

Politische Öffentlichkeit, Desinformation und das Problem von Deepfakes¹

Isabel Kusche

1. Einleitung: Was sind Deepfakes?

Im Jahr 2015 sorgte in Deutschland ein Video für Aufregung, in dem der damalige griechische Finanzminister Yanis Varoufakis Deutschland angeblich den Stinkefinger zeigte. Kurz nachdem der Moderator Günther Jauch Varoufakis in einer Talkshow mit diesem Video konfrontiert hatte, meldete sich der Satiriker Jan Böhmermann zu Wort und behauptete, der Stinkefinger sei eine mit Hilfe eines Special-Effects-Experten erstellte Montage, also eine Fälschung. Auch diese Behauptung wurde schnell in Zweifel gezogen, und die Frage der Echtheit des Videos wie des medialen Umgangs mit ihm beschäftigte die öffentliche Debatte in Deutschland noch tagelang (Plöchinger 2015).

Die Episode ist ein Indiz dafür, welche Bedeutung das Problem der Authentizität oder Inauthentizität von Videoaufnahmen prominenter und umstrittener Politiker in der Öffentlichkeit gewinnen kann. Was damals, auf dem Höhepunkt der griechischen Finanzkrise, wahlweise als Beleg der feindlichen Haltung eines griechischen Regierungsvertreters gegenüber Deutschland, als Indiz für die Oberflächlichkeit journalistischer Recherche oder als gelungene Satire gedeutet wurde, zeigt die Empfänglichkeit politischer Öffentlichkeit für Erregung und Chaos, wenn bei einem polarisierenden Thema in Frage steht, wer sich dazu in welcher Weise geäußert oder verhalten hat.

1 Große Teile dieses Beitrages sind im Rahmen von MOTRA (Monitoringsystem und Transferplattform Radikalisierung), einem durch das Bundesministerium für Bildung und Forschung geförderten Verbundprojekt, im Teilprojekt Technologiemonitoring entstanden, das am Institut für Technikfolgenabschätzung und Systemanalyse in Karlsruhe angesiedelt ist.

Nur wenige Jahre später steht inzwischen – verknüpft mit dem Stichwort Deepfakes – die Möglichkeit einer politischen Öffentlichkeit im Raum, in der sich viele Debatten künftig nicht oder nicht in erster Linie darum drehen werden, wie bestimmte politische Äußerungen oder Handlungen zu bewerten sind, sondern ob sie überhaupt so stattgefunden haben, wie es vorhandenes Audio- und Videomaterial zu dokumentieren scheint. Warnungen von den Folgen von Deepfakes verbreiteten sich erstmals 2018, als erste Beispiele Aufmerksamkeit erregten, die täuschend echt Aussagen oder Handlungen prominenter Personen präsentierten, ohne dass diese je stattgefunden hatten. Seit 2022 stehen mit Angeboten wie Dall-E, Midjourney oder Stable Diffusion Bildgeneratoren für ein breites Publikum zur Verfügung, und Anfang 2024 ist mit Sora ein Videogenerator hinzugekommen, wenn auch zunächst nur für ausgewählte Nutzerinnen und Nutzer. Parallel dazu hat sich die Faszination, aber auch die Sorge mit Blick auf Deepfakes verstärkt.

In einer frühen Einführung bezeichnet Tyler Bettilyon (2018) Deepfakes als »a frightening harbinger of the Desinformation Age«. Die in den folgenden beiden Jahren zahlreicher werdenden wissenschaftlichen Beiträge innerhalb und außerhalb Deutschlands bekräftigen, dass es unbedingt Grund zur Besorgnis gibt (z.B. Chesney/Citron 2018; Bovenschulte 2019; Farid/Schindler 2020). Befürchtet wird, dass manipuliertes Video- und Audiomaterial – Deepfakes – von nichtmanipuliertem Material so schwer zu unterscheiden ist, dass es politische Gegner diskreditieren und so insbesondere Wahlkämpfe beeinflussen kann. Komplementär dazu besteht demnach das Problem, dass nicht manipuliertes Material, das skandalträchtige Aussagen oder Handlungen politischer Akteure belegt, künftig als Deepfake abgetan werden könnte. Die möglichen Folgen für die politische Öffentlichkeit lassen sich daher nicht auf denkbare Effekte einzelner Deepfakes eingrenzen. Das grundlegendere Problem wird in ihrem Potential gesehen, die Idee von Wahrheit überhaupt zu untergraben. Zwar werden auch Gefahren für andere Aspekte des gesellschaftlichen Lebens thematisiert, etwa die Möglichkeiten, Deepfakes für private Rachezüge zu nutzen – insbesondere gegen Frauen, die mit pornografischem Material diskreditiert werden (Paris/Donovan 2019) – oder Aktienkurse von einzelnen Unternehmen, wenn nicht gar ganze Märkte zu beeinflussen (Bateman 2020). Im Mittelpunkt stehen Deepfakes jedoch meist als überlegenes Instrument politischer Desinformation.

Manipulationen von Audio- und Videoaufzeichnungen sind kein neues Phänomen. Als Deepfakes werden sie bezeichnet, sofern sie mittels Deep Learning hergestellt sind, einem Verfahren maschinellen Lernens, das in

besonderer Weise für die Potentiale künstlicher Intelligenz (KI) steht. Deepfakes gehören insofern zu den sogenannten synthetischen Medien – eine Sammelbezeichnung für Video-, Audio, Bild- und Textmaterial, das durch Verfahren künstlicher Intelligenz hergestellt wurde (Kalpokas 2020). Deepfakes verdanken sich Innovationen im Bereich des Deep Learning. Zunächst war es insbesondere der Ansatz der Generative Adversarial Networks (GANs), der die Herstellung täuschend echten Audio- und Videomaterials ermöglichte (Farid/Schindler 2020: 16f.): Im Fall von GANs werden zwei künstliche neuronale Netzwerke miteinander verkoppelt. Das erste, der Generator, wird darauf trainiert, Audio- oder Videodaten tatsächlich stattgefundenere Ereignisse so zu verändern, dass das Resultat sich von nichtmanipuliertem Material nicht unterscheiden lässt. Dieses Training wird dadurch optimiert, dass ein zweites künstliches neuronales Netzwerk, der Diskriminator, darauf trainiert wird, diese Unterscheidung – zwischen nichtmanipuliertem Originalmaterial und manipuliertem Material – zu treffen und das vom ersten Netz produzierte manipulierte Material zu prüfen. Das Ergebnis dieser Prüfung liefert wiederum neue Daten für das Training des ersten Netzwerks, so lange, bis der Diskriminator das vom Generator erstellte Material nicht mehr vom Originalmaterial unterscheiden kann.

