um die Bevölkerung so weit wie möglich gegen gezielte Desinformation etwa in Form der Ablehnung des wissenschaftlichen Konsenses zu immunisieren.

Ein erster Schritt zur Sensibilisierung könnte und sollte sein, Nutzer algorithmisch kuratierter Angebote überhaupt auf die Tatsache hinzuweisen, dass hier maschinelle Sortierprozesse am Werk sind. Zumindest in Bezug auf Facebook gibt es diverse empirische Hinweise, dass vielen Nutzern die Tatsache, dass ihr News Feed basierend auf bestimmten, für sie selbst intransparenten Kriterien sortiert wird, gänzlich unbekannt ist. So stellte Meredith Morris fest, dass manche Nutzer irritiert waren von der Häufigkeit, mit der sie Fotos neugeborener Babys in ihren News Feeds zu sehen bekamen und daraus den Schluss zogen, junge Mütter überfluteten Facebook mit Bildern ihres Nachwuchses (Morris 2014). Tatsächlich hielt diese Annahme einer empirischen Überprüfung nicht stand: Die wahrgenommene Häufigkeit von Babyfotos auf Facebook basierte auf der Tatsache, dass diese Bilder besonders viele Likes und Kommentare hervorriefen und daher vom Algorithmus häufig weit nach oben sortiert wurden.

Eslami et al. (2015a; 2015b) stellten in einer aufwendigen qualitativen Studie mit 40 Teilnehmern fest, dass weniger als die Hälfte sich der Existenz eines sortierenden Algorithmus bei Facebook bewusst war (siehe auch Abschnitt 3.2). Sie entwickelten eine Anwendung namens *FeedVis*, die es ermöglicht, neben der algorithmisch kuratierten Darstellung des persönlichen News Feed einer Person auch eine unkuratierte Darstellung sichtbar zu machen, in der die einzelnen Postings vollständig und in umgekehrt chronologischer Reihenfolge erscheinen.

FeedVis erlaubt dem Nutzer außerdem, selbst bestimmte Prioritäten festzulegen, nach denen die einlaufenden Posts sortiert werden sollen, etwa einzelne, bislang unterrepräsentierte Kontakte häufiger zu zeigen oder die Anzeigehäufigkeit anderer Kontakte zu verringern.

Eslami et al. (2015b) berichten, dass in einer ihrer Studien 83 Prozent der Teilnehmer im Anschluss an die Konfrontation ihres ungefilterten News Feed über FeedVis ihr Facebook-Verhalten verändert hätten. Mehrere begannen die Einstellungsmöglichkeiten des Netzwerks aktiv zu nutzen, andere änderten ihre Interaktionsmuster mit Kontakten auf der Plattform, um dem Algorithmus zu signalisieren, von welchen Personen oder Seiten sie gerne mehr Inhalte sehen wollten. Mehrere Teilnehmer hätten in Nachbefragungen beispielsweise erklärt, sie seien jetzt selektiver beim Klick auf den „Gefällt mir“-Button, „weil das ja Auswirkungen darauf hat, was ich in Zukunft sehe oder nicht sehe“. Eine Teilnehmerin habe sogar die Nutzung von Facebook vollständig eingestellt, weil sie sich nach den Erkenntnissen über die algorithmische Kuratierung der Inhalte „betrogen“ gefühlt habe (a. a. O.).

Derartige Gegenüberstellungen etwa im Rahmen von Unterrichtseinheiten oder Fortbildungsveranstaltungen wären eine einfach umzusetzende und vermutlich nachhaltig wirksame Methode der Aufklärung über die innerhalb solcher Angebote angewandten Mechanismen, die augenscheinlich sowohl Bewusstseins- als auch Verhaltensänderung befördern kann.

Eine nicht personalisierte, aber unmittelbar eindrucksvolle Methode, mögliche Verzerrungen durch Personalisierung sichtbar zu machen, sind Projekte wie Jon Keegans *Blue Feed, Red Feed* (Keegan 2016). Auf Basis der in Abschnitt 5.6 zitierten *Science*-Studie, die bei Facebook angestellte Wissenschaftler durchgeführt hatten, entwickelte Keegan für das Wall Street Journal zwei Facebook-Konten, die jeweils ausschließlich Artikel enthielten, die laut den Daten von Bakshy et al. (2015) von entweder „sehr konservativ“ oder „sehr liberal“ eingestellten Facebook-Accounts geteilt wurden. Das Ergebnis, das fortlaufend weiter gepflegt wird, ist ein thematisch sortierbarer Einblick in die mediale Wahrnehmungswelt zweier unterschiedlicher, fiktiver Facebook-Nutzer. Vergleichbare Demonstrationen ließen sich zweifellos auch für andere Länder und andere politische oder andere Ausrichtungen aufsetzen.

Ein weiterer Schritt, der sich in diesen Abschnitt einordnen lässt, sind die seit einigen Monaten von Facebook selbst angestellten Bemühungen, nachweisliche Falschmeldungen nach Prüfung durch externe Organisationen wie das unabhängige Recherchebüro *Correctiv* (Schraven 2017) von der Seite zu entfernen). Die gezielte Markierung nachweislich manipulativer Berichte könnte unter Umständen einen sensibilisierenden Effekt haben und Nutzer auch auf die Tatsache aufmerksam machen, dass die algorithmisch sortierte Zusammenstellung ihres News Feed auch Falschmeldungen enthalten kann. Welche Auswirkungen die Konfrontation mit faktischen Gegenargumenten jedoch tatsächlich hat, insbesondere bei bereits radikalisierten Personen, ist derzeit mindestens unklar (siehe Abschnitt 5.6).

8 Fazit

Im Kern beantwortet dieses Arbeitspapier drei zentrale Fragen der Öffentlichkeit im digitalen Zeitalter:

1. *Medienwandel: Wie verändert sich die Öffentlichkeit durch neue digitale Plattformen, über die heute viele Menschen gesellschaftlich relevante Informationen beziehen?*
 Von algorithmischen Prozessen gesteuerte Intermediäre wie Google oder Facebook haben – wenn man alle Altersgruppen zusammen betrachtet – im Vergleich zu redaktionell gesteuerten Medien wie dem Fernsehen einen großen, aber bislang nicht entscheidenden Einfluss auf die öffentliche Meinungsbildung. Sie bewerten die Relevanz von Inhalten viel stärker als redaktionelle Medien an unmittelbaren Reaktionen des Publikums.
2. *Gesellschaftliche Folgen: Sind Informationen, die Menschen auf diesen neuen Wegen erreichen, in ihrer Qualität und Vielfalt als Basis demokratischer Willensbildung und geeignet und teilhabeförderlich?*
 Die Nutzung der Intermediäre für öffentliche Meinungsbildung führt zu einem Strukturwandel der Öffentlichkeit. Zentral sind dabei algorithmische Prozesse als wesentliches Gestaltungsmittel und die bedeutende Rolle von Nutzerreaktionen als Input dieser Prozesse. Eine Reihe psychologischer Faktoren führt dazu, dass die ausgewerteten vor allem impulsiven Publikumsreaktionen schlecht geeignet sind, um Relevanz im Sinne der klassischen Leitwerte zu bewerten. Jener Leitwerte wie Wahrheit, Vielfalt oder gesellschaftliche Integration, die in Deutschland Grundlage der von redaktionellen Medien geschaffenen Öffentlichkeit sind.
3. *Lösungsansätze: Welche Ansatzpunkte sind denkbar, um die neuen digitalen Plattformen teilhabeförderlich zu gestalten?*
 Im Zentrum der komplexen Wechselwirkungen digitaler Öffentlichkeit stehen algorithmische Prozesse, die Inhalte sortieren und die Zusammenstellung personalisieren. Deshalb müssen Lösungen zuerst hier ansetzen. Zu den derzeit absehbaren wichtigsten Handlungsfeldern zählen die Ermöglichung externer Beforschbarkeit und Evaluation, Stärkung der Vielfalt algorithmischer Prozesse, Verankerung von Leitwerten, z. B. durch eine Professionsethik, und die Sensibilisierung des Publikums.

1. Medienwandel: Intermediäre sind relevant für die Meinungsbildung, aber nicht entscheidend.

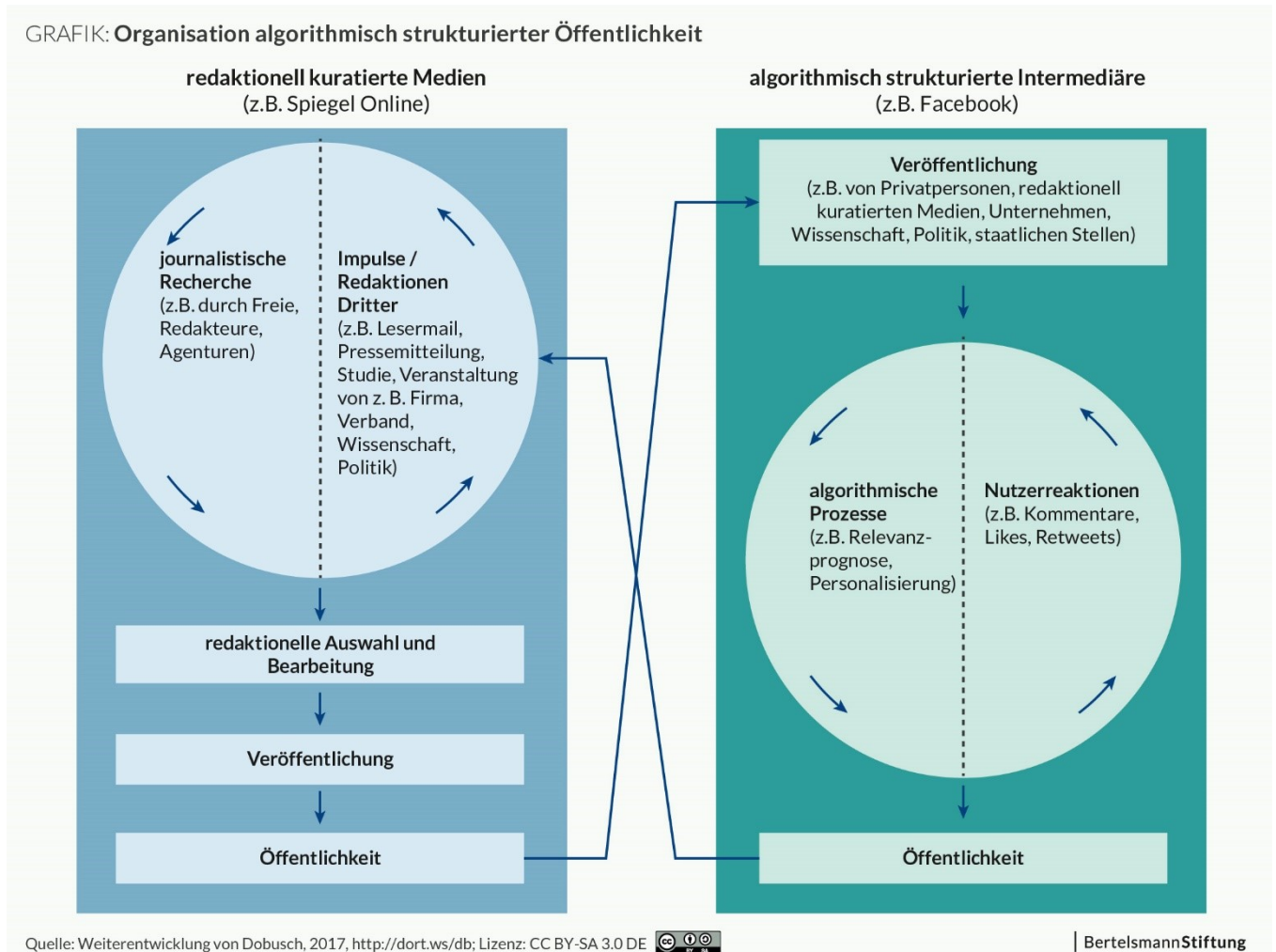
Dass sogenannte *Intermediäre* wie Google oder Facebook eine relevante Rolle für die Meinungsbildung spielen, auch hierzulande, zeigen zahlreiche Studien. 57 Prozent der deutschen Internetnutzer beziehen auch politisch-gesellschaftliche Informationen über Suchmaschinen oder soziale Netzwerke. Zwar ist der Anteil derer, die soziale Netzwerke als ihre wichtigste Nachrichtenquelle nennen, mit sechs Prozent aller Internetnutzer noch relativ klein – doch in den jüngeren Altersgruppen sind die Anteile deutlich höher. Insgesamt ist von einer zunehmenden Bedeutung solcher Plattformen auszugehen. Meinungsbildung sei „ohne Intermediäre nicht mehr denkbar“, so formulierten es 2016 Forscher vom Hamburger Hans-Bredow-Institut.

