

Deepfakes als Chance für die Demokratie?

Deepfakes – synthetische audio-visuelle Medien menschlicher Gesichter, Körper oder Stimmen, die meist mit Hilfe Künstlicher Intelligenz (KI) erstellt werden – gelten im öffentlichen Diskurs zunehmend als Bedrohung der „Wahrheit in der Politik“ (Galston 08.01.2020, eigene Übersetzung) oder gar der Demokratie (vgl. etwa Thomas 2020). Wissenschaftler*innen befürchten ihren Einsatz zur Wahlmanipulation, in der Vergangenheit insbesondere im Zusammenhang mit den US-amerikanischen Präsidentschaftswahlen 2020 (vgl. Diakopoulos/Johnson 2019). Deepfakes könnten auch für einen gezielten Angriff auf politische Gegner*innen eingesetzt werden oder gesellschaftliche Gräben vertiefen (vgl. Chesney/Citron 2019, S. 1776). Sie können sogar inner- oder zwischenstaatliche Konflikte befeuern, wie die Kontroverse um ein angeblich gefälschtes Video des Präsidenten von Gabun 2018 zeigt (vgl. Ajder et al. 2019, S. 10). Zunehmend werden Deepfakes auch in ausländischen Desinformationskampagnen eingesetzt (vgl. FBI 10.03.2021). Grundsätzlicher befürchten Beobachter*innen, sie könnten das Vertrauen der Bürger*innen in geteilte Fakten und Wahrheiten – und damit eine zentrale Grundlage für den demokratischen Diskurs – untergraben. Chesney/Citron (2019, S. 1785) identifizieren hier auch das Phänomen der „Lügner*innendividende“: Personen des öffentlichen Lebens können Kritik an eigenem Fehlverhalten von sich weisen, indem sie die Echtheit belastender Beweise anzweifeln.

Beispiele für einen politisch schädigenden Einsatz von Deepfakes mehrten sich seit 2018 mit ihrer zunehmenden Verfügbarkeit und Kommerzialisierung, ihrer steigenden Qualität und den sinkenden technischen Hürden für ihre Erstellung (vgl. Schreiner 21.02.2021). Doch auch „profane“ Anwendungen nehmen zu, etwa für Wirtschaft und Unterhaltung (Hao/Heaven 24.12.2020, eigene Übersetzung). Diese werden in der Debatte häufig vernachlässigt: Ein systematischer Überblick über die Deepfake-Forschung 2018–2020 zeigt, dass diese sich überwiegend mit den Risiken von Deepfakes beschäftigt. Nur sechs der 37 untersuchten Studien analysieren

Chancen der Technologie, etwa in den Bereichen Virtual Reality, Kunst oder Pornografie¹ (vgl. Godulla et al. 2021, 81, 83f).

Im Folgenden werden solche Chancen von Deepfakes genauer beleuchtet – in dem Bereich, in dem sie die größten Befürchtungen wecken: der Politik.² Im Zentrum steht die Frage, ob Deepfakes auch Potenziale für die Demokratie bergen. Diese Frage wurde bisher in der wissenschaftlichen Literatur noch nicht eingehend untersucht. Der Artikel trägt zur Schließung dieser Forschungslücke aus einer ethischen und sozialwissenschaftlichen Perspektive bei. Er basiert auf einer Literatur- und Dokumentenanalyse von über 300 Beiträgen aus Wissenschaft, Unternehmen, Medien und der Zivilgesellschaft.³

1. Was sind Deepfakes?

Grundlage des Beitrags ist ein weites Verständnis von „Deepfakes“ als synthetische und manipulierte Bilder, Audiodateien oder Videos von Menschen, ihren Gesichtern oder Stimmen. Dazu zähle ich neben „FaceSwap“-Videos (der ursprünglich als „Deepfakes“ bezeichneten Technologie), bei denen das Gesicht einer Person im Video mit dem einer anderen getauscht wird, auch Videos, bei denen die Bewegungen einer Person an die einer Darsteller*in oder an eine Audiodatei angepasst werden „LipSync“, „Facial Reenactment“ und „Puppeteering“, vgl. Thies 07.08.2020).

Im Gegensatz zu den Implikationen des Begriffs „Fake“ hat diese Definition keine negativen Implikationen und setzt keine Täuschungsabsicht voraus. Sie ist bewusst so weit gefasst und leitet sich aus dem Diskurs um Deepfakes ab: Zwar erhielt die Technologie ihren Namen 2017 im Kontext der Erstellung nicht-einvernehmlicher Pornografie. Der Begriff „Deepfake“ ist jedoch auch für andere Anwendungen weit verbreitet. So diskutieren

-
- 1 Es gibt beispielsweise einvernehmliche Deepfake-Pornografie, bei der die Betrachtenden die Darsteller*innen selbst aussuchen können. Die Darsteller*innen werden dann bei jedem „Einsatz“ in einem Deepfake-Porno vergütet.
 - 2 Kommerzielle Nutzungen oder die Anwendung zu Unterhaltungszwecken ohne politischen Bezug stehen nicht im Fokus dieses Beitrags.
 - 3 Die Literaturrecherche erfolgte von Februar 2020 bis Mai 2021. Wissenschaftliche Beiträge wurden über eine Stichwortsuche bei Google Scholar nach den Begriffen „deepfake“ und „deep fake“ sowie eine Auswertung der entsprechenden Literaturverzeichnisse gefunden. Berücksichtigt wurden so auch Berichte von Nichtregierungsorganisationen, politischen Körperschaften, Unternehmen und Policy-Analyst*innen. Entsprechende Beiträge sowie Medienberichte zu Deepfakes wurden zudem über eine systematische Analyse von Google Alerts nach den Suchbegriffen „deepfake“, „deep fake“, „synthetic media“ sowie entsprechenden deutschen Begriffen von Juni 2020 bis Mai 2021 einbezogen.