Mit der sogenannten Generativen KI sind in den letzten Jahren andere Verfahren zur Herstellung synthetischer Medien hinzugekommen, insbesondere sogenannte Diffusion Models (Corvi et al. 2023). Sie beruhen darauf, die im Trainingsdatensatz enthaltenen Bilder zunächst durch Hinzufügen von Bildrauschen zu zerstören, um anschließend künstliche neuronale Netze darauf zu trainieren, aus Bildrauschen Bilder zu rekonstruieren. Neben geeigneten Deep-Learning-Algorithmen hängt die Möglichkeit zur Herstellung von Deepfakes von weiteren Bedingungen ab (International Risk Governance Center 2019: 5–7; Bateman 2020: 5). Zum einen ist dies genügend Rechenkapazität, um die entsprechenden Algorithmen ausführen zu können. Zum anderen ist es die Verfügbarkeit von hinreichend vielen Daten in Form von Bild- bzw. Tonmaterial der Person, die in einem Deepfake auftreten soll, für das Training der Modelle. Diese Bedingungen greifen ineinander. Raffiniertere Algorithmen verlangen leistungsfähigere Computer; dank der rasanten Entwicklung bei GPUs und Cloud Computing wird es aber immer preisgünstiger, sich die entsprechende Rechenkapazität zu verschaffen. Gleichzeitig führt die Weiterentwicklung der Algorithmen dazu, dass sie mit weniger Trainingsdaten auskommen und die Herstellung von Deepfakes immer weniger davon abhängt, große Mengen an Video- und Audiomaterial zur Verfügung zu haben.

Waren es anfangs nur sehr bekannte und oft gefilmte Persönlichkeiten aus Unterhaltungsbranche und Politik, die für Deepfakes in Frage kamen, ist es daher zunehmend jede Person, von der es, dank Social Media, eine moderate Menge an Bildern im Internet gibt. Gleichzeitig ist die entsprechende Software und Rechnerkapazität immer breiter verfügbar (Farid/Schindler 2020: 20) und Unternehmen wie Stable Diffusion oder OpenAI bieten cloudbasierte Möglichkeiten, ihre Modelle zu nutzen, wobei die Eingabe von Textprompts genügt, um spezifische Bilder zu erzeugen

Die Bezeichnung »Deepfakes« umfasst mehrere typische Manipulationen von Videomaterial (Paris/Donovan 2019: 35–38; Farid/Schindler 2020: 15–17). Beim Face Swapping wird das Gesicht einer Person in einer Videoaufnahme durch das einer anderen ersetzt. Verbreitet ist dabei insbesondere die Nutzung pornografischen Materials, das so manipuliert wird, dass bekannte Schauspielerinnen darin scheinbar sexuelle Akte vollziehen. Ein zweiter Typ von Deepfakes manipuliert Lippenbewegungen und Sprache in Videos, so dass Personen bei Äußerungen gezeigt werden, die sie nie getätigt haben. Das bekannteste frühe Beispiel dafür ist ein 2018 von Jordan Peele und BuzzFeed produziertes Video, in dem Barack Obama vor den Gefahren von Deepfakes warnt, tatsächlich aber Peele Obamas Stimme imitiert und eine Software Obamas Lippenbewegungen an Peeles Sätze anpasst. Bei Puppet-Master-Deepfakes wird die Zielperson durch die Software quasi in eine Puppe verwandelt, deren Kopf- und Augenbewegungen sowie Mimik von einer anderen Person gesteuert werden, indem diese vor einer Kamera die entsprechenden Bewegungen ausführt.

Im Folgenden werden zunächst technische Gegenmaßnahmen thematisiert, die auf die Herausforderung reagieren, dass Deepfakes – im Unterschied zu älteren Verfahren der Manipulation von Bildern und Videos – durch das bloße Auge nicht mehr zu erkennen sind. Anschließend wird der Blick auf den Kontext politischer Öffentlichkeit gelenkt, in dem solche technisch avancierten Manipulationen nur das jüngste Beispiel einer tief verankerten Sorge darüber sind, dass die politische Öffentlichkeit auf unterschiedliche Weise bewusst getäuscht wird. Vor diesem Hintergrund lässt sich dann genauer bestimmen, ob und inwiefern Deepfakes tatsächlich eine besondere Herausforderung für die politische Öffentlichkeit darstellen – und das heißt, nicht im Hinblick auf die Einordnung technischer Artefakte als authentisch oder nicht-authentisch, sondern im Hinblick auf ihre Effekte im Kontext laufender politischer Kommunikation, in der sich politische Akteure ihren Publika unvermeidlich medial vermittelt präsentieren.

2. Gegenmaßnahmen: Technische Authentizitätsprüfungen

Das Funktionsprinzip von GANs beruht von vornherein auf der ständigen Verwendung einer Unterscheidung zwischen Originalmaterial und manipuliertem, durch GANs verändertem Material für die Produktion von weiterem verändertem Material. Es setzt insofern die Unterscheidung von Authentizität und Nichtauthentizität voraus, um mit ihrer Hilfe Material auf eine Weise zu prozessieren, die diese Unterscheidung schließlich unmöglich macht. Ein Deepfake ist daher nicht einfach ein manipuliertes und in diesem Sinne nicht-authentisches Bild- oder Tondokument; es ist im Zuge seiner technischen Herstellung bereits auf den Anschein von Authentizität technisch geprüft. Das schließt natürlich nicht aus, dass es forensischen Prüfmöglichkeiten unterworfen werden kann, die es als nichtauthentisch einstufen. Es bedeutet aber, dass solche technischen Prüfmöglichkeiten, sobald ihre Funktionsweise bekannt wird, in das Zusammenspiel von Generator und Diskriminator eingebaut werden können, um das künftige Unterscheiden zwischen authentischen und nichtauthentischen Aufnahmen erneut zu untergraben (International Risk Governance Center 2019: 17). Diffusion Models unterlaufen dagegen die Unterscheidung zwischen Original und manipuliertem Material, weil sie die in ihren riesigen Trainingsdatensätzen enthaltenen ursprünglichen Bilder im ersten Schritt zerstören, um auf dieser Basis scheinbar neue Bilder generieren zu können. In der Folge verliert die Frage danach, welches Original dem manipulierten Bild zugrunde liegt, ihren Sinn.

Die Erkennung von Deepfakes durch visuelle Prüfung war in deren Anfangszeit möglich, als sie sich durch gelegentliche kleine Fehler verrieten. Die inzwischen erreichbare Qualität von Deepfakes macht dagegen technische Lösungen unausweichlich, wenn das Ziel ist, Manipulationen in Video- und Audiomaterial zu erkennen. Auch hierbei kommen Varianten künstlicher neuronaler Netzwerke zum Einsatz, die darauf trainiert werden, bestimmte Hinweise auf Manipulationen zu erkennen. Solche Hinweise sind Fehler auf der Ebene einzelner Pixel, Inkonsistenzen zwischen Bild und vorhandenen Metadaten oder Unstimmigkeiten zwischen einzelnen Bereichen eines Gesichtes und dem Rest des Bildes (Farid/Schindler 2020: 18; Corvi et al. 2023). Andere Lösungen suchen Unstimmigkeiten in Mustern nicht auf der Ebene bestimmter Details von Bildern, sondern konzentrieren sich darauf, Vergleiche mit weiterem Material vorzunehmen, um Inkonsistenzen bei bestimmten Bewegungsabläufen, Verhaltensmustern oder auch Phonemen auszumachen (Farid/Schindler 2020: 19). Die forensische Suche nach dem Nichtauthentischen scheint aller-

dings auf ein Wettrennen von Hase und Igel hinauszulaufen, bei dem alle forensischen Bemühungen direkt die Weiterentwicklung von Deepfakes vorantreiben (Smith/Mansted 2020: 15).

