HELMUT SCHMIDT UNIVERSITÄT
Universität der Bundeswehr Hamburg

Künstliche Intelligenz und Regierungskommunikation

Charalampos-Babis Karpouchtsis
unter Mitarbeit von Gary S. Schaal
und in Kooperation mit Philip Jan Schäfer (Universität Bielefeld)

STATE OF THE ART REPORT

Stand November 2022

https://doi.org/10.24405/14980

State of the Art Report
Künstliche Intelligenz und Regierungskommunikation

Professur für Politikwissenschaft, insbesondere Politische Theorie
Helmut-Schmidt-Universität Hamburg/Universität der Bundeswehr

Projekt: Künstliche Intelligenz und öffentliche Kommunikation
Projektleitung: Prof. Dr. Gary Schaal
Projektkoordination: Charalampos-Babis Karpouchtsis
In Kooperation: Dr. Philip Jan Schäfer (Universität Bielefeld)

Autor: Charalampos-Babis Karpouchtsis
unter Mitarbeit von Gary S. Schaal und
in Kooperation mit Philip Jan Schäfer (Universität Bielefeld)

Korrespondenzadresse: karpoucc@hsu-hh.de

Stand November 2022

https://doi.org/10.24405/14980

Titelfoto: stock.adobe.com/Siarhei

Inhalt

Executive Summary 4

Regierungskommunikation und Künstliche Intelligenz – State of the Art Report 5

Künstliche Intelligenz: Begriffe, Verfahren, Anwendungen 8

Deliberative Demokratie im Umbruch: Herausforderungen des digitalen Wandels 11

Dimension der Vermittlung: wo und wie wird Kommunikation vermittelt? 13

Dimension der individuellen Disposition der Bürgerinnen und Bürger: Erwartungshaltung, Selbstwirksamkeitserwartung und Bias 22

Dimension des Deliberativen Prozesses: wie manifestieren sich KI-Herausforderungen und Chancen für die Demokratie 29

Limitationen, Ausblick und Forschungsbedarf 34

Literaturverzeichnis 36

Executive Summary

Ziel des Reports ist es den Forschungsstand zum Einsatz von Verfahren künstlicher Intelligenz in öffentlicher, politischer und Regierungskommunikation zu erfassen, systematisch einzuordnen und zu analysieren, um mögliche Anwendungsbereiche und Entwicklungstrends zu identifizieren und entsprechende Möglichkeiten zur Weiterentwicklung regierungsamtlicher Kommunikation im Speziellen und Regierungskommunikation im Allgemeinen vorzulegen.

Die Untersuchung ist anhand eines **analytischen Rasters** strukturiert, das den jeweiligen Forschungsstand in drei Dimensionen unterteilt und deren gegenseitige Wechselwirkungen aufzeigt. Zunächst wird **KI definiert**, als auch **Lern- und Anwendungsformen** von KI-Verfahren vorgestellt. Es folgt die Analyse entlang der drei Dimensionen, vor dem Hintergrund der **Herausforderungen des digitalen Wandels für die deliberative Demokratie**, welche die demokratietheoretische Grundlage der Bundesrepublik darstellt.

Die **Dimension der Vermittlung** konzentriert sich auf die Verortung der Kommunikation unter KI-Einsatz. Die **Dimension der individuellen Disposition** widmet sich der Kommunikationsbereitschaft und -erwartung der Bürgerinnen und Bürger, die sich durch Digitalisierung und KI-Einsatz verändern. Die **Dimension der Deliberation** erfasst Herausforderungen und Chancen für die deliberative Demokratie, die auf Dialog, Inklusion, Rationalität, Authentizität und Öffentlichkeit basiert.

Der Bericht geht nur begrenzt auf generative KI ein.

Regierungskommunikation und Künstliche Intelligenz – State of the Art Report

Ziele des Reports

Die Rahmenbedingungen effektiver Regierungskommunikation verändern sich im Zuge der digitalen Transformation der Öffentlichkeit (Habermas 2021) und der ubiquitären Nutzung von (schwacher) künstlicher Intelligenz grundlegend. Vor diesem Hintergrund verfolgt der vorliegende Bericht mehrere Ziele:

1. Die digitalen Transformationsprozesse sind nicht monolithischer Natur. Notwendig ist daher die Entwicklung eines Rasters, dass die Transformationsprozesse analytisch so weit ausdifferenziert, dass auf einer operativen Ebene kommunikative Gegenstrategien entwickelt werden können. Dieses analytische Raster zu entwickeln und es für die Rekonstruktion des aktuellen Forschungsstands zu nutzen ist das erste Ziel des Berichts.

2. Entlang dieses analytischen Rasters wird der aktuelle Stand der Forschung in jenen Bereichen zusammengefasst, die für effektive Regierungskommunikation thematisch/funktional einschlägig sind.

3. Auf Basis des aktuellen Forschungsstandes werden Möglichkeiten für die zukünftige Entwicklung effektiver Regierungskommunikation aufgezeigt.

Ein analytisches Raster für die Analyse der Veränderungen der Funktionsvoraussetzungen effektiver Regierungskommunikation

Regierungskommunikation kann in einer liberal-repräsentativen Demokratie wie der deutschen, als intentionales kommunikatives Handeln charakterisiert werden, das den öffentlichen politischen Diskurs beeinflussen soll. Sie kann u.a. der Information, Aufklärung oder der demokratischen Legitimation dienen. Funktional kann sie als ‘one-to-many’ Kommunikation beschrieben werden. Die Frage, ob die intendierten Effekte der Kommunikation realisiert werden, wird von vielen Erfolgsfaktoren beeinflusst, die wir in drei größere Dimensionen zusammenfassen wollen.

1. **Transformation der Kommunikationsvermittlung**

Die Kommunikationsinhalte müssen in einem technologischen Sinne übermittelt und von den Bürgerinnen und Bürgern wahrgenommen werden. Digitalisierungsprozesse im Allgemeinen und der Siegeszug digitaler Plattformen (insb. Social Media) transformieren die Vermittlungs- und Rezeptionswege von Regierungskommunikation. Deshalb gilt es, die Transformationen im Vermittlungsprozess zu analysieren.

2. **Transformation individueller Einstellungen und Dispositionen der Bürgerinnen und Bürger zu politischer Information und Deliberation:**

Deutschland folgt dem normativen Ideal einer deliberativen Demokratie. Deliberative Demokratie ist funktional extrem voraussetzungsvoll. Auf der individuellen Ebene muss zumindest folgendes gelten:

a. Epistemisch: Deliberation setzt voraus, dass Faktenwissen existiert und politische Präferenzen und Entscheidungen prinzipiell begründbar sind.

b. Motivation: Deliberation setzt voraus, dass Bürgerinnen und Bürger bereit sind, sich an öffentlichen politischen Diskussionen zu beteiligen und ihre politischen Positionen im Angesicht besserer Argumente zu verändern.

c. Kognition: Deliberation setzt voraus, dass Bürgerinnen und Bürger Wissen besitzen und kognitiv in der Lage sind, politische Prozesse zu verstehen.

d. Sozial: Deliberation setzt die wechselseitige soziale Anerkennung aller Diskursteilnehmerinnen und -teilnehmer als Gleiche voraus.

Die digitale Transformation und die Nutzung von KI verändert alle individuellen Voraussetzungen und Dispositionen deliberativer Demokratie: Sie unterminiert den Glauben an die Fakten, transformiert die Frustrationsresistenz im Diskurs mit anderen Meinungen durch die Herstellung von Konsenskommunikation in Echokammern, Filterblasen und mit Bots; sie vertieft vorgängig vorhandene soziale und kulturelle Spaltungen in der Gesellschaft, etc.

3. **Transformation der prozessualen Dimension deliberativer Öffentlichkeit und Demokratie**

Ergänzend zu den individuellen Voraussetzungen treten prozessuale: Deliberative Demokratie benötigt den inklusiven Austausch heterogener Positionen, d.h. Foren, an denen sich die Bürgerinnen und Bürger mit unterschiedlichen Meinungen treffen. Deliberative Demokratie benötigt die Differenzierung von privat und öffentlich, um einen Rollenwechsel bei den Bürgerinnen und Bürgern (z.B. hin zu gemeinwohlorientierter Problemwahrnehmung) zu motivieren. Schließlich benötigt deliberative Demokratie, dass kommunikative Konsense faktischer Natur sind; d.h., dass alle Betroffenen von einer Entscheidung inkludiert wurden, aber auch nur die Betroffenen. Dass der Konsens nach einer Diskussion das Ergebnis von Einsicht ist – d.h., dass die Diskursteilnehmerinnen und - teilnehmer in der Lage sein müssen, die Präferenzen zu ändern (was bei Bots nicht der Fall ist).

Die zentrale Aussage des Forschungsüberblicks lautet, dass alle Bereiche, die für erfolgreiche und effektive Regierungskommunikation notwendig sind, sich durch Digitalisierung und die Nutzung von KI in beschleunigter Transformation befinden, darauf jedoch operativ sehr unterschiedlich reagiert werden kann/muss: Die Vermittlung von Kommunikation ist etwas gänzlich anderes als veränderte Kommunikationserwartungen seitens der Bürgerinnen und Bürger. Die zentrale Herausforderung besteht nicht darin, die einzelnen Veränderungen zu identifizieren, sondern die emergenten Effekte aus deren Zusammenspiel zu modellieren.

Es existiert noch keine einschlägige Literatur, die sich dem Thema explizit und tiefgehend widmet und Regierungskommunikation und KI zusammenführt. Selbst größere Sammelbände, die Regierungskommunikation und Öffentlichkeitsarbeit (Borucki und Jun 2018; Raupp et al. 2018a) oder den digitalen Wandel der Öffentlichkeit (Eisenegger et al. 2021) behandeln schneiden das Thema nur bedingt an. Im Handbuch Politische Kommunikation findet sich nur ein einschlägiger Essay (Borucki et al. 2020a). Auch in anderen Ländern wird das Thema nur bedingt aus einer Makroperspektive betrachtet und werden viel mehr einzelne Bereiche und Teile der genannten Dimension, z.B. das Phänomen der Desinformation oder des Microtargeting und der algorithmischen Personalisierung behandelt (Bennett und Livingston 2021).

Der Bericht konzentriert sich auf deutsch- und englischsprachige Quellen, die mittels Schneeballmethodik identifiziert wurden. Hierfür wurden auch digitale KI-gestützte Tools herangezogen, vor allem die Plattformen „Connected Papers“ (*https://www.connectedpapers.com/*) und „Scholarometer“ (*https://scholarometer.indiana.edu/*), die Zitationen zählen und Verbindungen zu anderen Disziplinen, Autorinnen und Autoren und Journalen aufzeigen. Der vorliegende Bericht konzentriert sich auf Texte der Politik- und Kommunikationswissenschaft, der Psychologie sowie der Informatik. Viele der hier zitierten Quellen sind inter- und transdisziplinäre Produkte und weisen auf die facettenreiche Natur des Themas. Für relevante Fälle und Beispiele, die no ch nicht den Weg in die Forschung gefunden haben, werden Primärquellen als auch Expertiseblogs und Magazine herangezogen.

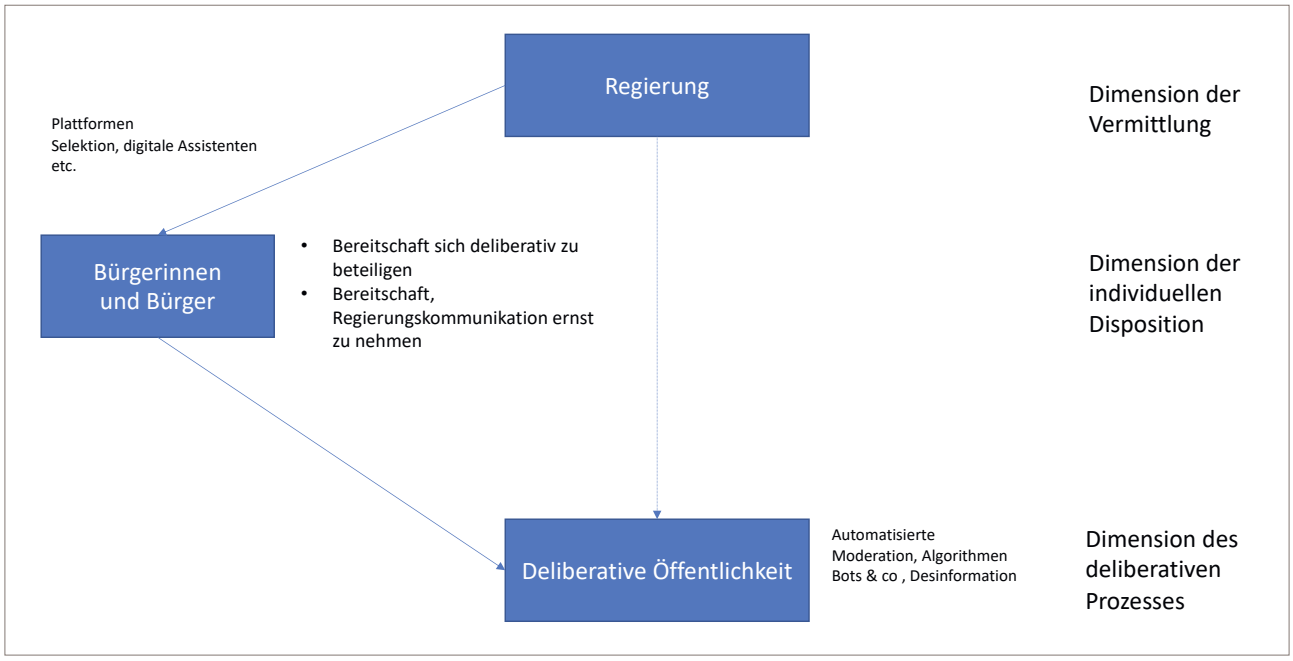

Abb. 1: Analytisches Raster – Dimensionen der Analyse, eigene Darstellung

Künstliche Intelligenz: Begriffe, Verfahren, Anwendungen

Mit dem Begriff „Künstliche Intelligenz“ (KI) umbeschreibt man eine Technologie, die in der Lage ist, ein Problem zu lösen, für das Intelligenz nötig ist. Auf der technischen Seite geht es um den Einsatz von Algorithmen, von unterschiedlichen Methoden des maschinellen und tiefen Lernens sowie um Big Data Analysen, die eine bisher nie dagewesene Abstraktionsebene, Muster- und Regelerkennung ermöglichen. Nach Kaplan (2016) manifestieren sich diese Technologien in der Anwendungsebene insbesondere in der Robotik, in der „Computer Vision“ (computerbasiertes Sehen), Spracherkennung und -verarbeitung (Natural Language Processing, (NLP)) sowie in der Problem- und Musterkennung. Operativ wird KI zur Reduzierung von Komplexität, Klassifizierung und Clusterung in einem Meer von Daten, zur Erfassung von Sprach- und Bildinformationen und zur algorithmischen Personalisierung (auch Individualisierung genannt, z.B. Recommendation-Systeme bzw. „Recommender“) eingesetzt (Kaplan 2016).

Definition von Künstlicher Intelligenz

Generell wird zwischen “schwacher KI”, “allgemeiner KI” und “starker KI” unterschieden. Schwache KI wird oft im englischen auch als “narrow” bezeichnet, also eine enge bzw. begrenzte KI. Diese Verfahren und Anwendungen künstlicher Intelligenz existieren. Der Begriff beschreibt Verfahren (maschinelles und tiefes Lernen), die der Bearbeitung einer bestimmten Aufgabe dienen, zur Lösung eines spezifischen Problems führen oder die Ausführung einer spezifischen, eingegrenzten Anwendung ermöglichen. Je nach Anwendungsbereich wird oft die menschliche Leistung um ein Vielfaches übertroffen. So zum Beispiel in der Bearbeitung von massiven Datensätzen oder in der Mustererkennung und Wiedergabe von Daten (z.B. Übersetzungstools wie DeepL, Natural Language Processing und weitere) oder beim Schachspielen. Diese “schwache KI” ist leistungsstark in einem Anwendungsbereich, aber auch nur in diesem einen, daher “schwach” oder “begrenzt”.

“Allgemeine KI” beschreibt ein theoretisches Konstrukt. Eine solche Intelligenz wäre in der Lage, weitgehend wie ein Mensch zu agieren, also multiple Bereiche zu bearbeiten und mit menschenähnlicher Intelligenz mehrere unterschiedliche Probleme zu lösen. KI-Assistenten, wie Siri, Google Assistant oder Alexa tangieren diese Ebene nur bedingt. Eine “starke KI”, ein weiteres theoretisches Konstrukt, wäre dem Menschen weit überlegen. Die letzten zwei Bereiche sind eine Zukunftsvision und werden in der vorliegenden Arbeit nicht weiterverfolgt.

Begriffe und Komponenten von KI

Oft werden die Begriffe „Künstliche Intelligenz”, „maschinelles Lernen” (machine learning), „neuronale Netzwerke” (neural networks) und „tiefes Lernen” bzw. „mehrschichtiges Lernen” (deep learning) austauschbar und/oder in unterschiedlichen Zusammenhängen verwendet. Die Begriffe sind miteinander verknüpft und stellen spezifischere Teile eines KI-Ver-

fahrens dar. Man kann sich den Zusammenhang wie eine Matroschka, eine Schachtelpuppe vorstellen.

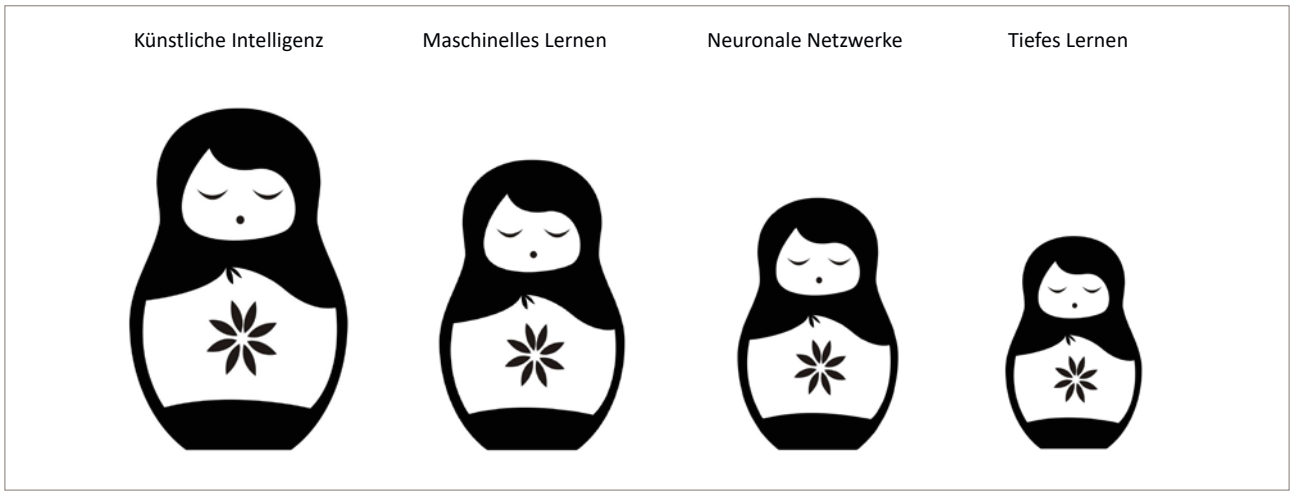

Abb. 2: KI-Begriffe als Schachtelpuppe, Abbildung anhand eines Beispiels von Kavlakoglou 2020.

KI ist der Oberbegriff. Maschinelles Lernen ist ein Anwendungsbereich von KI, der das Trainieren (das Lernen) eines Algorithmus befähigt. Tiefes Lernen ist ein Unterbereich von maschinellem Lernen, der ein genaueres und komplexeres Lernen und damit entsprechende Anwendung ermöglicht. Neuronale Netzwerke (künstliche neuronale Netzwerke) sind das Rückgrat für tiefes Lernen, denn nur durch sie wird „tiefes“ oder „mehrschichtiges“ Lernen möglich. Künstliche neuronale Netzwerke bestehen aus komplexen Algorithmen mit mehreren Parametern und Gewichtungen, die das menschliche Gehirn in seiner Komplexität imitieren sollen. Neuronale Netzwerke sind Algorithmen mit mehr als drei Schichten, also mit mehr als drei Parametern, die bei einer Entscheidungs- oder Musterfindung gewichtet werden (Trabold 2021; Kavlakoglu 2020).