bedeutende rechts-, sozial- und wirtschaftswissenschaftliche Beiträge unter dem Begriff „Deepfake“ böswillige und andere Anwendungen (vgl. bspw. Chesney/Citron 2019; Diakopoulos/Johnson 2019; Kietzmann et al. 2020). Auch die meisten Medienberichte definieren Deepfakes breit. Selbst die Bundesregierung erkennt den potenziellen Nutzen von Deepfakes z.B. für Kunst oder Satire an (vgl. Bundesregierung 02.12.2019, S. 4). Zwei Faktoren liegen dem zugrunde: Erstens basieren solche Anwendungen auf denselben Technologien wie ihre schädliche Nutzung. Zweitens wird der Begriff „Deepfake“ oft bewusst verwendet, um politische oder kommerzielle Aufmerksamkeit zu erregen. So erwähnen mediale Beiträge über das kommerzielle Potenzial von Deepfakes den ‚neutraleren‘ Oberbegriff der synthetischen Medien häufig nicht einmal (vgl. etwa Dietmar 21.05.2019).

Der Begriff „Deepfake“ ist also tief verankert im öffentlichen und politischen Diskurs um eine Vielzahl von synthetischen, den Menschen imitierenden audio-visuellen Medien in unterschiedlichen Kontexten. Ein breites Begriffsverständnis erlaubt es somit, auch die Chancen der Technologie zu betrachten.

2. Problem-orientierte und deliberative Demokratietheorien

Bieten Deepfakes auch Chancen für die Demokratie? Um dies zu beantworten, muss zunächst das zugrundeliegende Verständnis von Demokratie geklärt werden. Dazu greife ich auf systemische Demokratietheorien zurück. Diese definieren Kriterien, die ein System erfüllen muss, um als demokratisch zu gelten, ohne einem Kriterium Priorität einzuräumen (vgl. Tenove 2020, S. 521). Insbesondere beziehe ich mich auf die normative, problem-orientierte Demokratietheorie nach Warren, der zufolge ein System demokratisch ist, wenn es drei Kernfunktionen erfüllt: Es muss „empowered inclusion“ ermöglichen, also die gestärkte Inklusion aller Bürger*innen, sowie die kollektive Agenda- und Willensbildung und kollektive Entscheidungsfindung (Warren 2017, S. 43). Warren definiert die normativen Grundlagen dieser Funktionen genauer: ‚Empowered inclusion‘ basiert auf der Idee, dass diejenigen, die (potenziell) von einer Entscheidung betroffen sind, an ihr beteiligt sein sollten. Zudem ermächtigen Demokratien ihre Bürger*innen dazu, „ihre Inklusion einzufordern und durchzusetzen“, z.B. durch Wahlen, Repräsentation oder die Bildung von Organisationen (ebd., S. 44, eigene Übersetzung). Demokratien müssen zudem den Transfer der Präferenzen, Interessen und Werte von Individuen und Gruppen in eine kollektive Agenda ermöglichen. Dies basiert auf moralischen und ethischen Überlegungen wie der Notwendigkeit gegenseitiger Anerkennung, Respekt und Fairness. Schließlich müssen Demokratien Kollektive befähigen, Ent-

scheidungen zu relevanten Themen zu treffen. Dies dient ihrer Leistungs- oder Output-Legitimität (vgl. ebd., S. 45). Warren identifiziert zudem mehrere „soziale Praktiken“, die der Realisierung dieser demokratischen Kernfunktionen dienen. Dazu zählen unter anderem Deliberation, Wahlen und Repräsentation (ebd.).

Die Praktik der Deliberation als Austausch von Argumenten und als Aushandlungsprozess ist für eine Untersuchung von Deepfakes besonders relevant, da der Diskurs ihr Schädigungspotenzial für epistemische Sicherheit und deliberative Qualität betont (siehe oben). Deliberation dient laut Warren insbesondere der Stärkung der kollektiven Agenda- und Willensbildung (ebd., S. 48). Diese Einschätzung teilen auch Theorien der deliberativen Demokratie, etwa bei Habermas (2005, S. 287). Mansbridge et al. (2012, 11f) spezifizieren zudem eine „epistemische“, wissensbezogene Funktion deliberativer Systeme: Sie erlauben es, „Präferenzen, Meinungen und Entscheidungen zu produzieren, die angemessen durch Fakten und Logik informiert sind und das Ergebnis einer substanziellen und sinnvollen Abwägung relevanter Gründe sind“ (eigene Übersetzung). Zudem erfüllt Deliberation eine „ethische Funktion“, indem sie den gegenseitigen Respekt und die Anerkennung unter den Bürger*innen fördert. Dies ist gleichzeitig ein intrinsisches Element der Deliberation und eine Voraussetzung für „effektive Kommunikation“ (ebd.).