Die Gestaltungsprinzipien dieser Intermediäre führen zu einem Strukturwandel der Öffentlichkeit. Zentrale Aspekte sind:

- *Entkopplung von Veröffentlichung und Reichweite:* Jeder kann veröffentlichen. Aber nicht jeder findet ein Publikum. Aufmerksamkeit entsteht erst durch das Zusammenwirken von Menschen und ADM-Prozessen.
- *Entbündelung von Publikationen:* Reichweite wird auf Beitragsebene ausgehandelt.
- *Personalisierung:* Nutzer erfahren mehr über ihre Interessengebiete
- *Größerer Einfluss des Publikums auf Reichweiten:* Nutzerreaktionen beeinflussen ADM-Prozess insgesamt und die Reichweite jedes Beitrags.
- *Zentralisierung der Auswahlinstanzen:* Es gibt bei Intermediären eine deutlich geringere Vielfalt als bei redaktionell kuratierten Medien.
- *Wechselwirkung zwischen redaktioneller und maschineller Kuratierung:* Redaktionell kuratierte Medien verbreiten Inhalte über Intermediäre und nutzen Reaktionen in diesen als Signal für Publikumsinteresse.

Der Vergleich der beiden Öffentlichkeitsverläufe in Abbildung 2 zeigt die neue, zentrale Rolle von Nutzerreaktionen und algorithmischen Prozessen. Beide bestimmen die Verteilung von Aufmerksamkeit über Intermediäre. Unsere Hypothese ist, dass Nutzerreaktionen und ADM-Prozesse sich dabei nicht eindeutig in eine lineare Kausalkette bringen lassen.

Abbildung 2: Organisation algorithmisch strukturierter Öffentlichkeit



Google, Facebook und Co. spielen im öffentlichen Diskurs also bereits jetzt wichtige Rollen. Dabei wurden diese Plattformen ursprünglich nicht primär dafür konstruiert, Medieninhalte von journalistischen Organisationen an Konsumenten weiterzureichen. Sie benutzen technische Systeme, um zu entscheiden, ob ein bestimmter Inhalt aus einem gewaltigen Fundus für einen bestimmten Nutzer interessant, relevant sein könnte oder nicht. Diese Systeme waren ursprünglich aber eher darauf ausgelegt, etwa – im Falle von Suchmaschinen – Webseiten auszuwerfen, die eine bestimmte Information enthalten, oder – im Falle von sozialen Netzwerken – besonders interessante Wortmeldungen oder Fotos aus dem eigenen Freundeskreis prominent zu platzieren. Sie sortieren Inhalte deshalb nach teilweise völlig anderen Kriterien, als beispielsweise die Redakteure einer Tageszeitung oder eines Magazins das tun würden. Relevanz bedeutet für Google etwas anderes als für Facebook und beide verstehen unter dem Begriff etwas anderes als die Redaktionen beispielsweise von SPIEGEL ONLINE oder Sueddeutsche.de.

Die Intermediäre erfassen zur Errechnung solcher Relevanzwerte für einzelne Inhalte eine Vielzahl von Variablen. Diese Signale reichen von basalen Verhaltensmaßen wie der Scrollgeschwindigkeit oder der Verweildauer auf einzelnen Seiten bis hin zum Grad der Interaktion zwischen mehreren Nutzern eines sozialen Netzwerks. Wenn eine Person, mit der man bei Facebook schon öfter kommuniziert hat, einen Inhalt postet, ist die Wahrscheinlichkeit höher, dass man diesen Inhalt zu sehen bekommt als bei einer anderen Person, mit der

man zwar theoretisch digital verknüpft ist, praktisch aber nie in Kontakt tritt. Auch die Signale, die andere Nutzer – oft unwissentlich – aussenden, gehen in die Relevanzbewertung mit ein, seien es Verlinkungen, Klicks auf Links oder den „Gefällt mir“-Button, Weiterreichungen, sogenannte *Shares*, oder die Anzahl der Kommentare, die ein bestimmter Inhalt hervorruft.

2. Gesellschaftliche Folgen: Die heute von den für die Meinungsbildung besonders relevanten Intermediären genutzten algorithmischen Systeme werten menschliche Reaktionen auf Inhalte aus. Sie fördern und verstärken einen für Verzerrungen anfälligen menschlichen Kognitionsstil und sind für gezielte technische Manipulationen anfällig.

Die Relevanzsignale, über die die Plattformbetreiber aus Wettbewerbs- und anderen Gründen nur ungern detaillierte Auskünfte geben, sind potenziell problematisch. Zunächst einmal deshalb, weil die Betreiber der entsprechenden Plattformen sie selbst ununterbrochen verändern: Systeme wie die von Google oder Facebook sind in einem permanenten Wandel begriffen, an nahezu jedem Aspekt der Benutzeroberfläche und anderen Eigenschaften der Plattformen wird experimentiert und geschraubt, um bestimmte Effekte wie zum Beispiel Interaktionsintensität zu erzielen. Jede dieser Veränderungen beeinflusst ihrerseits potenziell die Relevanzsignale, die die Plattformen selbst messen.

Ein gutes Beispiel ist die „People you may know“-Funktion bei Facebook, die auf Basis von Netzwerkauswertungen mögliche Bekannte des jeweiligen Nutzers als zusätzliche Kontakte vorschlägt. Als dieses Empfehlungssystem eingeführt wurde, verdoppelte sich auf einen Schlag die Anzahl der Facebook-internen Verknüpfungen, die jeden Tag hinzukommen. Die innerhalb von Netzwerkplattformen abgebildeten Beziehungsnetzwerke hängen also von den Angeboten ab, die die Betreiber machen. Gleichzeitig gehen die so katalogisierten Bekanntschaftsnetzwerke als Variablen in die Relevanzbestimmung ein. Wer zusätzliche „Freunde“ hat, bekommt womöglich auch andere Medieninhalte zu sehen.

Ein weiteres Problem der von den Plattformbetreibern erfassten Signale hat mit der Art von Interaktion zu tun, für die solche Plattformen optimiert sind: Ein zentrales Credo der Gestaltung lautet, dass Interaktionen möglichst einfach und mühelos sein sollten, um ihre Wahrscheinlichkeit zu maximieren. Auf den „Gefällt mir“-Button oder einen Link zu klicken, erfordert keinerlei kognitive Anstrengung. Und von dieser Anstrengungslosigkeit machen viele Nutzer augenscheinlich intensiv Gebrauch: So legen empirische Untersuchungen nahe, dass viele Artikel, die in sozialen Netzwerken mit einem Klick an den eigenen digitalen Freundeskreis weitergereicht werden, zuvor nicht gelesen worden sein können. Nutzer verbreiten also Medieninhalte weiter, von denen sie selbst nur die Überschrift und den Anreißertext kennen. Sie gaukeln dem Algorithmus und damit ihren „Freunden und Followern“ gewissermaßen nur vor, einen Text gelesen zu haben.

Die Leichtigkeit der Interaktion begünstigt zudem kognitive Verzerrungen, die in der Sozialpsychologie schon seit vielen Jahren bekannt sind. Ein gutes Beispiel ist die Verfügbarkeitsheuristik: Wenn ein Ereignis oder eine Erinnerung sich leicht aus dem Gedächtnis abrufen lässt, wird es oder sie als besonders wahrscheinlich oder häufig eingeschätzt. Die Konsequenz: Nicht gelesene, aber aufgrund einer Überschrift besonders häufig weitergereichte Medieninhalte begegnen Nutzern eines sozialen Netzwerks oft – und werden deshalb im Nachhinein als „wahr“ oder „wahrscheinlich“ erinnert. Das gilt auch dann, wenn der Text selbst womöglich klarmachen würde, dass die Überschrift eine groteske Übertreibung oder schlicht irreführend ist.

Auch eine Reihe anderer psychologischer Faktoren spielt hier eine wichtige Rolle, etwa die Tatsache, dass Nutzer gerade soziale Medien nicht nur zu informativen Zwecken, sondern auch als Werkzeug des Identitätsmanagements einsetzen: Manche Medieninhalte werden womöglich nur weitergereicht, um die eigene Zugehörigkeit etwa zu einem politischen Lager zu demonstrieren. Zudem begünstigt das Design vieler digitaler Plattformen flüchtige, emotionale Auseinandersetzung mit Inhalten explizit und absichtlich. Tatsächlich zeigen einschlägige Studien, dass besonders *emotionalisierende Inhalte* auf Netzwerkplattformen besonders häufig kommentiert und weitergereicht werden – vor allem dann, wenn es sich um negative Emotionen handelt.