Anwendungsbereiche und Lernformen

Der Anwendungsbereich (welches Ziel soll ein Algorithmus erfüllen) entscheidet über das Verfahren (wie und was der Algorithmus lernt) und wie komplex er sein muss, also darüber, ob maschinelles oder tiefes Lernen nötig ist. Sollte zum Beispiel eine Maschine nur erkennen, ob ein Bild einen Hund oder keinen Hund beinhaltet, so genügt ein Verfahren von maschinellem Lernen. Sollen jedoch komplexere Muster, wie mehrere Tiere und ihre Anzahl auf einem Bild erkannt werden, so wird tiefes Lernen und dafür ein neuronales Netz nötig. Für beide Verfahren sind größere Datenmengen nötig, um den Algorithmus so gut wie möglich zu trainieren. Als Faustregel gilt, je mehr Lerndaten ein Algorithmus erhält, desto genauer sind die gelieferten Ergebnisse.

Es gibt verschiedene Lernverfahren, die für KI-Verfahren von Bedeutung sind. Hier werden sie kurz vorgestellt.

Das “unüberwachte Lernen”: Ein Algorithmus, der unüberwacht lernt, bekommt währenddessen kein Feedback. Er versucht in diesem Fall, Muster in den Daten zu finden und daraus zu lernen. Zu den unüberwachten Lernmethoden gehören alle ähnlichkeitsbasierten Verfah-

ren, Clusteranalyse eingeschlossen, welche ähnliche Objekte zu Gruppen zusammenfasst, oder einige Verfahren der Dimensionsreduktion.

Das “überwachte Lernen”: Beim überwachten Lernen ist eine Rückmeldung für jedes Trainingsbeispiel möglich. Das Feedback wird vom Algorithmus automatisch aus den Trainingsdaten gesammelt. Dies setzt voraus, dass die Trainingsdaten so beschaffen sind, dass die richtige Antwort für jeden Datenpunkt bekannt ist. Ein klassisches Beispiel für überwachtes Lernen ist das Training von Klassifikationsmodellen.

Das „semi-überwachte Lernen”: Semi-überwachtes Lernen ist ein Lernverfahren, in dem viele der Trainingsdaten bzw. Lerndaten nicht gekennzeichnet sind. Der Algorithmus kann diese zusätzlichen nicht gelabelten Daten nutzen, um die Form der zugrunde liegenden Datenverteilung besser zu erfassen und besser auf neue Stichproben zu verallgemeinern. Diese Algorithmen können gute Ergebnisse erzielen, wenn wir eine sehr kleine Menge an gelabelten und eine große Menge an nicht gelabelten Trainingsdaten haben.

Das „reinforcement Learning“, zu Deutsch auch bestärkendes Lernen, unterscheidet sich von den genannten Typen dadurch, dass nicht zu jedem Datenpunkt ein Feedback erzeugt wird, sondern erst nach dem Durchlaufen einer Reihe von Feedback-Schleifen. Das ist der Fall bei Spielen, wo es für Gewinnen ein positives und für Verlieren ein negatives Signal gibt. Gesamtfeedback gibt es am Ende des Spiels, wenn alle Züge zum Verlieren oder Gewinnen geführt haben. Die Züge werden zurückgerechnet. Damit der Algorithmus eine gute Strategie lernen kann, muss ein Spiel mehrmals gespielt werden.

Das aktive Lernen: In einigen Anwendungsfällen ist es unmöglich oder sehr teuer, die richtigen Antworten für alle Datenpunkte zu beziehen. Empfehlungssysteme sind ein klassisches Beispiel für den Einsatz von aktivem Lernen. Ohne das Feedback der Nutzerinnen und Nutzer können solche Systeme nicht funktionieren. Das System analysiert die Präferenzen der Nutzenden und passt die Vorschläge entsprechend an. Durch die aktive Auswahl der Vorschläge wird das System zu einer wichtigen Entscheidungshilfe für Kundinnen und Kunden. Aktives Lernen dient zum Beispiel der Personalisierung von Algorithmen und zum Training von „Recommender“-Systemen (Vorschlagssystemen).

Deliberative Demokratie im Umbruch: Herausforderungen des digitalen Wandels

Das disruptive Potential der Digitalisierung wurde bereits früh erkannt, aber das Ausmaß und die Auswirkungen für den öffentlichen Dialog sind erst später systematisch erfasst worden (Habermas 2021). In dieser Sektion betrachten wir insbesondere die Rolle der deliberativen Demokratie und die Disruption der Digitalisierung.

Das Konzept der deliberativen Demokratie stellt für das demokratische System die kommunikative Dimension in den Mittelpunkt. Damit ist das Konzept nützlich, um öffentliche und politische Kommunikation als Teil des demokratischen Systems und der demokratischen Willensbildung zu verstehen. Deliberation, also der öffentliche Austausch von faktenbasierten Argumenten, ist der Kern der öffentlichen Debatte, sei es auf grassroots-Ebene, im Parlament eines Landes oder im Bundestag. Der öffentliche Austausch von Informationen und Argumenten dient zur Willensfindung und -bildung und hat eine rationalisierende Wirkung auf die Beteiligten. Massenmedien haben diesen Prozess weitgehend möglich gemacht und beeinflusst (Habermas 1989).

Abb. 3: Dimensionen der Digitalisierung der politischen Kommunikation nach Vowe 2020, Abb. ibd. Vowe 2020, S. 3

Die Dynamik der Digitalisierung hat einen „erneuten Strukturwandel der politischen Öffentlichkeit“ ausgelöst. Diesen zu begreifen ist nicht nur für die deliberative Demokratie und das demokratische System wichtig, sondern auch, um in der politischen und öffentlichen Kommunikation zu navigieren. Habermas betont, dass Digitalisierung und digitale Medien die rationalisierende Wirkung der Öffentlichkeit minimiert haben, weil das Öffentliche und das Private weitgehend untrennbar geworden sind (Habermas 2021). Das Verschmelzen von Privatem und Öffentlichem ist nur einer von mehreren Effekten des digitalen Wandels. Zur

Kategorisierung der Effekte der Digitalisierung auf Medien und Öffentlichkeit dient Vowe's Oktogon, das den strukturellen Wandel treffend zusammenfasst (Vowe 2020).

Vowe spricht dabei von der „Algorithmisierung" der politischen Kommunikation, denn das gesamte Oktogon, das mit „Digitalisierung" beginnt, wird erst durch den Einsatz von Algorithmen möglich. Diese Algorithmen sind allgegenwärtig und ermöglichen mithilfe der digitalen Plattformen und des Internets (Gillespie 2018; Gorwa 2019) alle hier genannten Teile des strukturellen Wandels der politischen Kommunikation. Dabei sind bestimmte Teile für unser Verständnis von politischer Kommunikation von Bedeutung:

a. Die Ökonomisierung der Kommunikation bedeutet, dass der Algorithmus effizient die Aufmerksamkeit der Nutzerinnen und Nutzer aufrecht erhalten muss – ein Phänomen auch beschrieben als Attention Economy – oder Aufmerksamkeitsökonomie (Crogan und Kinsley 2012; Davenport und Beck 2001). Dies drückt sich z.B. in der durch Algorithmen gesteuerten Monetarisierung von Aufmerksamkeit in Form des individualisierten Ausspielens von Aufmerksamkeit generierenden Inhalten durch Social Media Plattformen aus.

b. Die Ökonomisierung und Dynamisierung gehen Hand in Hand, denn der Algorithmus erkennt die jeweiligen Nutzerinnen und Nutzer und ihr Verhalten, und bietet immer weiter ähnliche Inhalte zum richtigen Zeitpunkt an, um die Aufmerksamkeit zu halten (Ciampaglia et al. 2015; Harsin 2015).

c. Pluralisierung bezieht sich auf die große Anzahl von Teilnehmenden, wobei es um algorithmische Steuerung geht, die bestimmten Personen, Themen, Artikeln, Videos etc. mehr oder weniger Sichtbarkeit bietet, die die entsprechende Aufmerksamkeit generiert.

d. Der Begriff 'Hybridisierung' greift im Kontext der Dimensionen der Digitalisierung politischer Kommunikation zu weit, um ihn sinnvoll zu operationalisieren. Es handelt sich um das Verschmelzen von Kommunikation und Medien in der digitalen und der analogen Welt. Die Hybridisierung beinhaltet auch das Verschmelzen privater und öffentlicher Person durch die Nutzung sozialer Medien. Dabei werden Daten von Plattformen erhoben, die die „öffentliche" und „private" Person verschmelzen lassen. Im theoretischen Idealbild der deliberativen Demokratie ist die öffentliche und die private Person einer Bürgerin oder eines Bürgers voneinander getrennt.

e. Spezifizierung beschreibt den Kern des Problems eines lernenden Algorithmus, der seine Nutzerinnen und Nutzer immer besser und tiefer kennt, und ihnen entsprechende Inhalte (content) anbietet, um deren Aufmerksamkeit aufrechtzuerhalten oder auf bestimmte Produkte zu lenken. Wir sprechen dabei von „Personalisierung" – den „richtigen" Content an die „richtige" Person zu lenken.

Dimension der Vermittlung: wo und wie wird Kommunikation vermittelt?

Neben den analogen Nachrichten in Zeitungen, Fernsehen und Radio gewinnen digitale Medien an Bedeutung, besonders bei Personen unter 30 Jahren. Hier erlangen Nutzerinnen und Nutzer die Informationen über soziale Medien, Apps, Nachrichtenseiten und weitere digitale Quellen, sei es über Messenger-Apps (Telegram, Whatsapp, Viber, etc.) oder über Webseiten. Dabei steigt die Komplexität und die Diffusion zwischen klassischen Nachrichten (z.B. Online-Zeitungen und ihre Outlets in sozialen Medien), algorithmisch kuratierten Vorschlägen aus einem Pool von Nachrichtenseiten (z.B. über Google News) und Informationen aus einzelnen Profilen bzw. Quellen ohne journalistische Redaktion, z.B. Profile von (oft nur angenommenen) Expertinnen und Experten, Influencerinnen und Influencern, Blogs und (nichtnachweisbaren) Profilen oder Quellen (inklusive Bots) (Neuberger 2014, 2009).

Es ist wichtig zu differenzieren: nicht alles, was digital ist, basiert auf oder nutzt extensiv KI-Verfahren. Einschlägiges Beispiel ist die Messenger-App Telegram. Sie nutzt KI-gestützte NLP zur Übersetzung von Nachrichten und ermöglicht es, Chatbots in Gruppen und Kanälen einzusetzen. Allerdings verfügt sie über keine algorithmische Personalisierung: Es werden keine Vorschläge gemacht, wem zu folgen oder was zu betrachten wäre. Das ist ein Unterschied zu Twitter, das direkt von den bereits abonnierten Nutzerinnen und Nutzern und dem Engagement mit bestimmten Posts „lernt“ und weitere Profile vorschlägt, denen gefolgt werden könnte. Facebook und Instagram gehen ähnlich vor und schlagen dabei den Beitritt zu bestimmten Gruppen, oder das „Liken“ einer bestimmten Seite vor. Youtube und Tik Tok empfehlen Videos – oder spielen diese direkt ab. Die Empfehlungen resultieren dabei aus dem bisherigen Nutzungsverhalten, mit dem Ziel Userinnen und User am Bildschirm zu halten. Selbst Webseiten, die als statisch wahrgenommen werden, können zum Teil dynamisch sein und durch eine „Vorschlagsfunktion“ die Aufmerksamkeit der Userinnen und User auf weitere Inhalte lenken, so zum Beispiel *www.medium.com* oder *www.bild.de*. Diese Webseiten schlagen weitere Artikel, die die Lesenden interessieren könnten, vor. Der Vorschlag entsteht aus den Informationen, die die Webseite zu den Lesenden besitzt.

Hier lassen sich zwei verschiedene Phänomene beobachten: Das erste Phänomen ist, dass ähnlich Gesinnte sich gegenseitig folgen und so in Gruppen zusammenfinden. Das schafft eine sogenannte „Echochamber“ – eine Echokammer. Der erste, der diese These aufgebaut und bestätigt hat ist der US-amerikanische Wirtschaftsverhaltenswissenschaftler Cass Sunstein (Sunstein 2001, 2007). Die Nutzerinnen und Nutzer der Gruppe befinden sich in einer solchen virtuellen „Echokammer“, weil ähnliche Interessen, Meinungen und Weltbilder sie dorthin geführt haben. Die sich in einem geteilten Raum befindenden Personen sind weitgehend gleicher Meinung. Sie konsumieren ähnliche Medien und Nachrichten und folgen einander, reproduzieren dabei ihre bereits bestehenden Vorstellungen und Voreingenommenheiten. Hier geht es um „Nodes“, um Knotenpunkte, die die Nutzerinnen und Nutzer immer stärker verbinden. Diese Knotenpunkte führen zu Personalisierungseffekten und zum Personenzentrismus im Sinne der personenzentrierten Kommunikation (zum Beispiel das Folgen von Präsident Donald Trump) – in dem Fall der Glaube an Persönlichkeiten und ihre Seriosität (Influencerinnen und Influencer), die als wahrgenommene Expertinnen und Experten Nachrichten (egal ob wahr, unwahr, faktisch oder imaginär) verbreiten. So wird Kommu-

nikation über bestimmte einflussreiche Menschen vermittelt. Das gilt für Inhalte als auch für die Art und Weise der Kommunikation (personenzentrierte Kommunikation).

Das zweite Phänomen sind algorithmisch kuratierte Inhalte, die anhand des bisherigen Verhaltens als relevant eingestuft und vorgeschlagen werden. So schlägt der Algorithmus einer Plattform immer ähnlichere Nachrichten vor, mit Inhalten, die sich durch die Filtereinstellungen des Nutzers als auch den genutzten Algorithmus immer weiter personalisieren und eine „Blase“ um den Nutzer herum bilden. Man spricht von der sogenannten „Filterblase“. Diese These hat der US-amerikanische Vorstand einer der größten politischen Organisationen der USA Eli Pariser aufgestellt (Pariser 2011). Ein einschlägiges Beispiel von Pariser für die Filterblase war die Tatsache, dass verschiedene Menschen beim Googeln des gleichen Begriffs unterschiedliche Ergebnisse erhalten. Dabei geht es nicht nur darum, was die Suchenden sehen, sondern was ausgeblendet wird, ohne selbst zu wissen was und warum dies ausgefiltert wird. Der Bericht konzentriert sich zunächst auf die genannten Phänomene, um auf die algorithmische Personalisierung einzugehen.

Beide Phänomene, Echokammern und Filterblasen, unterliegen algorithmischer Kuratierung und beeinflussen öffentliche Kommunikation. Dies geschieht nicht nur, weil Kuratierung genutzt wird, um Inhalte zu vermitteln, sondern auch weil sie auf weitere Kommunikation unbemerkt einwirkt und den Content, die die Nutzerinnen und Nutzer konsumieren, mitbestimmt. Neben den Plattformen, in denen Information direkt (über persönliche Profile) fließt, spielen auch algorithmisch personalisierte Nachrichten (APN) eine Rolle. Nachrichten und ihre Vermittlung sind unabdingbarer Bestandteil einer deliberativen Demokratie (Neuberger 2009). Ihre Vermittlung ändert sich und verändert die Einstellung der Bürgerinnen und Bürger. Inzwischen existieren Nachrichten-Apps, als auch Nachrichtenvorschlagsfunktionen von unterschiedlichen Herstellern, Social Media Plattformen und Medienanbietern. Sie sind bereits gängige Kanäle der Informationsvermittlung. Daher verdient der Einsatz von Nachrichtenpersonalisierungsalgorithmen spezielle Beachtung. Jeder kennt es: das Handy gibt ein Signal und man wird über bestimmte Nachrichten informiert. Der personalisierte Nachrichtenalgorithmus sucht die Aufmerksamkeit der Nutzerinnern und Nutzer und bietet gleichzeitig Informationen an. Die Hypothese, dass genau diese Funktion zur einseitigen Informiertheit und zum Verlust der Öffentlichkeit oder gar zur Radikalisierung führen kann ist nicht neu, wurde bereits 2014 im deutschsprachigen Raum identifiziert (Neuberger 2014).

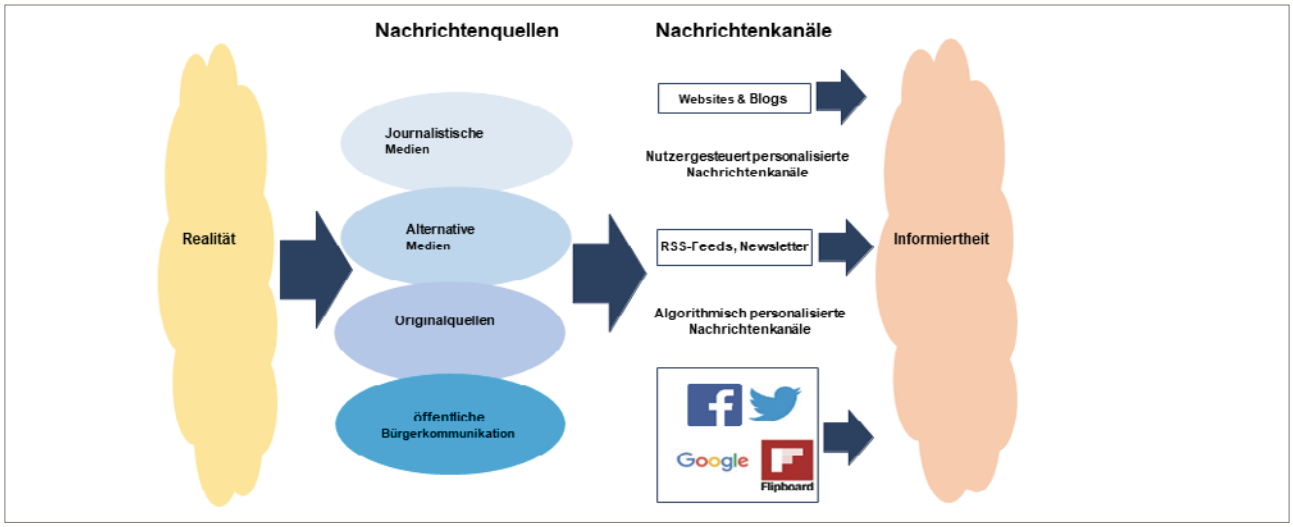

Abb. 4.: Nachrichten und intermediäre Medien aus Schweiger et al. 2019, S. 14

Basierend auf einer repräsentativen Untersuchung (n=1005) haben Schweiger et al. (2019) die Einstellungen anhand einer Vielzahl von Charakteristika der Befragten wie Alter, Bildungsniveau, Wohnort, Geschlecht, Einkommen und politische Einstellung untersucht. Zudem wurden psychosoziale Merkmale der Untersuchten einbezogen, wie Selbstwahrnehmung und Selbsterwartungshaltung. Schweiger et al. kommen zum Schluss, dass APN die „Nachrichtennutzung, Informiertheit und Meinungsbildung vieler Bürgerinnen und Bürger in den letzten Jahren erheblich verändert" haben (Schweiger et al. 2019, S. 109).