Wie können Deepfakes zu den demokratischen Kernfunktionen und insbesondere zur kollektiven Agenda- und Willensbildung durch Deliberation beitragen – und welche Wertkonflikte bestehen? Diesen Fragen widmen sich die nachfolgenden Kapitel, die vier unterschiedliche Anwendungskontexte untersuchen: Satire, politische Kampagnen und Aktivismus, politische Kunst und Bildung.

3. Satire

Deepfakes werden zunehmend für politische Parodien und Satire genutzt. Dazu gehören Deepfake-Musikvideos von Politiker*innen, beispielsweise des ehemaligen österreichischen Kanzlers Sebastian Kurz im Kontext einer Affäre um die Staatsholding ‚Österreichische Beteiligungs AG‘ (ÖBAG) mit seinem Vertrauten, dem ÖBAG-Vorsitzenden Thomas Schmid (vgl. OE24 07.04.2021). Im Oktober 2020 starteten die Schöpfer der Serie South Park mit „Sassy Justice“ zudem die erste Serie, die ganz auf Deepfakes etwa von Donald Trump basiert (vgl. Hao 28.10.2020).

Deepfake-Satire kann die Mächtigen kritisieren und Fehlverhalten anprangern. Besonders bedeutsam ist sie als Instrument für marginalisierte Bevölkerungsgruppen (vgl. Collins-Dexter in WITNESS 2020c). Damit

fördert sie „empowered inclusion“ (vgl. Warren 2017, S.48). Indem sie auf unterhaltsame Weise den Fokus auf gewisse politische Gegebenheiten richtet, kann Satire zudem die epistemische Qualität der Deliberation erhöhen und dazu beitragen, dass der öffentliche Diskurs auf „gesamtgesellschaftliche Problemlagen“ reagiert (Habermas 2005, S.290). Deepfakes sind dabei ein audio-visuell potenziell besonders wirkungsvolles und zunehmend zugängliches Werkzeug, das Satiriker*innen nutzen können.

Gleichzeitig kann Deepfake-Satire die epistemische Qualität der Deliberation auch bedrohen, wenn sie ihr Publikum täuscht. Dies ist der Fall, wenn Satire aus ihrem ursprünglichen Kontext heraus transferiert wird und dann nicht mehr als solche erkennbar ist – eine Gefahr, die bei qualitativ hochwertigen Deepfakes besonders ausgeprägt ist, da die audio-visuelle Manipulation oder synthetische Erzeugung nicht mehr durch den Menschen erkannt werden kann. Beispielsweise wurden Ausschnitte einer satirische Fernsehshow mit Deepfakes des ehemaligen italienischen Premierminister Matteo Renzi, in denen dieser andere Politiker*innen beleidigte, auf sozialen Medien geteilt und dort häufig für echt gehalten (vgl. Menses 2021, S.4). Zudem gibt es bestimmte Bevölkerungsgruppen, die Satire etwa aufgrund mangelnder Sprachkenntnisse oder eingeschränkter kognitiver Fähigkeiten nicht erkennen können (vgl. Bieß/Pawelec 2020).

Häufig verschwimmen auch die Grenzen zwischen Satire, politischer Meinungsäußerung, und dem Versuch der Schädigung. Böswillige Deepfakes werden zunehmend als Satire bezeichnet, um einer Löschung etwa durch soziale Medien zu entgehen (vgl. Lytvynenko in WITNESS 2020c). Mitunter bleibt die Intention bestimmter Deepfakes auch unklar. So tauchten im März 2021 (vor einer Abstimmung über die Gründung einer Gewerkschaft) gefälschte Twitter-Profilen angeblicher Amazon-Mitarbeiter*innen mit Deepfake-Profilfotos auf. Sie twitterten „ziemlich lächerliche, übertriebene Verteidigungen von Amazons Arbeitspraktiken“ und sorgten für Verwirrung: Ursprünglich wurde vermutet, Amazon stecke selbst dahinter; dies wurde später jedoch entkräftet (vgl. Hao 31.03.2021). Politisch brisant ist auch der Fall angeblicher Deepfake-Videokonferenzen europäischer Abgeordneter mit dem fingierten Stabschef des russischen Oppositionellen Alexej Nawalny im Frühjahr 2021. Nach breiter medialer Berichterstattung über die Täuschung der Abgeordneten durch Deepfakes stellte sich heraus, dass sich ein russischer Komiker lediglich als Stabschef verkleidet hatte. Der Fall wurde dennoch insbesondere von baltischen Abgeordneten als Angriff Russlands gewertet – auch aufgrund der engen Verbindungen des Komikers zum Kreml (vgl. Rötzer 28.04.2021). Dieses Beispiel zeigt das Potenzial (angeblicher) Deepfake-Satire, politische Unsicherheit und Unruhe zu stiften.

Deepfake-Satire kann somit die Verantwortlichkeit mächtiger Personen und die Qualität kollektiver Agenda- und Willensbildung erhöhen. Gleichzeitig können satirische Deepfakes Bürger*innen täuschen und politischen Schaden anrichten – insbesondere dann, wenn Satire als Vorwand für beleidigende und irreführende Deepfakes genutzt wird. Solche Deepfakes untergraben zudem gegenseitigen Respekt und die Zivilität politischer Deliberation. Doch auch unbeabsichtigt können satirische Deepfakes täuschen und verwirren, wenn sie nicht klar gekennzeichnet sind.