Ein derartig emotionaler Umgang mit Nachrichteninhalten kann zu einer stärkeren *gesellschaftlichen Polarisierung* führen. Insbesondere in den USA gibt es für diese These auch erste empirische Belege. Allerdings scheinen solche Polarisierungseffekte von einer Reihe weiterer Faktoren abzuhängen, etwa dem Wahlsystem eines Landes: Gesellschaften mit Mehrheitswahlrecht wie die der USA sind womöglich anfälliger für extreme politische Polarisierung als solche mit Verhältniswahlrecht, in denen wechselnde Koalitionen regieren und ein Mehrparteiensystem institutionalisiert Interessenausgleich begünstigt. Vermutlich besteht auch eine Wechselwirkung zwischen bereits erfolgter Polarisierung und den Ergebnissen algorithmischer Sortierung von Medieninhalten. Eine Studie zeigt zum Beispiel, dass sich Anhänger von Verschwörungstheorien bei Facebook im Lauf der Zeit immer stärker der eigenen Verschwörungstheoretiker-Community zuwenden. Intensiviert wird dieser Prozess womöglich durch einen Algorithmus, der ihnen entsprechende Inhalte immer häufiger vor Augen führt. Zumindest bei Menschen mit extremen Ansichten könnten diese Systeme also tatsächlich die Entstehung sogenannter *Echokammern* befördern.

Auch *technische Manipulationen* können die Signale beeinflussen, die Intermediäre zur Relevanzbeurteilung einsetzen. Sogenannte *Bots*, das sind teilautonom agierende Softwarelösungen, die etwa in sozialen Netzwerken als echte Nutzer getarnt werden, können zumindest das Volumen der digitalen Kommunikation über bestimmte Themen massiv verzerren. Einer Studie zufolge waren etwa im Umfeld der US-Präsidentenwahlen bei Twitter 400.000 solcher Bots im Einsatz, die etwa ein Fünftel der gesamten Konversation über die TV-Debatten zwischen den Kandidaten bestritten. Es ist unklar, inwieweit solche automatisierten Systeme Menschen tatsächlich in ihrer Wahlentscheidung beeinflussen können. Klar ist dagegen, dass die von ihnen produzierten Signale – Klicks, Likes, Shares – in die Relevanzbeurteilungen der algorithmischen Entscheidungssysteme eingehen. Bots können also einen Artikel so interessant erscheinen lassen, dass der Algorithmus ihn menschlichen Nutzern zeigt.

Zusammenfassend lässt sich sagen, dass die Relevanzbewertungen algorithmischer Sortiersysteme für Medieninhalte nicht notwendigerweise gesellschaftlich wünschenswerten Kriterien folgen: Leitwerte wie Orientierung an der Wahrheit oder gesellschaftliche Integration spielen keine Rolle. Primär zählt die Optimierung der Interaktionswahrscheinlichkeit und der Verweildauer auf der jeweiligen Plattform. Interessierte Parteien, denen an gezielter Desinformation gelegen ist, können diese Mechanismen für ihre Zwecke ausnutzen: Eine kreative, zielgerichtete Lüge kann im Zweifel emotional aktivierender und damit innerhalb solcher Systeme erfolgreicher und damit reichweitenstärker sein als die langweilige Wahrheit.

3. Lösungsansätze: Im Zentrum aller komplexen Wechselwirkungen digitaler Öffentlichkeit steht die algorithmische Sortierung von Inhalten. Hier müssen Lösungen ansetzen.

Der letzte Abschnitt dieses Arbeitspapiers enthält eine Reihe von möglichen Lösungsansätzen für diese Herausforderungen. Ein erstes, vergleichsweise einfach zu erreichendes Ziel ist die Sensibilisierung der Nutzer für die genannten Prozesse und Mechanismen. Studien zeigen, dass Nutzer sozialer Netzwerkplattformen sich nicht einmal der bloßen Existenz eines Sortieralgorithmus bewusst sind, geschweige denn seiner Funktionsweise. Hier könnten Bildung und Fortbildung ebenso ansetzen wie bei einer verstärkten Sensibilisierung für und Immunisierung gegen Desinformationsversuche, etwa mithilfe einer faktenorientierten Aufklärung.

Über im Zweifel noch effektivere Interventionsmöglichkeiten verfügen die Plattformbetreiber selbst: zum Beispiel, indem sie bestimmte Leitwerte wie Angemessenheit, Verantwortlichkeit und Kompetenz, etwa bei den Entwicklung und Gestaltung der entsprechenden Systeme, stärker verankern. Mittelfristiges Ziel könnte eine branchenweite Professionsethik für Entwickler algorithmischer Entscheidungssysteme sein.

Aber auch Fachleute, die nicht für die Plattformbetreiber selbst arbeiten, könnten und sollten in die Lage versetzt werden, die Auswirkungen der dort getroffenen Entscheidungen wissenschaftlich zu begleiten und zu erforschen. Derzeit ist der Zugriff auf dafür notwendige Daten, die den Betreibern selbst in gewaltiger Zahl vorliegen, für externe Wissenschaftler oder Regulierungsbehörden mühselig bis unmöglich. Sowohl die Designentscheidungen der Plattformbetreiber als auch ihre Auswirkungen für individuelle Nutzer sind weitgehend intransparent.

Systematische Verzerrungen etwa in einer bestimmten politischen Richtung ließen sich auf Basis der derzeit verfügbaren Daten kaum erkennen. Mehr Transparenz durch eine Kombination aus freiwilligen Selbstverpflichtungen und notwendigenfalls auch regulatorischen Maßnahmen könnte unabhängiges Wissen über die tatsächlichen gesellschaftlichen Auswirkungen algorithmischer Sortierung von Medieninhalten ermöglichen und mögliche Gefahren frühzeitig erkennbar machen. Eine bessere Beforschbarkeit würde eine sachliche und lösungsorientierte Debatte fördern und könnte neue Lösungsansätze aufzeigen. Eine solche Entwicklung könnte die teilhabeförderliche Gestaltung und Wirkung algorithmischer Systeme stärken. Das würde eine differenzierte Betrachtung algorithmischer Prozesse fördern und könnte das Vertrauen in die zum Nutzen der gesamten Gesellschaft gestalteten Systeme stärken.

9 Literatur

- Alter, Adam L., Daniel M. Oppenheimer, Nicholas Epley und Rebecca N. Eyre (2007). „Overcoming intuition: metacognitive difficulty activates analytic reasoning“. *Journal of Experimental Psychology. General* (136) 4. 569–576. <https://doi.org/10.1037/0096-3445.136.4.569> (Download 2.6.2017).
- An, Jisun, Daniele Quercia und Jon Crowcroft (2013). „Fragmented social media: a look into selective exposure to political news“. *Proceedings of the 22nd International Conference on World Wide Web*. New York NY: ACM. 51–52. <http://dl.acm.org/citation.cfm?id=2487807> (Download 2.6.2017).
- Ananny, Mike, und Kate Crawford (2016). „Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability“. *New Media & Society* 15 (7). 1005–1021. <https://doi.org/10.1177/1461444816676645> (Download 2.6.2017).
- Association for Computing Machinery US Public Policy Council (USACM). (2017). „Statement on Algorithmic Transparency and Accountability“. 12.1.2017. https://www.acm.org/binaries/content/assets/public-policy/2017_usacm_statement_algorithms.pdf (Download 2.6.2017).
- Backstrom, Lars. (2013). „News Feed FYI: A Window Into News Feed“. 6.3.2013. <https://www.facebook.com/business/news/News-Feed-FYI-A-Window-Into-News-Feed> (Download 1.3.2017).
- Bakshy, Eytan (2014). „Big experiments: Big data’s friend for making decisions“. 3.4.2014. <https://www.facebook.com/notes/facebook-data-science/big-experiments-big-datas-friend-for-making-decisions/10152160441298859/> (Download 2.6.2017).
- Bakshy, Eytan, Solomon Messing und Lada A. Adamic (2015). „Exposure to ideologically diverse news and opinion on Facebook“. *Science* (348) 6239. 1130–1132. <https://doi.org/10.1126/science.aaa1160> (Download 2.6.2017).
- Beirat Integration. (2013). „Soziale Teilhabe‘ Handlungsempfehlungen des Beirats der Integrationsbeauftragten“. Hrsg. Die Beauftragte der Bundesregierung für Migration, Flüchtlinge und Integration. 22.2.2013. <http://www.bagiv.de/pdf/soziale-teilhabe-empfehlungen-beirat.pdf> (Download 2.6.2017).
- Benkler, Yochai, Robert Faris, Hal Roberts und Ethan Zuckerman (2017). „Breitbart-led right-wing media ecosystem altered broader media agenda“. *Columbia Journalism Review*. <http://www.cjr.org/analysis/breitbart-media-trump-harvard-study.php> (Download 2.6.2017)
- Bentele, Günter (2016). „Wahrheit“. *Handbuch Medien-und Informationsethik*. Hrsg. Jessica Heesen. Stuttgart. 59–66. http://link.springer.com/chapter/10.1007/978-3-476-05394-7_3 (Download 2.6.2017).

- Bertelsmann Stiftung (2011). *Soziale Gerechtigkeit in der OECD – Wo steht Deutschland? Sustainable Governance Indicators 2011*. Gütersloh. http://news.sgi-network.org/uploads/tx_amsgistudies/SGI11_Social_Justice_DE.pdf (Download 2.6.2017).
- Bessi, A., & Ferrara, E. (2016). Social bots distort the 2016 US Presidential election online discussion. *First Monday*, 21(11). Retrieved from <http://firstmonday.org/ojs/index.php/fm/article/view/7090>
- Boland, Brian (2014). „Organic Reach on Facebook: Your Questions Answered“. 5.6.2014. <https://www.facebook.com/business/news/Organic-Reach-on-Facebook> (Download 6.4.2017).
- Bond, Robert M., Christopher J. Fariss, Jason J. Jones, Adam D. I. Kramer, Cameron Marlow, Jaime E. Settle und James H. Fowler (2012). „A 61-million-person experiment in social influence and political mobilization“. *Nature* (489) 7415. 295–298. <https://doi.org/10.1038/nature11421> (Download 2.6.2017).
- Borgesius, Frederik J. Zuiderveen, Damian Trilling, Judith Moeller, Balázs Bodó, Cales H. de Vreese und Natali Helberger (2016). *Should We Worry About Filter Bubbles?* (SSRN Scholarly Paper No. ID 2758126). Rochester NY: Social Science Research Network. <https://papers.ssrn.com/abstract=2758126> (Download 2.6.2017).
- Bundesgesetzblatt (1966). „Internationaler Pakt über wirtschaftliche, soziale und kulturelle Rechte“. Bundesgesetzblatt (BGBl) 1976 II, 428. 19.12.1966. http://www.institut-fuer-menschenrechte.de/fileadmin/user_upload/PDF-Dateien/Pakte_Konventionen/ICESCR/icescr_de.pdf (Download 8.5.2017).
- Bundesprüfstelle für jugendgefährdende Medien (BPjM) (o. J.). „BPjM-Modul“. <http://www.bundespruefstelle.de/bpjm/Aufgaben/Listenfuehrung/bpjm-modul.html> (Download 2.6.2017).
- Bundesverfassungsgericht (1981). „BVerfGE 57, 295 - 3. Rundfunkentscheidung“. (16. Juni 1981). <http://sorminiserv.unibe.ch:8080/tools/ainfo.exe?Command=ShowPrintVersion&Name=bv057295> (Download 2.6.2017).
- Burri, Mira (2013). „Contemplating a ‘Public Service Navigator’: In Search of New (and Better) Functioning Public Service Media“. 5.5.2015. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2364951 (Download 2.6.2017).
- Cheng, Justin, Michael Bernstein, Cristian Danescu-Niculescu-Mizil und Jure Leskovec (2017). „Anyone Can Become a Troll: Causes of Trolling Behavior in Online Discussions“. <https://pdfs.semanticscholar.org/3427/1ca4d5c91e258c2e5bf67a5f4d4698dd1885.pdf> (Download 2.6.2017).