Vor diesem Hintergrund konzentriert sich eine zweite Gruppe von technischen Lösungen darauf, die Herkunft von Bild- und Tonmaterial zu prüfen und zu authentifizieren. Ein Vorschlag setzt auf die Hardware, also die Einführung von digitalen Signaturen für Kameras, die die mit ihnen aufgenommenen Bilder als Originale markieren, mit denen spätere Versionen dieser Bilder verglichen werden können. Das setzt allerdings voraus, dass solche Metadaten beim Hochladen auf Plattformen nicht verloren gehen, was ebenfalls abgesichert werden muss (International Risk Governance Center 2019: 18). Im Sinne eines *digital fingerprinting* ist das Ziel hier, Authentizität dadurch zu sichern, dass Material zurückverfolgt werden kann, und nicht rückverfolgbares Material als der Nichtauthentizität verdächtig zu behandeln (Daly 2020). Umgekehrt gibt es im Fall Generativer KI Initiativen, ihre Produkte durch unsichtbare digitale Wasserzeichen als solche kenntlich zu machen (OpenAI 2024). Auch Blockchain-Anwendungen werden als vielversprechend betrachtet, um die Herkunft von Bildern nachvollziehen zu können (Farid/Schindler 2020, 39; Smith/Mansted 2020, 16). Ein anderer Ansatz setzt darauf, öffentliche Suchmaschinen und gebührenpflichtige Medienarchive auf Bilder hin zu durchsuchen, die dem zu prüfenden Material ähneln, und auf diese Weise Deepfakes zu erkennen, die manipulierte Versionen dieser Bilder sind (International Risk Governance Center 2019: 18).

Radikalere Überlegungen laufen darauf hinaus, dass es in einer Welt, in der es nicht mehr möglich ist, authentische von nichtauthentischen Bildern zu unterscheiden, zumindest für politisch oder anderweitig exponierte Personen attraktiv werden könnte, sämtliche ihrer Handlungen und Äußerungen, auch in nichtöffentlichen Situationen, kontinuierlich, z. B. mittels Bodycams, zu dokumentieren (Chesney/Citron 2018). Im Falle der Verbreitung eines Deepfakes, so das Argument, wären Personen auf diese Weise in der Lage, dessen Nichtauthentizität nachzuweisen. Dahinter steht offensichtlich der Gedanke, dass für bestimmte Personen, insbesondere politische Akteure, Deepfakes besonders problematisch sind. Damit ist angedeutet, dass die Frage nach der Authentizität medialer Artefakte je nach kommunikativem Kontext, in dem auf diese Artefakte Bezug genommen wird, unterschiedlich bedeutsam ist.

3. Der Blick zurück

Das Problem, das in Deepfakes gesehen wird, lässt sich folgendermaßen zusammenfassen:

»In the digital world, a picture, a video, or a bit of audio isn't assessed on its own terms. You are no longer responsible for believing your own eyes or ears; only if you trust the person who produced the photograph should you conclude that it shows what it purports to show. Otherwise, you can guiltlessly conclude that it's a fake.« (Manjoo 2008: 84)

Der Umstand, dass diese Problembeschreibung bereits aus dem Jahr 2008 stammt, ist allerdings ein Hinweis darauf, dass Deepfakes möglicherweise eine weniger einschneidende Entwicklung sind, als es scheint. Die Diagnose, mit der Digitalisierung sei eine postfaktische Gesellschaft entstanden (ebd., passim), gab es schon zehn Jahre, bevor die ersten Deepfakes veröffentlicht wurden. Ganz ähnliche Sorgen wie jene im Zusammenhang mit Deepfakes verbanden sich mit Blick auf mediale Artefakte insbesondere mit Bildbearbeitungsprogrammen wie Photoshop (Ricchiardi 2007). Britt Paris und Joan Donovan (2019) richten den Blick noch weiter zurück und machen darauf aufmerksam, dass die Authentizität audiovisuellen Materials im Sinne von »Wahrheit« niemals selbstevident gewesen ist, sondern sich stets institutionellen Arrangements verdankt hat, die diese Wahrheit bezeugen. Sie unterscheiden zwischen Beweis- (*evidence*) und Ausdruckskraft (*expression*) von medialen Artefakten (ebd.: 25). Ihre Beweiskraft spielt dann eine Rolle, wenn es darum geht zu bestimmen, was wahr ist und was nicht. Ob ein bestimmtes Material eine Wahrheit beweist oder nicht, wurde immer schon durch Instanzen wie Gerichte, Journalistinnen und Journalisten oder Museen verhandelt und entschieden; die Beweiskraft medialer Artefakte war nie unvermittelt gegeben. Zur Veranschaulichung führen Paris und Donovan (ebd.: 17f.) das Beispiel der Fotografie an, die Ende des 19. Jahrhunderts nur zögerlich als Beweismittel vor Gerichten anerkannt wurde. Bilder sprechen nie für sich selbst, sondern werden – mit Blick auf politische Ereignisse traditionell von Journalistinnen und Journalisten – gedeutet, eingeordnet und auf ihre Aussagekraft hin ausgewählt oder verworfen.

Darüber hinaus machen Paris und Donovan auf Arten von Manipulationen aufmerksam, die lange vor dem Aufkommen von Deepfakes und zuweilen höchst wirkungsvoll Videomaterial mit einfachen Mitteln verändert ha-

ben. Ein Zusammenschnitt von Material aus verschiedenen alten Quellen kann ausreichen, um als vermeintlicher Beleg für ein aktuelles Ereignis, z.B. eine Demonstration oder einen Unglücksfall, präsentiert zu werden. Ebenso kann das verlangsamte Abspielen authentischen Materials Effekte erzeugen und Interpretationen nahelegen, die bei Normalgeschwindigkeit abwegig erscheinen würden. Ein Beispiel aus dem Jahr 2019 ist der Fall des manipulierten Videos, das eine betrunken wirkende oder jedenfalls merkwürdig schleppend sprechende Nancy Pelosi, die damalige Sprecherin des US-Repräsentantenhauses, zeigte. Der Erfolg dieses Videos beruhte gerade nicht auf einem raffinierten Deepfake, sondern auf einer einfachen und relativ einfach durchschaubaren Manipulation – zur Kontrastierung auch Cheapfake genannt.

Neu an Deepfakes ist in erster Linie, dass die Manipulationen die im Video gezeigten Körper selbst, deren Aussehen und Stimme, tiefgreifend verändern (ebd.: 7). Hinzu kommt, dass sie – ebenso wie simplere Manipulationen oder Falschmeldungen in Textformen – im Zeitalter von Internet und Social Media sich in zuvor ungekannter Schnelligkeit und potentieller Reichweite verbreiten lassen. Die große Aufmerksamkeit für die Möglichkeiten von Deepfakes ist insofern zunächst damit zu erklären, dass sie ein besonders exponiertes Beispiel für Desinformationen darstellen, die in einem Mediensystem, das nicht länger von den traditionellen Massenmedien Fernsehen, Radio und Zeitung dominiert ist, zu einem besonderen Problem geworden zu sein scheinen (Benkler/Faris/Roberts 2018).