4. Politische Kampagnen und Aktivismus

Deepfakes werden auch für politische Kampagnen und Aktivismus eingesetzt. Die Grenzen zur politischen Kunst sind fließend. Dieses Kapitel konzentriert sich daher auf den Einsatz durch zivilgesellschaftliche Organisationen, politische Bewegungen sowie Behörden, insbesondere zur Aufklärung und Mobilisierung, aber auch zur politischen Opposition.

So war Fußball-Star David Beckham 2019 in einer Aufklärungskampagne gegen Malaria dank Deepfake-Technologie in neun verschiedenen Sprachen zu sehen (vgl. Kietzmann et al. 2020, S. 143). Im Februar 2020 ließen eine mexikanische Nichtregierungsorganisation (NRO) und ein Medienkonzern einen Journalisten Gerechtigkeit für seine eigene Ermordung fordern – unter anderem per Video-Deepfake an den mexikanischen Präsidenten auf allen Fernsehkanälen (vgl. Tyukavkin 02.12.2020). Auch im Rahmen der US-Wahlen wurden Deepfakes für politische Kampagnen genutzt. Die Antikorruptions-Organisation RepresentUS warnte mit Deepfakes von Kim Jong-un und Vladimir Putin vor einer Bedrohung der US-Demokratie und rief Bürger*innen zu mehr Beteiligung auf (vgl. Haasch 29.09.2020). Im Oktober 2020 startete zudem eine Kampagne, die einen beim Parkland-Massaker ermordeten Jugendlichen per Deepfake ‚wiederbelebte‘, um strengere Waffengesetze zu fordern (vgl. Diaz 02.10.2020). Weitere Kampagnen mit Deepfakes Verstorbener liefen 2020, als eine kanadische NRO Deepfakes von Josef Stalin erstellte, um die Aufnahme des Worts ‚Holodomor‘ für Genozid durch Hunger in englischsprachige Wörterbücher zu fordern (vgl. HNAT 21.07.2020), sowie Ende 2020 in Australien, um Polizist*innen-Suizide zu verhindern (vgl. Cameron 16.12.2020) und im März 2021 mit dem 1989 verstorbenen Künstler Salvador Dalí für mehr Forschung zu neurodegenerativen Erkrankungen (vgl. Gülen 01.04.2021). Anfang 2021 warb zudem eine französische NRO mit Deepfakes gegen niederschweligen Rassismus (vgl. Spary 09.02.2021).

Deepfakes erlauben es politischen Organisationen, Kampagnen günstiger zu produzieren oder in verschiedenen Sprachen auszustrahlen. Denkbar

ist auch ihre Nutzung in repressiven Staaten, um oppositionelle Botschaften zu verbreiten (vgl. Thies 07.08.2020) oder Gesichtserkennungssoftware zu umgehen (vgl. Rajagopalan in WITNESS 2020a). Solche Deepfakes können ‚empowered inclusion‘ fördern, indem (marginalisierten) politischen Organisationen und Anliegen mehr Gehör verschafft wird und Mächtige zur Verantwortung gezogen werden. Aufklärungskampagnen in seltenen Sprachen sind zudem mehr Menschen zugänglich. Deepfakes, insbesondere mit berühmten oder verstorbenen Personen, ermöglichen auch neue Inhalte und damit neue Formen der Mobilisierung. Beides kann dazu beitragen, Inklusion und die epistemische Qualität der Deliberation zu erhöhen, indem mehr Meinungen im Diskurs vertreten sind und auf Missstände hingewiesen wird.

Gleichzeitig besteht erneut die Gefahr der Täuschung, insbesondere gegenüber bestimmten Personengruppen oder bei einem Kontexttransfer. Deepfake-Kampagnen werfen zudem schwierige politische, ethische und rechtliche Fragen auf, wenn Personen ohne deren Zustimmung abgebildet werden.⁴ Dies gilt umso mehr für Verstorbene. Zentral sind hier etwa Fragen nach deren Würde und postmortalem Persönlichkeitsrecht, der Bedeutung ihrer Zustimmung zu Lebzeiten oder der Zustimmung ihrer Angehörigen, und möglicher emotionaler Belastungen für diese. Diese Fragen müssen – ebenso wie die Gefahr der Täuschung – gegen die Potenziale solcher Deepfakes für die Demokratie abgewogen werden.