- Churchill, Winston (1943). „House of Commons Rebuilding“. *Hansard*, 28.10.1943. *HC* (393). 403–473.
<http://hansard.millbanksystems.com/commons/1943/oct/28/house-of-commons-rebuilding> (Download 2.6.2017).
- Clark, Jack (2015). „Google Turning Its Lucrative Web Search Over to AI Machines“. *Bloomberg.com*.
26.10.2015. <https://www.bloomberg.com/news/articles/2015-10-26/google-turning-its-lucrative-web-search-over-to-ai-machines> (Download 2.6.2017).
- Constine, Josh (2016). „How Facebook News Feed Works“. 9.9.2016.
<http://social.techcrunch.com/2016/09/06/ultimate-guide-to-the-news-feed/> (Download 9.9.2016).
- Costa, Albert, Alice Foucart, Sayun Hayakawa, Malina Aparici, Jose Apesteguia, Joy Heafner und Boaz Keysar (2014). „Your Morals Depend on Language“. *PLOS ONE* (9) 4, e94842.
<https://doi.org/10.1371/journal.pone.0094842> (Download 2.6.2017).
- Council of Europe (2012a). „Recommendation CM/Rec(2012)3 of the Committee of Ministers to member States on the protection of human rights with regard to search engines“. 4.4.2012.
https://search.coe.int/cm/Pages/result_details.aspx?ObjectID=09000016805caa87 (Download 2.6.2017).
- Council of Europe (2012b). „Recommendation CM/Rec(2012)4 of the Committee of Ministers to member States on the protection of human rights with regard to social networking services“.
https://search.coe.int/cm/Pages/result_details.aspx?ObjectID=09000016805caa9b (Download 2.6.2017).
- Cutts, Matt (2011). „Finding more high-quality sites in search (2011)“. 24.2.2011.
<https://googleblog.blogspot.de/2011/02/finding-more-high-quality-sites-in.html> (Download 2.6.2017).
- Dearringer, Jeremy (2011). „Mission ImposSERPble: Establishing Click-through Rates“. 25.7.2011.
<https://moz.com/blog/mission-imposserpble-establishing-clickthrough-rates> (Download 2.6.2017).
- Del Vicario, Michaela, Allesandro Bessi, Fabiana Zollo, Fabio Petroni, Antonio Scala, Guido Caldarelli, H. Eugene Staley und Walter Quattrociocchi (2016). „The spreading of misinformation online“. *Proceedings of the National Academy of Sciences* (113) 3. 554–559. <https://doi.org/10.1073/pnas.1517441113> (Download 2.6.2017).
- Deutscher Presserat (2017). „Politische Grundsätze (Presskodex). Richtlinien für die publizistische Arbeit nach den Empfehlungen des Deutschen Presserates. Beschwerdeordnung“. Berlin.
http://www.presserat.de/fileadmin/user_upload/Downloads_Dateien/Pressekodex2017_web.pdf
(Download 2.6.2017).
- Diakopoulos, Nicholas (2016a). „Accountability in Algorithmic Decision Making“. *Communications of the ACM* (59) 2. 56–62. <https://doi.org/10.1145/2844110> (Download 2.6.2017).

- Diakopoulos, Nicholas (2016b). „Artificial Moderation: A Reading List“. 29.3.2017.
<https://blog.coralproject.net/artificial-moderation-a-reading-list/> (Download 5.1.2017).
- die medienanstalten (2016). „Staatsvertrag für Rundfunk und Telemedien (Rundfunkstaatsvertrag - RStV) vom 31. August 1991“. http://www.die-medienanstalten.de/fileadmin/Download/Rechtsgrundlagen/Gesetze_aktuell/19_RfAendStV_medienanstalten_Layout_final.pdf (Download 2.6.2017).
- Dobusch, Leonhard (2015). „Projekt ‚aula‘ sucht Schulen, die mit Liquid Democracy experimentieren wollen“. 23.11.2017. <https://netzpolitik.org/2015/projekt-aula-sucht-schulen-die-mit-liquid-democracy-experimentieren-wollen/> (Download 27.4.2017).
- Dobusch, Leonhard (2017). „Die Organisation der Digitalität: Zwischen grenzenloser Offenheit und offener Exklusion“. 1.2.2017. <https://netzpolitik.org/2017/die-organisation-der-digitalitaet-zwischen-grenzenloser-offenheit-und-offener-exklusion/> (Download 9.2.2017).
- Donges, Patrick (2016). „Funktionsaufträge des Rundfunks“. *Handbuch Medien- und Informationsethik*. Hrsg. Jessica Heesen. Stuttgart. 89–104. http://link.springer.com/chapter/10.1007/978-3-476-05394-7_4 (Download 2.6.2017).
- Dörr, Dieter, und Richard Deicke (2015). *Positive Vielfaltsicherung. Bedeutung und zukünftige Entwicklung der Fensterprogramme für die Meinungsvielfalt in den privaten Fernsehprogrammen*. Mainz.
<http://www.mainzer-medieninstitut.de/dokumente/Studie%20-%20Positive%20Vielfaltsicherung.pdf> (Download 2.6.2017).
- Ecke, Oliver (2016). *Wie häufig und wofür werden Intermediäre genutzt?* Berlin. http://www.die-medienanstalten.de/fileadmin/Download/Veranstaltungen/Pr%C3%A4sentation_Intermedi%C3%A4re/TNS_Intermedi%C3%A4re_und_Meinungsbildung_Pr%C3%A4si_Web_Mappe_final.pdf (Download 2.6.2017).
- El-Arini, Khalid, und Joyce Tang (2014). „Click-baiting“. 25.8.2014. <https://newsroom.fb.com/news/2014/08/news-feed-fyi-click-baiting/> (Download 2.3.2017).
- Epstein, Robert, und Ronald E. Robertson (2015). „The search engine manipulation effect (SEME) and its possible impact on the outcomes of elections“. *Proceedings of the National Academy of Sciences of the United States of America* (112) 33. 4512–4521. <https://doi.org/10.1073/pnas.1419828112> (Download 2.6.2017).
- Eslami, Montahhare, Amirhossein Aleyasen, Karrie Karahalios, Kevin Hamilton und Christian Sandvig (2015a). FeedVis: A Path for Exploring News Feed Curation Algorithms. *Proceedings of the 18th ACM Conference*

Companion on Computer Supported Cooperative Work & Social Computing. New York NY: ACM. 65–68.
<https://doi.org/10.1145/2685553.2702690> (Download 2.6.2017).

Eslami, Montahhare, Aimee Rickman, Kristen Vaccaro, Amirhossein Aleyasen, Andy Vuong, Karrie Karahalios, Kevin Hamilton und Christian Sandvig (2015b). „I always assumed that I wasn't really that close to [her]": Reasoning about Invisible Algorithms in news feeds“. Gehalten auf der CHI 2015, Crossings, Seoul, Korea. 153–162. <https://doi.org/10.1145/2702123.2702556> (Download 2.6.2017).

Etzioni, Amitai, und Oren Etzioni (2017). „Incorporating Ethics into Artificial Intelligence“. *The Journal of Ethics* März 2017. 1–16.

Eslami, M., Karahalios, K., Sandvig, C., Vaccaro, K., Rickman, A., Hamilton, K., & Kirlik, A. (2016). First I “Like” It, then I Hide It: Folk Theories of Social Feeds. In Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (pp. 2371–2382). New York, NY, USA: ACM.
<https://doi.org/10.1145/2858036.2858494>

Eyal, Nir (2014). *Hooked: how to build habit-forming products*. New York NY: Portfolio/Penguin.

Facebook (2006). „Facebook Gets a Facelift“. 5.9.2006. <https://www.facebook.com/notes/facebook/facebook-gets-a-facelift/2207967130/> (Download 2.3.2017).

Facebook (2016a). „How News Feed Works“. 17.6.2017. <http://nonprofits.fb.com/2016/06/17/how-news-feed-works/> (Download 21.2.2017).

Facebook (2016b). „Werte des News Feeds“. 28.6.2017. <https://newsfeed.fb.com/values/> (Download 1.3.2017).

Facebook (o. J.). „Gemeinschaftsstandards“. <https://www.facebook.com/communitystandards> (Download 8.4.2017).

Fairness, Accountability, and Transparency in Machine Learning (FAT/ML) (2016). „Principles for Accountable Algorithms and a Social Impact Statement for Algorithms“. <http://www.fatml.org/resources/principles-for-accountable-algorithms> (Download 2.6.2017).

Festinger, Leon (2001). *A theory of cognitive dissonance*. (Reissued by Stanford Univ. Press in 1962, renewed 1985 by author, [Nachdr.]). Stanford: Stanford Univ. Press.

Fetzer, Thomas (2015). *Effektive Vielfaltssicherung im Internet*. Düsseldorf.
https://mbem.nrw/sites/default/files/asset/document/effektive_vielfaltssicherung_fetzer_finalcc_by_nd_30_de.pdf (Download 2.6.2017).

Filloux, Frederic (2017). „What Web Page Structure Reveals on News Quality – Monday Note“. 10.4.2017.
<https://mondaynote.com/what-web-page-structure-reveals-on-news-quality-67b845e230db> (Download 13.4.2017).