Andererseits ist diese Aufmerksamkeit aber eben nur die jüngste Iteration der Sorge darüber, dass die politische Öffentlichkeit bewusst falsch informiert wird. Der zu Beginn des Abschnitts zitierte Farhad Manjoo gab seinem Buch schon 2008 den Titel »True Enough. Learning to Live in a Post-Fact Society«. Geht man noch weiter zurück, nämlich in die 1990er und frühen 2000er Jahre, trifft man im Zusammenhang mit Politik ebenfalls auf die Schwierigkeit, Wahrheit und Lüge zu unterscheiden. Damals standen für dieses Problem Begriffe wie Mediendemokratie oder *packaging politics* (Franklin 1994; Meyer 2002a), und die Problemdiagnose konzentrierte sich auf die Abhängigkeit der Politik von der massenmedialen Berichterstattung und die professionalisierte politische PR, die darauf reagierte. Politischer Journalismus berichtet demnach zum überwiegenden Teil über Ereignisse, die in Erwartung massenmedialer Berichterstattung von politischen Akteuren und ihren Beraterinnen und Beratern geplant und vorbereitet worden sind (Kepplinger 2007). Damit geht gleichzeitig der Verdacht einher, dass das, was in der Politik tatsächlich passiert, in den Massenmedien und den von ihnen produzierten Bildern gerade

nicht vorkommt, womit der Öffentlichkeit insofern ein falsches Bild nicht einzelner politischer Akteure, sondern des gesamten politischen Prozesses präsentiert wird (Meyer 2002b: 7). Befürchtet wird, dass Herstellung und Darstellung von Politik zunehmend auseinanderfallen (Sarcinelli 1994).

Die Beschreibung des Verhältnisses von Politik und Massenmedien als Mediendemokratie stellte grundlegend in Frage, dass Massenmedien und die von ihnen produzierten Artefakte – insbesondere Fernsehbilder – authentische politische Ereignisse zeigen. Die Rolle politischer PR führt zur Diagnose der Nichtauthentizität der Berichterstattung, weil diese sich auf Ereignisse konzentriert, die ohne die antizipierte Beobachtung durch Massenmedien gar nicht stattgefunden hätten, z.B. Pressekonferenzen, oder jedenfalls nicht so stattgefunden hätten, sondern stattdessen weniger theatralisch und mit mehr inhaltlicher Tiefe. Polemisch zugespitzt brachte das ein Gastbeitrag, der 2002 in der Süddeutschen Zeitung aus Anlass des Bundestagswahlkampfes und des Fernsehduells zwischen Gerhard Schröder und Edmund Stoiber erschien, folgendermaßen auf den Punkt:

»Wir leben in einer Mediendemokratie, will in etwa heißen, die Medien simulieren Demokratie und üben sie dabei de facto aus. Mediendemokratie, das ist Verpackungsdemokratie; wer die Korsettstrukturen von Fernsehduellen geißelt und durch »echte«, womöglich »substantielle Debatten« ersetzt sehen möchte, der hat einfach noch nicht begriffen, dass wir längst im Zeitalter der Wirkungsästhetik angekommen sind.« (Politycki 2002)

Die Idee der authentischen, wahren Darstellung ist also gerade im Zusammenhang mit Politik schon lange prekär gewesen (Kusche 2006). Vor der Verbreitung von Internet und Social Media war es insbesondere die massenmediale Berichterstattung im Fernsehen, die im Verdacht der Nichtauthentizität stand. Ihre Bilder galten zwar als authentisch im technischen Sinne, wurden gleichzeitig aber dahingehend als nichtauthentisch kritisiert, als sie ein vereinfachtes, emotionalisiertes und insofern falsches Bild der Politik zeichnen. Mit Hilfe der Unterscheidung von Herstellung und Darstellung wurde ein Realitätsgefälle markiert, dem zufolge das Herstellen kollektiv verbindlicher Entscheidungen das eigentlich Politische ist, während die mediale Darstellung von Politik auf stets defizitäre Weise darüber lediglich informiert (Japp/Kusche 2004, 515–518).

Aus differenzierungstheoretischer Perspektive gibt es hingegen keinen Anlass, politischen Operationen einen privilegierten ontologischen Status

zuzusprechen. Der Unterschied zwischen Politik und Massenmedien liegt nicht in ihrem Realitätsgehalt, sondern in der Orientierung an Gelegenheiten, kollektiv verbindliche Entscheidungen herzustellen, einerseits (Luhmann 2000) und der Orientierung an neuen Informationen andererseits (Luhmann 1996). Wahrheit spielt in beiden Zusammenhängen eine Rolle als Wert, was bedeutet, dass es riskant ist, diesen durch Lügen und Täuschung zu verletzen, weil dies zum Verlust von Reputation führen *kann*. Das ist aber in der Politik wie in den Massenmedien nur insoweit der Fall, wie sich das Vorliegen einer Lüge bzw. Täuschung als wahrer Sachverhalt feststellen lässt.

Die Überlegungen von Paris und Donovan (2019) zur Authentizität medialer Artefakte müssen vor diesem Hintergrund radikalisiert werden: Für Kommunikationen in der Politik (wie in den Massenmedien) gilt generell, unabhängig vom Medium, in dem sie erfolgen, dass es von institutionellen Arrangements abhängt, ob die Wahrheit von politischen Aussagen bezeugt, bezweifelt oder überhaupt zum Thema gemacht wird, also die Beweiskraft medialer Artefakte ins Spiel kommt. Die Unterscheidung zwischen Herstellung und Darstellung von Politik markiert, dass sowohl von Seiten der Politik als auch von Seiten der über sie berichtenden Massenmedien immer schon damit gerechnet wird, dass Wahrheit oder Authentizität nicht selbstevident sind. Schon jede Auskunft der Beteiligten darüber, was in einem bestimmten Meeting gesagt oder nicht gesagt wurde, ist eine Darstellung, die unvermeidlich vereinfachen, verschönern oder auch dramatisieren wird. Im strengen Sinne widerlegen lässt sie sich nicht; allenfalls besondere Verfahren, wie z.B. Untersuchungsausschüsse, können das leisten, aber ebenfalls nicht im Sinne einer authentischen Rekonstruktion des Gesagten und Vorgefallenen, sondern nur als ein besonderes institutionelles Arrangement, das in Ausnahmefällen herangezogen wird, um den Spielraum von Darstellungen so weit wie möglich einzuengen (vgl. Japp/Kusche 2004: 526). Jenseits davon bleibt nur die öffentliche Meinung, in der politische Akteure sich mehr oder weniger überzeugend präsentieren können und im Zuge dessen gegebenenfalls anderen Akteuren auch Täuschung oder Lügen vorwerfen. Wohlgemerkt beruht Überzeugungskraft aber nicht primär auf der Wahrheit von politischen Äußerungen, denn politische Kommunikation ist nicht auf deren Überprüfung angelegt; sie ist darauf angelegt, Politik so darzustellen, dass künftige Machtchancen, d.h. in erster Linie die Chance, Wählerstimmen zu gewinnen, gewahrt und ausgebaut werden. Überzeugungskraft kann sich daher auch aus der Ausdruckskraft politischer Äußerungen ableiten, also daraus, dass eine emotionale Sprache, patriotische Appelle oder auch die Verunglimpfung bestimmter Gruppen oder