5. Politische Kunst

Auch in der Kunst erzeugen Deepfakes Aufmerksamkeit für gesellschaftliche Themen und Missstände und stoßen politische Debatten an – teils über Deepfakes selbst. So warnte BuzzFeed News mit einem Deepfake-Clip von Barack Obama schon 2018 vor den Gefahren der Technologie (vgl. Chesney/Citron 2019, S. 1770); ein Deepfake der Queen Ende 2020 hatte dieselbe Intention (vgl. BBC 23.12.2020). Auch das Kunstprojekt ‚Spectre‘ mit Deepfakes unter anderem von Mark Zuckerberg verweist auf die Gefahren von Deepfakes sowie auf Überwachung und Desinformation (vgl. Ajder et al. 2019, S. 12). Auf Überwachung, aber auch auf die Schließung der EU-Außengrenzen aufmerksam machte das Künstlerkollektiv ‚Peng!‘, als es 2018 erfolgreich einen Pass mit einem gefälschten Foto beantragte, das das Bild einer Künstlerin mit dem der damaligen EU-Außenbeauftragten Mogherini morphete (Thelen/Horchert 22.09.2018). Anonyme erstellten darüber

4 Dies gilt auch für Satire; hier ist eine Zustimmung der Abgebildeten – anders als bei politischen Kampagnen – aber ohnehin selten gegeben.

hinaus 2019 eine Webseite, auf der Besucher*innen Audio-Deepfakes eines pro-patriarchalischen Professors erstellen konnten. Die Seite wurde genutzt, um ihn feministische Texte ‚lesen zu lassen‘ und erst gelöscht, als er mit rechtlichen Schritten drohte (vgl. Kietzmann et al. 2020, S. 143). Ende 2020 startete zudem das Kunstprojekt ‚DeepReckonings‘, in dem mächtige Personen wie Mark Zuckerberg imaginär Rechenschaft über ihre Fehler und Sinneswandel ablegen (Lepp in WITNESS 2020b). Die wohl umfangreichste künstlerische Nutzung von Deepfakes bisher erfolgte 2020 im Dokumentarfilm ‚Welcome to Chechnya‘ über die Verfolgung von Homosexuellen in Tschetschenien. Deepfakes verschleiern hier zu deren Schutz die Identität der Protagonist*innen (France in WITNESS 2020d).⁵

Künstler*innen nutzen Deepfakes mit unterschiedlichen Intentionen: Daniel Howe (‚Spectre‘) kämpft nach eigenen Aussagen für „durchschnittliche Bürger*innen, die weniger mächtig sind, und für die Demokratie“ (in WITNESS 2020b, eigene Übersetzung). Der Regisseur von ‚Welcome to Chechnya‘ sieht in Deepfakes eine Möglichkeit, die „Menschlichkeit“ der Verfolgten wiederherzustellen und ihnen eine Stimme zu verleihen (France in WITNESS 2020d, eigene Übersetzung). Stephanie Lepp (‚DeepReckonings‘) möchte mit „prosozialen“ Deepfakes die Mächtigen zum Umdenken bewegen (in WITNESS 2020b, eigene Übersetzung).

Künstlerische Deepfakes können somit die Mächtigen kritisieren und den Einbezug marginalisierter Stimmen fördern – und damit ‚empowered inclusion‘. Auch fördern sie die epistemische Qualität der Deliberation, wenn sie Missstände aufdecken (oder dies ermöglichen) oder Lösungsvorschläge aufzeigen. Voraussetzung ist erneut Transparenz über die Tatsache der Fälschung. Der Fall des genannten Professors, ebenso wie die teils empörten Reaktionen auf den Deepfake der Queen Ende 2020, verweisen zudem wieder auf die Notwendigkeit, die Rechte der Abgebildeten gegen dieses demokratiefördernde Potenzial – sowie die künstlerische Freiheit – abzuwägen.

6. Bildung

Der Einsatz von Deepfakes für Bildung ist einer der „offensichtlichsten [...] prosozialen“ Nutzungen von Deepfakes (Chesney/Citron 2019, S. 1768, eige-

5 Anders als bei den anderen Kunstprojekten werden Deepfakes hier nicht für Fiktion eingesetzt. Die Grenzen zum politischen Aktivismus sind fließend. Der Dokumentarfilm wurde hier jedoch dem Anwendungskontext der Kunst zugeordnet, da er nicht von einer politischen Organisation erstellt wurde und keine eindeutige politische Agenda und Mobilisierungsabsicht verfolgt.

ne Übersetzung). Zur Abgrenzung von Kunst und Aktivismus konzentriert sich dieses Kapitel auf Deepfakes, die von Einrichtungen mit Bildungsauftrag wie Schulen oder Museen geschaffen werden, sowie Deepfakes mit historischem Bezug, aber ohne eindeutige politische Agenda.

Potenzial bergen Deepfakes für die Museumspädagogik. Deepfakes von Zeitzeug*innen oder Künstler*innen können Geschichte und Kunst intensiver erfahrbar machen. So können Besucher*innen in der Dalí Lives-Ausstellung in St. Petersburg, USA mit dem verstorbenen Künstler ‚interagieren‘ (vgl. Kwok/Koh 2020, S. 3). Andere Installationen etwa zur Interaktion ‚mit‘ Holocaust-Überlebenden könnten in Zukunft von der Technologie profitieren.