Die APN-Nutzung der Altersgruppen unter 30 Jahren ist unabhängig von weiteren Merkmalen (soziale, Bildung, ökonomischer Status) und weit stärker ausgeprägt als bei den über 30-Jährigen. APN führen dabei zu einer Polarisierung und zu einer Abnahme von Toleranz von anderen Meinungen (mehr dazu im Kapitel „Disposition"). Die Forschungsgruppe stellte fest, dass Echokammern und Filterblasen existieren und politisch relevant sind. Der Polarisierungseffekt von algorithmischen Vorschlägen im Fall der Knotenpunkte (algorithmische Vorschläge, bestimmten Personen zu folgen oder bestimmten Gruppen beizutreten) wird von Santos, Elkes et al. (Santos et al. 2021) durch ein mathematisches Modell algorithmisch rekonstruiert. Sie können zeigen, dass es tatsächlich einen Polarisierungseffekt geben kann, wenn Recommendersysteme keine Öffnung zu anderen Meinungen und Informationen ermöglichen. Das gilt insbesondere, wenn ein Recommendersystem anhand bestehender Kontakte agiert. Die amerikanische Forschungsgruppe zeigt auch, dass ein leichtes Tweaken des Algorithmus eine leicht erhöhte Offenheit für nicht-vereinheitlichte Sichtweisen bietet. Hier wird eine Abnahme der Polarisierung bei den meisten Nutzerinnen und Nutzern beobachtet. Das heißt, wenn der Algorithmus weniger personalisierend agiert, wird auch weniger Polarisierung beobachtet.

Festzuhalten ist: Kommunikation wird über Apps, Social Media und Webseiten vermittelt und unterliegt algorithmischen Prozessen, die veränderbar und manipulierbar sind. Regierungsnahe und regierungsferne Akteure unterschiedlicher Bereiche nutzen den Personalisierungsmechanismus, um ihre Inhalte zu verbreiten, eigene Kommunikation und damit Politik zu verfolgen. Drei KI-Verfahren sind hier von Bedeutung: **Government Data-Mining** als Voraussetzung für Datenerfassung und Clusterung in Anlehnung zum Data-Mining, das private Anbieter im Netz betreiben, und die Verfahren **Nudging** und **Microtargeting**, die auf Datenerhebung und -erfassung basieren.

Government Data-Mining: Datenerhebung und -erfassung für politische Kommunikation?

Bis zur Digitalisierung wurden Daten analog in Papier- oder anderen Formaten generiert und gespeichert, während viele Daten gar nicht generiert wurden (wie z.B. GPS-Daten, die jetzt über die Ortung des Handys ständig generiert werden). Große Datenmengen zu durchforsten und auszuwerten, nennt man **Data-Mining**. Das hauptsächliche Ziel von Data-Mining ist aus ungeordneten Daten (raw data) strategische Intelligenz zu generieren, die für die Ziele einer Organisation relevant sind (Gandy 2019). Data-Mining ist explorativ – es beschreibt Zusammenhänge und wird oft zur Vorschau genutzt (nicht immer mit Erfolg) (Slobogin 2008).

Ähnlich den Marketingfirmen, die anhand von Such- und Konsumverhalten personalisierte Werbung anbieten, können politische Akteure vorgehen. Große Datenmengen, die eine einzelne Person oder die gesamte Bevölkerung erzeugen, können durch Algorithmen erfasst

werden. Das erlaubt Big Data Analysen, um zum Beispiel zu sehen, wie viele Menschen um eine bestimmte Uhrzeit bestimmte Nachrichten lesen oder welche Altersgruppen bestimmte Medien nutzen oder Videos konsumieren. So lassen sich Individualanalysen, Kategorisierungen und Topografien erstellen. Im Transfer könnte man zum Beispiel im Rahmen einer Informationskampagne altersgerecht in bestimmten Ortschaften kommunizieren, die als relevant erkannt wurden.

Um Government Data-Mining zu betreiben, muss eine Regierung die Daten über Jahre digital gespeichert haben und Zugang zu den Daten besitzen (digital-warehousing) (Rakesh K. und Manoj, K., Gupta 2017). Eine der gängigsten Data-Mining Lernformen ist unter dem Begriff „association analysis model“ bekannt und erlaubt durch tiefes Lernen Datenzusammenhänge zu erkennen, zu gewichten und zu kategorisieren (Yuan Tongwei und Chen 2012). So hat im Zuge der Terroranschläge von 9/11 die USA angefangen, größere Datensätze der eigenen Bevölkerung zu erforschen, um Muster zu erkennen, bestimmte Gruppen zu identifizieren, oder gar Krisen und andere Geschehnisse zu antizipieren (Gandy 2007). Diese Möglichkeiten sind in Deutschland im Rahmen der Datenschutzgrundverordnung nicht so umfassend und weitgreifend möglich. Die Idee, dass Big Data von Bürgerinnen und Bürgern für und von der Politik gesammelt werden, wird auch kritisch gesehen. Das sogenannte „demos scraping“, (Ulbricht 2020b), das Sammeln von Daten um bessere Policy zu gestalten, kann Politik delegitimieren, da sie den deliberativen Prozess überspringen kann. Hierbei geht es aber nicht um die kommunikativen Möglichkeiten, sondern um Policy-Implementierung. Im Bereich der Kommunikation kann Data-Mining, im Rahmen des rechtlich Möglichen, präzisere Informationsarbeit ermöglichen.

Massive Datensätze sind für Kommunikationsstrategie von besonderem Interesse, weil sie Informationen zu einzelnen Personen, zu Gruppen sowie öffentlichen Prozessen beinhalten. Unabhängig von Fragen des Datenschutzes und personenbezogener Daten, kann Government Data-Mining die Kommunikation von Regierung zum Bürger revolutionieren. Geographische, altersspezifische oder sozioökonomische Merkmale können eine Kommunikationsstrategie verbessern. Sie können auch ermöglichen, frühzeitig Krisenkommunikation oder gar -prävention zu betreiben, sollten z.B. gehäuft GPS-Bewegungen wahrgenommen werden, aufgrund derer man sich auf Krisenkommunikation einstellt und vorbereitet. Offene Verwaltungsdaten ermöglichen weitere kommunikative Strategien. Dies betrachten wir in den nächsten Kapiteln. Bereiche in denen Staaten und Regierungen Data-Mining betreiben können sind zur Erkennung von möglichen terroristischen Anschlägen (Amoore und Goede 2021), zur gezielten Information für Bürger in bestimmten Regionen, bestimmten (Minderheiten-) Sprachen oder in der gezielten Bedienung von Bürgeranliegen, wobei eine Kombination von mehreren KI-gestützten Verfahren zum Zuge kommen.

Government Data-Mining ist insbesondere in den USA anders geregelt als in der EU und in Deutschland. Die Datenschutzgrundverordnung (DSGVO) lässt die Nutzung von Daten nur für den erhobenen Grund zu. Jedoch existieren Konzepte zu Open e-Governance und Open e-Government, die sich explizit mit dem Thema „offene Verwaltungs- und offenere Regierungsdaten“ auseinandersetzen. Die Bundesrepublik ist in diesem Rahmen Mitglied des „Open Government Partnerships“ und realisiert in bestimmten Bereichen eine Öffnung von öffentlichen Datenbeständen, die dem Gemeinwohl dienen sollen, z.B. bei Forschungsdaten öffentlicher Einrichtungen (Lucke und Gollasch 2022a, S. 168). Zudem besteht die Möglichkeit eines „opt-in“ zur freiwilligen Datenerhebung der Bürgerinnen und Bürger. Die Bundesregierung hat mit der „Datenstrategie“ im Jahr 2021 die Nutzung von personenbezogenen,

nicht-personenbezogenen und erweiterten Roh- oder verarbeiteten Daten unter Berücksichtigung geltenden Rechts und Schutzes der Privatsphäre und des Urheberschutzes beschlossen (Die Bundesregierung 2021). Die Strategie sieht keine explizite Nutzung für Informationsarbeit vor, doch das könnte im Rahmen des Nutzungstransfers erfolgen. Wenn Daten im Rahmen der offenen Verwaltungs- und Regierungsdaten anonymisiert und ohne Rückverfolgung erhoben, erfasst und fürs Mining genutzt werden dürfen, so kann das auch zum Transfer für Verwaltungskommunikation und Regierungskommunikation genutzt werden, so das Beispiel aus einem anderen EU-Land.

In Griechenland hat eine Forschungsgruppe in enger Kooperation mit dem Finanzministerium durch eine Kombination von KI-Verfahren den erfolgreichen Einsatz von einem „AI-guided“ Chatbot erprobt, wobei der Chatbot wie ein vereinfachter Bürgerservicepoint agierte. Es kamen mehrere KI-Verfahren zum Einsatz, um komplexe verwaltungstechnische Fragen präzise zu beantworten. Einerseits konnte Data-Mining helfen Gruppen zu clustern und entsprechend die Kommunikation des Chatbots anzupassen und zu verbessern, als auch die Dienstleistung von staatlicher Seite aus zu verbessern und mit Information zu unterstützen (Androutsopoulou et al. 2019). Dieses Beispiel wäre so in Deutschland nicht umsetzbar, da Informationsarbeit und Regierungskommunikation streng getrennt von Verwaltungskommunikation sind. Allerdings dient dieses Beispiel der Verbildlichung der Vorteile von Government Data-Mining, da durch die Auswertung von bestehender Kommunikation, zukünftiger Kommunikationsbedarf und Themen zum Teil vorausgesehen werden können. Um den Chatbot zu trainieren, benötigt man Data-Mining. Zudem können KI-gesteuerte Chatbots bestimmte Prozesse der Informationsarbeit automatisieren. Welche das sein könnten ist eine politische, juristische und soziale Frage, die geprüft werden müsste.

Algorithmisch Entscheidungen „anstoßen“?

Der Wirtschaftspsychologe und Nobelpreisträger Richard Thaler hat mit seinem bereits erwähnten Kollegen Cass Sunstein das Konzept „Nudge“ zur positiven Verhaltensänderung vorgestellt (Thaler und Sunstein 2009). Das Konzept stammt aus dem Bereich der Verhaltenswissenschaften. Es geht darum, Menschen bei Entscheidungen zu beeinflussen und in die „richtige“ Richtung zu lenken: es wird davon ausgegangen, dass Menschen teilrational handelnde Akteure sind, aber durch ein „Nudging“ in die richtige Richtung „bessere“ Entscheidungen treffen können. Das geschieht durch Eingriffe in die Entscheidungsarchitektur, oder durch die Art und Weise wie Information dargestellt wird. So das Beispiel, dass man gesünder essen wird, wenn in der Kantine gesundes Essen etwas günstiger ist – man hat immer noch die Wahl, aber man wird zur „gesünderen Entscheidung“ angestoßen. In diesem Abschnitt wird die Literatur zum kommunikativen und algorithmischen Nudging diskutiert.

Während sich das Konzept in der analogen Welt auf Eingriffe in die Umwelt, also im Design einer Entscheidungsfindung bezieht, wurde es ab 2012 auch im digitalen Raum erprobt. Dafür wurden Datenerhebung, -erfassung und -analyse mit digitalen und analogen Kommunikationsmaßnahmen kombiniert. Eine der bekanntesten und weitgehend offen agierende Regierungsbehörde, die es als erste tat, ist das Behavioral Insight Team (BIT) aus Großbritannien (bekannt auch als „Nudging Unit“). BIT wurde bereits 2010 von der britischen Regierung ins Leben gerufen. Mehrere Regierungen, darunter aus Frankreich, Kanada, Australien und den USA haben das BIT eingeladen, um Unterstützung in Policy-Umsetzung zu erhalten. Nudgen erscheint für bestimmte Fälle als attraktive Alternative zur Deliberation oder als

schlaue Kombinationslösung neben traditionellen gesellschaftlichen Prozessen (Benartzi et al. 2017).

2015 hat die Bundesregierung das Politiklabor Wirksam regieren geschaffen, das weiterhin besteht und wissenschaftliche Erkenntnisse in praktische Politik zu übersetzen sucht (Die Bundesregierung 2022). Das Labor, das im Bundeskanzleramt angesiedelt ist und als ein Pendant zum BIT gesehen werden könnte, beschäftigt sich mit Gesundheitsförderung, Energieeffizienz und weiteren Themen. Zu Wirksam regieren konnte der Schreibende nur eine begrenzte Zahl von wissenschaftlichen Beiträgen finden, die die politische Dimension tangieren (Weber and Schäfer 2017; Krisam et al. 2017).

Nudgen durch Information ist anders als Nudgen durch Design. Es geht nämlich um eine Verhaltensänderung durch Kommunikation, nicht durch Regulationen und Gesetze oder durch Designeingriffe. Ein Beispiel aus der Finanzverwaltung Großbritanniens zeigt, wie Kommunikation Verhalten verändern kann, ohne auf Strafandrohung zu setzen: Die Finanzverwaltung hat Menschen, die ihrer Steuerschuld nicht nachkommen, kontaktiert, mit der Nachricht, dass die Mehrheit der Bewohnerinnen und Bewohner ihrer Stadt bereits bezahlt hat. Das führte zu einer Zahlungssteigerung von 15%. Dem BIT nach hatte das eine positive Wirkung. Der Vorgang wäre sogar auf die gesamte Bevölkerung skalierbar (The Guardian 2013). Das BIT ist heute weiter hin aktiv und wird in der ökonomischen Verhaltensänderung positiv gewertet (The Behavioural Insights Team 2023). Für die Erfolgswahrnehmung des BIT spricht seine Aufnahme in Großbritanniens Innovationsstiftung Nesta im Dezember 2021.

Mit Blick auf Nudging werden neue Technologien und neue Designelemente, wie Kinetics, Virtual Reality und das Internet der Dinge interessant. Denn hier werden sich neue Design- und Nudging-Möglichkeiten bieten, die Entscheidungen beeinflussen werden, daher lohnt sich insbesondere Forschung im digitalen Nudging (Weinmann et al. 2016). Die Effektivität von Nudging wurde hinterfragt und dafür wurden Publikationen qualitativ und quantitativ verglichen (Hummel und Maedche 2019). Hummel und Maedche haben zeigen können, dass Nudging vom Anwendungsfall abhängig ist (insbesondere im Bereich Gesundheit) und dass Nudgen im digitalen Bereich ähnlich zum analogen ist. Sie bemerken, dass hier viel Raum für weitere Forschung besteht, insbesondere mit Blick auf weiteren Zukunftstechnologien. KI kann Nudgen weiter verbessern, da der Algorithmus durch Nutzungsfeedback seine Wirkung messen kann, um die Nutzerinnen und Nutzer in die erwünschte Richtung zu „schubsen“ (Wagner 2021). Wir müssen hier verschiedene Elemente voneinander trennen: Nudgen als Regulation bzw. Design, KI im Nudge-Verfahren als Monitoring und Nudging durch Kommunikation in Kombination mit Monitoring, um zu sehen, ob die Kommunikationsmaßnahme greift. In der Kommunikation und mit Blick auf das genannte Feld der APN kann man sich fragen, ob APN das Verhalten einer Nutzerin oder eines Nutzers beeinflusst – quasi in eine bestimmte Informationsrichtung „nudged“.

Jesse und Jannach (2021) haben eine vergleichende Studie von „recommender systems“ gemacht (gemeint sind algorithmisch kuratierte Vorschläge) und sind zum Ergebnis gekommen, dass sie stark wirken und die Nutzer in die „gewünschte“ Richtung lenken (Jesse und Jannach 2021). Unabhängig von den ethischen Aspekten der Manipulation, der Autonomie oder der erwünschten Stärkung des Gemeinwohls (Schmidt und Engelen 2020; Sunstein 2015; Wilkinson 2013), sorgt die Anwendung von Nudging in der Kombination mit Government Data und Algorithmen im realen wie im digitalen Bereich auch im Vereinigten König-

reich für Aufsehen. In einem kritischen Artikel der The Guardian aus 2021 wird die mögliche Wirkung wie folgt dargestellt:

> *While targeted advertising is common across business, one researcher argues that the government using it to drive behavioural change could create a perfect feedback loop.*
>
> *"With the government, you've got access to all this data where you can see pretty much in real time who you need to talk to demographically, and then on the other end you can actually see, well, 'did this make a difference?'," said Ben Collier, of the University of Edinburgh. "The government doing this supercharges the ability of it to actually work." (The Guardian 2021)*

Nudgen ist hier nicht doppeldeutig: es geht um die kommunizierte Nachricht und das Monitoring der Effektivität der Maßnahme. Wurde durch die Kommunikationsmaßnahme das Verhalten der Bürgerin oder des Bürgers verändert, so lässt sich das durch digitale Instrumente und Algorithmen nachvollziehen. Durch Big Data lässt sich das auf die gesamte Bevölkerung modellieren und digital „reproduzieren". Im theoretischen Sinne ermöglicht Digitalität eine Individualisierung und Skalierung einer Kommunikationsmaßnahme auf die gesamte Bevölkerung. Dass Nudging in der politischen Kommunikation eine Rolle spielen kann, wurde auch in der Regierungskommunikation Großbritanniens während der Covid-Pandemie sichtbar, wo die Nudge-unit auf „scare tactics" setzte, um die Bürgerinnen und Bürger zu beeinflussen und zur Einhaltung von Maßnahmen zu bringen (Yorke 2022). Es ist ein Unterschied, ob man auf Vernunft und Information setzt, oder durch Kommunikationsmaßnahmen Menschen in eine Richtung bewegen möchte. Inzwischen wird Nudging in verschiedenen Formen in mehr als 130 Ländern angewendet und der Trend soll zunehmen (Whitehead et al. 2014).

Wenn Nudge-Maßnahmen auf die gesamte Bevölkerung zielen, so werden Big Data mit Nuding kombiniert. Man spricht von „Big Nudging" (Irgmaier und Ulbricht 2017). Die WZB-Mitarbeiterin und der Mitarbeiter kommen zum Schluss: „Rechtliche Regulierung bietet die Möglichkeit, nützliche Formen datengestützter Verhaltensbeeinflussung zu fördern und potenziell schädliche Anwendungen auszuschließen." Gleichzeitig sehen sie die Gefahr die Selbstbestimmung der gesamten Gesellschaft zu untergraben (Irgmaier und Ulbricht 2017). Ähnlich sehen es weitere Wissenschaftlerinnen und Wissenschaftler. Big Nudging kann nicht die Antwort für große Herausforderungen sein, da wir nicht vorhersehen können, ob das Rezept überhaupt zum Ziel führt (Helbing 2017). Hinzu kommt die Entmündigung der Bürgerinnen und Bürger (Hofstetter 2016; Helbing et al. 2018). Was sichtbar wird: Big Data, und damit auch Data-Mining sind für mehrere KI-Verfahren relevant. Inzwischen zeichnet sich ein neues Konzept ab, das Big Nudging und KI zusammenführt, das „algorithmic nudging" (Möhlmann 2021).

Es muss unterschieden werden, ob es darum geht, Menschen zum Händewaschen zu motivieren oder darum, politisches Verhalten und öffentliche Policy zu verändern. Nudgen findet im freien Markt, im Internet und anderen monetarisierten Bereichen ständig statt. Problematischer ist der Einsatz im Politischen, als Teil öffentlicher Policy, oder als subtile Kommunikationsmaßnahme mit politischen Zielen. Nudgen ist keineswegs deliberativ. Ausnahmen wären Änderungen im Design, die öffentlich debattiert werden, z.B. ob alle Menschen zur Organspende verpflichtet sind, wenn sie es nicht explizit verneinen.

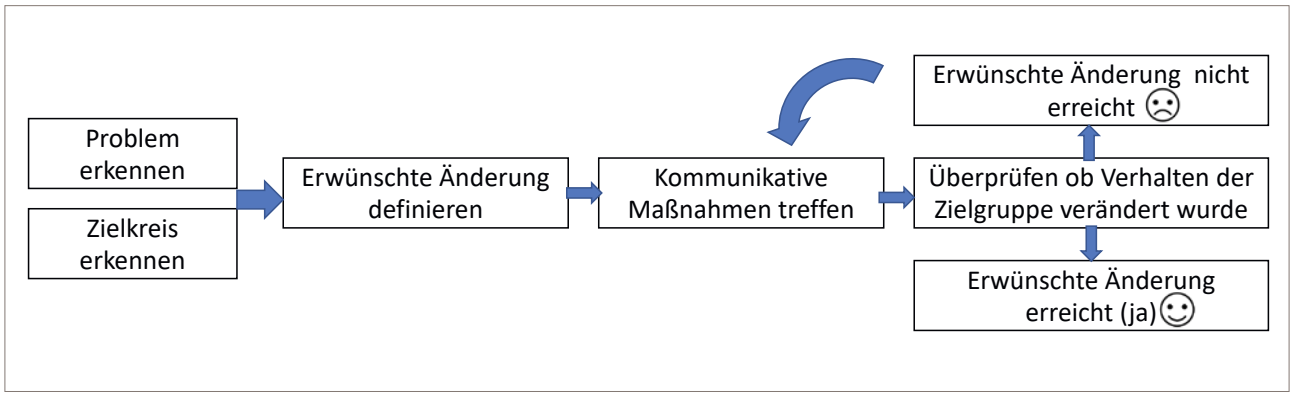

Abb. 5: Der „Nudging" Prozess mit Kommunikationsmaßnahmen und algorithmischer Überprüfung, eigene Darstellung.