politischer Gegner für Teile des politischen Publikums anschlussfähig sind, weil sie Identifikationspotentiale schaffen (Hamlin/Jennings 2011: 653). Der Unterscheidung zwischen Beweis- und Ausdruckskraft medialer Artefakte (Paris/Donovan 2019) entspricht folglich in der politischen Kommunikation eine Unterscheidung zwischen Wahrheit und Ausdruckskraft politischer Äußerungen. Beide Unterscheidungen sind erst einmal symmetrisch, so dass es vom Kontext abhängt, ob die eine oder die andere Seite präferiert wird. In der durch Massenmedien dominierten politischen Kommunikation ermöglicht allerdings die Unterscheidung von Herstellung und Darstellung dank der in sie eingebauten Asymmetrie, an der Idee einer Differenz zwischen Realität und Inszenierung und damit einer Differenz zwischen Wahrheit und Lüge bzw. Authentizität und Nichtauthentizität festzuhalten, obwohl Politik gerade nicht über einen Mechanismus verfügt, um die entsprechende Unterscheidung routinemäßig zu treffen.

In einem hybriden Mediensystem (Chadwick 2017), in dem neben Massenmedien Social Media eine zentrale Rolle spielen, funktioniert die Unterscheidung von Herstellung und Darstellung nicht mehr in dieser Weise. Es sind solche hybriden Mediensysteme, in denen Deepfakes mit Blick auf die Unterscheidung zwischen Authentizität und Nichtauthentizität gegebenenfalls zum Problem werden.

4. Mobile Kommunikation und Social Media

Im Vergleich zu dem, was Anfang der 2000er Jahre als Mediendemokratie beschrieben wurde, scheint sich mit der Verbreitung von Smartphones der Darstellungsspielraum politischer Kommunikation zu verringern. Im Internet verbreitete Videos, die nicht von Journalistinnen oder Journalisten, sondern von beliebigen Personen bei Politikerauftritten abseits der massenmedialen Öffentlichkeit aufgenommen wurden, stellen zunächst mehr Authentizität in Aussicht (Mann/Ferenbok 2013). Allerdings antizipieren professionalisierte Politik- und Wahlkampfberater auch diese Möglichkeit und passen die politische PR und das entsprechende Training politischer Akteure an. Die Erwartungen an deren professionelles Auftreten weiten sich in der Folge zeitlich und räumlich aus (Serazio 2015).

»Candidates have to be more disciplined, because events are taped that can go online, and you can't go out to the gun club and tell them, ›Oh,

I love guns.« And then go to the library and say that you're for gun control, which a lot of people probably did in the past.« (Interview mit einem US-amerikanischen Politikberater 2012, zit. in Serazio 2015: 1917)

Die Antizipation, dass nichtprofessionelle Akteure mit einfachsten Mitteln nahezu jede Situation, in der sich Politikerinnen und Politiker wiederfinden, in eine öffentliche Situation verwandeln können, führt also zu erhöhten Konsistenzanforderungen. Das betrifft sowohl den Sachgehalt von politischen Äußerungen als auch ihre Expressivität. Mit dem relativen Bedeutungsgewinn von Social Media im Verhältnis zu traditionellen Massenmedien geht einher, dass die Unterscheidung von Herstellung und Darstellung für die Beobachtung von Politik an Bedeutung verliert. Wenn aus Sitzungen heraus getwittert werden kann, wird die Behauptung einer Differenz zwischen »eigentlicher« Politik und ihrer, stets defizitären, Darstellung prekär; jedenfalls spielt sie in gegenwärtigen Beschreibungen demokratischer Politik praktisch keine Rolle. Auch Politikerinnen und Politiker können nun direkt mediale Artefakte für ein generalisiertes Publikum produzieren, so dass augenfällig wird, was im Grunde immer schon der Fall war, nämlich dass jede politische Kommunikation einen Herstellungs- und einen Darstellungsaspekt hat (Japp/Kusche 2004).

In der Ära der Massenmedien konnte es plausibel erscheinen, Darstellung als nachträgliche Vermittlung und Inszenierung genuiner, mit der Herstellung politischer Entscheidungen befasster Politik zu begreifen (ebd.: 515). Mit der Verbreitung von Social Media wird es dagegen zunehmend abwegig, Darstellungszwänge politikextern einem System der Massenmedien und seinen professionalisierten Akteuren zuzurechnen. Sowohl Politikerinnen und Politiker als auch massenmediale wie politische Laien machen von den Möglichkeiten mobiler Kommunikation und Social Media Gebrauch, um neue Informationen mit Politikbezug zu verbreiten. Im Zuge dessen verliert die Zuschreibung eines irgendwie höheren Realitätsgehaltes auf Seiten der Politik, die mit dem Gebrauch der Unterscheidung von Herstellung und Darstellung einherging, an Plausibilität. Die Antizipation nahezu ständiger Beobachtbarkeit untergräbt die Vorstellung, es gäbe einen geschützten Bereich, in dem Politik inszenierungsfrei einfach »gemacht« würde und der insofern authentischer sei als das, was man von der Politik – in Social Media und anderswo – zu lesen, zu hören und zu sehen bekommt. Die zeit- und ortsunabhängige Konsistenz von Äußerungen, die demselben politischen Akteur zugeschrieben werden, wird damit zum zentralen Ausweis von Authentizität (Luebke 2020; Friedman/Kampf

2020), sowohl mit Blick auf den Sachgehalt oder die Wahrheit von Aussagen als auch mit Blick auf ihre Ausdruckskraft.

Vor diesem Hintergrund lässt sich nun präzisieren, worin genau die spezifische Problematik von Deepfakes in der Politik liegt. In Anbetracht des Primats der Konsistenz lassen sich auf Basis der Listen zu möglichen böswilligen Verwendungsweisen von Deepfakes (z.B. Chesney/Citron 2018) zwei Arten von erwarteten Effekten unterscheiden. Die erste Möglichkeit besteht darin, dass Politikerinnen und Politiker bei Fehlritten und kommunikativen Entgleisungen gezeigt werden, die ihnen tatsächlich nicht unterlaufen sind. Der erwartete Effekt besteht hier in der Untergrabung jener Authentizität, die in Zeiten von Social Media maßgeblich auf der zeit- und ortsunabhängigen Konsistenz der Handlungen und Kommunikationen politischer Akteure beruht. Ein Beispiel dafür, wenn auch in Form eines Cheapfakes, ist das oben erwähnte Video Pelosis, das den Eindruck erweckte, sie wäre betrunken oder aus anderen Gründen ihrer Sprache nicht mächtig. Die zweite Möglichkeit beruht dagegen darauf, ein Deepfake zu präsentieren, das weitestgehend den Anschein konsistenter politischer Kommunikation hat, aber etwas verkündet, das massive Reaktionen anderer Akteure hervorruft. Ein fiktives Beispiel dafür wäre ein scheinbar offizielles Statement des amerikanischen Präsidenten, ein Angriff auf das mit Atomwaffen ausgestattete Nordkorea wäre gerade im Gange.