Auch jenseits von Museen machen Deepfakes historische Ereignisse erlebbarer. So erstellte die Zeitung ‚The Times‘ 2019 mit einer Werbeagentur einen Audio-Deepfake der Rede, die John F. Kennedy am Tag seiner Ermordung gehalten hätte (vgl. Bundesregierung 02.12.2019, S. 2). Ähnliche Beispiele sind ein Virtual-Reality-Erlebnis von Martin Luther Kings ‚I Have a Dream‘-Rede (vgl. Hao 28.08.2020), und ein Deepfake-Video von Richard Nixons Rede im Falle eines Scheiterns der Apollo-11 Mission (vgl. DelViscio 20.07.2020). In der formalen Bildung könnten Deepfakes zudem den Schulunterricht spannender und interaktiver gestalten (vgl. Chesney/Citron 2019, S. 1769) oder die Kosten und den Zeitaufwand für professionelle Online-Lehrveranstaltungen senken (Respeecher 02.02.2021).⁶

Deepfakes im Bildungsbereich können das historische Bewusstsein von Bürger*innen schärfen – und damit ihr Verständnis von und Interesse an Politik und Gesellschaft. Dies ist essenziell für Inklusion und Partizipation. Solche Deepfakes inspirieren zudem Debatten nicht nur über Geschichte und ihren Einfluss auf die heutige Politik, sondern auch über das Verhältnis von Technik und Gesellschaft (vgl. Citron in DelViscio 20.07.2020). Damit fördern sie die kollektive Agenda- und Willensbildung. Die Vermittlung geschichtlicher Ereignisse setzt jedoch ihre Interpretation voraus und ist damit politisch. Daher ist es bei immersiven Erfahrungen aufgrund ihrer großen Wirkung besonders entscheidend, zu verdeutlichen, dass es sich um Deepfakes handelt (vgl. Zittrain in DelViscio 20.07.2020) – ebenso wie im Kontext der Bildung von Kindern und Jugendlichen.

6 Bisher (Stand: Juni 2021) scheint die Technologie hier jedoch noch keine breite Anwendung zu finden.

7. Fazit und Ausblick

Die Befürchtungen in Bezug auf Deepfakes und die Zukunft der Demokratie sind groß. Gleichzeitig bergen Deepfakes großes Potenzial, marginalisierten Stimmen und Themen mehr Gehör zu verschaffen, Fehlverhalten mächtiger Personen zu thematisieren und politische Bildung zu fördern. Dieses Potenzial verdient größere Beachtung – zumal ein allgemeines Verbot von Deepfakes angesichts ihrer legitimen Nutzungen sowie der Implikationen für die Meinungsfreiheit ethisch nicht wünschenswert ist (vgl. Bieß/Pawelec 2020).

Wie kann also eine Gesellschaft vom Potenzial von Deepfakes profitieren und gleichzeitig ihre schädlichen Wirkungen eindämmen? Die Antwort darauf ist komplex und vielschichtig, denn Deepfakes werden von verschiedenen Akteur*innen in unterschiedlichen Kontexten eingesetzt. Daher bedarf es kontextspezifischer Regulierungsdiskussionen (vgl. Bieß/Pawelec 2020). Entscheidend ist jedoch, prosoziale (ebenso wie kommerzielle und unterhaltende) Deepfakes zu kennzeichnen, um Manipulation und Täuschung entgegenzuwirken (ebd.). Entsprechend wurden die Deepfakes in allen oben genannten politischen Kampagnen, Bildungs- und Kunstprojekten bis auf die des Kollektivs ‚Peng!‘ eindeutig gekennzeichnet; Künstlerin Stephanie Lepp bezeichnet dies als die „einfachste, ethische Linie“, die es zu ziehen gelte (in Di Placido 03.11.2020, eigene Übersetzung).

Eine Markierung ist keine unzulässige Einschränkung der künstlerischen oder Meinungsfreiheit. Im Gegenteil: Sie schützt die Meinungsfreiheit, „indem sie die Rahmenbedingungen für eine vertrauenswürdige Medienkommunikation schafft“ (Heesen in Science Media Centre 22.04.2021). Daher sollte die EU-Kommission etwa ihren KI-Regulierungsvorschlag vom April 2021 anpassen, der Ausnahmen von der Markierungspflicht für Deepfakes zum Schutz der Meinungs- und künstlerischen Freiheit vorsieht (vgl. European Commission 21.04.2021, S. 69). Nur wenn prosoziale (sowie kommerzielle und unterhaltende) Deepfakes eindeutig gekennzeichnet werden, können sie ihre positiven Wirkungen – auch für die Demokratie – entfalten, ohne das Vertrauen in die epistemische Qualität der Deliberation und damit den demokratischen Diskurs zu schwächen.

Literatur

- Ajder, H.; Patrini, G.; Cavalli, F.; Cullen, L. (2019): The State of Deepfakes. Landscape, Threats, and Impact. Deeptrace
- BBC (23.12.2020): Deepfake queen to deliver Channel 4 Christmas message. In: BBC; <https://www.bbc.com/news/technology-55424730> [aufgesucht am 07.06.2021]