In den genannten Artikeln von Wager (2021), Hummel und Maedche (2019) und Jesse und Janach (2021) werden Formen der Erforschung von Nudging als auch Zukunftsbereiche, wie Recommendation-Systeme, und neue Technologien (virtual reality) vorgestellt, die erforscht werden sollen. Besonders im letzten Teil können alle Maßnahmen des Nudgen zusammengeführt werden, inklusive Eingriff in die Umwelt, ins Design und in der Kommunikation. Zudem kann „nudging" auf sozialen Plattformen durch weitere Apps integriert werden. Wenn man die Menschen zum Überprüfen einer Information animiert, kann man damit zum Faktencheck zur Fakes-Erkennung beitragen (Thornhill et al. 2019). Damit zeigt sich eine Lücke, insbesondere in Deutschland, wo Nudgen zwar von Marketingseite aus stattfindet, aber im politischen Kommunikationsbereich und in der Regierungskommunikation zumindest im digitalen Format fast gänzlich abwesend ist, oder entsprechende Forschung von den Autoren dieses Berichts nicht gefunden werden konnte. Jedoch ist hier Vorsicht geboten: Big Nudging ist schwer mit Deliberation vereinbar, wenn nicht komplett konträr zum deliberativen demokratischen Prozess. Die subtile Lenkung von Bürgerinnen und Bürgern in eine bestimmte Richtung ohne einen offenen Dialog über Richtung und Lenkung kann unvorhersehbare Folgen haben. Was in anderen Ländern längst umgesetzt wird, ist in Deutschland so – zumindest in der nach Innen gerichteten Regierungskommunikation – nicht möglich.

Zielgruppenkommunikation und themendatenbasierte Zielgruppenansprache

Bekannt aus dem Marketing und in direktem Zusammenhang zu Data-Mining, mit dem Ziel präzise zu vermarkten, steht „Targeting" und „Microtargeting". Targeting ist die datenbasierte Zielgruppenansprache. Durch das Sammeln von Daten auf digitalen Plattformen, sozialen Medien und Webseiten können Zielgruppen identifiziert werden und so gezielt von Marketingkampagnen angesprochen werden. Microtargeting ist noch enger gefasst. Hier werden noch präzisere Personenkreise definiert und passgenau ihren Präferenzen entsprechend angesprochen (Leitherer 2019). Es kommt zur Kombination von mehreren Verfahren (Mensch-KI-Mensch-Interaktion), wobei die Zielgruppe über Menschen definiert (nicht nur), aber durch Algorithmen erkannt wird und entsprechend mit kommunikativen Elementen von einer KI bespielt wird, deren Inhalte jedoch von Menschen bestimmt sind und durch KI-Verfahren weiter verfeinert werden können. Die Firma Cambridge Analytica hat behauptet, mit Microtargeting die US-Wahlkampagne beeinflusst und zum Wahlerfolg von Donald Trump beigetragen zu haben. Während es schwer nachweisbar ist, ob Microtargeting in dem Fall

so effektiv war, ist die Natur der Anwendung für kommunikative Zwecke interessant, da sie gezielte und passgenaue Kommunikation ermöglicht (Leitherer 2019). Unabhängig von der Legitimität der Anwendung für politische Zwecke, ist hier die Anwendung für Regierungskommunikation von besonderem Interesse mit viel Raum zur Forschung.

Microtargeting für politische Kampagnen wird weiterwachsen und bringt Vor- und Nachteile mit sich (Zuiderveen Borgesius et al. 2018). Zu den Erkennungsmethoden des Microtargetings gehören „Personality-Profiling Algorithmen", die nach Kriterien gewichtet und angepasst werden können (Zarouali et al. 2020). Interessanter noch ist die Frage, was Microtargeting für Deutschland bedeutet (Papakyriakopoulos et al. 2017). Papakyriakopoulos et al. haben zeigen können, wie sie aus politischen Diskussionen auf Facebook durch algorithmische Anwendungen und maschinellem Lernen Zielgruppen identifizieren konnten. Diese wurden durch weitere Charakteristika und Präferenzen analysiert, die über Facebook sichtbar waren. Diese Daten wurden weiter bereichert und nach politischer Präferenz klassifiziert. Im Paper zeigt sich, wie Microtargeting zur zielgenauen politischen Kommunikation angewendet werden kann, indem freiwillig offengestellte Daten genutzt werden. Sie zeigen, dass es möglich ist – unabhängig von den kognitiven, ethischen, sozialen und psychologischen Dimensionen, die Microtargeting tangiert (Papakyriakopoulos et al. 2018). Die genauen Effekte des Microtargeting sind davon abhängig, wie dieses wahrgenommen wird – ob als personalisierte Werbung oder als genuine Nachricht. Das lässt Raum zur weiteren Forschung offen (Kruikemeier et al. 2016), da man sich fragen kann, wie Nachrichten gekennzeichnet werden müssten, um als Werbung oder Information erkannt zu werden. Kruikemeier et al. tangieren hier die kognitive Ebene, um zu verstehen, wann Microtargeting funktioniert.

In Deutschland ist politisches Microtargeting von Regierungsseite aus nicht möglich, da Daten nur zweckgebunden erhoben und verarbeitet werden dürfen. Was denkbar wäre, wäre ein allgemeineres Targeting zu Informationszwecken, durch die Nutzung von KI-Verfahren und durch die Schaffung von Topografien. Zum Beispiel könnte man zum Katastrophenschutz Informationen an Anwohnerinnen und Anwohner von bestimmten Regionen über verschiedene digitale Kanäle kommunizieren. Man könnte Informationen an in Krisengebiete Ausreisende automatisieren, wenn sie ein bestimmtes Flughafen-Gate betreten. Solche Funktionen würden natürlich ein entsprechendes Einverständnis fordern. Dies sind nur Transfermöglichkeiten von Targeting der politischen Informationsarbeit und dienen als Beispiele, wenn es eine Opt-in Option dafür gäbe. Ein zum Teil erfolgreicher Fall für Krisenkommunikation durch Opt-in war und ist die Corona Warnapp der Bundeszentrale für Gesundheitliche Aufklärung. Die Nutzung basiert auf Freiwilligkeit, die Informationen sind nicht politischer, sondern gesundheitlicher Natur, dienen aber dem Schutz der gesamten Bevölkerung und ermöglichen weitere Kommunikation.

Im nächsten Schritt des Papiers wird die Wirkung bzw. Veränderung in der Dimension der Erwartungshaltung der Bürgerinnen und Bürger gegenüber politischer Öffentlichkeit und Regierungskommunikation behandelt.

Dimension der individuellen Disposition der Bürgerinnen und Bürger: Erwartungshaltung, Selbstwirksamkeitserwartung und Bias

Dynamisierung, Monetarisierung und Aufmerksamkeitsmaximierung dominieren die Logik der Plattformen und vieler online-Dienste, die Algorithmen benutzen. Wie im Kapitel „Dimension der Vermittlung“ gezeigt, kann ein Algorithmus die Nutzerin und den Nutzer sehr gut kennen und bietet immer weiter personalisierte Inhalte an, um sie und ihn engagiert zu halten. Inzwischen ist bekannt, dass die großen „Datenkraken“ (Schröder und Schwanebeck 2019) und die Interessen der “Big Nine”[1] (Webb 2019) als auch deren kleineren Verwandten ein “race to the bottom of the brainstem” (Harris 2017) ausgelöst haben. Mit dem Begriff hat der ehemalige Google Mitarbeiter, Ethiker und Computerwissenschaftler Tristan Harris die Situation beschrieben, um zu zeigen, wie Plattformen und Apps die tiefsten, ursprünglichen Gefühle, Instinkte und Emotionen der Menschen zu berühren suchen, um Aufmerksamkeit zu generieren. Dabei basiert die Technologie und das Design auf Erkenntnissen aus der Kognitionswissenschaft und der Psychologie.

Die Konzerne zielen auf die Gefühlswelt der Nutzerinnen und Nutzer. Sie beeinflussen sowohl analoge als auch online Kommunikation sowie die Wahrnehmung beider Domänen (Hendricks und Mehlsen 2022). Dagegen stellen sich inzwischen Fragen des ethischen Designs: eine Plattform/App und der Algorithmus dahinter sollen Rücksicht nehmen und keine suchterzeugende Wirkung haben (Saura et al. 2021; Baldini et al. 2018). Die Vielfalt der „AI-Ethical Standards“ ist einerseits nicht bindend und andererseits nicht durch demokratische Prozesse geschaffen oder gar von Regierungen, sondern viel mehr von Firmen, Privatpersonen und Interessensgruppen erarbeitet worden, was eine Diskussion über die KI-Ethik ausgelöst hat (Fukuda-Parr und Gibbons 2021; Hagendorff 2020; Wallach und Asaro 2020; Bryson und Winfield 2017).

Zu den befürchteten Effekten dieses Rennens, das durch Monetarisierung entsteht, zählt die sogenannte Feedback-Loop. Der Algorithmus bietet etwas und lernt immer weiter Ähnliches zu bieten, um die Nutzerinnen und Nutzer engagiert zu halten. Dabei wird angenommen, dass die Inhalte so stark zugeschnitten werden, dass die Userin und der User nicht in Kontakt mit nicht-beliebten oder als nicht-relevant gestuften Nachrichten, Informationen, Sichtweisen, Meinungen und Profilen kommen. Damit könnten sie von vielen Teilen der breiten Öffentlichkeit abgeschnitten werden. So könnte die Erwartung entstehen, die eigene Sichtweise gespiegelt zu bekommen, während die Toleranz für andere Meinungen und die Bereitschaft zur Auseinandersetzung mit Andersdenkenden weiter abnehmen würde. Aus der Perspektive fragmentiert sich die Öffentlichkeit, sodass Deliberation und Dialog schwieriger oder gar unmöglich sein können. Die Hypothese ist, dass eine veränderte Erwartungshaltung

1 Mit „Big Nine“ beschreibt Webb die großen Techfirmen, die im gesamten Big Data Bereich den Weltmarkt dominieren und das sind: Amazon, Google, Facebook (jetzt META), Tencent, Baidu, Alibaba, Microsoft, IBM und Apple. Die Amerikanerin diskutiert sie als „KI-Götter“ mit der Möglichkeit die Welt zu verschlechtern oder zu verbessern.

an (Regierungs-)Kommunikation entsteht, die weitgehend die Sicht der einzelnen Personen auf einer bestimmten Weise ansprechen muss, um diese zu erreichen.

Algorithmisch personalisierte Nachrichten: Digitale Amplifikationen und die Veränderung der Erwartungshaltung der Bürgerinnen und Bürger

Eine Untersuchung im bereits angesprochenen Feld der algorithmisch personalisierten Nachrichten (APN), hat gezeigt, dass sie die Disposition der Bürgerinnen und Bürger verändern:

> *„Der befürchtete Zusammenhang zwischen einer intensiven Nachrichtennutzung über algorithmisch personalisierte Kanäle und polarisierten Meinungslagern existiert in Deutschland tatsächlich. Er tritt aber nur ab einem bestimmten Alter auf, das in unserer Studie bei 30 Jahren beginnt."*
> *(Schweiger et al. 2019, S. 116)*

Dementsprechend bleibt die Entwicklung für die jüngere Generation noch abzuwarten. Andere Studien hinterfragen Schweigers Ergebnisse im Spezifischen und die Wirkung der Echokammern und Filterblasen im allgemeinen (Stark et al. 2021; Rau und Stier 2019). Stark, Magnin et al. wählen sogar den herausfordernden Titel „Maßlos überschätzt", um Filterblasen und Echokammern zu beschreiben, und argumentieren, dass diese nur bedingt existieren, da Menschen insbesondere auf sozialen Medien auch flüchtig mit vielen anderen Meinungen und Nachrichten in Kontakt kommen. Das heißt jedoch nicht, dass die Gefahren für eine weitere Fragmentierung der Öffentlichkeit in der Zukunft heruntergespielt werden sollen, ganz im Gegenteil, man benötigt weitere empirische Untersuchungen (Stark et al. 2021). Ähnlich argumentieren Rau und Stier, die insbesondere auf den Mangel größerer und spezifischer empirischer Daten hinweisen (Rau und Stier 2019). Beide Studien kommen zum Schluss, dass trotz der „übertriebenen" Metapher der Echokammern und Filterblasen, *die eine* gemeinsame Öffentlichkeit für die Bürgerinnen und Bürger schon längst nicht mehr *durch eine Fernsehsendung* geschaffen werden kann und dass Teilöffentlichkeiten existieren. Wichtiger noch ist die Warnung vor Teilöffentlichkeiten und Themen, die auf sozialen Medien große Resonanz erzeugen und den Anschein hoher gesellschaftlicher Relevanz erhalten, obwohl es nicht der mehrheitlichen Realität entspricht (Papacharissi und Trevey 2018). Die Fragmentierung und Polarisierung der Öffentlichkeit ist in Europa geringer als in den USA (Silver 2021). Die Wirkung von Echokammern und Filterblasen ist noch nicht so stark ausgeprägt wie oft erwartet oder in warnenden Texten dargestellt (Ross Arguedas et al. 2022).

Geht man von den genannten Effekten der Echokammern und Filterblasen als auch von den Vermittlungsformen des vorigen Kapitels aus, so existiert inzwischen eine veränderte Erwartungshaltung auf der Empfängerseite (Bürgerin und Bürger) als auch auf der Senderseite (Regierung in unserem Fall). Hier müssen die Bürgerin und der Bürger als Nutzerinnen und Nutzer unterschiedlicher Plattformen und (sozialer) Medien erkannt werden, die einer Individualisierung und Personalisierung gewöhnt sind. So müssten sie auf der jeweiligen Plattform im richtigen Stil/Ton/Format angesprochen werden. Eine Spiegelung von Inhalten „à la copy paste" auf allen Kanälen ist dabei das Gegenteil von kanalkonformer und passender Kommunikation. Unterschiedliche Plattformen nutzen unterschiedliche Algorithmen mit ver-

schiedenen Effekten. Das bedeutet auch, dass sie eine andere Erwartung bei den Nutzerinnen und Nutzern auslösen (Santos et al. 2021). Daher gilt es auch die Form (Länge, Text, Bild, Video, Meme, Artikel oder Kurzzitat) der Regierungskommunikation auf die jeweilige Plattform anzupassen, um überhaupt die Zielgruppen erreichen zu können.

Die divergierenden Erkenntnisse schließen sich nicht aus, sie weisen genau auf die in der Einführung genannten emergenten Effekte zwischen der *Vermittlung* und *Verortung* von Kommunikation und der *Kommunikationserwartung* der Bürgerinnen und Bürger hin. Es wird sichtbar, dass es eine Lücke in der Forschung gibt, aber auch Hinweise dafür, dass in verschiedenen Gesellschaften Polarisierung auf unterschiedliche Art, Weise und Intensität stattfindet. Die kritische Analyse der Konzepte und Paper zu „Filterblasen“ und „Echokammern“ des Australiers Axel Bruns zeigt sogar eine andere Richtung auf. Nämlich, dass nicht per se soziale Medien die Treiber von Polarisierung und Populismus sind, sondern die Menschen an sich. So schließt er: “if support for illiberal and extreme political actors increases even in spite of the diverse information encountered by social media users, then the causes for this are more likely to be psychological than technological” (Bruns 2021).

Nutzerinnen und Nutzer von algorithmisch gesteuerten, dynamischen Webseiten (in der Regel soziale Medien) unterliegen der ordnenden Macht des jeweiligen Algorithmus (Katzenbach 2018). Die Kommunikation wird algorithmisch beeinflusst, wenn nicht direkt gesteuert. Menschen verändern ihr Verhalten und ihren kommunikativen Anspruch entsprechend den algorithmischen Vorgängen der jeweiligen Plattform (Schrape 2021a). Hierbei geht es um die Plattformisierung von Kommunikation (Schrape 2021b) und eine entsprechende Kommunikationserwartung. Erreichen Posts auf bestimmten sozialen Medien zu bestimmten Uhrzeiten mehr Menschen, so fangen Menschen an, um diese Uhrzeit zu posten, um ihre Nachricht zu kommunizieren. Hierbei passen Nutzerinnen und Nutzer ihr Verhalten an die jeweilige „algorithmic governance“ der Plattform an (Katzenbach und Ulbricht 2019). Ähnlich ist es bei Plattformen wie Youtube. Hier prüft der Algorithmus, welche Art von Videos mehr Engagement bei den Zuschauerinnen und Zuschauern schaffen. Dann schlägt die Plattform den Youtuberinnen und Youtubern (den Videomachenden) anhand von Big Data Analysen ihrer Videostatistiken zur Maximierung ihrer Visibilität Änderungen vor. Um weiter relevant zu bleiben, befolgen viele Videomachende diese Tipps, unabhängig davon, ob es ihrer Kreativität oder ihrer Vorstellung entspricht. An den Beispielen wird sichtbar, dass der Algorithmus sowohl „Senderseite“ als auch „Empfängerseite“, deren Verhalten und Erwartungshaltung verändert.

Bias, Voreingenommenheit und psychologische Beeinflussung

Ein Faktor, der die Polarisierung durch die Nutzung von APN verstärkt, ist die politische Wirksamkeitserwartung (political efficacy), unterschieden in „internal“, als politische Selbstwirksamkeitserwartung und „external“ – externe politische Wirksamkeitserwartung (OECD 2021). Die politische Selbstwirksamkeitserwartung ist hoch, wenn man glaubt, Einfluss auf das politische Geschehen nehmen zu können. Die externe politische Wirksamkeitserwartung ist hoch, wenn man glaubt, dass die politischen und institutionellen Handlungen z.B. der Regierung oder anderer Institutionen wirksam sind und die Bürgerinnen und Bürger zufrieden stellen (Niemi et al. 1991). Die jeweilige Wahrnehmung ist subjektiv. Schweiger et al. sehen hohes Polarisierungspotential und erhöhte APN-Nutzung bei Menschen mit ho-

her interner und niedriger externer politischer Wirksamkeitserwartung. Das bedeutet, dass diese einerseits glauben, die politischen Probleme und ihre Lösung zu kennen und durch eigenes Verhalten/Handeln die Gesellschaft verbessern zu können. Andererseits glauben sie, dass andere Akteure aus Politik, Journalismus und Wissenschaft etc. die Probleme ignorieren oder nicht verstehen und Lösungsansätze nicht kennen (low external political efficacy). „Diese Mischung aus politischer Selbstüberschätzung, Politisierung und Parteien- bzw. Politikerverdrossenheit {ist, d. A.} das Resultat und die Ursache einer intensiven Nutzung sozialer und alternativer Medien“ (Schweiger et al. 2019, S. 117). Schweiger et al. erkennen, dass das Problem vielschichtiger und nicht allein durch APN-Nutzung erklärbar ist. Jedoch gehen sie von einer Spirale aus, in der die Nutzenden aktiv nach alternativen oder kritischen Stimmen zu bestimmten Themen suchen, sich mit diesen verbinden und durch die APN sich immer weiter mit ähnlich gesinnten Nachrichten, Personen und Meinungen umgeben.