Allerdings hängt es von den Erwartungen der jeweiligen Beobachter ab, welches Verhalten eines politischen Akteurs als konsistent mit vergangenem Verhalten eingeordnet wird. Wer Pelosi schon immer für inkompetent oder gar Schlimmeres gehalten hat, wird ein Video, das sie scheinbar in betrunkenem Zustand zeigt, eher als Bestätigung der eigenen Erwartungen zu dieser Politikerin betrachten. Erste empirische Untersuchungen zu den Wirkungen von Deepfakes deuten darauf hin, dass individuelle Einschätzungen zu ihrer Glaubwürdigkeit nicht von ihrer technischen Güte abhängen, sondern vom vorhandenen Wissen über die gezeigte Person und von der politisch-ideologischen Affinität zu ihr (Hameleers/van der Meer/Dobber 2023; 2024). Die möglichen Effekte von Deepfakes lassen sich daher nicht isoliert mit Blick auf das mediale Artefakt verstehen, sondern nur, wenn dieses Artefakt als ein kommunikatives Ereignis begriffen wird, das an vorangegangene Kommunikationen anschließt, auf mögliche künftige Kommunikationen verweist und im Zuge dessen Erwartungen bestätigt oder enttäuscht (Luhmann 1984). Ein Deepfake, das über Social Media verschickt wird, ist nur als Element einer Kette aneinander anknüpfender Kommunikationen interpretierbar und wird unterschied-

lich interpretiert werden, je nachdem, wie es sich mit anderen Kommunikationen verknüpfen lässt.

Kommunikationen in Social Media sind nun – anders als in Massenmedien – an individuelle Nutzerinnen und Nutzer angepasste Kommunikationen. Algorithmen personalisieren kommunikative Anschlüsse, indem sie aus vergangenen Online-Kommunikationen des jeweiligen Nutzers und anderer Nutzer, die sie in bestimmten Hinsichten als ähnlich kategorisieren, ableiten, welche Kommunikationen – sowohl Nachrichten als auch Werbung – wahrscheinlich kommunikative Anschlüsse bei dem betreffenden Nutzer erzeugen werden (Soffer 2019; Kusche 2020: 5–6). Das sogenannte Microtargeting ermöglicht es zahlenden Klienten, Werbeinhalte ausschließlich jenen Nutzerinnen und Nutzern zu zeigen, denen aufgrund der über sie und ihre Online-Aktivitäten gesammelten Daten bestimmte Merkmale und Interessen algorithmisch zugeschrieben werden. Als politisches Microtargeting funktioniert dies auch für politische Werbung, also bezahlte politische Kommunikation, die bestimmten, gegebenenfalls auch sehr kleinen, Zielgruppen präsentiert wird. Im Falle sogenannter organischer, das heißt von Nutzerinnen und Nutzern selbst erzeugter Inhalte sind es Empfehlungsalgorithmen, die darüber (mit-)bestimmen, wem welche geteilten Nachrichten wie prominent angezeigt werden.

Vor diesem Hintergrund erscheint die Kombination von Deepfakes und politischem Microtargeting auf den ersten Blick als eine besonders effektive Form der Desinformation. Erste wissenschaftliche Experimente (Dobber et al. 2020) dazu kommen allerdings zu dem Schluss, dass man sich zwar über Deepfakes im Allgemeinen Sorgen machen sollte, weil sie von Probandinnen und Probanden selbst bei mäßiger Qualität oft nicht erkannt werden. Das Microtargeting von Deepfakes scheitert dagegen zumindest derzeit an mangelnder Präzision, so dass die Personen, denen ein Deepfake gezielt gezeigt wird, tatsächlich nur selten die Merkmale aufweisen, die sie für das Deepfake vermeintlich besonders empfänglich machen.

Auch ohne Microtargeting impliziert allerdings die Personalisierung von Inhalten, die Nutzerinnen und Nutzer auf Social-Media-Plattformen gezeigt bekommen, dass die der betreffenden Nutzerin algorithmisch zugeschriebenen Interessen und Vorlieben zu einer Konsistenz der Inhalte über die Zeit führen. Dem entspricht die weitverbreitete Besorgnis über Echokammern (Sunstein 2017) und Filterblasen (Pariser 2012), die davon ausgeht, dass Social-Media-Plattformen maßgeblich dazu beitragen, dass Nutzerinnen und Nutzer fast nur noch mit Kommunikation zu tun haben, die ihre bestehenden Ansichten bestätigt. Wie stark solche Filterblaseneffekte tatsächlich sind und wer

von ihnen betroffen ist, ist allerdings umstritten (Dubois/Blank 2018). Bisherige empirische Befunde deuten darauf hin, dass es sich allenfalls um eine, gemessen an allen Nutzerinnen und Nutzern, sehr kleine Gruppe handelt, die in medialen Echokammern isoliert ist (Barberá et al. 2015; Eady et al. 2019).

Wenn die Erwartungen eines Beobachters dafür ausschlaggebend sind, ob das Verhalten eines in einem Deepfake dargestellten Akteurs als konsistent mit dessen früherem Verhalten betrachtet wird, diese Erwartungen aber nach wie vor nicht überwiegend das Ergebnis eines einseitigen personalisierten Konsums von Social-Media-Posts sind, verschärfen Social Media das Problem von Deepfakes als Kommunikationsereignis nicht prinzipiell. Wer durch die Darstellung seine Erwartungen bestätigt sieht, wird mit weiteren Kommunikationen – z. B. durch Teilen oder Kommentieren – anders anschließen als jemand, der vom Inhalt überrascht ist. Da ein Deepfake nicht ohne Weiteres als solches zu erkennen ist, wird die individuelle Nutzerin sich die Frage, ob ein Video manipuliert ist oder nicht, entsprechend nur vor dem Hintergrund ganz bestimmter Erwartungen stellen. Durch bloße Betrachtung des medialen Artefaktes lässt sich die Frage, sofern sie gestellt wird, nicht beantworten; eine Antwort wird die Nutzerin daher nur finden, wenn es wiederum institutionelle Arrangements (Paris/Donovan 2019) gibt, die den Anspruch haben, die Authentizität solcher Artefakte zu prüfen, *und* sie diesen Arrangements vertraut.