- Flaxman, Seth, Sharad Goel und Justin M. Rao (2016). „Filter bubbles, echo chambers, and online news consumption“. *Public Opinion Quarterly* (80) S1. 298–320.
- Fogg, B. J. (2003). *Persuasive technology: using computers to change what we think and do*. Amsterdam, Boston MA: Morgan Kaufmann Publishers.
- Future of Life Institute (2017). „Asilomar AI Principles“. <https://futureoflife.org/ai-principles/> (Download 15.4.2017).
- Gabelkov, Maksym, Arthi Ramachandran, Augustin Chaintreau und Arnaud Legout (2016). „Social Clicks: What and Who Gets Read on Twitter?“. *ACM SIGMETRICS/IFIP Performance 2016*. <https://hal.inria.fr/hal-01281190/file/sigm095-gabelkov.pdf> (Download 2.6.2017).
- Galtung, Johan, und Mari Holmboe Ruge (1965). „The Structure of Foreign News“. *Journal of Peace Research* (2) 1. 64–91.
- Garb, Rachel (2008). „More transparency in customized search results“. 30.6.2008. <https://googleblog.blogspot.de/2008/07/more-transparency-in-customized-search.html> (Download 2.6.2017).
- Geipel, Janet, Constantinos Hadjichristidis und Luca Surian (2015). „How foreign language shapes moral judgment“. *Journal of Experimental Social Psychology* 59. 8–17. <https://doi.org/10.1016/j.jesp.2015.02.001> (Download 2.6.2017).
- Gigerenzer, Gerd, und Wolfgang Gaissmaier (2011). „Heuristic Decision Making“. *Annual Review of Psychology* (62) 1. 451–482. <https://doi.org/10.1146/annurev-psych-120709-145346> (Download 2.6.2017).
- Gillespie, T. (2014). The Relevance of Algorithms. In T. Gillespie, P. Boczkowski, & K. Foot (Eds.), *Essays on Communication, Materiality, and Society*. (pp. 167–194). Cambridge.
- Goodman, Bryce, und Seth Flaxman (2016). „EU regulations on algorithmic decision-making and a ‚right to explanation‘“. *WHI* 2016. 26–30. <http://arxiv.org/abs/1606.08813> (Download 2.6.2017).
- Google (2016). „Blitzschnelle Suche“. <https://static.googleusercontent.com/media/www.google.com/de//insidesearch/howsearchworks/assets/searchInfographic.pdf> (Download 2.6.2017).
- Google (2017). „General Guidelines“. 14.3.2017. <https://static.googleusercontent.com/media/www.google.com/de//insidesearch/howsearchworks/assets/searchqualityevaluatorguidelines.pdf> (Download 2.6.2017).
- Gottfried, J., & Shearer, E. (2016). News use across social media platforms 2016. Pew Research Center, 26. Retrieved from http://assets.pewresearch.org/wp-content/uploads/sites/13/2016/05/PJ_2016.05.26_social-media-and-news_FINAL-1.pdf

- Habermas, Jürgen (1981). *Theorie des kommunikativen Handelns*. Frankfurt am Main.
- Habermas, Jürgen (2008). *Ach, Europa: Kleine Politische Schriften XI* (Originalausgabe). Frankfurt am Main.
- Hasebrink, Uwe, Jan-Hinrik Schmidt und Lisa Merten (2016). *Wie fließen Intermediäre in die Meinungsbildung ein? Die qualitative Perspektive der Vertiefungsstudie*. Berlin. http://www.die-medienanstalten.de/fileadmin/Download/Veranstaltungen/Pr%C3%A4sentation_Intermedi%C3%A4re/HL_Intermedi%C3%A4re_und_Meinungsbildung_Pr%C3%A4si_Web_Mappe_final.pdf (Download 2.6.2017).
- Heesen, Jessica (2016). „Freiheit“. *Handbuch Medien-und Informationsethik*. Hrsg. Jessica Heesen. Stuttgart. 51–58. http://link.springer.com/chapter/10.1007/978-3-476-05394-7_3 (Download 2.6.2017).
- Hegelich, Simon, und Dietmar Janetzko (2016). „Are Social Bots on Twitter Political Actors? Empirical Evidence from a Ukrainian Social Botnet“. <http://www.aai.org/ocs/index.php/ICWSM/ICWSM16/paper/viewPDFInterstitial/13015/12793> (Download 2.6.2017).
- Herbold, Astrid (2016). „Hatespeech: Argumente sind kein Allheilmittel“. *Die Zeit* 4.2.2014. <http://www.zeit.de/digital/internet/2016-02/hatespeech-counterspeech-facebook-forschung/komplettansicht> (Download 2.6.2017).
- Hern, Alex (2014). „Why Google has 200m reasons to put engineers over designers. *The Guardian* 5.2.2014. <https://www.theguardian.com/technology/2014/feb/05/why-google-engineers-designers> (Download 2.6.2017).
- Horling, Bryan, und Robby Bryant (2009). „Personalized Search for everyone“. 4.12.2009. <https://googleblog.blogspot.com/2009/12/personalized-search-for-everyone.html> (Download 6.4.2017).
- Howard, Philip N., und Bence Kollanyi (2016). „Bots, #StrongerIn, and #Brexit: Computational Propaganda during the UK-EU Referendum“. *arXiv:1606.06356 [physics]*. <http://arxiv.org/abs/1606.06356> (Download 2.6.2017).
- Iredale, Sophie, und Aleksej Heinze (2016). „Ethics and Professional Intimacy Within the Search Engine Optimisation (SEO) Industry“. *Technology and Intimacy: Choice or Coercion*. Bd. 474. Hrsg. David Kreps, Gordon Fletcher und Marie Griffiths. Cham. 106–115. https://doi.org/10.1007/978-3-319-44805-3_9 (Download 2.6.2017).
- Iyengar, Shanto, und Kyu S. Hahn (2009). „Red media, blue media: Evidence of ideological selectivity in media use“. *Journal of Communication* (59) 1. 19–39.

- Iyengar, Shanto, und Sean J. Westwood (2015). „Fear and Loathing across Party Lines: New Evidence on Group Polarization: FEAR AND LOATHING ACROSS PARTY LINES“. *American Journal of Political Science* (59) 3. 690–707. <https://doi.org/10.1111/ajps.12152> (Download 2.6.2017).
- Jenyns, David (2013). „How To Recover From A Google Panda/Penguin Spanking“. 6.1.2013
<http://www.melbourneseoservices.com/how-to-recover-from-google-panda/> (Download 11.4.2017).
- Justitia US Supreme Court (1927). „Whitney v. California 274 U.S. 357 (1927)“.
<https://supreme.justia.com/cases/federal/us/274/357/case.html> (Download 2.6.2017).
- Kacholia, Varun (2013). „Showing More High Quality Content“. 23.8.2013.
<https://newsroom.fb.com/news/2013/08/news-feed-fyi-showing-more-high-quality-content/> (Download 2.3.2017).
- Kahan, Dan M., Asheley Landrum, Katie Carpenter, Laura Helft und Kathleen Hall Jamieson (2016). *Science Curiosity and Political Information Processing*. (SSRN Scholarly Paper No. ID 2816803). Rochester NY: Social Science Research Network. <https://papers.ssrn.com/abstract=2816803> (Download 2.6.2017).
- Kahneman, Daniel (2012a). *Schnelles Denken, langsames Denken*. (T. Schmidt, Übers.). München.
- Kahneman, Daniel (2012b). *Thinking, fast and slow*. London: Penguin Books.
- Keegan, Jon (2016). „Blue Feed, Red Feed. See Liberal Facebook and Conservative Facebook, Side by Side“. *The Wall Street Journal online* 18.5.2016. <http://graphics.wsj.com/blue-feed-red-feed/> (Download 15.4.2017).
- Klayman, Joshua (1995). „Varieties of Confirmation Bias“. *Psychology of Learning and Motivation* Bd. 32. 385–418). [https://doi.org/10.1016/S0079-7421\(08\)60315-1](https://doi.org/10.1016/S0079-7421(08)60315-1) (Download 2.6.2017).
- Kollanyi, Bence, Philip N. Howard und Samuel C. Woolley (2016). „Bots and Automation over Twitter during the U.S. Election“. <http://politicalbots.org/wp-content/uploads/2016/10/Data-Memo-First-Presidential-Debate.pdf> (Download 2.6.2017).
- Kosinski, Michal, Sandra C. Matz, Samuel D. Gosling, Vesselin Popov und David Stillwell (2015). „Facebook as a research tool for the social sciences: Opportunities, challenges, ethical considerations, and practical guidelines“. *American Psychologist* (70) 6. 543.
- Kramer, Adam D., Jamie E. Guillory, Jeffrey T. Hancock (2014). „Experimental evidence of massive-scale emotional contagion through social networks“. *Proceedings of the National Academy of Sciences* (111) 24. 8788–8790. <https://doi.org/10.1073/pnas.1320040111> (Download 2.6.2017).

- Lada, A., Li, J., & Ding, S. (2017, January 31). New Signals to Show You More Authentic and Timely Stories. Retrieved June 15, 2017, from <https://newsroom.fb.com/news/2017/01/news-feed-fyi-new-signals-to-show-you-more-authentic-and-timely-stories/>
- Lazer, David. (2015). „The rise of the social algorithm“. *Science* (348) 6239. 1090–1091. <https://doi.org/10.1126/science.aab1422> (Download 2.6.2017).
- Leslie, Ian (2016). „The scientists who make apps addictive“. *The Economist* / 1843 Oct./Nov. 2016. <https://www.1843magazine.com/features/the-scientists-who-make-apps-addictive> (Download 2.6.2017).
- Levinovitz, Alan (2017). „William Shatner’s Tweets Are a Classic Case of Misinformation Spread“. 6.4.2017. http://www.slate.com/articles/health_and_science/science/2017/04/what_we_can_learn_from_william_shatner_s_twitter_meltdown.html (Download 8.4.2017).
- Levy, Steven (2016). „How Google is Remaking Itself as a ‚Machine Learning First‘ Company“. *Backchannel* 22.6.2016. <https://backchannel.com/how-google-is-remaking-itself-as-a-machine-learning-first-company-ada63defcb70> (Download 7.4.2017).
- Lichterman, Joseph (2017). „This site is ‚taking the edge off rant mode‘ by making readers pass a quiz before commenting“. *NiemanLab* 1.3.2017. <http://www.niemanlab.org/2017/03/this-site-is-taking-the-edge-off-rant-mode-by-making-readers-pass-a-quiz-before-commenting/> (Download 2.6.2017).
- Lill, Tobias, Martin U. Müller, Felix Scheidl und Hilmar Schmudt (2012). „Falsche Fans“. *Der Spiegel* 30. <http://www.spiegel.de/spiegel/print/d-87482751.html> (Download 2.6.2017).
- Lippmann, Walter, und Elisabeth Noelle-Neumann (1990). *Die öffentliche Meinung: Reprint des Publizistik-Klassikers*. Bochum.
- Lischka, Konrad, und Anita Klingel (2017). *Wenn Maschinen Menschen bewerten*. Bertelsmann Stiftung. Gütersloh. <https://doi.org/10.11586/2017025> (Download 2.6.2017).
- Lobo, Sascha, und Christopher Lauer (2015). *Aufstieg und Niedergang der Piratenpartei: Versuch der Erklärung eines politischen Internetphänomens* (1. Aufl.). Hamburg.
- Lomborg, Stine, und Anja Bechmann (2014). „Using APIs for Data Collection on Social Media“. *The Information Society* (30) 4. 256–265. <https://doi.org/10.1080/01972243.2014.915276> (Download 2.6.2017).
- Morganti, Luciano, Kristina Irion, Natali Helberger, Katharina Kleinen-von Königslöw und Rob van der Noll (2015). „Regulating the new information intermediaries as gatekeepers of information diversity“. *info* (17) 6. 50–71.