Was Schweiger und seine Kolleginnen und Kollegen nicht ansprechen, aber ausschlaggebend ist, ist ein Befund, der die Diskussion um „Bias“ und Voreingenommenheit revolutioniert hat, und zwar der „myside bias“. Dieser existiert online und offline, wird aber in der und durch die Online-Welt verstärkt. „Myside bias“ wurde durch den viel zitierten und inzwischen emeritierten Psychologieprofessor der Toronto University E. Stanovich bekannt (Stanovich und West 2007) und wird oft als Unterkategorie des „confirmation bias” (Bestätigungsvoreingenommenheit) verstanden, geht aber weiter und ist präziser. In einem analytischen Paper von Wang und Jeon, das die Literatur zusammenfasst, wird myside bias wie folgt definiert:

> *“Myside bias is a common type of cognitive bias where people process information in a manner biased toward their own prior beliefs, opinions, and attitudes. It occurs when people seek, interpret, and remember evidence to confirm their preexisting beliefs and to refute opinions different from their own. (...) Myside bias reflects a deficiency in critical or rational thinking rather than general cognitive abilities or intelligence.”*
> *(Wang und Jeon 2020)*

Dieser Bias existiert bei den meisten Personen, selbst bei den von der Gesellschaft wahrgenommenen „kognitiven Eliten“, unabhängig vom Bildungs- und Intelligenzgrad einer Person (Stanovich et al. 2013). Nach Stanovich sind es nicht nur Desinformationen, Halbwahrheiten und ihre Amplifikation durch soziale Medien und Algorithmen, die Meinungsbilder und öffentliche Kommunikation beeinflussen, sondern ein tiefliegender „myside bias“ (Stanovich 2021). Stanovichs empirische Studien zeigen, dass das Phänomen des „myside bias“ einen starken Einfluss auf die Wahrnehmung der Argumente anderer, als auch die nicht-Wahrnehmung der eigenen Voreingenommenheiten besitzt. Seine Analyse bestätigt die Befürchtung einer Fragmentierung der Öffentlichkeit durch die Abnahme der Bereitschaft, sich mit anderen Meinungen und Argumenten auseinander zu setzen mit großer Genauigkeit, die auf mehrere Jahrzehnte psychologischer Experimente und Untersuchungen beruht. Auch das „neue“ Mistrauen gegenüber wissenschaftlichen Erkenntnissen erklärt er durch myside bias (Stanovich 2021, S. 74–80). Die Sichtbarkeit dieser „Antiwissenschaftlichkeit“ hat sich nur vergrößert, weil sich Gruppen über soziale Medien besser vernetzen und sich organisieren können. Dabei beschreibt er gesellschaftliche und (partei-)politische Auseinandersetzungen in den USA als „tribal“ – als stammesähnliches Verhalten. Dass algorithmische Personalisierung dabei uralte Instinkte bedient, ist unstrittig (Stanovich 2021, S. 114–118). Er rät insbesondere zur Überprüfung des eigenen myside bias durch kontrastierende Einstellungen zu gleichen oder ähnlichen Themen, die von „eigenen“ oder „fremden“ Personen kommen. So würde das bedeuten, dass man ein Argument von „einer gleichgesinnten“

Person annimmt, aber das gleiche Argument von einer „fremden“ Person nicht annimmt. Zudem unterstützt Stanovich die Ansicht, dass insbesondere identitätspolitische Sprache als Kommunikationsstil vermieden werden sollte, da dieser auf „myside bias“ beruht. Bisher existieren keine Analysen, die den Zusammenhang zwischen dem Einsatz von künstlicher Intelligenz in der öffentlichen Kommunikation und den „myside bias“ näher beleuchten.

Die Ergebnisse von Schweiger et al. mit den Erkenntnissen von Stanovich zusammen zu führen, um den Zusammenhang zwischen APN, myside bias und political efficacy zu zeigen, würde weitere Einblicke in die genannte Dynamik bieten. Schließlich beruhen die Effekte auf dem Konzept der kognitiven Dissonanztheorie (von Festinger in der Psychologie seit 1957 vertreten), da ein Mensch das unangenehme Gefühl von entgegenwirkenden Aktionen/Emotionen nicht lange aushalten kann und den Einklang zwischen Emotion, Kognition und Gedankenwelt sucht (Festinger 2009). So wird die kognitive Dissonanztheorie insbesondere im deutschen Raum zusammengefasst, um den Diskurs im Bereich der politischen Kommunikation zu avancieren (Jaster und Lanius 2020). Dissonanz ist insbesondere für die Öffentlichkeit und die öffentliche Kommunikation charakteristisch (Pfetsch et al. 2018) – insbesondere vor dem Hintergrund einer hybriden Öffentlichkeit, wo offline, online und vernetzt kommuniziert wird und Nachrichten und Neuigkeiten verbreitet werden. Die Berliner Kommunikationswissenschaftlerin und ihre Kollegen sehen im Konzept der dissonanten Öffentlichkeiten einen theoretischen Ansatz, der zur Analyse von Diskursveränderungen und dahinter liegende Dynamiken beitragen kann (Pfetsch et al. 2018). Sie weisen darauf hin, dass durch den Einsatz von Technologien der künstlichen Intelligenz Dissonanzen in der Öffentlichkeit und ihre Komplexität zunehmen werden.

Zurück zu Stanovich, der klarstellt, dass die zunehmende Komplexität und Dissonanz für den „average citizen“ unübersichtlicher wird. Das heißt, Menschen werden weiterhin Expertise und damit auch Journalistinnen und Journalisten benötigen, um zu erfahren, was wahr ist. Doch wie man die Expertinnen und Experten wählt, denen man folgen möchte, bleibt eine Frage des myside bias (Stanovich 2021, S. 118–119). Folgt man dem Argument ist es nicht der Algorithmus per se, der die Kommunikationserwartung der Nutzerinnen und Nutzer ändert, sondern die Natur des Menschen an sich, die vom Algorithmus erkannt, überreizt und ausgespielt wird. Das führt zurück zum Argument des ethical designs und der ethischen Aspekte im Einsatz von KI Verfahren (Baldini et al. 2018; Fukuda-Parr und Gibbons 2021; Buxmann und Schmidt 2021; Zeilinger 2019).

Hieraus ist zu erkennen, dass Menschen sich weiterhin auf „Expertise“ von Expertinnen und Experten verlassen, die Wissen und Information vermitteln. Problematisch ist es allerdings, wenn die eigenen Expertinnen und Experten oder gar die eigene Regierung Falschinformationen verbreitet oder direkte wissenschaftliche Erkenntnisse unterminiert (Rußmann 2018). Ein Beispiel für diese Praxis wäre der ehemalige US-Präsident Donald Trump während der Corona-Pandemie. Hier manifestieren sich zwei gegenläufige Erwartungen: a) die Erwartung, dass Regierungskommunikation wahr ist und b), dass Regierungskommunikation als wahr angenommen wird, wenn sie der eigenen Vorstellung entspricht. Das zweite bestätigt den myside bias. In dem Fall sind aber nicht Algorithmen, sondern Menschen am Zug. Die kommunikative Reichweite dieser Menschen und ihrer (angebliche) Expertise sind Ergebnisse der Digitalisierung, die Vowe unter „Dynamisierung“ und „Pluralisierung“ subsummiert (Vowe 2020). Plakativ gesagt: es ist nicht der Algorithmus per se, der Dysfunktionalität erzeugt, sondern Menschen, die ihn radikalisierend nutzen und eigene Voreingenommenheit reproduzieren. So heißt es in einem Artikel von Foreign Affairs:

„Although leaks from inside Facebook suggest that algorithms can indeed guide people toward radical content, recent research indicates that the algorithms don't themselves change what people are looking for. People who search for extreme YouTube videos are likely to be guided toward more of what they want, but people who aren't already interested in dangerous content are unlikely to follow the algorithms' recommendations. If feedback in democratic societies were to become increasingly deranged, machine learning would not be entirely at fault; it would only have lent a helping hand." (Farrell et al. 2022)

Daher kann man vermuten, warum in größeren sozialen Medien (wie Twitter), die Heterogenität und Konfrontation erzeugen, extreme Meinungen unterrepräsentiert sind. Diese Stimmen verlassen den öffentlichen sozialen Raum und sammeln sich in geschlossenen Chat- und Medien Gruppen. Dort bestätigen sie sich gegenseitig. Die individuelle Erwartung, die eigene Meinung gespiegelt zu bekommen, kann zum Rückzug aus dynamischen, algorithmisch gesteuerten Medien führen und in semi-geschlossene Teilöffentlichkeiten münden. Studien zu Teilöffentlichkeiten im digitalen, analogen und hybriden öffentlichen Raum belegen deren Existenz, ihre emotionale und emotionalisierende als auch weltumspannende Wirkung (siehe dafür die Werke von Papacharissi und Trevey; Papacharissi (2018; 2015)). Eine tiefere Studie zu dieser „Rückzugs-Annahme" aus mainstream social media von Andersgesinnten existiert noch nicht. Sie basiert auf der Beobachtung der Telegram-Kanäle von Qanon, Querdenkern und Reichsbürgern[2]. Hier existiert eine weitere Forschungslücke – herauszufinden, ob das was, in beliebten Medien abwesend ist, sich auf anderen Plattformen konzentriert, wo keine Recommender-Algorithmen genutzt werden (wie auf Telegram).

Mit Bots in die neue Epoche der Regierungskommunikation?

Eine weiterer aufkommender Trend ist der Einsatz von KI-Technologien in der öffentlichen Verwaltung, um die Kommunikation zwischen Bürger, Regierung und Staat zu verbessern (Hill 2020). Das bedeutet implizit eine Öffnung im (öffentlichen) Datenbestand (Lucke und Gollasch 2022b), als auch die Entstehung der digitalen Kommunikation mit/zu der Verwaltung (Lucke 2020). Die veränderte Erwartungshaltung gegenüber der öffentlichen Verwaltung wurde in manchen Ländern bereits angegangen. Für Deutschland benötigt man weitere empirische Daten, um die hiesige Erwartungshaltung genauer bestimmen zu können. Vor dem Hintergrund, dass durch die Digitalisierung sich die gesamte kommunikative Erwartungshaltung der Bürgerinnen und Bürger verändert (Raupp et al. 2018b), ist davon auszugehen, dass eine durch KI-Systeme verbesserte und aufgewertete Kommunikation zwischen Verwaltung und Bürgerinnen und Bürgern einen positiven Effekt hat. Die Lehren daraus können dem Transfer dienen, und könnten auch zur Verbesserung der Informationsarbeit der Regierung genutzt werden.

Im Rahmen einer „digitalisierten Demokratie" wird davon ausgegangen, dass Anwendungen der künstlichen Intelligenz den Weg in die politischen Prozesse und in die politische und Regierungskommunikation finden werden (Borucki et al. 2020b, 2020c). Borucki und ihre Kolleginnen und Kollegen sehen hierbei weiteren Forschungsbedarf in der Analyse der Anwendungsbereiche, als auch deren methodologischen Untersuchung und ihrer Effekte. Ein

2 Diese Annahme wurde im Gespräch vom Datenwissenschaftler und Co-direktor des Center for Monitoring, Analysis, and Strategy (CeMAS) geäußert (*https://cemas.io/about-cemas/*). Sie wurde noch nicht empirisch und theoretisch untersucht, sondern stellt den Eindruck des CeMAS-Redners dar.

im Text bereits genanntes Werk von einem griechischen Forschungsteam hat gezeigt, dass KI-gestützte Kommunikation im Bereich der öffentlichen Verwaltung einen positiven Effekt für die Bürgerinnen und Bürger hatte (Androutsopoulou et al. 2019). Hier ging es um einen KI-gesteuerten Chatbot der Verwaltung, um Anfragen zu klären, zu beantworten und ihre Lösung bzw. Antragsstellung einzuleiten. Das Forschungsteam hat dafür mehrere KI-Verfahren kombiniert, darunter Data und Literatur-Mining Verfahren, NLP-gestützte Chatbot- und Recommendersysteme und hat eine positive Transformation der „citizen-government" Kommunikation festgestellt. Die Studie trennt nicht scharf zwischen Regierung und Verwaltung, weil das Finanzministerium als Teil der Regierung beteiligt war, aber auch, weil die Forschenden mit dem politischen Verständnis des „government services" arbeiten. In Deutschland wäre die Trennung zwischen Regierung und Verwaltung klarer, aber die Transferleistung wäre ähnlich. Statt einer statischen Webseite könnte ein „Bundesministeriums"-Chatbot Informationen auf Anfrage bieten. Durch das Einverständnis der Nutzenden zur Personalisierung könnte sich der Chatbot anpassen, z.B. um altersgerecht oder regionsabhängig zu kommunizieren. So könnte eine ältere Person bei der Frage „warum sind wir in Mali" die gleiche Antwort erhalten wie eine jüngere, wobei die Tonalität oder die Sprachfarbe dem Alter angepasst wären. Ein Pilotprojekt in dem Rahmen wäre äußerst interessant.

Was angesichts der Chatbots noch nicht erforscht aber erprobt wurde, ist deren Anwendung in der parteipolitischen Kommunikation. So hat zum Beispiel die SPD ihre Wahlkampagne 2021 revolutioniert, indem sie einen Wahlkampfkanal mit einem semi-interaktiven Chatbot zu Informations-, Organisations-, und Mobilisierungszwecken schuf. Hier wurde zwar kein Recommender-System, kein Data-Mining und kein lernender Algorithmus angewendet, sondern ein semi-interaktiver Chatbot, der lediglich als Informationsabrufsystem fungierte. Leider existieren keine Analysen dazu, aber ein Versuch und ein Transfer, ein ähnliches Tool für Informationsarbeit aufzubauen, wäre denkbar. Es gilt zwischen Regierungs- und Parteikommunikation zu unterscheiden, doch die Art und Weise der Kommunikation wird sich verändern, wenn mehrere Parteien auf automatisierte und semipersonalisierte Informationskanäle zurückgreifen.

Zurzeit existieren relevante nicht-interaktive Kanäle auf Telegram. Das Bundesministerium für Gesundheit verschickt darüber Informationen zur Covid-Pandemie, die Bundeswehr Neuigkeiten zur Bundeswehr. Es gibt auch zwei Landesregierungskanäle von Baden-Württemberg und Hessen, die Allgemeininformationen versenden. Diese nutzen keine KI-Verfahren und fungieren als einseitige Sender von Nachrichten. Der Transfer in eine Anwendung, wie Androutsopoulou und ihre Kolleginnen und Kollegen sie in Griechenland erprobt haben, wäre denkbar und möglich. Dafür müsste man nicht Telegram oder eine marktübliche App benutzen, sondern könnte nach Überprüfung die eigenen Webseiten der Ministerien, als auch Kanäle von Regierungspersonal (Profile von gewählten Vertreterinnen und Vertretern) oder Presse- und Informationsstellen nutzen.

Dimension des Deliberativen Prozesses: wie manifestieren sich KI-Herausforderungen und Chancen für die Demokratie

Der Einsatz von Technologien der künstlichen Intelligenz ist nicht neutral und nicht frei von Voreingenommenheit und Verzerrungen. Wie sich die Technologie auswirkt, hängt von der jeweiligen Art und Weise ihrer Funktion, ihrer Lernform und den verfolgten Zielen ab. Beim maschinellen Lernen kann ein Bias entstehen, wenn die Maschine aus Daten lernt, die selbst einen Bias aufweisen. Das führt dazu, dass der Algorithmus danach keine „guten" Ergebnisse bietet, wenn es z.B. um Sprach- oder Bilderkennung geht. Der Bias in Prozessen des maschinellen Lernens wurde schon früh erkannt (Mooney 1996) und ist inzwischen als „Algorithmic Bias" ein fester Bestandteil der KI-Literatur (Hajian et al. 2016). Dem kann man begegnen, indem mehrere Lernalgorithmen gleichzeitig genutzt und „lernen", mehrere relevante Datensätze genutzt werden und auch, indem die Ergebnisse von Menschen überprüft werden, um den Algorithmus zu verfeinern (Mehrabi et al. 2021). „Human in the loop" ist für eine menschenzentrierte KI scheinbar noch notwendig.

In diesem Abschnitt geht der Bericht auf Themen ein, die den deliberativen Prozess und entsprechend die Regierungskommunikation beeinflussen.

KI-Gefahr und Chance: (Social) Bots und Recommender Systeme

Ferrara und seine Kollegen aus der Indiana University haben bereits 2016 den Aufstieg der Social Bots[3] beschrieben, deren Nützlichkeit und vor allem Gefahr aufgezeigt (Ferrara et al. 2016). Die Idee ist, dass Social Bots, also automatisierte Software, in sozialen Medien wie ein Mensch auftreten können. Sie können Inhalt generieren, an Diskussionen teilnehmen und diese wahrscheinlich beeinflussen. Ein Bot kann Themen platzieren und „chatter" (Geräusch) darüber kreieren, somit Relevanz und Sichtbarkeit schaffen – die nicht der Realität entsprechen. Andere Beispiele beinhalten die Beeinflussung der Börse oder den Wert bestimmter Aktien und Finanzprodukte durch bot-generierten Online-Chatter. Ferrara et al. schlagen Methoden zur Bot-Erkennung vor und räumen ein, dass noch keine von diesen Methoden mit absoluter Sicherheit Bots erkennen kann, zumindest nicht ohne massive Kostensteigerung oder durch Übergriffe in die privaten Daten von social Media Accounts. Dies wurde von weiteren Teams bestätigt (Cresci et al. 2017). Wichtiger noch als die Funktion der social Bots selbst, ist der „Puppet Master", der Puppenspieler, der die Bots kontrolliert, programmiert und einsetzt (Ferrara et al. 2016). Die Anzahl der social Bots im Netz ist nicht wirklich quantifizierbar; keiner weiß, wie viele genutzt werden, aber sie stellen definitiv eine Herausforderung (sogar eine Gefahr) für Plattformen und soziale Medien dar, als auch für das Vertrauen in Medien und im online-Dialog insgesamt dar.

3 In der relevanten „Bot-Literatur" wird oft zwischen Chatbots, Social Bots, Spam Bots etc. unterschieden. Die Unterscheidung beruht auf die präzise Anwendung, ob es nur um Chatten, um den Einsatz in sozialen Medien oder Spam geht. Die Technologie dahinter bleibt weitgehend ähnlich und fließt in den Anwendungsbereichen über und ein.

Wichtiger noch als ein social Bot ist ein Netzwerk von Bots. So können sie als Gruppe koordiniert agieren, Anweisungen folgen und ausführen (Zhang et al. 2018). Zhang et al. zeigen, wie Botnetzwerke in der Lage sind, massiven Einfluss auf Twitter zu erhalten. Dafür haben sie selbst ein solches Netzwerk geschaffen und zur Spamverteilung eingesetzt, um die Ergebnisse auszuwerten. Durch Algorithmen und Anweisungen verfolgen die Botnetzwerke bestimmte Ziele und können diese auch erreichen. Das früherkannte Problem scheint auch für dieses Forschungsteam weiter ungelöst: es wird immer schwieriger einen Bot von einem Menschen zu unterscheiden.