5. Schlussfolgerungen

Angesichts der inzwischen erreichten Qualität von Deepfakes lässt sich auf Manipulationen in solchem Video- und Audiomaterial nur mit Hilfe von technischen Lösungen schließen. Diese zielen entweder auf eine forensische Prüfung oder auf die Nachvollziehbarkeit der Herkunft des Materials ab. Für ein Publikum werden entsprechende Prüfergebnisse aber nur relevant, wenn sie es in jene Medien schaffen, die dieses Publikum nutzt, und das Publikum den Informationen mehr vertraut als dem Augenschein eines Deepfake und dem kommunikativen Kontext, in den es eingebettet ist. Damit ist das Problem von Deepfakes mit Blick auf die politische Öffentlichkeit weder größer noch kleiner als das Problem mit anderen Formen der Desinformation. Hier ist bislang keineswegs klar, in welchem Maße Effekte von Falschmeldungen darauf beruhen, dass das Publikum den Quellen zu sehr vertraut, oder darauf, dass das Publikum allen Quellen so stark misstraut, dass Ausdruckskraft statt Evi-

denz zum ausschlaggebenden Kriterium wird, wenn es darum geht, an welchen Nachrichten man sich orientiert (MacMullen 2020).

Deepfakes sind insofern nicht die Vorboten eines Zeitalters der Desinformation (Bettilyon 2018), sondern ein Symbol für die neuen Möglichkeiten zur Desinformation, die sich in einem hybriden Mediensystem (Chadwick 2017) ergeben. Der Fokus auf technische Lösungen kann daher in die Irre führen. Er suggeriert, dass das zentrale Problem darin liege, dass Deepfakes nicht als solche erkannt werden. Dieses Problem der Kognition stellt sich eindeutig für die etablierten institutionellen Arrangements journalistischer Evidenzprüfung, für die Deepfakes natürlich eine zusätzliche Herausforderung sind (Silverman 2020). Es stellt sich gegebenenfalls auch für im Entstehen begriffene institutionelle Arrangements zur Moderation der Inhalte von Social-Media-Plattformen, sofern diese darauf abzielen, mehr Verantwortung für Inhalte zu übernehmen. Technische Lösungen können in solche Arrangements eingebaut werden; ihre Beweiskraft in öffentlichen Debatten hängt dann vom Vertrauen in diese Arrangements ab, also in die Organisationen, die solche technischen Lösungen anwenden oder anbieten. Ihre Beweiskraft hängt damit aber auch davon ab, ob das Publikum an Wahrheit überhaupt maßgeblich interessiert ist. Ein Publikum, das die Orientierung an Wahrheit in der politischen Kommunikation wertschätzt, ist ein Publikum, das vom Augenschein eines Deepfakes gegebenenfalls getäuscht werden kann, aber gleichzeitig institutionelle Arrangements zur Evidenzprüfung wertschätzt und diesen – in Maßen – vertraut. Ein Publikum dagegen, das ausschließlich die Ausdruckskraft politischer Kommunikation wertschätzt, ist ein Publikum, für das es keinen Unterschied macht, ob ein Deepfake authentisch oder nicht authentisch im Sinne von Wahrheit oder Herkunft ist, und für das institutionelle Arrangements zur Evidenzprüfung daher irrelevant sind. Das Besondere an der Sorge um Deepfakes scheint zu sein, dass sie scheinbar die erste Art von Publikum im Auge hat, sich aber tatsächlich aus der Angst vor der zweiten Art von Publikum speist.

Literatur

Barberá, Pablo et al. (2015): »Tweeting From Left to Right: Is Online Political Communication More Than an Echo Chamber?«, in: *Psychological Science* 26.10, S. 1531–1542, DOI: 10.1177/0956797615594620.

- Bateman, Jon (2020): Deepfakes and Synthetic Media in the Financial System: Assessing Threat Scenarios, online verfügbar unter <https://carnegieendowment.org/2020/07/08/deepfakes-and-synthetic-media-in-financial-system-assessing-threat-scenarios-pub-82237>, abgerufen am 26.04.2024.
- Benkler, Yochai/Faris, Rob/Roberts, Hal (2018): *Network propaganda: manipulation, disinformation, and radicalization in American politics*, New York: Oxford University Press.
- Bettilyon, Tyler Elliot (2018): Deep Fakes and The Future of Propaganda, online verfügbar unter <https://medium.com/@TebbaVonMathenstien/deep-fakes-and-the-future-of-propaganda-18a257026571>, abgerufen am 26.04.2024.
- Bovenschulte, Marc (2019): *Deepfakes – Manipulation von Filmsequenzen*, Berlin: TAB – Büro für Technikfolgen-Abschätzung beim Deutschen Bundestag.
- Chadwick, Andrew (2017): *The hybrid media system: Politics and power*, New York: Oxford University Press.
- Chesney, Bobby/Citron, Danielle (2018): »Deepfakes: A Looming Crisis for National Security, Democracy and Privacy?«, in: *California Law Review* 107, S. 1753–1820.
- Corvi, Riccardo et al. (2023): »Intriguing Properties of Synthetic Images: From Generative Adversarial Networks to Diffusion Models«, in: 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Vancouver, B.C.: IEEE, S. 973–982, DOI: 10.1109/CVPRW59228.2023.00104.
- Daly, Kyle (2020): »System for Thwarting Deepfakes Unveiled«, in: *Axios* vom 03.08.2020, online verfügbar unter <https://www.axios.com/system-for-thwarting-deepfakes-unveiled-d8fc760c-9a44-43fb-a92e-8a86498414d6.html>, abgerufen am 26.04.2024.
- Dobber, Tom et al. (2020): »Do (Microtargeted) Deepfakes Have Real Effects on Political Attitudes?«, in: *The International Journal of Press/Politics* 1940161220944364, DOI: 10.1177/1940161220944364.
- Dommett, Katharine/Temple, Luke (2018): »Digital Campaigning: The Rise of Facebook and Satellite Campaigns«, in: *Parliamentary Affairs* 71, S. 189–202.
- Dubois, Elizabeth/Blank, Grant (2018): »The echo chamber is overstated: the moderating effect of political interest and diverse media«, in: *Information, Communication & Society* 21.5, S. 729–745, DOI: 10.1080/1369118X.2018.1428656.