- Lumb, David (2015). „Why Scientists Are Upset About The Facebook Filter Bubble Study“. *Fast Company* 5.8.2015. <https://www.fastcompany.com/3046111/fast-feed/why-scientists-are-upset-over-the-facebook-filter-bubble-study> (Download 2.6.2017).
- Malik, Momin. M., und Jürgen Pfeffer (2016). „Identifying Platform Effects in Social Media Data“. http://mominmalik.com/malik_icwsm2016.pdf (Download 2.6.2017).
- Manjoo, Farhad (2013). „You Won't Finish This Article“. *Slate* 6.6.2013. http://www.slate.com/articles/technology/technology/2013/06/how_people_read_online_why_you_won_t_finish_this_article.html (Download 2.6.2017).
- Meyer, Thomas (2016). „Gleichheit – warum, von was und wie viel?“ *Neue Gesellschaft/Frankfurter Hefte* 11. 42–46.
- Mierau, Caspar Clemens (2016). „Fake News zum Holocaust sind noch immer Top-Treffer auf Google“. *Motherboard* 15.12.2016. <https://motherboard.vice.com/de/article/holocaust-leugnungen-google> (Download 6.4.2017).
- Mittelstadt, Brent (2016a). „Auditing for Transparency in Content Personalization Systems“. *International Journal of Communication* (10). 4991–5002.
- Mittelstadt, Brent (2016b). „Auditing for Transparency in Content Personalization Systems“. *International Journal of Communication* (10). 4991–5002.
- Morris, Meredith Ringel (2014). „Social networking site use by mothers of young children“. *ACM Digital Library*. 1272–1282. <https://doi.org/10.1145/2531602.2531603> (Download 2.6.2017).
- Napoli, Philip M. (2014a). „Automated Media: An Institutional Theory Perspective on Algorithmic Media Production and Consumption: Automated Media“. *Communication Theory* (24) 3. 340–360. <https://doi.org/10.1111/comt.12039> (Download 2.6.2017).
- Napoli, Philip M. (2014b). „On Automation in Media Industries: Integrating Algorithmic Media Production into Media Industries Scholarship“. *Media Industries* (1) 1. 33–38. <http://www.mediaindustriesjournal.org/index.php/mij/article/view/14> (Download 2.6.2017).
- Oremus, Will (2016). „Who Controls Your Facebook Feed“. *Slate* 3.1.2016. http://www.slate.com/articles/technology/cover_story/2016/01/how_facebook_s_news_feed_algorithm_works.html (Download 2.6.2017).
- Pariser, Eli (2011). *The filter bubble: what the Internet is hiding from you*. New York: Penguin Press.
- Paus-Hasebrink, Ingrid, Jan-Hinrik Schmidt und Uwe Hasebrink (2009). „Zur Erforschung der Rolle des Social Web im Alltag von Heranwachsenden“. *Heranwachsen mit dem Social Web. Zur Rolle von Web 2.0-*

Angeboten im Alltag von Jugendlichen und jungen Erwachsenen. Hrsg. Jan-Hinrik Schmidt, Ingrid Paus-Hasebrink und Uwe Hasebrink. Berlin. 13–40.

Perset, Karine (2010). „The Economic and Social Role of Internet Intermediaries“. *OECD Digital Economy Papers* No. 171. Paris: OECD Publishing. http://www.oecd-ilibrary.org/science-and-technology/the-economic-and-social-role-of-internet-intermediaries_5kmh79zszs8vb-en (Download 2.6.2017).

Puschmann, C., & Burgess, J. (2013). *The Politics of Twitter Data* (SSRN Scholarly Paper No. ID 2206225). Rochester, NY: Social Science Research Network. Retrieved from <https://papers.ssrn.com/abstract=2206225> (Download 17.6.2017).

Peters, Bernhard (2002). „Die Leistungsfähigkeit heutiger Öffentlichkeiten – einige theoretische Kontroversen“. *Integration und Medien*. Hrsg. Kurt Imhof, Otfried Jarren und Roger Blum. Wiesbaden. 23–35. http://link.springer.com/chapter/10.1007/978-3-322-97101-2_3 (Download 2.6.2017).

Rainie, Lee, Janna Anderson und Jonathan Albright (2017). „The Future of Free Speech, Trolls, Anonymity and Fake News Online“. *PewResearchCenter* 29.3.2017. <http://www.pewinternet.org/2017/03/29/the-future-of-free-speech-trolls-anonymity-and-fake-news-online/> (Download 2.6.2017).

Ratcliff, Christopher (2016). „WebPromo’s Q&A with Google’s Andrey Lipattsev“. 6.4.2016. <https://searchenginewatch.com/2016/04/06/webpromos-qa-with-googles-andrey-lipattsev-transcript/> (Download 7.4.2017).

Reinbold, Fabian (2017). „Propaganda in der Ära Trump: Die Wahrheit über Fake News“. *Spiegel Online* 12.1.2017. <http://www.spiegel.de/netzwelt/web/donald-trump-die-wahrheit-ueber-fake-news-a-1129628.html> (Download 13.4.2017).

Reinbold, Fabian, und Marcel Rosenbach (2016). „Hetze im Netz. Facebook löscht Hass-Kommentare von Berlin aus“. *Spiegel Online* 15.1.2016. <http://www.spiegel.de/netzwelt/web/facebook-neues-loesch-team-geht-gegen-hasskommentare-vor-a-1072175.html> (Download 26.5.2017).

Rosania, Paul (2015). „While you were away ... “. 21.1.2017. <https://blog.twitter.com/2015/while-you-were-away-0> (Download 2.3.2017).

Sandvig, Christian, Kevin Hamilton, Karrie Karahalios und Cedric Langbort (2014). „Auditing algorithms: Research methods for detecting discrimination on internet platforms“. Paper presented to „*Data and discrimination: converting critical concerns into productive inquiry*“, a preconference at the 64th Annual Meeting of the International Communication Association. 22. Mai 2014. Seattle WA, USA.

<https://pdfs.semanticscholar.org/b722/7cbd34766655dea10d0437ab10df3a127396.pdf> (Download 2.6.2017).

Schäfer, Mirko Tobias, und Karin van Es (2017). *The datafied society: Studying culture through data*. Amsterdam: University Press Amsterdam. <http://dare.uva.nl/aup/en/record/624657> (Download 2.6.2017).

Schmidt, Jan-Hinrik (2016). „Ethik des Internets“. *Handbuch Medien-und Informationsethik*. Hrsg. Jessica Heesen. Stuttgart. 283–292. http://link.springer.com/chapter/10.1007/978-3-476-05394-7_8 (Download 2.6.2017).

Schmidt, Jan.-Hinrik, Isabelle Petric, Amelie Rolfs, Uwe Hasebrink und Lisa Merten (2017). *Zur Relevanz von Online-Intermediären für die Meinungsbildung*. Arbeitspapiere des Hans-Bredow-Instituts Nr. 40. Hamburg. http://www.hans-bredow-institut.de/webfm_send/1172 (Download 2.6.2017).

Schraven, David (2017). „Warum wollt ihr auf Facebook Fakten checken?“. *Correctiv.org* 20.1.2017. <https://correctiv.org/blog/2017/01/20/warum-wollt-ihr-fuer-facebook-fakten-checken/> (Download 2.6.2017).

Schulz, Winfried (1976). *Die Konstruktion von Realität in den Nachrichtenmedien: e. Analyse der aktuellen Berichterstattung*. 1. Auflage. Freiburg i. Br., München.

Schulz, Winfried, und Kevin Dankert (2016). „Die Macht der Informationsintermediäre. Erscheinungsformen, Strukturen und Regulierungsoptionen“. Hrsg. Friedrich-Ebert-Stiftung. <http://library.fes.de/pdf-files/akademie/12408.pdf> (Download 2.6.2017).

Silverman, Craig (2016). „Here’s How Fake Election News Outperformed Real Election News On Facebook“. *BuzzFeed News* 16.11.2016. <https://www.buzzfeed.com/craigsilverman/viral-fake-election-news-outperformed-real-news-on-facebook> (Download 14.12.2016).

Silverman, Craig, und Lawrence Alexander (2016). „How Teens In The Balkans Are Duping Trump Supporters With Fake News“. *Buzzfeed News*. 4.11.2016. https://www.buzzfeed.com/craigsilverman/how-macedonia-became-a-global-hub-for-pro-trump-misinfo?utm_term=.ckNZqrx7Ln#.vjYnOeyGbZ (Download 2.6.2017).

Simon, Herbert A. (1979). „Rational Decision Making in Business Organizations“. *The American Economic Review* (69) 4. 493–513.

Spiegel Online (2017). „Google will Fake News in Suchergebnissen bekämpfen“. 25.4.2017. <http://www.spiegel.de/netzwelt/web/project-owl-google-will-fake-news-in-suchergebnissen-bekaempfen-a-1144797.html> (Download 27.4.2017).

- Stieglitz, Stefan, und Linh Dang-Xuan (2012). „Impact and Diffusion of Sentiment in Public Communication on Facebook“. *ECIS 2012 Proceedings*. 98. <http://aisel.aisnet.org/ecis2012/98> (Download 2.6.2017).
- Stieglitz, Stefan, und Linh Dang-Xuan (2013). „Emotions and Information Diffusion in Social Media – Sentiment of Microblogs and Sharing Behavior“. *Journal of Management Information Systems* (29) 4. 217–248. <https://doi.org/10.2753/MIS0742-1222290408> (Download 2.6.2017).
- Su, Ao-Jan, Y. Charlie Hu, Aleksandar Kuzmanovic, und Cheng-Kok Koh (2014). „How to improve your search engine ranking: Myths and reality“. *ACM Transactions on the Web (TWEB)* (8) 2. 8.
- Sullivan, Danny (2017). „Google launches new effort to flag upsetting or offensive content in search“. *Search Engine Land* 14.3.2017. <http://searchengineland.com/google-flag-upsetting-offensive-content-271119> (Download 2.6.2017).
- Sunstein, Cass R. (2001). *Republic.com*. Princeton NJ: Princeton University Press.
- Sunstein, Cass R. (2008). *Infotopia: how many minds produce knowledge*. New York NY, Oxford: Oxford University Press.
- Taber, Charles. S., Damin Cann und Simona Kucsova (2009). „The Motivated Processing of Political Arguments“. *Political Behavior* (31) 2. 137–155. <https://doi.org/10.1007/s11109-008-9075-8> (Download 2.6.2017).
- The Coral Project (2016). „About the Coral Project“. <http://coralproject.net/about> (Download 24.4.2017).
- Trilling, Damian, Marijn van Klingereren und Yarif Tsfati (2016). „Selective Exposure, Political Polarization, and Possible Mediators: Evidence From the Netherlands“. *International Journal of Public Opinion Research* edw003. <https://doi.org/10.1093/ijpor/edw003> (Download 2.6.2017).
- Tufekci, Zeynep (2015). „How Facebook’s Algorithm Suppresses Content Diversity (Modestly) and How the Newsfeed Rules Your Clicks“. 5.5.2015. <https://medium.com/message/how-facebook-s-algorithm-suppresses-content-diversity-modestly-how-the-newsfeed-rules-the-clicks-b5f8a4bb7bab> (Download 14.4.2017).
- Tutt, Andrew (2016). *An FDA for Algorithms*. (SSRN Scholarly Paper No. ID 2747994). Rochester NY: Social Science Research Network. <https://papers.ssrn.com/abstract=2747994> (Download 2.6.2017).
- Tversky, Amos, und Daniel Kahneman (1974). „Judgment under Uncertainty: Heuristics and Biases“. *Science* (185) 4157. 1124–1131.
- Twitter (2017). „About your Twitter timeline“. <https://support.twitter.com/articles/164083#settings> (Download 2.6.2017).