Cresci et al. vertreten die Ansicht, dass die Bekämpfung von Social Bots ein „arms' race" ist, das gewonnen werden muss, um deren Einsatz unter Kontrolle zu bringen (Cresci et al. 2017). Neben dem Mehrwert von Zhangs et al. Algorithmendarstellung ist Crescis et al. Datenset zur Analyse von Twitteraktivitäten einschlägig. Beide unterstreichen die Nützlichkeit von großen Datensätzen, als auch die wissenschaftliche Darlegung der eingesetzten Methoden. Durch die Analyse von massiven Datensätzen können Bots und Botnetzwerke besser erkannt werden, dafür sind bestimmte KI-Anwendung am besten geeignet. Allerdings bleibt bis heute das Problem der präzisen Boterkennung ungelöst (Martini et al. 2021). Die existierenden Ansätze beziehen sich hauptsächlich auf Twitter, da Twitter besonders für die US-Politik großen Stellenwert besitzt. Einflussreiche Wissenschaftlerinnen und Wissenschaftler aus den USA schlagen Tools zur Boterkennung vor. Sie raten zur Auseinandersetzung aller Bürgerinnen und Bürger mit diesen Tools (Yang et al. 2022). Yang und Ferrara argumentieren, dass Politik-, Sozial- und Kommunikationswissenschaftlerinnen und -wissenschaftler, die sich mit KI, Digitalität und sozialen Netzwerken auseinandersetzen und diese erforschen, stets zumindest die Basis technischen Wissens besitzen sollten. Dafür haben sie auch den Kurzführer zum „Social bot practicum for computational social scientists" herausgegeben (Yang et al. 2022). Hier raten Sie zur Nutzung des Werkzeugs „Botometer", das inzwischen präzise Einschätzungen von „bot-ähnlichem" Verhalten abgeben kann. Doch eine wichtige Limitation ist, dass die Spracherkennung des Tools durch die gefütterten Daten trainiert wird. Damit leiden Sprachen, die auf Twitter weniger Gewicht haben, unter schlechterer Erkennungsqualität. Eine weitere Limitation ist, dass Botometer nur für Twitter geeignet ist.

In Deutschland haben sich mehrere Forscherinnen und Forscher mit der Thematik der Bots und ihren Einfluss auf Wahlen beschäftigt. Die Erkenntnisse sind eher ernüchternd und anders als die Ergebnisse in den USA. So kommen die Forscherin Ulrike Klinger und ihr Kollege zum Ergebnis, dass Bots auf Twitter durch die Analyse von mehreren, längerfristigen Daten in Kombination mit Metadaten erkennbar sind (Keller und Klinger 2021). D.h. KI-Verfahren können zur Erkennung von Bots und Botnetzwerken beitragen, weil sie in der Lage sind, Big Data Analysen durchzuführen. Für einzelne Menschen sind Bots bei gängigen Twitter-Interaktionen oft nicht erkennbar (Martini et al. 2021). Die Plattformen haben eigene Methoden und KI-Tools entwickelt, um Bots zu erkennen – so z.B. auch die Technologie „Deep Entity Classification" von Facebook, die eingesetzt wird, um Fakeprofile und Bots zu entfernen. Die Datenhoheit bleibt bei den Plattformen. „Der eingeschränkte Datenzugang bedeutet letztlich auch, dass eine zentrale Funktion von Öffentlichkeit derzeit beschnitten ist: die Selbst-Beobachtung der Gesellschaft" (Klinger 2019). Der deutsche Informationswissenschaftler Florian Gallwitz und der Datenjournalist Michael Kreil haben die Studien von Botometer in Kombination mit anderen Verfahren ausgewertet. Sie kommen zum Schluss, dass zu viele falsch-positiv Beispiele genutzt werden. Wichtiger noch, Bots und Botnetzwerke existieren, jedoch gibt es keine einschlägigen Nachweise für erfolgreiche politische social Bots. Die erfolgreichsten waren lediglich menschgesteuerte Profile, oder wenn, dann semiautomati-

sierte Bots, die als „social bots“ falsch positiv identifiziert wurden (Gallwitz und Kreil 2021, 2022). Das wirft auch die Frage auf, wie man Akteure in einem Diskurs bestimmen kann, unabhängig davon, ob es sich um einen Menschen, oder ein (semi-)automatisiertes Profil handelt (Dreesen und Krasselt 2022). Die Meinungen zur Effektivität von social Bots sind allerdings nicht ganz einheitlich, da neue Erkenntnisse aus der Bundestagswahl 2021 Gefahren für die kommende Zeit aufzeigen (Sorge 2022). Der Technikfolgenabschätzungsdienst des Bundestages hat sich bereits 2017 mit dem Thema auseinandergesetzt und erkannt, dass Boterkennung noch nicht genug entwickelt war, aber dass Bots das Potential zur „Meinungsmache“ haben können (Kind et al. 2017). Damals sprach das Forschungsteam eine langfristige Handlungsempfehlung aus, und zwar die Stärkung der Medienkompetenz der Bürgerinnen und Bürger ab dem Schulalter.

Die bisherigen einschlägigen Ansätze lassen große Plattformen wie Facebook, Instagram, TikTok, als auch Chatplattformen wie Whatsapp, Telegram, Snapchat und viele weitere, als auch die Kommentarsektionen auf anderen Plattformen unberührt und unerforscht. Im Endeffekt bleibt Boterkennung Angelegenheit der jeweiligen Plattform. Ausnahmen sind Untersuchungen, in denen die Forscherinnen und Forscher selbst Data-Mining von öffentlichen Social Media Daten betreiben (Papakyriakopoulos et al. 2017). Inzwischen werden Bots als ausschlaggebend für und im politischen Diskurs gesehen, da sie eine Diskussion beeinflussen und Content weiterverbreiten können (Hagen et al. 2022). Ähnlich steht es mit der Möglichkeit, Graswurzelbewegungen durch Bots vorzutäuschen (auf Englisch „astroturfing“ genannt) (García-Orosa 2021).

Denken wir an neue NLP-Anwendungen mit KI-Systemen, die die Möglichkeit der kreativen, argumentativen und emotionalen Textproduktion haben, wie z.B. GPT3 von OpenAI, so wird Bot-Erkennung immer schwieriger und ihr Einsatz immer attraktiver. Zudem öffnen sich neue Möglichkeiten der Maschinenkommunikation, die über die politische Öffentlichkeit hinaus bis hin in die Kunst- und Literaturproduktion reichen können.

Eine Art Regulierung zur Markierung von Bots, kommunikativen KI-Anwendungen und nichtmenschlichen Kommunikatoren auf Plattformen wird immer dringender. Auch dazu lautet die Lösung „KI“. In den Worten Marc Zuckerbergs während seiner Kongressbefragung „AI will fix this“ (Katzenbach 2021). Um Content zu moderieren und Bots zu erkennen wird KI angewendet. Inwieweit das erfolgreich sein wird ohne Einschnitte in Meinungsfreiheit oder einer klaren Fehlerkennung ist fragwürdig. Inhaltsmoderation ist nicht nur eine Frage der Technologie, sondern der Meinungsfreiheit, deren Schicksal inzwischen in den Händen der Plattformen liegt. Die Entwicklung hin zu einer „Private Speech Governance“ lässt viel Raum für Regulierung offen (Wong und Floridi 2022; Robyn Caplan 2021; Balkin 2017). Hier sind Plattformen Hauptakteure und Regulierungsgegenstand. Ein Dialog mit den Plattformen wäre nützlich, wenn nicht imperativ, um Freiheiten aufrechtzuerhalten, als auch um kommunikationsrelevantes Know-how zu bekommen. Regulierung lässt die Frage unbeantwortet: wie erkennen wir eine kommunizierende Maschine, die wie ein Mensch kommuniziert und wenn, wann macht das einen Unterschied?

Für eine Veränderung in der Wahrnehmung spricht das Beispiel der Anwendung LaMDA[4] von Google: hier behauptete die Maschine von sich selbst, Empfindung, Gefühle, Ängste und Bewusstsein zu haben. Dies hat viele, darunter auch den zuständigen Ingenieur, über-

4 LaMDA steht für Language Model for Dialogue Applications und ist eine KI-Technologie, die menschenähnlich kommunizieren kann.

zeugt (Lemoine 2022). Die Herausforderung bleibt bestehen: je besser man lernt Bots zu erkennen, desto besser lernen Bots menschähnlicher zu werden, um nicht von Menschen unterscheidbar zu sein (Oberer et al. 2019). Weitere Herausforderungen stellen die „Puppenmeister“, die Strippenzieher der Bots dar, da Menschen hinter deren Einsatz agieren. Es ist aber auch vorstellbar, Bots in Redaktionen einzusetzen und damit auch journalistische Arbeit und öffentliche Kommunikation insgesamt positiv zu gestalten (Kadel 2022).

Für die Deliberation haben social Bots, die am politischen Dialog teilnehmen, mehrere Nachteile: A) sie haben keine Authentizität, denn sie sind keine Menschen, sondern semigesteuerte Programme. B) Sie verzerren den kommunikativen Konsens, da sie keine Akteure mit einem genuinen politischen Interesse sind. C) Sie besitzen (noch) nicht die Fähigkeit, ihre Meinung und Präferenz im Rahmen einer argumentativen Auseinandersetzung zu ändern und haben damit keine Möglichkeit der Einsicht. Wichtig für die Deliberation ist insbesondere der erste und zweite Punkt, weil sie die Öffentlichkeit verzerren. Bots können in so großen Zahlen „Lärm“ produzieren, dass sie Themen, die keine reale Mehrheit haben, auf sozialen Medien nach oben pushen. Sie können auch den gegenseitigen Effekt erzielen, nämlich durch „Überschwemmung“ oder „Überflutung“, um Themen, die einer größeren Gruppe wichtig sind, zu übertönen und deren Sichtbarkeit und Relevanzzuschreibung zu verhindern (bildlich heißt das Konzept „Information-Flooding“). Allerdings können Bots auch eingesetzt werden, um einen Dialog zu moderieren oder um Journalisten bei ihrer redaktionellen Arbeit zu unterstützen (Neuberger 2017; Meier et al. 2021) .

Für den Einsatz von Bots in der Regierungskommunikation ist das jeweilige politische System, sei es autoritär, liberal, demokratisch etc. ausschlaggebend. Es gibt Länder, deren Regierungen explizit auf Flooding durch Bots setzen, um Aktivismus und journalistische Arbeit oder soziale Bewegungen zu marginalisieren oder zu unterdrücken (Han 2015). Sie setzen online-Trolls (Menschen) und Bots ein, um Angst zu schüren und Stimmung zu machen (Ergin Bulut und Erdem Yörük 2017; Saka 2018). Andere schaffen Einheiten, die gezielt Themen und Diskussionen platzieren, ändern oder steuern wollen, auch durch den Einsatz von Botnetzwerken (Bradshaw und Howard 2017) und das auch in der Außenpolitikkommunikation (Woolley 2016). Während solche Taktiken oft von autokratischen Regierungen im inneren eingesetzt werden, stehen demokratische Regierungen vor der Herausforderung, im „dunklen“ zu kommunizieren. So kann es sein, dass eine Politikerin, ein Politiker oder ein Regierungsamt auf sozialen Medien Hassnachrichten, schlechte Kritik oder gar „normale“ Kommunikation erhalten, welche zuerst auf ihre Authentizität geprüft werden muss. Das setzt die Kommunikationsstrategie unter Druck, da die Geschwindigkeit der sozialen Medien enorm ist. Doch NLP-KI und Sentiment-Analyse-KI können zur Klarheit beitragen (Kanakaraj und Guddeti 2015; Tetsuya Nasukawa und Jeonghee Yi 2003), benötigen aber auch menschlichen Input. Auch hier hilft die „human in the loop“ – Methode, also algorithmische Lernprozesse integriert mit dem Feedback und der Aufsicht von Menschen.

Ein weiteres Problem für die deliberative Demokratie stellen die Effekte der Recommendersysteme und der damit einhergehenden Bereitschaft (oder derer Abwesenheit), sich mit anderen Meinungen objektiv, argumentativ und offen auseinander zu setzen, dar. Tribalism, stammesähnliches Verhalten ist nicht neu und nicht nur durch Algorithmen beeinflusst. Computer vermittelte Kommunikation hat von Anfang an die online-Diskussionskultur verändert (Postmes et al. 1998). Aber KI und Recommendersysteme können Deliberation erleichtern. Hier besteht im Rahmen einer eingegrenzten Nutzerinnen- und Nutzerzahl die Möglichkeit der algorithmischen Moderation einer Diskussion – sodass Gewichtungen im

Dialog und in der Wahrnehmung als auch in der Visualisierung sichtbar werden können (Alnemr 2021; Kim et al. 2021; Helbig 2018). Solche Verfahren werden nicht alle Nutzerinnen und Nutzer zum vernünftigen Dialog zwingen, aber die Rolle der konstruktiven „Dialogbefähigung“ (facilitator) übernehmen können. Dafür ist die Aufsicht von Menschen über den Algorithmus nötig (Human in/next to the Loop). Solche Verfahren können auch von Regierungen in Konsultationsprozessen eingesetzt werden oder in Beratungsrunden auf der regional- oder lokalpolitischen Ebene.

Die hier aufgezeigten Trends einer gespaltenen Öffentlichkeit in mehrere Teilöffentlichkeiten sind eine Herausforderung für Deliberation, da theoretisch alle Bürgerinnen und Bürger für bestimmte Fragen zur Meinungs- und Willensbildung am demokratischen Prozess beteiligt sein sollten. Hierfür werden bisher zwei Lösungen gesehen: Die erste besteht darin, Algorithmen und ihre Governance offen zu legen und entsprechend zu verbessern, sodass ein Algorithmus trotz Personalisierung nie gänzlich die Nutzenden von bestimmten Nachrichten abschottet. Das ist einerseits durch Regulierung (von staatlicher Seite aus) (Cuéllar und Huq 2022) und andererseits durch ethische Regelwerke, die Nutzenden, Techfirmen und Staaten im Rahmen des technogesellschaftlichen Wandels bestimmen, möglich (Wallach und Asaro 2020; Fukuda-Parr und Gibbons 2021). Diese Erwartung ist schwer zu erfüllen, denn der goldene Schnitt zwischen Fakenews- und Hatespeech-Bereinigung und Meinungsfreiheit birgt viele Herausforderungen, welche auch viel Raum zur Diskursanalyse und zur empirischen Forschung bieten (Gorwa et al. 2020). Für Regierungskommunikation bedeutet das, dass sie ein technisches und gesellschaftliches Verständnis von der kommunikationsregulierenden Kraft der unterschiedlichen Algorithmen besitzen muss, um sie entsprechend zu durchdringen und Teilöffentlichkeiten zur erreichen oder gar zu verbinden. Ein runder Tisch, an dem Informationsarbeiterinnen und -arbeiter auf Regierungsseite und Plattformen zusammenkommen und im Dialog ein besseres Verständnis von und für beide Seiten entwickeln, ist zu empfehlen.

Es ist davon auszugehen, dass nicht alle KI-gestützten Verfahren Deliberation fördern: geht man zu weit mit der algorithmischen Moderation, kann es zu Einschnitten in der Meinungsfreiheit kommen (Katzenbach 2022). Nichtsdestotrotz können Algorithmen immer besser und immer genauer Falsch- und Desinformation erkennen, was aber dann die weitere Verfeinerung solcher malignen Verfahren zur Folge hat. Ähnlich verhält es sich mit der Erkennung von Bots, als auch mit der Erkennung von deep-fake Videos. Es gibt Methoden, Deepfakes zu erkennen und zwar durch Bildfrequenzanalyseanwendungen (Frank et al. 2020) oder durch biometrische Analysen (Agarwal et al. 2020). Doch auch hier ist davon auszugehen, dass Deepfakes besser und ihre Erkennung schwieriger werden. Daher muss für Regierungskommunikation zu einem gewissen Grad die Authentizität der Kommunikation innerhalb der „Agora“ gewährleistet werden und dafür werden human in the loop Verfahren und ein gewisses Grad an Identifizierbarkeit nötig bleiben.

Limitationen, Ausblick und Forschungsbedarf

Dieser Report folgt dem vorgestellten analytischen Raster und schafft in dem Bereich einen präzisen und konzisen Überblick von relevanter Literatur. Es werden Anwendungen und Beispiele vorgebracht, die die Tangierung von Künstlicher Intelligenz mit Regierungskommunikation aufzeigen. Das Analyseraster schafft einen analytischen Mehrwert, da dieses als Vorlage für weitere Untersuchungen verwendet, ergänzt und erweitert werden kann. Neben dem inhaltlichen Nutzen hat der Report auch gewisse Limitationen, die wichtig zu benennen sind, weil sie den Stand der Forschung im deutschsprachigen Raum aus einer anderen Perspektive aufzeigen: die Diskursanalyse aus kommunikationswissenschaftlicher Sicht.

Vom mechanischen Riesen „Thalos“ der altgriechischen Mythologie bis hin zum japanischen Manga „Astro Boy“ der 1950er Jahre sind KI-Narrative und gesellschaftliche Bilder von „intelligenten Maschinen“ allgegenwertig (Cave et al. 2018). Sie durchdringen nicht nur Politik und Gesellschaft, sondern auch Einzelpersonen und Gruppen. Ihre Ideen und Vorstellungen von Gesellschaft und Zukunft demonstrieren, wie wir uns künstliche Intelligenz vorstellen und was wir von ihr erwarten, hoffen oder befürchten (Cave et al. 2020). Diese Erwartungen sind von Gesellschaft zu Gesellschaft und Land zu Land unterschiedlich, was sich in nationalen KI-Strategien niederschlägt (Bareis und Katzenbach 2021; Katzenbach und Bächle 2019). Das Gleiche demonstriert die Erwartung, dass große Techfirmen in der Lage sein werden, durch KI Meinungsfreiheit und Desinformation erfolgreich zu moderieren (Katzenbach 2021; Bächle 2015; Ulbricht 2020a). Ebenso wirken Bilder, wie das des „Terminator“ und ähnlicher tödlicher KI (Cave et al. 2019), die aktuelle Politik und Regulierungsanstrengungen beeinflussen, obwohl die existierenden KI-Verfahren keine Möglichkeit einer „Terminator“ Maschine haben (Thimm und Bächle 2019a, 2019b). Solche Narrativanalysen bieten die Möglichkeit, die Regierungskommunikation aus inhaltlicher und langangelegter Sicht zu veranschaulichen, ihre Wahrnehmung in der Gesellschaft zu untersuchen und verdient einen eigenen weiterführenden Bericht.

Der vorliegende Report bietet auch **einen Ausblick und zeigt Möglichkeiten zur weiteren Erforschung** und **Analyse auf**. Hier sind die Forschungsmöglichkeiten aufgelistet:

- Es bleibt zu sehen, wie neue Technologien wie Kinetics, das Internet der Dinge und Virtual Reality genutzt werden und für die Regierungskommunikation relevant sein werden.
- Es bleibt weiter zu untersuchen, ob APN Menschen in bestimmten politischen Informationsbereichen „nudgen“ können.
- Im Rahmen der politischen Kommunikation bleibt viel Raum zur Erforschung des „Algorithmic Nudgens“ in Verbindung mit Kommunikationsmaßnahmen.
- Es bleibt zu analysieren, ob Microtargeting wirkt, wenn Nachrichten entsprechend ihrer Quelle gekennzeichnet sind und wenn, dann wie und wann.

- Bisher existieren keine Analysen, die den Zusammenhang zwischen dem Einsatz von künstlicher Intelligenz in der öffentlichen Kommunikation und den „myside bias“ näher beleuchten – das könnte in Kombination mit political efficacy erforscht werden.

- Die Annahme, dass radikalisierte Personen mainstream Medien verlassen, weil diese Heterogenität besitzen und zu „algorithmus-freien“ Medien wechseln, bleibt zu überprüfen.

- Eine Analyse zur Nutzung von Chatbots in der Parteienkommunikation könnte neue Erkenntnisse für den digitalen Wandel der Demokratie bieten, als auch Kommunikationsstrategien für Parteien, Wählerinnen und Wähler, als auch für Regierungen hervorbringen.

- Die aufgezeigten Anwendungs- und Forschungsmöglichkeiten müssen im Rahmen der Entwicklung generativer KI weitergedacht und untersucht werden.

Literaturverzeichnis

Agarwal, Shruti; Farid, Hany; El-Gaaly, Tarek; Lim, Ser-Nam (2020): Detecting Deep-Fake Videos from Appearance and Behavior. In: 2020 IEEE International Workshop on Information Forensics and Security (WIFS), S. 1–6.

Alnemr, Nardine (2021): Citizens and intelligent machines : algorithms in deliberative democracy. University of Canberra.

Amoore, Louise; Goede, Marieke de (2021): Datawars: reflections twenty years after 9/11. In: *Critical Studies on Terrorism* 14 (4), S. 425–429. DOI: 10.1080/17539153.2021.1982117.