- Bieß, C.; Pawelec, M. (2020): Schaden Deepfakes (wirklich) der Demokratie? Warum die Debatte um Deepfakes in der Politik oft zu kurz greift. *BedenkZeiten Blog: Internationales Zentrum für Ethik in den Wissenschaften*; <https://uni-tuebingen.de/de/199237> [aufgesucht am 13.04.2021]
- Bundesregierung (02.12.2019): Beschäftigung der Bundesregierung mit Deepfakes. Antwort der Bundesregierung auf die Kleine Anfrage der Abgeordneten Manuel Höferlin, Frank Sitta, Grigorios Aggelidis, weiterer Abgeordneter und der Fraktion der FDP -Drucksache 19/15210; <http://dip21.bundestag.de/dip21/btd/19/156/1915657.pdf> [aufgesucht am 28.05.2020]
- Cameron, N. (16.12.2020): Deepfake tech used to share powerful message for Victoria Police. In: CMO; <https://www.cmo.com.au/article/685229/deepfake-tech-used-share-powerful-message-victoria-police/> [aufgesucht am 07.06.2021]
- Chesney, R.; Citron, D. (2019): Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security. In: *California Law Review*, Nr. 107, S. 1753–1819
- DeViscio, J. (20.07.2020): A Nixon Deepfake, a ‘Moon Disaster’ Speech and an Information Ecosystem at Risk. A new video re-creates a history that never happened, showing the power of AI-generated media. In: *Scientific American*; <https://www.scientificamerican.com/article/a-nixon-deepfake-a-moon-disaster-speech-and-an-information-ecosystem-at-risk/> [aufgesucht am 25.08.2020]
- Di Placido, D. (03.11.2020): Is It Possible To Create A Benevolent Deepfake? In: *Forbes*; <https://www.forbes.com/sites/danidiplacido/2020/11/03/is-it-possible-to-create-a-benevolent-deepfake/> [aufgesucht am 05.12.2020]
- Diakopoulos, N.; Johnson, D. (2019): Anticipating and Addressing the Ethical Implications of Deepfakes in the Context of Elections. In: *SSRN Electronic Journal*, S. 1–26
- Diaz, A.-C. (02.10.2020): Parkland victim Joaquin Oliver comes back to life in heartbreaking plea to voters. In: *AdAge*; <https://adage.com/article/advertising/parkland-victim-joaquin-oliver-comes-back-life-heartbreaking-plea-voters/2285166> [aufgesucht am 05.12.2020]
- Dietmar, J. (21.05.2019): GANs And Deepfakes Could Revolutionize The Fashion Industry. In: *Forbes*; <https://www.forbes.com/sites/forbestechcouncil/2019/05/21/gans-and-deepfakes-could-revolutionize-the-fashion-industry/#40d9f8dc3d17> [aufgesucht am 11.02.2021]
- European Commission (21.04.2021): Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts. COM(2021) 206 final; <https://ec.europa.eu/newsroom/dae/redirection/document/75788> [aufgesucht am 12.05.2021]
- FBI (10.03.2021): Malicious Actors Almost Certainly Will Leverage Synthetic Content for Cyber and Foreign Influence Operations. Private Industry Notification. Federal Bureau of Investigation; <https://assets.documentcloud.org/documents/20509703/fbipin-3102021.pdf> [aufgesucht am 11.03.2021]
- Galston, W. A. (08.01.2020): Is seeing still believing? The deepfake challenge to truth in politics. *Brookings*; <https://www.brookings.edu/research/is-seeing-still-believing-the-deepfake-challenge-to-truth-in-politics/> [aufgesucht am 19.04.2021]
- Godulla, A.; Hoffmann, C. P.; Seibert, D. (2021): Dealing with deepfakes – an interdisciplinary examination of the state of research and implications for communication studies. In: *Studies in Communication and Media*, Nr. 10(1), S. 72–96

- Gülen, K. (01.04.2021): Deepfake Salvador Dalí macht auf neurodegenerative Erkrankungen aufmerksam. In: TechBriefly; <https://de.techbriefly.com/deepfake-salvador-dali-macht-auf-neurodegenerative-erkrankungen-aufmerksam-tech-29253/> [aufgesucht am 07.06.2021]
- Haasch, P. (29.09.2020): A creepy presidential debate ad shows a deepfake of Putin telling Americans they're ruining their own democracy. In: Insider; <https://www.insider.com/presidential-debate-ad-putin-kim-jong-un-video-deepfake-democracy-2020-9> [aufgesucht am 05.12.2020]
- Habermas, J. (2005): Die Einbeziehung des Anderen. Studien zur politischen Theorie. Frankfurt am Main
- Hao, K. (28.08.2020): Memers are making deepfakes, and things are getting weird. In: MIT Technology Review; <https://www.technologyreview.com/2020/08/28/1007746/ai-deepfakes-memes/> [aufgesucht am 05.12.2020]
- Hao, K. (28.10.2020): The creators of South Park have a new weekly deepfake satire show. In: MIT Technology Review; <https://www.technologyreview.com/2020/10/28/1011336/ai-deepfake-satire-from-south-park-creators/> [aufgesucht am 18.03.2021]
- Hao, K. (31.03.2021): Deepfake "Amazon workers" are sowing confusion on Twitter. That's not the problem. In: MIT Technology Review; <https://www.technologyreview.com/2021/03/31/1021487/deepfake-amazon-workers-are-sowing-confusion-on-twitter-thats-not-the-problem/> [aufgesucht am 06.04.2021]
- Hao, K.; Heaven, W.D. (24.12.2020): The year deepfakes went mainstream. In: MIT Technology Review; <https://www.technologyreview.com/2020/12/24/1015380/best-ai-deepfakes-of-2020/> [aufgesucht am 19.04.2021]
- HNAT – Holodomor National Awareness Tour (21.07.2020): Turning Fake News on its Head. Stalin Lives Again to Reveal Deep Truth. Toronto. In: Cision Canada; <https://www.newswire.ca/news-releases/turning-fake-news-on-its-head-stalin-lives-again-to-reveal-deep-truth-850507473.html> [aufgesucht am 05.12.2020]
- Kietzmann, J.; Lee, L.W.; McCarthy, I.P.; Kietzmann, T.C. (2020): Deepfakes. Trick or treat? In: Business Horizons, Nr. 63(2), S. 135–146
- Kwok, A.O.J.; Koh, S.G.M. (2020): Deepfake: a social construction of technology perspective. In: Current Issues in Tourism, S. 1–5
- Mansbridge, J.; Bohman, J.; Chambers, S.; Christiano, T.; Fung, A.; Parkinson, J.; Thompson, D.F.; Warren, M.E. (2012): A systemic approach to deliberative democracy. In: Parkinson, J.; Mansbridge, J. (Hg.): Deliberative Systems. Cambridge, S. 1–26
- Meneses, J.P. (2021): Deepfakes and the 2020 US elections: what (did not) happen; <https://arxiv.org/pdf/2101.09092> [aufgesucht am 02.04.2021]
- OE24 (07.04.2021): Causa ÖBAG: Über dieses Video lacht das Netz. In: OE24; <https://www.oe24.at/oesterreich/chronik/causa-oebag-ueber-dieses-video-lacht-das-netz/472608766> [aufgesucht am 01.06.2021]
- Respeecher (02.02.2021): Deepfake Voice Technology: The Good. The Bad. The Future. In: EconoTimes; <https://www.econotimes.com/Deepfake-Voice-Technology-The-Good-The-Bad-The-Future-1601278> [aufgesucht am 02.02.2021]