- Underwood, Mimi (2015). „Updating Our Search Quality Rating Guidelines“. *Google Webmaster Central Blog* 19.11.2015. <https://webmasters.googleblog.com/2015/11/updating-our-search-quality-rating.html> (Download 6.4.2017).
- US Supreme Court (1969). „Brandenburg v. Ohio 395 U.S. 444“. 9.6.1969. <https://supreme.justia.com/cases/federal/us/395/444/> (Download 2.6.2017).
- van der Linden, Sander, Anthony Leiserowitz, Seth Rosenthal, und Edward Maibach (2017). „Inoculating the Public against Misinformation about Climate Change“. *Global Challenges* 27.2.2017. <https://doi.org/10.1002/gch2.201600008> (Download 2.6.2017).
- van Deursen, Alexander J. A. M., und Jan A. G. M. van Dijk (2009). „Using the Internet: Skill related problems in users' online behavior“. *Interacting with Computers* (21) 5–6. 393–402. <https://doi.org/10.1016/j.intcom.2009.06.005> (Download 2.6.2017).
- Vike-Freiberga, Vaira, Herta Däubler-Gmelin, Ben Hammersley und Luis Miguel Poiarés Pessoa Maduro (2013). „The Report of the High Level Group on Media Freedom and Pluralism“. <https://ec.europa.eu/digital-single-market/sites/digital-agenda/files/HLG%20Final%20Report.pdf> (Download 2.6.2017).
- Vieth, Kilian, und Ben Wagner. (2017). *Teilhabe, ausgerechnet*. Hrsg. Bertelsmann Stiftung. Gütersloh. <https://doi.org/10.11586/2017027> (Download 2.6.2017).
- Volokh, Eugene, und Donald M. Falk (2012). *First Amendment Protection for Search Engine Search Results -- White Paper Commissioned by Google*. (SSRN Scholarly Paper No. ID 2055364). Rochester NY: Social Science Research Network. <https://papers.ssrn.com/abstract=2055364> (Download 2.6.2017).
- Weischenberg, Siegfried (2001). *Nachrichten-Journalismus: Anleitungen und Qualitäts-Standards für die Medienpraxis*. Wiesbaden. <http://link.springer.com/openurl?genre=book&isbn=978-3-322-80408-2> (Download 2.6.2017).
- Williams, Raymond (1974). *Television Technology and Cultural Form*. London: Fontana.
- Wojcieszak, Magdalena (2010). „„Don't talk to me'': effects of ideologically homogeneous online groups and politically dissimilar offline ties on extremism“. *New Media & Society* 28.1.2010. <http://nms.sagepub.com/content/early/2010/01/28/1461444809342775.abstract> (Download 2.6.2017).
- Xu, Jie, Akos Lada und Vibhi Kant (2016). „Showing You More Personally Informative Stories“. 11.8.2016. <http://newsroom.fb.com/news/2016/08/news-feed-fyi-showing-you-more-personally-informative-stories/> (Download 1.3.2017).

- Zarsky, Tal (2016). „The Trouble with Algorithmic Decisions: An Analytic Road Map to Examine Efficiency and Fairness in Automated and Opaque Decision Making“. *Science, Technology, & Human Values* (41) 1. 118–132. <https://doi.org/10.1177/0162243915605575> (Download 2.6.2017).
- Zeifman, Igal (2017). „Bot Traffic Report 2016“. 24.1.2017. <https://www.incapsula.com/blog/bot-traffic-report-2016.html> (Download 2.6.2017).
- Zhang, Cheng, und Si Chen (2016). „Using Qualitative Feedback to Show Relevant Stories“. 1.2.2016. <https://newsroom.fb.com/news/2016/02/news-feed-fyi-using-qualitative-feedback-to-show-relevant-stories/> (Download 2.3.2017).
- Zollo, Fabiana, Alessandro Bessi, Michaela Del Vicario, Antonio Scala, Guido Caldarelli, Louis Shekhtman, Shlomo Havlin und Walter Quattrociocchi (2015). „Debunking in a World of Tribes“. *arXiv preprint arXiv:1510.04267*. <https://arxiv.org/abs/1510.04267> (Download 2.6.2017).
- Zuckerberg, Mark. (2016, November 12). I want to share some thoughts on Facebook and the election. Retrieved June 15, 2017, from <https://www.facebook.com/zuck/posts/10103253901916271>
- Zweig, Katharina Anna (2017a). „Watching the watchers: Epstein and Robertson’s, Search Engine Manipulation Effect“. *Algorithm Watch* 7.4.2017. <https://algorithmwatch.org/watching-the-watchers-epstein-and-robertsons-search-engine-manipulation-effect/> (Download 7.4.2017).
- Zweig, Katharina Anna (2017b, im Erscheinen). *Wo maschinelle Prognosen irren können*. Bertelsmann Stiftung. Gütersloh.

10 Executive Summary (English)

Essentially this paper answers three key questions relating to public communication and opinion in the digital age:

4. *Media transformation: How is public discourse changing because of the new digital platforms through which many people now receive socially relevant information?*

When all age groups are considered, intermediaries driven by algorithmic processes, such as Google and Facebook, have had neither a large nor defining influence on how public opinion is formed, compared to editorially driven media such as television. These intermediaries judge the relevance of content based on the public's immediate reaction to a much greater degree than do traditional media.

5. *Social consequences: In terms of quality and diversity, is the information that reaches people via these new channels suitable for the formation of basic democratic attitudes and does it promote participation?*

Use of these intermediaries for forming public opinion is leading to a structural change in public life. Key factors here are the algorithmic processes used as a basic formational tool and the leading role that user reactions play as input for these processes. Since they are the result of numerous psychological factors, these digitally assessed and, above all, impulsive public reactions are poorly suited to determining relevance in terms of traditional social values. These values, such as truth, diversity and social integration, have served in Germany until now as the basis for public opinion and public life as formed by editorial media.

6. *Solutions: How can the new digital platforms be designed to ensure they promote participation?*

Algorithms that sort content and personalize how it is assembled form the core of the complex, interdependent process underlying public communication and the formation of public opinion. That is why solutions must be found here first. The most important areas that will need action in the foreseeable future are facilitating external research and evaluation, strengthening the diversity of algorithmic processes, anchoring basic social values (e.g. by focusing on professional ethics) and increasing awareness among the public.

1. Media transformation: Intermediaries are a relevant but not determining factor for the formation of public opinion.

Numerous studies have shown that so-called intermediaries such as Google and Facebook play a role in the formation of public opinion in numerous countries including Germany. For example, 57 percent of German Internet users receive politically and socially relevant information via search machines or social networks. And although the share of users who say that social networks are their most important news source is still relatively small at 6 percent of all Internet users, this figure is significantly higher among younger users. It can thus be assumed that these types of platforms will generally increase in importance. The creation of public opinion is "no longer conceivable without intermediaries," as researchers at the Hamburg-based Hans Bredow Institute put it in 2016.

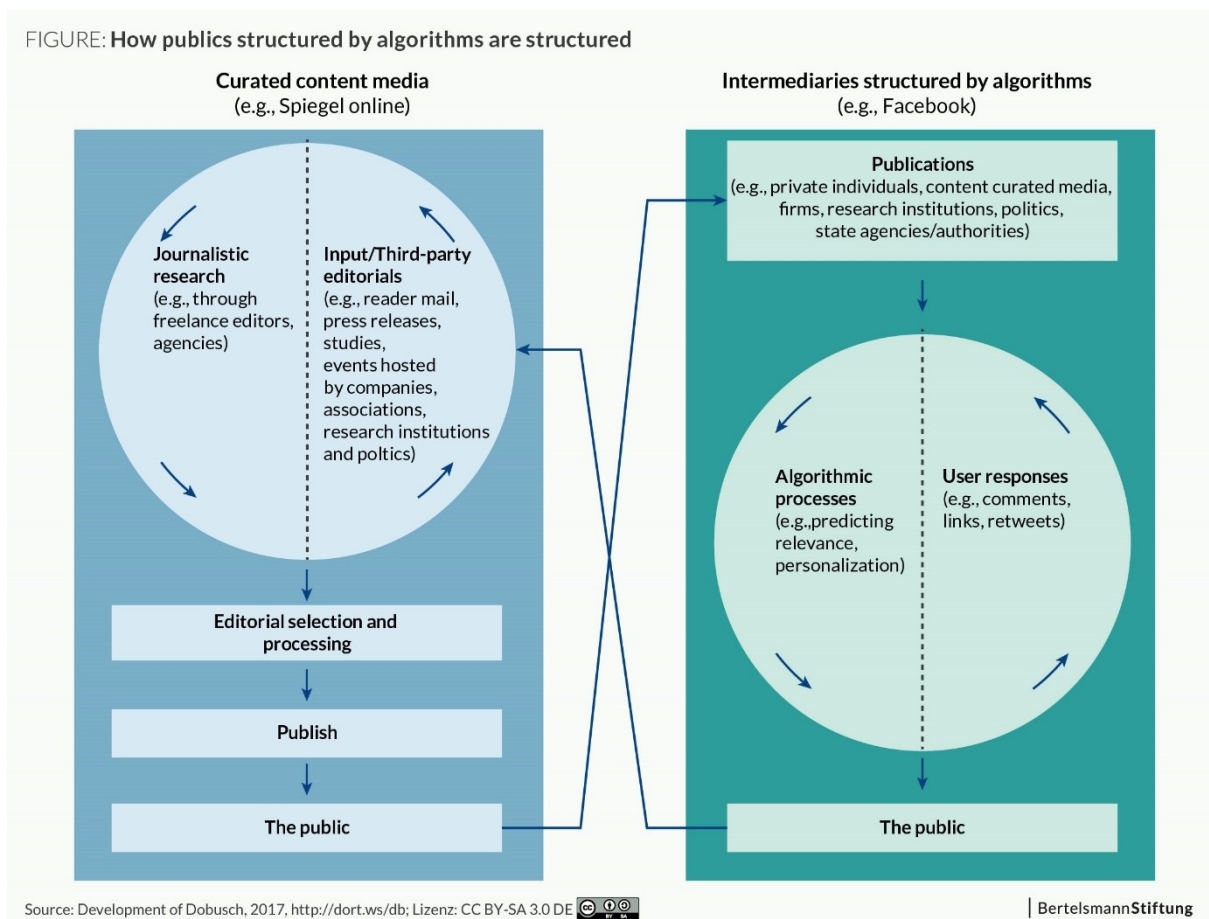
The principles that these intermediaries use for shaping content are leading to a structural shift in public opinion and public life. Key aspects are:

- **Decoupling of publication and reach:** Anyone can make information public, but not everyone finds an audience. Attention is only gained through the interaction of people and algorithmic decision-making (ADM) processes.
- **Detachment from publications:** Each item or article has its own reach. Inclusion in a publication reduces the chances of becoming part of the content provided by intermediaries.
- **Personalization:** Users receive more information about their specific areas of interest.
- **Increased influence of public on reach:** User reactions influence ADM processes in general and the reach of each item or article.
- **Centralization of selection agents:** Intermediaries offer much less diversity than do traditionally curated media.