Androutsopoulou, Aggeliki; Karacapilidis, Nikos; Loukis, Euripidis; Charalabidis, Yannis (2019): Transforming the communication between citizens and government through AI-guided chatbots. In: *Government Information Quarterly* 36 (2), S. 358–367. DOI: 10.1016/j.giq.2018.10.001.

Bächle, Thomas Christian (2015): Mythos Algorithmus. Die Fabrikation des computerisierbaren Menschen. Wiesbaden: Springer VS (Research).

Baldini, Gianmarco; Botterman, Maarten; Neisse, Ricardo; Tallacchini, Mariachiara (2018): Ethical Design in the Internet of Things. In: *Sci Eng Ethics* 24 (3), S. 905–925. DOI: 10.1007/s11948-016-9754-5.

Balkin, Jack M. (2017): Free Speech in the Algorithmic Society: Big Data, Private Governance, and New School Speech Regulation. In: *SSRN Journal*. DOI: 10.2139/ssrn.3038939.

Bareis, Jascha; Katzenbach, Christian (2021): Talking AI into Being: The Narratives and Imaginaries of National AI Strategies and Their Performative Politics. In: Science, Technology, & Human Values. DOI: 10.1177/01622439211030007.

Benartzi, Shlomo; Beshears, John; Milkman, Katherine L.; Sunstein, Cass R.; Thaler, Richard H.; Shankar, Maya et al. (2017): Should Governments Invest More in Nudging? In: *Psychological science* 28 (8), S. 1041–1055. DOI: 10.1177/0956797617702501.

Bennett, W. Lance; Livingston, Steven (Hg.) (2021): The disinformation age. Politics, technology, and disruptive communication in the United States. New York NY: Cambridge University Press (SSRC anxieties of democracy).

Borucki, Isabelle; Jun, Uwe (2018): Regierungskommunikation im Wandel – Politikwissenschaftliche Perspektiven. In: Juliana Raupp, Jan Niklas Kocks und Kim Murphy (Hg.): Regierungskommunikation und staatliche Öffentlichkeitsarbeit: Springer VS, Wiesbaden, S. 25–46. Online verfügbar unter https://link.springer.com/chapter/10.1007/978-3-658-20589-8_3.

Borucki, Isabelle; Kleinen-von Königslöw, Katharina; Marschall, Stefan; Zerback, Thomas (Hg.) (2020a): Handbuch Politische Kommunikation. Living Reference Work, continuously updated edition. Wiesbaden: Springer Fachmedien Wiesbaden.

Borucki, Isabelle; Michels, Dennis; Marschall, Stefan (2020b): Die digitalisierte Demokratie. Ein Überblick. In: *Z Politikwiss* 30 (2), S. 163–169. DOI: 10.1007/s41358-020-00224-5.

Borucki, Isabelle; Michels, Dennis; Marschall, Stefan (2020c): Die Zukunft digitalisierter Demokratie – Perspektiven für die Forschung. In: *Z Politikwiss* 30 (2), S. 359–378. DOI: 10.1007/s41358-020-00234-3.

Bradshaw, S.; Howard, P. (2017): Troops, Trolls and Troublemakers: A Global Inventory of Organized Social Media Manipulation. In: *Computational Propaganda Research Project*. Online verfügbar unter https://ora.ox.ac.uk/objects/uuid:cef7e8d9-27bf-4ea5-9fd6-855209b3e1f6.

Bruns, Axel (2021): Beyond the Bubble: A Critical Review of the Evidence for Echo Chambers and Filter Bubbles. ECREA. 07. September 2021. In: *European Communication Conference*, https://snurb.info/node/2620.

Bryson, Joanna; Winfield, Alan (2017): Standardizing Ethical Design for Artificial Intelligence and Autonomous Systems. In: *Computer* 50 (5), S. 116–119. DOI: 10.1109/MC.2017.154.

Buxmann, Peter; Schmidt, Holger (2021): Ethische Aspekte der Künstlichen Intelligenz. In: Künstliche Intelligenz: Springer Gabler, Berlin, Heidelberg, S. 215–229. Online verfügbar unter https://link.springer.com/chapter/10.1007/978-3-662-61794-6_13#citeas.

Cave, Stephen; Coughlan, Kate; Dihal, Kanta (2019): "Scary Robots". In: Vincent Conitzer (Hg.): Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society. AIES '19: AAAI/ACM Conference on AI, Ethics, and Society. Honolulu HI USA, 27 01 2019 28 01 2019. New York,NY,United States: Association for Computing Machinery (ACM Digital Library), S. 331–337.

Cave, Stephen; Craig, Claire; Dihal, Kanta; Dillon, Sarah; Montgomery, Jessica; Singler, Beth; Taylor, Lindsay (2018): Portrayals and perceptions of AI and why they matter. Unter Mitarbeit von Apollo - University of Cambridge Repository. The Royal Society. Online verfügbar unter https://www.repository.cam.ac.uk/handle/1810/287193.

Cave, Stephen; Dihal, Kanta; Dillon, Sarah (2020): AI narratives: A history of imaginative thinking about intelligent machines: Oxford University Press.

Ciampaglia, Giovanni Luca; Flammini, Alessandro; Menczer, Filippo (2015): The production of information in the attention economy. In: *Scientific reports* 5 (1), S. 1–6.

Cresci, Stefano; Di Pietro, Roberto; Petrocchi, Marinella; Spognardi, Angelo; Tesconi, Maurizio (2017): The Paradigm-Shift of Social Spambots. In: *Proceedings of the 26th International Conference on World Wide Web Companion*, S. 963–972. DOI: 10.1145/3041021.3055135.

Crogan, Patrick; Kinsley, Samuel (2012): Paying attention: Towards a critique of the attention economy. In: *Culture Machine* 13.

Cuéllar, Mariano-Florentino; Huq, Aziz Z. (2022): The Democratic Regulation of Artificial Intelligence. A case for focusing on forward-looking policy considerations rather than a rights framework in regulating "AI systems". 31 January 2022. Online verfügbar unter https://knightcolumbia.org/content/the-democratic-regulation-of-artificial-intelligence, zuletzt aktualisiert am 26.04.2022, zuletzt geprüft am 26.04.2022.

Davenport, Thomas H.; Beck, John C. (2001): The attention economy. In: *Ubiquity* 2001 (May), 1-es.

Die Bundesregierung (2021): Datenstrategie der Bundesregierung. Presse- und Informationsamt der Bundesregierung (BPA). Eine Innovationsstrategie für gesellschaftlichen Fort-

schritt und nachhaltiges Wachstum - Kabinettfassung, 27. Januar 2021. 27 Januar 2021. Die Bundesregierung. Online verfügbar unter https://www.bundesregierung.de/breg-de/suche/datenstrategie-der-bundesregierung-1845632, zuletzt aktualisiert am 20.10.2022, zuletzt geprüft am 20.10.2022.

Die Bundesregierung (2022): Wirksam regieren - Mit Bürgern für Bürger. Online verfügbar unter https://www.bundesregierung.de/breg-de/themen/wirksam-regieren, zuletzt geprüft am 15.01.23.

Dreesen, Philipp; Krasselt, Julia (2022): Social Bots als Stimmen im Diskurs. In: Eva Gredel (Hg.): Diskurse - digital. Boston: De Gruyter (Diskursmuster / discourse patterns), S. 271–282.

Eisenegger, Mark; Prinzing, Marlis; Ettinger, Patrik; Blum, Roger (2021): Digitaler Strukturwandel der Öffentlichkeit. Historische Verortung, Modelle und Konsequenzen. Wiesbaden: Springer VS Springer Fachmedien Wiesbaden GmbH (Research (Wiesbaden, Germany)).

Ergin Bulut; Erdem Yörük (2017): Mediatized Populisms| Digital Populism: Trolls and Political Polarization of Twitter in Turkey. In: *International Journal of Communication* 11 (0), S. 25. Online verfügbar unter https://ijoc.org/index.php/ijoc/article/view/6702.

Farrell, Henry; Newman, Abraham; Wallace, Jeremy (2022): Spirals of Delusion: How AI Distorts Decision-Making and Makes Dictators More Dangerous. In: *Foreign Affairs Magazine*, 31.08.2022. Online verfügbar unter https://www.foreignaffairs.com/world/spirals-delusion-artificial-intelligence-decision-making, zuletzt geprüft am 01.09.2022.

Ferrara, Emilio; Varol, Onur; Davis, Clayton; Menczer, Filippo; Flammini, Alessandro (2016): The rise of social bots. In: *Commun. ACM* 59 (7), S. 96–104. DOI: 10.1145/2818717.

Festinger, Leon (2009): A theory of cognitive dissonance. Renewed 1985 by author. Stanford, Calif.: Stanford University Press.

Frank, Joel; Eisenhofer, Thorsten; Schönherr, Lea; Fischer, Asja; Kolossa, Dorothea; Holz, Thorsten (2020): Leveraging Frequency Analysis for Deep Fake Image Recognition. In: Hal Daumé III und Aarti Singh (Hg.): Proceedings of the 37th International Conference on Machine Learning, Bd. 119: PMLR (Proceedings of Machine Learning Research), S. 3247–3258. Online verfügbar unter https://proceedings.mlr.press/v119/frank20a.html.

Fukuda-Parr, Sakiko; Gibbons, Elizabeth (2021): Emerging Consensus on 'Ethical AI': Human Rights Critique of Stakeholder Guidelines. In: *Glob Policy* 12 (S6), S. 32–44. DOI: 10.1111/1758-5899.12965.

Gallwitz, Florian; Kreil, Michael (2021): The Rise and Fall of 'Social Bot' Research.

Gallwitz, Florian; Kreil, Michael (2022): Investigating the Validity of Botometer-Based Social Bot Studies. In:. Multidisciplinary International Symposium on Disinformation in Open Online Media: Springer, Cham, S. 63–78. Online verfügbar unter https://link.springer.com/chapter/10.1007/978-3-031-18253-2_5.

Gandy, Oscar (2019): 15. Data Mining, Surveillance, and Discrimination in the Post-9/11 Environment. In: Kevin D. Haggerty und Richard V. Ericson (Hg.): The new politics of surveillance and visibility: University of Toronto Press, S. 363–384.

Gandy, Oscar H. (2007): Data mining and surveillance in the post 9/11 environment. In: *The surveillance studies reader*, S. 147–157.

García-Orosa, Berta (2021): Disinformation, social media, bots, and astroturfing: the fourth wave of digital democracy. In: *EPI*, Artikel e300603. DOI: 10.3145/epi.2021.nov.03.

Gillespie, Tarleton (2018): Custodians of the internet. Platforms, content moderation, and the hidden decisions that shape social media. New Haven: Yale University Press.

Gorwa, Robert (2019): What is platform governance? In: *Information, Communication & Society* 22 (6), S. 854–871. DOI: 10.1080/1369118X.2019.1573914.

Gorwa, Robert; Binns, Reuben; Katzenbach, Christian (2020): Algorithmic content moderation: Technical and political challenges in the automation of platform governance. In: *Big Data & Society* 7 (1), 205395171989794. DOI: 10.1177/2053951719897945.

Habermas, Jürgen (1989): The structural transformation of the public sphere. An inquiry into a category of bourgeois society. Cambridge, Mass.: MIT Press.

Habermas, Jürgen (2021): Überlegungen und Hypothesen zu einem erneuten Strukturwandel der politischen Öffentlichkeit. In: Martin Seeliger und Sebastian Sevignani (Hg.): Ein neuer Strukturwandel der Öffentlichkeit? 1. Auflage. Baden-Baden: Nomos (Leviathan. Sonderband, 37), S. 470–500.

Hagen, Loni; Neely, Stephen; Keller, Thomas E.; Scharf, Ryan; Vasquez, Fatima Espinoza (2022): Rise of the Machines? Examining the Influence of Social Bots on a Political Discussion Network. In: *Social Science Computer Review* 40 (2), S. 264–287. DOI: 10.1177/0894439320908190.

Hagendorff, Thilo (2020): The Ethics of AI Ethics: An Evaluation of Guidelines. In: *Minds and Machines* 30 (1), S. 99–120. DOI: 10.1007/s11023-020-09517-8.

Hajian, Sara; Bonchi, Francesco; Castillo, Carlos (2016): Algorithmic Bias. In: Balaji Krishnapuram, Mohak Shah, Alex Smola, Charu Aggarwal, Dou Shen und Rajeev Rastogi (Hg.): Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. KDD '16: The 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. San Francisco California USA, 13 08 2016 17 08 2016. New York, NY: ACM, S. 2125–2126.

Han, Rongbin (2015): Manufacturing consent in cyberspace: China's "fifty-cent army". In: *Journal of Current Chinese Affairs* 44 (2), S. 105–134.

Harris, Tristan (2017): The Eyeball Economy: How Advertising Co-Opts Independent Thought. In: *Big Think*, 11.04.2017. Online verfügbar unter https://bigthink.com/high-culture/tristan-harris-the-attention-economy-a-race-to-the-bottom-of-the-brain-stem/, zuletzt geprüft am 09.08.2022.

Harsin, Jayson (2015): Regimes of Posttruth, Postpolitics, and Attention Economies. In: *Communication, Culture & Critique* 8 (2), S. 327–333. DOI: 10.1111/cccr.12097.

Helbig, Karoline (2018): Der Einfluss von Algorithmen auf demokratische Deliberation. In: Resa Mohabbat-Kar (Hg.): (Un)berechenbar? Algorithmen und Automatisierung in Staat und Gesellschaft. 1. Auflage. Berlin: Kompetenzzentrum Öffentliche IT, S. 339–363. Online verfügbar unter https://www.econstor.eu/handle/10419/191920.

Helbing, Dirk (2017): »Big Nudging« – zur Problemlösung wenig geeignet. In: Carsten Könneker (Hg.): Unsere digitale Zukunft: In welcher Welt wollen wir leben? Berlin, Heidelberg: Springer Berlin Heidelberg, S. 49–52.

Helbing, Dirk; Frey, Bruno S.; Gigerenzer, Gerd; Hafen, Ernst; Hagner, Michael; Hofstetter, Yvonne et al. (2018): Will Democracy Survive Big Data and Artificial Intelligence? In: Dirk Helbing (Hg.): Towards digital enlightenment. Essays on the dark and light sides of the digital revolution / Dirk Helbing, editor. Cham: Springer, S. 73–98.

Hendricks, Vincent F.; Mehlsen, Camilla (2022): The ministry of truth. BigTech's influence on facts, feelings and fictions. Cham: Springer. Online verfügbar unter https://ebookcentral.proquest.com/lib/kxp/detail.action?docID=7013196.

Hill, Hermann (2020): Moderne Verwaltungskommunikation. In: Verständliche Verwaltungskommunikation in Zeiten der Digitalisierung: Nomos Verlagsgesellschaft mbH & Co. KG, S. 77–96.

Hofstetter, Yvonne (2016): Das Ende der Demokratie. Wie die künstliche Intelligenz die Politik übernimmt und uns entmündigt. 1. Auflage. München: C. Bertelsmann.

Hummel, Dennis; Maedche, Alexander (2019): How effective is nudging? A quantitative review on the effect sizes and limits of empirical nudging studies. In: *Journal of Behavioral and Experimental Economics* 80, S. 47–58. DOI: 10.1016/j.socec.2019.03.005.

Irgmaier, Florian; Ulbricht, Lena (2017): Big Data und Nudging. Kommt der digitale Überwachungsstaat. In: *WZB Mitteilungen* 158, S. 15–17.

Jaster, Romy; Lanius, David (2020): Die Wahrheit schafft sich ab. Wie Fake News Politik machen : [Was bedeutet das alles?]. 5. Auflage. Ditzingen: Reclam (Reclams Universal-Bibliothek, 19608).

Jesse, Mathias; Jannach, Dietmar (2021): Digital nudging with recommender systems: Survey and future directions. In: *Computers in Human Behavior Reports* 3, S. 100052. DOI: 10.1016/j.chbr.2020.100052.

Kadel, Ramón (2022): Warum redaktionelle Bots die Zukunft sind. In: RAMN KADEL (Hg.): REDAKBOT. Includes digital download. [S.l.]: MORGAN KAUFMANN, S. 7–25.

Kanakaraj, Monisha; Guddeti, Ram Mohana Reddy (2015): Performance analysis of Ensemble methods on Twitter sentiment analysis using NLP techniques. In: Proceedings of the 2015 IEEE 9th International Conference on Semantic Computing (IEEE ICSC 2015): IEEE.

Kaplan, Jerry (2016): Artificial intelligence. What everyone needs to know. Oxford: Oxford University Press.

Katzenbach, Christian (2018): Die Ordnung der Algorithmen - zur Automatisierung von Relevanz und Regulierung gesellschaftlicher Kommunikation. In:, S. 15. Online verfügbar unter https://www.ssoar.info/ssoar/bitstream/document/57154/1/ssoar-2018-katzenbach-Die_Ordnung_der_Algorithmen_-.pdf.

Katzenbach, Christian (2021): "AI will fix this" – The Technical, Discursive, and Political Turn to AI in Governing Communication. In: *Big Data & Society* 8 (2), 20539517211046182. DOI: 10.1177/20539517211046182.

Katzenbach, Christian (2022): Der „Algorithmic turn" in der Plattform-Governance. In: *Köln Z Soziol*, S. 1–23. DOI: 10.1007/s11577-022-00837-4.

Katzenbach, Christian; Bächle, Thomas Christian (2019): Defining concepts of the digital society. In: *Internet Policy Review* 8 (4). DOI: 10.14763/2019.4.1430.

Katzenbach, Christian; Ulbricht, Lena (2019): Algorithmic governance. In: *Internet Policy Review* 8 (4). DOI: 10.14763/2019.4.1424.

Kavlakoglu, Eva (2020): AI vs. Machine Learning vs. Deep Learning vs. Neural Networks: What's the Difference?, 27.05.2020. Online verfügbar unter https://www.ibm.com/cloud/blog/ai-vs-machine-learning-vs-deep-learning-vs-neural-networks, zuletzt geprüft am 30.08.2022.

Keller, Tobias R.; Klinger, Ulrike (2021): Chapter Two The Needle in the Haystack: Finding Social Bots on Twitter. In: Eszter Hargittai (Hg.): Research exposed. How empirical social science gets done in the digital age / edited by Eszter Hargittai. New York: Columbia University Press, S. 30–49.

Kim, Soomin; Eun, Jinsu; Seering, Joseph; Lee, Joonhwan (2021): Moderator Chatbot for Deliberative Discussion. In: *Proc. ACM Hum.-Comput. Interact.* 5 (CSCW1), S. 1–26. DOI: 10.1145/3449161.

Kind, Sonja; Jetzke, Tobias; Weide, Sebastian; Ehrenberg-Silies, Simone; Bovenschulte, Marc (2017): Social Bots. TA-Vorstudie.

Klinger, Ulrike (2019): Bürger oder Bots? Automatisierte Kommunikation im Bundestagswahlkampf 2017. In: *Bundeszentrale für politische Bildung*, 02.05.2019. Online verfügbar unter https://www.bpb.de/themen/medien-journalismus/digitale-desinformation/290557/buerger-oder-bots-automatisierte-kommunikation-im-bundestagswahlkampf-2017/, zuletzt geprüft am 25.10.2022.

Krisam, Mathias; Philipsborn, Peter von; Meder, Björn (2017): Nudging in der Primärprävention: Eine Übersicht und Perspektiven für Deutschland. In Gesundheitswesen 79 (2), pp. 117–123. DOI: 10.1055/s-0042-121598.

Kruikemeier, Sanne; Sezgin, Minem; Boerman, Sophie C. (2016): Political Microtargeting: Relationship Between Personalized Advertising on Facebook and Voters' Responses. In: *Cyberpsychology, behavior and social networking* 19 (6), S. 367–372. DOI: 10.1089/cyber.2015.0652.

Leitherer, Johanna (2019): Kompakt erklärt: Was ist Microtargeting? In: *springerprofessional.de*, 09.09.2019. Online verfügbar unter https://www.springerprofessional.de/en/digitales-marketing/marketingkommunikation/kompakt-erklaert--was-ist-microtargeting-/17149634, zuletzt geprüft am 31.08.2022.