- Rötzer, F. (28.04.2021): Deep Fake: Russische Prankster führen Politiker und Prominente vor. In: Telepolis; <https://www.heise.de/tp/features/Deep-Fake-Russische-Prankster-fuehren-Politiker-und-Prominente-vor-6030802.html> [aufgesucht am 21.05.2021]
- Schreiner, M. (21.02.2021): Deepfakes: Wie alles begann – und wohin es führen könnte. In: Mixed; <https://mixed.de/geschichte-der-deepfakes-so-rasant-geht-es-mit-ki-fakes-voran/> [aufgesucht am 26.03.2021]
- Science Media Centre (22.04.2021): EU will Künstliche Intelligenz regulieren; <https://www.sciencemediacenter.de/alle-angebote/rapid-reaction/details/news/eu-will-kuenstliche-intelligenz-regulieren/> [aufgesucht am 19.05.2021]
- Spary, S. (09.02.2021): French Anti-Racism PSA Uses Deepfake Tech to Show Many Faces of Prejudice. In: Adweek; <https://www.adweek.com/creativity/french-anti-racism-psa-uses-deepfake-tech-to-show-many-faces-of-prejudice/> [aufgesucht am 07.06.2021]
- Tenove, C. (2020): Protecting Democracy from Disinformation: Normative Threats and Policy Responses. In: The International Journal of Press/Politics, Nr. 25(3), S. 517–537
- Thelen, R.; Horchert, J. (22.09.2018): Aktivisten schmuggeln Fotomontage in Reisepass. In: Der Spiegel; <https://www.spiegel.de/netzwelt/netzpolitik/biometrie-im-reisepass-peng-kollektiv-schmuggelt-fotomontage-in-ausweis-a-1229418.html> [aufgesucht am 07.06.2021]
- Thies, J. (07.08.2020): Interview mit Dr. Justus Thies, Mitglied des Visual Computing Lab an der TU München, Entwickler der Methode „Face2Face: Real-time Facial Reenactment“. Interviewende: Bieß, C.; Pawelec, M.
- Thomas, D. (2020): Deepfakes: A threat to democracy or just a bit of fun? In: BBC; <https://www.bbc.com/news/business-51204954> [aufgesucht am 18.03.2020]
- Tyukavkin, A. (02.12.2020): From deepfake to deep truth: how we helped a killed Mexican journalist address the president. Publicis Riga; https://www.linkedin.com/pulse/from-deep-fake-deep-truth-how-we-helped-killed-mexican-tyukavkin/trk-public_profile_article_view [aufgesucht am 07.06.2021]
- Warren, M.E. (2017): A Problem-Based Approach to Democratic Theory. In: American Political Science Review, Nr. 111(1), S. 39–53
- WITNESS (2020a): Boundary lines? Deepfakes weaponized against journalists and activists. Samantha Cole (Vice) and Nina Schick in conversation with Assia Boundaoui (MIT Co-Creation Studio), Deepfakery Webtalk series, Nr. 4
- WITNESS (2020b): Faking the Powerful. Bill Posters and Daniel Howe (Spectre Project) and Stephanie Lepp (Deep Reckonings) in conversation with Sam Gregory (WITNESS), Deepfakery Webtalk series, Nr. 1
- WITNESS (2020c): Not funny anymore: Deepfakes, manipulated media, parody and mis/disinformation. Jane Lytvynenko (Buzzfeed News), Karen Hao (MIT Tech Review) and Brandi Collins-Dexter (Color of Change) in conversation with Corin Faife (WITNESS), Deepfakery Webtalk series, Nr. 2
- WITNESS (2020d): Using AI-generated Face Doubles in Documentary: Welcome to Chechnya. David France (Welcome to Chechnya) in conversation with Kat Cizek (MIT Co-Creation Studio), Deepfakery Webtalk series, Nr. 3