- **Interplay of human and machine-based curation:** Traditionally curated media disseminate content via intermediaries and use the resulting reactions on intermediary sites as a measure of public interest.

The comparison of processes in the graphic below show the key role now being played by user reactions and algorithmic processes in public communication and the formation of public opinion. The user reactions and algorithmic processes determine the attention received via intermediaries. Our hypothesis is that these reactions and processes cannot be definitively included in a linear causal chain.

Illustration 1: Organization of an algorithmically structured public (based on Dobusch, 2017)



Google, Facebook and other intermediaries are already playing an important role in the public discourse, even though as a rule these platforms were not originally conceived to supply consumers with media content from journalistic organizations. They use technical systems instead to decide whether certain content taken from a huge pool of information could be interesting or relevant to a specific user. These systems were initially intended – in the case of search engines – to identify websites, for example, that contain certain information or – in the case of social networks – to display in a prominent position particularly interesting messages or photos from a user’s own circle of friends. In many cases they therefore sort content according to completely different criteria than editors at a daily newspaper or a magazine would. “Relevant” means something different to Google than it does to Facebook, and both understand it differently than the editors at SPIEGEL ONLINE or Sueddeutsche.de.

The intermediaries use numerous variables to calculate the relevance of individual items. These variables range from basic behavioral metrics such as scrolling speed or how long a page is viewed to the level of interaction among multiple users in a social network. When someone with whom a user has repeatedly communicated on Facebook posts content, the probability is higher that the user will be shown this content than they would if someone posts with whom the user has, theoretically, a digital connection, but with whom the user has never truly had contact. The inputs that other users provide – often unknowingly – are also included in assessments of

relevance, whether they be links, clicks on links, clicks on the “like” button, forwards, shares or the number of comments that certain content receives.

2. Social consequences: As used by the most relevant intermediaries currently forming public opinion, algorithmic processes evaluate user reactions to content. They promote and reinforce a manner of human cognition that is susceptible to distortion, and they themselves are susceptible to technical manipulation.

The metrics signaling relevance – about which platform operators are hesitant to provide details because of competition-related factors or other reasons – are potentially problematic. This is true, first, because the operators themselves are constantly changing them. Systems such as those used by Google and Facebook are being altered on an ongoing basis; the operators experiment with and tweak almost every aspect of the user interface and other platform features in order to achieve specific goals such as increased interactivity. Each of these changes can potentially impact the metrics that the platforms themselves capture to measure relevance.

A good example is the “People you may know” feature used by Facebook which provides users with suggestions for additional contacts based on assessments of possible acquaintances within the network. When this function was introduced, the number of links made every day within the Facebook community immediately doubled. The network of relationships displayed on such platforms is thus dependent on the products and services that the operators offer. At the same time, the networks of acquaintances thus captured also become variables in the metrics determining what is relevant. Whoever makes additional “friends” is thus possibly shown different content.

A further problem of the metrics collected by the platform operators has to do with the type of interaction for which such platforms are optimized. A key design maxim is that interactions should be as simple and convenient as possible in order to maximize the probability of their taking place. Clicking on a “like” button or a link demands no cognitive effort whatsoever, and many users are evidently happy to indulge this lack of effort. Empiric studies suggest, for example, that many articles in social networks forwarded with a click to the user’s circle of friends could not possibly have been read. Users thus disseminate media content after having seen only the headline and introduction. To some extent they deceive the algorithm and, with it, their “friends and followers” into believing that they have read the text.

The ease of interaction also promotes cognitive distortions that have been known to social psychologists for years. A prime example is the availability heuristic: If an event or memory can easily be recalled, it is assumed to be particularly probable or common. The consequence is that users frequently encounter unread media content that has been forwarded due to a headline, and the content is later remembered as being “true” or “likely.” This is also the case when the text itself makes it clear that the headline is a grotesque exaggeration or simply misleading.

A number of other psychological factors also play an important role here, for example the fact that people use social media in particular not only for informational purposes, but also as a tool for identity management, with some media content being forwarded only to demonstrate the user’s affiliation with a certain political camp, for example. Moreover, the design of many digital platforms explicitly and intentionally encourages a fleeting, emotional interaction with content. Studies of networking platforms indeed show that **content which rouses emotion** is commented on and forwarded particularly often – above all when negative emotions are involved.

Such an emotional treatment of news content can lead to increased **societal polarization**, a hypothesis for which initial empirical evidence already exists, especially in the United States. At the same time, however, such polarizing effects seem to be dependent on a number of other factors such as a country’s electoral system. Societies with first-past-the-post systems such as the US are potentially more liable to extreme political polarization than those with proportional systems, in which ruling coalitions change and institutionalized multiparty structures tend to balance out competing interests. Existing societal polarization presumably influences and is influenced by the algorithmic sorting of media content. For example, one study shows that Facebook users who believe in conspiracy theories tend over time to turn to the community of conspiracy theorists holding the same

view. This process is possibly exacerbated by algorithms that present them more and more often with the relevant content. These systems could in fact result in the creation of so-called **echo chambers**, at least among people with extremist views.

Technical manipulation can also influence the metrics that intermediaries use to ascertain relevance. So-called bots – partially self-acting software applications that can be disguised as real users, for example in social networks – can massively distort the volume of digital communication occurring around certain topics. According to one study, during the recent presidential election in the US, 400,000 such bots were in use on Twitter, accounting for about one-fifth of the entire discussion of the candidates' TV debates. It is not clear to which extent these types of automated systems actually influence how people vote. What is clear is that the responses they produce – clicks, likes, shares – are included in the relevance assessments produced by ADM systems. Bots can thus make an article seem so interesting that an algorithm will then present it to human users.

In sum, it can be said that the relevance assessments that algorithmic systems create for media content do not necessarily reflect criteria that are desirable from a societal perspective. Basic values such as truthfulness or social integration do not play a role. The main goal is to increase the probability of an interaction and the time users spend visiting the relevant platform. Interested parties whose goal is a strategic dissemination of disinformation can use these mechanisms to further their cause: A creative, targeted lie can, on balance, prove more emotionally “inspiring” and more successful within such systems – and thus have a greater reach – than the boring truth.

3. Solutions: Algorithmic sorting of content is at the heart of the complex interdependencies affecting public life in the digital context. This is where solutions must be applied.

The last section of this paper contains a series of possible solutions for these challenges. A first goal, one that is comparatively easy to achieve, is making users more aware of the processes and mechanisms described here. Studies show that users of social networking platforms do not even know that such sorting algorithms exist, let alone how they work. Educational responses, including in the area of continuing education, would thus be appropriate, along with efforts to increase awareness of disinformation and decrease susceptibility to it, for example through the use of fact-based materials.

Platform operators themselves clearly have more effective possibilities for intervention. For example, they could do more to ensure that values such as appropriateness, responsibility and competency are adhered to when the relevant systems are being designed and developed. A medium-term goal could be defining industry-wide professional ethics for developers of ADM systems.

Moreover, specialists who do not work for platform operators themselves should be in a position to evaluate and research the impact being made by the operators' decisions. Until now it has been difficult if not impossible for external researchers or government authorities to gain access to the required data, of which operators have vast amounts at their disposal. Neither the design decisions made by platform operators nor the impacts of those decisions on individual users are transparent to any significant degree. Systematic distortions, for example in favor of one political viewpoint or another, are difficult to identify using currently available data. More transparency – through a combination of industry self-regulation and, where necessary, legislative measures – would make it possible to gain an unbiased understanding of the actual social consequences of algorithmic sorting and to identify potential dangers early on. Making it easier to conduct research would also stimulate an objective, solution-oriented debate of the issue and could identify new solutions. Measures like these could also make it easier to design algorithmic systems that increase participation. This would foster a more differentiated view of algorithmic processes. At the same time, it could increase trust in systems designed to benefit all of society.

11 Über die Autoren

Konrad Lischka (Jahrgang 1979) schreibt seit 1999 über die digitale Gesellschaft – Bücher, Essays und Blogs. Nach dem Diplomstudium der Journalistik und der Ausbildung an der Deutschen Journalistenschule arbeitete er als Chefredakteur des bücher Magazins und stellvertretender Ressortleiter Netzwelt bei Spiegel Online. Danach Wechsel in die Medien- und Netzpolitik als Referent Digitale Gesellschaft in der Staatskanzlei Nordrhein-Westfalen, seit 2016 Projektmanager bei der Bertelsmann Stiftung im Projekt Teilhabe in einer digitalisierten Welt.

Prof. Dr. Christian Stöcker (Jahrgang 1973) studierte Psychologie in Würzburg und Bristol und promovierte 2003 in kognitiver Psychologie. In München studierte er anschließend an der bayerischen Theaterakademie Kulturkritik und schrieb parallel unter anderem für die Süddeutsche Zeitung, Die Zeit und Spiegel Online. Seit Februar 2005 arbeitet er bei Spiegel Online in den Ressorts Wissenschaft und Netzwelt, seit 2011 Ressortleiter Netzwelt, seit 2016 Professor für Digitale Kommunikation an der Hochschule für Angewandte Wissenschaften (HAW) Hamburg.

Adresse | Kontakt

Bertelsmann Stiftung
Carl-Bertelsmann-Straße 256
33311 Gütersloh
Telefon +49 5241 81-81216

Konrad Lischka
Taskforce Digitalisierung
Telefon +49 5241 81-81216
konrad.lischka@bertelsmann-stiftung.de

www.bertelsmann-stiftung.de