Lemoine, Blake (2022): Is LaMDA Sentient? — an Interview - Blake Lemoine - Medium. In: *Medium*, 11.06.2022. Online verfügbar unter https://cajundiscordian.medium.com/is-lamda-sentient-an-interview-ea64d916d917, zuletzt geprüft am 08.08.2022.

Lucke, Joern von; Gollasch, Katja (Hg.) (2022a): OPEN GOVERNMENT. Offenes regierungsundverwaltungshandeln - leitbilder, ziele und... methoden. [S.l.]: SPRINGER GABLER.

Lucke, Jörn von (2020): Zukünftige Formen der digitalen Kommunikation mit der Verwaltung. In: Rudolf Fisch (Hg.): Verständliche Verwaltungskommunikation in Zeiten der Digitalisierung. Konzepte - Lösungen - Fallbeispiele. 1. Edition. Baden-Baden: Nomos (Verwaltungsressourcen und Verwaltungsstrukturen, Band 35), S. 109–122.

Lucke, Jörn von; Gollasch, Katja (2022b): Offene Daten und offene Verwaltungsdaten – Öffnung von Datenbeständen. In: Open Government: Springer Gabler, Wiesbaden, S. 49–73. Online verfügbar unter https://link.springer.com/chapter/10.1007/978-3-658-36795-4_3.

Martini, Franziska; Samula, Paul; Keller, Tobias R.; Klinger, Ulrike (2021): Bot, or not? Comparing three methods for detecting social bots in five political discourses. In: *Big Data & Society* 8 (2), 205395172110335. DOI: 10.1177/20539517211033566.

Mehrabi, Ninareh; Morstatter, Fred; Saxena, Nripsuta; Lerman, Kristina; Galstyan, Aram (2021): A Survey on Bias and Fairness in Machine Learning. In: *ACM Comput. Surv.* 54 (6), S. 1–35. DOI: 10.1145/3457607.

Meier, Klaus; Graßl, Michael; Schützeneder, Jonas (2021): KI als Anwendung im Journalismus: zwischen Misstrauen und Aufklärung. In: *KI-Campus Blog*. Online verfügbar unter https://edoc.ku.de/id/eprint/26572/.

Möhlmann, M. (2021): Algorithmic nudges don't have to be unethical. In: *Harvard Business Review* 22.

Mooney, Raymond J. (1996): Comparative Experiments on Disambiguating Word Senses: An Illustration of the Role of Bias in Machine Learning: arXiv.

Neuberger, Christoph (2009): Internet, Journalismus und Öffentlichkeit. In: Christoph Neuberger (Hg.): Journalismus im Internet. Profession - Partizipation - Technisierung. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 19–105.

Neuberger, Christoph (2014): Die Identität und Qualität des Journalismus im Internet aus der Sicht des Publikums. In: Journalismus und (sein) Publikum: Springer VS, Wiesbaden, S. 229–251. Online verfügbar unter https://link.springer.com/chapter/10.1007/978-3-531-19821-7_13.

Neuberger, Christoph (2017): Journalismus und Digitalisierung: Profession, Partizipation und Algorithmen.

Niemi, Richard G.; Craig, Stephen C.; Mattei, Franco (1991): Measuring Internal Political Efficacy in the 1988 National Election Study. In: *The American Political Science Review* 85 (4), S. 1407–1413. DOI: 10.2307/1963953.

Oberer, Birgit; Erkollar, Alptekin; Stein, Anna (2019): Social Bots – Act Like a Human, Think Like a Bot. In: Marcus Stumpf (Hg.): Digitalisierung und Kommunikation. Konsequenzen der digitalen Transformation für die Wirtschaftskommunikation. Wiesbaden: Springer VS (Research (Wiesbaden, Germany), Band 31), S. 311–327.

OECD (2021): Internal and external political efficacy.

Papacharissi, Zizi (2015): Affective publics. Sentiment, technology, and politics / Zizi Papacharissi. Oxford, New York, NY: Oxford University Press (Oxford studies in digital politics).

Papacharissi, Zizi; Trevey, Meggan Taylor (2018): Affective publics and windows of opportunity : Social media and the potential for social change. In: Graham Meikle (Hg.): The Routledge companion to media and activism. 1st. London: Routledge (Routledge media and cultural studies companions), S. 87–96. Online verfügbar unter https://www.taylorfrancis.com/chapters/edit/10.4324/9781315475059-9/affective-publics-windows-opportunity-zizi-papacharissi-meggan-taylor-trevey.

Papakyriakopoulos, Orestis; Hegelich, Simon; Shahrezaye, Morteza; Serrano, Juan Carlos Medina (2018): Social media and microtargeting: Political data processing and the consequences for Germany. In: *Big Data & Society* 5 (2), 205395171881184. DOI: 10.1177/2053951718811844.

Papakyriakopoulos, Orestis; Shahrezaye, Morteza; Thieltges, Andree; Medina Serrano, Juan Carlos; Hegelich, Simon (2017): Social Media und Microtargeting in Deutschland. In: *Informatik-Spektrum* 40 (4), S. 327–335. DOI: 10.1007/s00287-017-1051-4.

Pariser, Eli (2011): The filter bubble. What the Internet is hiding from you / Eli Pariser. London: Viking.

Pfetsch, Barbara; Löblich, Maria; Eilders, Christiane (2018): Dissonante Öffentlichkeiten als Perspektive kommunikationswissenschaftlicher Theoriebildung. In: *Publizistik* 63 (4), S. 477–495. DOI: 10.1007/s11616-018-0441-1.

Postmes, O. M.; Spears, R.; Lea, M. (1998): Breaching or Building Social Boundaries?: SIDE-Effects of Computer-Mediated Communication. In: *Communication Research* 25 (6), S. 689–715. DOI: 10.1177/009365098025006006.

Rakesh K., Arora; Manoj, K., Gupta (2017): e-Governance using Data Warehousing and Data Mining. In: *IJCA* 169 (8), S. 28–31. DOI: 10.5120/ijca2017914785.

Rau, Jan Philipp; Stier, Sebastian (2019): Die Echokammer-Hypothese: Fragmentierung der Öffentlichkeit und politische Polarisierung durch digitale Medien? In: *Z Vgl Polit Wiss* 13 (3), S. 399–417. DOI: 10.1007/s12286-019-00429-1.

Raupp, Juliana; Kocks, Jan Niklas; Murphy, Kim (Hg.) (2018a): Regierungskommunikation und staatliche Öffentlichkeitsarbeit: Springer VS, Wiesbaden.

Raupp, Juliana; Kocks, Jan Niklas; Murphy, Kim (Hg.) (2018b): Regierungskommunikation und staatliche Öffentlichkeitsarbeit. Implikationen des technologisch induzierten Medienwandels. Wiesbaden: Springer VS (Springer VS research).

Robyn Caplan (2021): 6. The Artisan and the Decision Factory: The Organizational Dynamics of Private Speech Governance. In: *Digital Technology and Democratic Theory*, S. 167–190. Online verfügbar unter https://www.degruyter.com/document/doi/10.7208/9780226748603-007/html?lang=en.

Ross Arguedas, A.; Robertson, C.; Fletcher, R.; Nielsen, R. (2022): Echo chambers, filter bubbles, and polarisation: a literature review: Reuters Institute for the Study of Journalism.

Rußmann, Uta (2018): Output-orientierte Forschungsansätze zur Analyse(digitaler) Regierungskommunikation und staatlicher Öffentlichkeitsarbeit – Herausforderungen und Möglichkeiten. In: Juliana Raupp, Jan Niklas Kocks und Kim Murphy (Hg.): Regierungskommunikation und staatliche Öffentlichkeitsarbeit. Implikationen des technologisch induzierten Medienwandels. Wiesbaden: Springer VS (Springer VS research), S. 167–183.

Saka, Erkan (2018): Social Media in Turkey as a Space for Political Battles: AKTrolls and other Politically motivated trolling. In: *Middle East Critique* 27 (2), S. 161–177. DOI: 10.1080/19436149.2018.1439271.

Santos, Fernando P.; Lelkes, Yphtach; Levin, Simon A. (2021): Link recommendation algorithms and dynamics of polarization in online social networks. In: *Proceedings of the National Academy of Sciences of the United States of America* 118 (50). DOI: 10.1073/pnas.2102141118.

Saura, José Ramón; Palacios-Marqués, Daniel; Iturricha-Fernández, Agustín (2021): Ethical design in social media: Assessing the main performance measurements of user online beha-

vior modification. In: *Journal of Business Research* 129, S. 271–281. DOI: 10.1016/j.jbusres.2021.03.001.

Schmidt, Andreas T.; Engelen, Bart (2020): The ethics of nudging: An overview. In: *Philosophy Compass* 15 (4). DOI: 10.1111/phc3.12658.

Schrape, Jan-Felix (2021a): Öffentliche Kommunikation in der digitalisierten Gesellschaft. In: Kathrin Braun und Cordula Kropp (Hg.): In digitaler Gesellschaft, Bd. 2. Bielefeld, Germany: transcript Verlag (Politik in der digitalen Gesellschaft), S. 249–274.

Schrape, Jan-Felix (2021b): Platformization, Pluralization, Synthetization. Public Communication in the Digital Age. In: *SSRN Journal*. DOI: 10.2139/ssrn.3911301.

Schweiger, Wolfgang.; Weber, Patrick.; Prochazka, Fabian.; Brückner, Lara. (2019): Algorithmisch personalisierte Nachrichtenkanäle. Begriffe, Nutzung, Wirkung. Wiesbaden: Springer VS.

Silver, Laura (2021): Ideological divisions over cultural issues are far wider in the U.S. than in the UK, France and Germany. In: *Pew Research Center*, 05.05.2021. Online verfügbar unter https://www.pewresearch.org/fact-tank/2021/05/05/ideological-divisions-over-cultural-issues-are-far-wider-in-the-u-s-than-in-the-uk-france-and-germany/, zuletzt geprüft am 29.09.2022.

Slobogin, Christopher (2008): Government Data Mining and the Fourth Amendment. In: *The University of Chicago Law Review* 75 (1), S. 317–341. Online verfügbar unter http://www.jstor.org/stable/20141910.

Sorge, Aden (2022): Digitaler Bundestagswahlkampf 2021: Desinformation, „Microtargeting“ und „Social-Bots“. 31-53 Seiten / Zeitschrift für Parteienwissenschaften, Nr. 1 (2022): Zeitschrift für Parteienwissenschaften. DOI: 10.24338/MIP-202231-53.

Stanovich, Keith E. (2021): The bias that divides us. The science and politics of myside thinking. Cambridge, Massachusetts: The MIT Press.

Stanovich, Keith E.; West, Richard F. (2007): Natural myside bias is independent of cognitive ability. In: *Thinking & reasoning* 13 (3), S. 225–247. DOI: 10.1080/13546780600780796.

Stanovich, Keith E.; West, Richard F.; Toplak, Maggie E. (2013): Myside Bias, Rational Thinking, and Intelligence. In: *Curr Dir Psychol Sci* 22 (4), S. 259–264. DOI: 10.1177/0963721413480174.

Stark, Birgit; Magin, Melanie; Jürgens, Pascal (2021): Maßlos überschätzt. Ein Überblick über theoretische Annahmen und empirische Befunde zu Filterblasen und Echokammern. In: Mark Eisenegger, Marlis Prinzing, Patrik Ettinger und Roger Blum (Hg.): Digitaler Strukturwandel der Öffentlichkeit. Historische Verortung, Modelle und Konsequenzen. Wiesbaden: Springer VS Springer Fachmedien Wiesbaden GmbH (Research (Wiesbaden, Germany)), S. 303–321.

Sunstein, Cass R. (2001): Echo chambers. Bush v. Gore, impeachment, and beyond. Princeton, N.J.: Princeton University Press.

Sunstein, Cass R. (2007): Republic.com 2.0: Princeton University Press. Online verfügbar unter http://www.jstor.org/stable/j.ctt7tbsw.

Sunstein, Cass R. (2015): The Ethics of Nudging. In: *Yale Journal on Regulation* 32 (2), S. 413–450. Online verfügbar unter https://heinonline.org/HOL/P?h=hein.journals/yjor32&i=419.

Tetsuya Nasukawa; Jeonghee Yi (2003): Sentiment analysis: capturing favorability using natural language processing. In: Proceedings of the 2nd international conference on Knowledge capture. Sanibel Island, FL, USA: Association for Computing Machinery, S. 70–77.

Thaler, Richard H.; Sunstein, Cass R. (2009): Nudge [electronic resource]. Improving decisions about health, wealth, and happiness / Richard H. Thaler, Cass R. Sunstein. Revised and expanded edition. New York: Penguin Books.

The Behavioural Insights Team (2023): Reports. Online verfügbar unter https://www.bi.team/our-work-2/publications/, zuletzt geprüft am 15.01.23.

The Guardian (2013): The nudge unit – has it worked so far? In: *The Guardian*, 02.05.2013. Online verfügbar unter https://www.theguardian.com/politics/2013/may/02/nudge-unit-has-it-worked, zuletzt geprüft am 31.08.2022.

The Guardian (2021): Study finds growing government use of sensitive data to 'nudge' behaviour. In: *The Guardian*, 08.09.2021. Online verfügbar unter https://www.theguardian.com/technology/2021/sep/08/study-finds-growing-government-use-of-sensitive-data-to-nudge-behaviour, zuletzt geprüft am 30.08.2022.

Thimm, Caja; Bächle, Thomas Christian (2019a): Autonomie der Technologie und autonome Systeme als ethische Herausforderung. In: Maschinenethik: Springer VS, Wiesbaden, S. 73–87. Online verfügbar unter https://link.springer.com/chapter/10.1007/978-3-658-21083-0_5.

Thimm, Caja; Bächle, Thomas Christian (2019b): Die Maschine: Freund oder Feind? In: Die Maschine: Freund oder Feind?: Springer VS, Wiesbaden, S. 1–13. Online verfügbar unter https://link.springer.com/chapter/10.1007/978-3-658-22954-2_1.

Thornhill, Calum; Meeus, Quentin; Peperkamp, Jeroen; Berendt, Bettina (2019): A Digital Nudge to Counter Confirmation Bias. In: *Frontiers in big data* 2, S. 11. DOI: 10.3389/fdata.2019.00011.

Trabold, Daniel (2021): Welche Arten von Maschinellem Lernen gibt es? - ML2R-Blog. Online verfügbar unter https://machinelearning-blog.de/grundlagen/welche-arten-von-maschinellem-lernen-gibt-es/, zuletzt aktualisiert am 10.05.2022, zuletzt geprüft am 29.08.2022.

Ulbricht, Lena (2020a): Algorithmen und Politisierung. In: Andreas Schäfer und David Meiering (Hg.): (Ent-)Politisierung? Die demokratische Gesellschaft im 21. Jahrhundert. 1. Auflage. Baden-Baden: Nomos (Nomos eLibrary Politikwissenschaft, 35), S. 253–278.

Ulbricht, Lena (2020b): Scraping the demos. Digitalization, web scraping and the democratic project. In: *Democratization* 27 (3), S. 426–442. DOI: 10.1080/13510347.2020.1714595.

Vowe, Gerhard (2020): Digitalisierung als grundlegender Veränderungsprozess der politischen Kommunikation. In: Isabelle Borucki, Katharina Kleinen-von Königslöw, Stefan Marschall und Thomas Zerback (Hg.): Handbuch Politische Kommunikation. Living Reference Work, continuously updated edition. Wiesbaden: Springer Fachmedien Wiesbaden, S. 1–15.

Wagner, Dirk (2021): On the emergence and design of AI nudging: the gentle big brother? In: *ROBONOMICS: The Journal of the Automated Economy* 2, S. 18. Online verfügbar unter https://journal.robonomics.science/index.php/rj/article/view/18.

Wallach, Wendell; Asaro, Peter (2020): Machine ethics and robot ethics. 1st. London: Routledge (The library of essays on the ethics of emerging technologies).

Wang, Qi; Jeon, Hee Jin (2020): Bias in bias recognition: People view others but not themselves as biased by preexisting beliefs and social stigmas. In: *PloS one* 15 (10), e0240232. DOI: 10.1371/journal.pone.0240232.

Webb, Amy (2019): The big nine. How the tech titans and their thinking machines could warp humanity / Amy Webb. New York: PublicAffairs.

Weber, Franziska; Schäfer, Hans-Bernd (2017): „Nudging“, Ein Spross der Verhaltensökonomie. Überlegungen zum liberalen Paternalismus auf gesetzgeberischer Ebene. In Der Staat 56 (4), pp. 561–592. DOI: 10.3790/staa.56.4.561.

Weinmann, Markus; Schneider, Christoph; vom Brocke, Jan (2016): Digital Nudging. In: *Bus Inf Syst Eng* 58 (6), S. 433–436. DOI: 10.1007/s12599-016-0453-1.

Whitehead, M.; Lilley, R.; Howell, R.; Jones, R.; Pykett, J. (2014): “Nudging all over the World”: Report on the ESRC transforming behaviours project: assessing the global impact of the behavioural sciences on public policy. Online verfügbar unter https://research.birmingham.ac.uk/en/publications/nudging-all-over-the-world-report-on-the-esrc-transforming-behavi.

Wilkinson, T. M. (2013): Nudging and Manipulation. In: *Political Studies* 61 (2), S. 341–355. DOI: 10.1111/j.1467-9248.2012.00974.x.

Wong, David; Floridi, Luciano (2022): Meta’s Oversight Board: A Review and Critical Assessment. In: *Minds and Machines*. DOI: 10.1007/s11023-022-09613-x.

Woolley, Samuel C. (2016): Automating power: Social bot interference in global politics. In: *FM*. DOI: 10.5210/fm.v21i4.6161.

Yang, Kai-Cheng; Ferrara, Emilio; Menczer, Filippo (2022): Botometer 101: Social bot practicum for computational social scientists. In: *ArXiv* abs/2201.01608.

Yorke, Harry (2022): Downing Street’s controversial ‘Nudge Unit’ accused of exploiting scare tactics during Covid crisis. 13 January 2022. In: *The Telegraph*, 13.01.2022. Online verfügbar unter https://www.telegraph.co.uk/politics/2022/01/13/downing-streets-controversial-nudge-unit-accused-exploiting/, zuletzt geprüft am 31.08.2022.

Yuan Tongwei; Chen, Peng (2012): Data Mining Applications in E-Government Information Security. In: *Procedia Engineering* 29, S. 235–240. DOI: 10.1016/j.proeng.2011.12.700.

Zarouali, Brahim; Dobber, Tom; Pauw, Guy de; Vreese, Claes de (2020): Using a Personality-Profiling Algorithm to Investigate Political Microtargeting: Assessing the Persuasion Effects of Personality-Tailored Ads on Social Media. In: *Communication Research*, 009365022096196. DOI: 10.1177/0093650220961965.

Zeilinger, Thomas (2019): Brauchen Algorithmen eine eigene Ethik? In: Michael Schröder und Axel Schwanebeck (Hg.): BIG DATA - IN DEN FANGEN DER DATENKRAKEN. [Place of publication not identified]: NOMOS VERLAGSGESELLSCHAFT, S. 139–156.

Zhang, Jinxue; Zhang, Rui; Zhang, Yanchao; Yan, Guanhua (2018): The Rise of Social Botnets: Attacks and Countermeasures. In: *IEEE Trans. Dependable and Secure Comput.* 15 (6), S. 1068–1082. DOI: 10.1109/TDSC.2016.2641441.

Zuiderveen Borgesius, Frederik J.; Möller, Judith; Kruikemeier, Sanne; Ó Fathaigh, Ronan; Irion, Kristina; Dobber, Tom et al. (2018): Online Political Microtargeting: Promises and Threats for Democracy. In: *ULR* 14 (1), S. 82. DOI: 10.18352/ulr.420.

Professur für Politikwissenschaft, insbesondere Politische Theorie
Helmut-Schmidt-Universität Hamburg/Universität der Bundeswehr

https://doi.org/10.24405/14980