- Eady, Gregory et al. (2019): »How Many People Live in Political Bubbles on Social Media? Evidence From Linked Survey and Twitter Data«, in: *SAGE Open* 9.1, 2158244019832705, DOI: 10.1177/2158244019832705.
- Farid, Hany/Schindler, Hans-Jakob (2020): *Deep Fakes: On the Threat of Deep Fakes to Democracy and Society*, online verfügbar unter <https://www.kas.de/de/einzeltitel/-/content/die-gefahr-von-deep-fakes-fuer-unsere-demokratie>, abgerufen am 26.04.2024.
- Franklin, Bob (1994): *Packaging Politics: Political Communication in Britain's Media Democracy*, London: Edward Arnold.
- Friedman, Elie/Kampf, Zohar (2020): »To Thine Own Self Be True«: The Perceived Meanings and Functions of Political Consistency«, in: *Language in Society* 49.1, S. 89–113, DOI: 10.1017/S004740451900068X.
- Hameleers, Michael/van der Meer, Toni G. L. A./Dobber, Tom (2023): »They Would Never Say Anything Like This! Reasons To Doubt Political Deep-fakes«, in: *European Journal of Communication* 39.1, S. 56–70, DOI: 10.1177/02673231231184703.
- (2024): »Distorting the Truth versus Blatant Lies: The Effects of Different Degrees of Deception in Domestic and Foreign Political Deepfakes«, in: *Computers in Human Behavior* 152, 108096, DOI: 10.1016/j.chb.2023.108096.
- Hamlin, Alan/Jennings, Colin (2011): »Expressive Political Behaviour: Foundations, Scope and Implications«, in: *British Journal of Political Science* 41.3, S. 645–670.
- International Risk Governance Center (2019): *Forged Authenticity: Governing Deepfake Risks*, online verfügbar unter <https://www.epfl.ch/research/domains/irgc/specific-risk-domains/projects-cybersecurity/forging-authenticity-governing-deepfake-risks/>, abgerufen am 26.04.2024.
- Japp, Klaus P./Kusche, Isabel (2004): »Die Kommunikation des politischen Systems: Zur Differenz von Herstellung und Darstellung im politischen System«, in: *Zeitschrift für Soziologie* 33.6, S. 511–531.
- Kalpokas, Ignas (2020): »Problematising Reality: The Promises and Perils of Synthetic Media«, in: *SN Social Sciences* 1.1, 1, DOI: 10.1007/s43545-020-00010-8.
- Keplinger, Hans Mathias (2007): »Reciprocal Effects. Toward a Theory of Mass Media Effects on Decision Makers«, in: *Press/Politics* 12.2, S. 3–23.
- Klonick, Kate (2021): »Inside the Making of Facebook's Supreme Court«, in: *The New Yorker* vom 12.02.2021, online verfügbar unter <https://www.newyorker.com/tech/annals-of-technology/inside-the-making-of-facebooks-supreme-court>, abgerufen am 26.04.2024.

- Kusche, Isabel (2006): »Wa(h)re Politik? Politische Kommunikationsforschung und Marketingtheorie«, in: Jörn Lamla/Sighard Neckel (Hg.), *Politisierter Konsum – konsumierte Politik*, Wiesbaden: VS, S. 281–299.
- (2020): »The Old in the New: Voter Surveillance in Political Clientelism and Datafied Campaigning«, in: *Big Data & Society* 7.1, 2053951720908290, DOI: 10.1177/2053951720908290.
- Luebke, Simon M. (2020): »Political Authenticity: Conceptualization of a Popular Term«, in: *The International Journal of Press/Politics* 26.3, S. 635–653, DOI: 10.1177/1940161220948013.
- Luhmann, Niklas (1984): *Soziale Systeme*, Frankfurt a.M.: Suhrkamp.
- (1996): *Die Realität der Massenmedien*. 2., erw. Aufl., Opladen: Westdeutscher Verlag.
- (2000): *Die Politik der Gesellschaft*, Frankfurt a.M.: Suhrkamp.
- MacMullen, Ian (2020): »Survey Article: What Is ›Post-factual‹ Politics?«, in: *Journal of Political Philosophy* 28.1, S. 97–116, DOI: 10.1111/jopp.12212.
- Manjoo, Farhad (2008): *True enough: learning to live in a post-fact society*, Hoboken, N.J.: Wiley.
- Mann, Steve/Ferenbok, Joseph (2013): »New Media and the Power Politics of Sousveillance in a Surveillance-Dominated World«, in: *Surveillance & Society* 11.1-2, S. 18–34, DOI: 10.24908/ss.v11i1/2.4456.
- Matney, Lucas (2021): »Twitter’s Decentralized Future«, in: *TechCrunch* vom 15.01.2021, online verfügbar unter <https://techcrunch.com/2021/01/15/twitters-vision-of-decentralization-could-also-be-the-far-rights-internet-endgame/>, abgerufen am 04.07.2024.
- Meyer, Thomas (2002a): *Media Democracy. How the Media Colonize Politics*, Cambridge: Polity.
- (2002b): »Mediokratie – Auf dem Weg in eine andere Demokratie?«, in: *Aus Politik und Zeitgeschichte* 15–16, S. 7–14.
- OpenAI (2024), *G2PA in DALL·E 3*, online verfügbar unter <https://help.openai.com/en/articles/8912793-c2pa-in-dall-e-3>, abgerufen am 26.04.2024.
- Paris, Britt/Donovan, Joan (2019): *Deepfakes and Cheap Fakes*, online verfügbar unter https://datasociety.net/wp-content/uploads/2019/09/DS_Deepfakes_Cheap_FakesFinal-1.pdf, abgerufen am 26.04.2024.
- Pariser, Eli (2012): *The Filter Bubble: What the Internet Is Hiding from You*, London: Penguin.
- Plöchinger, Stefan (2015): »Böhmermann zeigt Erregungsdeutschland den Stinkefinger«, in: *Süddeutsche Zeitung* vom 19.03.2015, online verfügbar unter <https://www.sueddeutsche.de/medien/varoufakis-video-boeh>

- mermann-zeigt-erregungsdeutschland-den-stinkefinger-1.2400610, abgerufen am 26.04.2024.
- Politycki, Matthias (2002): »Der Tag, an dem die Tantentäuscher rotweiße Krautwatten trugen. Schreiber oder Stöder – welchen dieser mehr oder weniger willigen Provinzfürsten wollen wir denn haben?«, in: *Süddeutsche Zeitung* vom 10.09.2002.
- Ricchiardi, Sherry (2007): »Distorted picture: thanks to Photoshop, it's awfully easy to manipulate photographs, as a number of recent scandals make painfully clear. Misuse of the technology poses a serious threat to photojournalism's credibility«, in: *American Journalism Review* 29.4, S. 36–44.
- Sarcinelli, Ulrich (1994): »Mediale Politikdarstellung und politisches Handeln: analytische Anmerkungen zu einer notwendigerweise spannungsreichen Beziehung«, in: *Politische Kommunikation in Hörfunk und Fernsehen. Gegenwartskunde Sonderheft 8*, Opladen: Leske + Budrich, S. 35–50.
- Serazio, Michael (2015): »Qualitative Political Communication | Managing the Digital News Cyclone: Power, Participation, and Political Production Strategies«, in: *International Journal of Communication* 9, S. 1907–1925.
- Silverman, Craig (Hg.) (2020): *Verification Handbook – Das Handbuch zur Überprüfung von Desinformation und Medienmanipulation*, übers. v. Marcus Engert, Düsseldorf: Landesanstalt für Medien NRW, online verfügbar unter https://www.medienanstalt-nrw.de/fileadmin/user_upload/NeueWebsite_0120/Themen/Desinformation/VerificationHandbook_deutsch_2102_LFMNRW.pdf, abgerufen am 26.04.2024.
- Smith, Hannah/Mansted, Katherine (2020): *Weaponised Deep Fakes: National Security and Democracy*, Barton, A.C.T.: Australian Strategic Policy Institute.
- Soffer, Oren (2019): »Algorithmic Personalization and the Two-Step Flow of Communication«, in: *Communication Theory* 31.3, S. 297–315, DOI: 10.1093/ct/qtz008.
- Sunstein, Cass R. (2017): *#Republic: divided democracy in the age of social media*, Princeton, N.J./Oxford: Princeton University Press.