

# Newsletter

## Deepfakes und manipulierte Realitäten

■ Von Laetitia Ramelet\*

Wäre es mitunter nicht praktisch, politischen Rivalen vor der gesamten digitalen Öffentlichkeit peinliche Aussagen in den Mund legen zu können? Mithilfe der Künstlichen Intelligenz (KI) wird das zum Kinderspiel. Videos können immer effektiver manipuliert und verändert werden. Diese sogenannten Deepfakes sind kaum noch von echten Aufnahmen zu unterscheiden – und die dazu benötigte Software steht jedermann zur Verfügung.

Bisher ist die übergrosse Mehrheit der im Internet zirkulierenden Deepfakes pornographischen Inhaltes: Das Gesicht einer Frau – die nichts davon weiss – wird auf den Körper einer Pornodarstellerin montiert. Nicht von ungefähr geht der Begriff «Deepfake» auf das Pseudonym eines Nutzers zurück, der auf Reddit pornografische Deepfakes von berühmten Schauspielerinnen hochgeladen hatte – mit einem Code, der es auch anderen Usern ermöglichen sollte, solche Bildmanipulationen zu fabrizieren. Die Aussagen zahlreicher Opfer belegen, wie erniedrigend und traumatisch die Verbreitung solcher Fakes für die betroffenen Personen sein können. Und wie schwerwiegend die Folgen für ihr Privat- und Berufsleben.

Inzwischen sind Deepfakes auch zu einem gefährlichen Betrugs- oder Erpressungsinstrument geworden, und stellen damit für viele Unternehmen ein ernstzunehmendes Risiko dar. Sie machen es beispielsweise möglich, durch gefälschte Stimmen und Gesichter an sensible Daten zu gelangen oder Geld zu erpressen. In Europa sind bereits mehrere solche Fälle von Betrug und Identitätsdiebstahl dokumentiert.

Doch trotz dieser eher düsteren Anfänge beginnen Deepfakes durchaus auch gesellschaftliche Chancen zu eröffnen, die es zu nutzen gilt (mehr dazu auf Seite 2). Die Studie von TA-SWISS wird sich also mit unterschiedlichsten Anwendungen und Szenarien befassen. Im Bereich der Politik sticht insbesondere das Desinformations- und Manipulationspotenzial von Deepfakes ins Auge. Deepfakes scheinen sich ideal dafür zu eignen, gegnerische Kandidaten im Wahlkampf zu diskreditieren oder Falschinformationen zu verbreiten, indem man sich des Gesichts, der Stimme und

damit der Glaubwürdigkeit politischer Leader bedient. Ein drastisches Beispiel, das vor Kurzem um die Welt ging, war ein Deepfake, in dem Präsident Selenskyj das ukrainische

Volk dazu aufruft, die Waffen niederzulegen. Längerfristig könnte die Verbreitung gefälschter, auf Anhieb aber glaubwürdig erscheinender Videos zu einem allgemeinen Misstrauen gegenüber allen Informationsquellen führen und den Prozess der Meinungsbildung – in einer Demokratie ein zentraler Prozess – beeinträchtigen.

Der Umgang mit dem Phänomen der Deepfakes ist mit vielen neuen Herausforderungen verbunden – die TA-SWISS-Studie zielt darauf ab, mögliche Strategien aufzuzeigen. Wie können sich die Opfer von Deepfakes wehren und wie kann die Justiz dieser neuen Form von virtuellem, gleichzeitig aber auch sehr greifbarem Unrecht begegnen? Wie soll sich ein Unternehmen gegen Deepfakes wappnen? Was müssen Politik, Bevölkerung und Medien tun, um die Auswirkung dieser KI-generierten Manipulationen auf unser

öffentliches Leben einzugrenzen? Und wie lässt sich verhindern, dass die mit den Deepfakes verbundenen Risiken am Ende auch ihre Chancen zunichtemachen?

\*Die politische Philosophin Dr. Laetitia Ramelet arbeitet als wissenschaftliche Mitarbeiterin bei TA-SWISS.



### Interdisziplinäre Abschätzung der Chancen und Risiken

Die Deepfake-Studie wird von einem interdisziplinären Team durchgeführt. Es umfasst Forschende des Fraunhofer-Instituts für System- und Innovationsforschung (ISI) in Karlsruhe, des Departements für Kommunikationswissenschaft und Medienforschung (DCM) und des Instituts für Europarecht der Universität Freiburg i. Üe. sowie des Instituts für Kommunikationswissenschaft und Medienforschung (IKMZ) der Universität Zürich. Die Resultate werden Anfang 2024 vorliegen.

Weitere Informationen: [www.ta-swiss.ch/deepfakes](http://www.ta-swiss.ch/deepfakes)

## Ja, es gibt durchaus auch Chancen!

*TA: Deepfakes haben einen schlechten Ruf. Zu Recht?*

Murat Karaboga: In der Literatur werden unter den Gesichtspunkten der individuellen, institutionellen und gesellschaftlichen Risiken sowohl das reale als auch das hypothetische zukünftige Schadenspotenzial von Deepfakes diskutiert. Wir haben uns anhand einer empirischen Medienberichterstattungsanalyse auch angeschaut, wo Deepfakes heute schon sinnvoll genutzt werden.

*In welchen Bereichen?*

Viele Beispiele gibt es in Entertainment und Werbung. In der Filmindustrie können Deepfakes teure Spezialeffekte unnötig machen, etwa um Menschen älter oder jünger aussehen zu lassen, oder sie generieren auf Basis alter Aufnahmen ganz neue Inhalte. Das reduziert natürlich die Produktionskosten, aber nicht nur das: Gewisse Deepfake-Varianten von Star-Wars-Filmen gelten sogar als überzeugender als die mittels «klassischer» 3D-Computergrafik (CGI) generierten Szenen. Auch in der Musikindustrie entsteht ein lukrativer Zukunftsmarkt: Die koreanische K-Pop-Girlgroup Eternity beispielsweise existiert nicht wirklich: Ihr Song («I am real») und ihr Musikvideo sind Deepfakes – und die Produktionskosten entsprechend tief. Im Gaming können Spieler einer der Figuren ihre Gesichtszüge verleihen, auf TikTok können Fans Kurzvideos heraufladen, die sie mit einer berühmten Persönlichkeit zeigen. Apps wie My Heritage, die alte Familienfotos zum Leben erwecken, versprechen nostalgisches Entertainment – mit genügend Daten trainiert, tun sie dies auf sehr überzeu-

gende Weise. In der Werbung schliesslich schaffen Deepfakes durch die Personalisierung und Individualisierung von Medieninhalten neue Arten der Kundenbindung: Das können auf jede angesprochene Person zugeschnittene Werbespots sein oder die Möglichkeit zu sehen, wie mir die angebotenen Outfits, Frisuren oder Make-ups stehen würden.

*Gibt es auch nicht-kommerzielle Use-Cases, wo Deepfakes von Nutzen sind?*

Im Bereich der Bildung ermöglichen sie innovative Formen der Unterrichtsgestaltung: Lerninhalte können animiert, historische Figuren als virtuelle Zeitzeugen zum Leben erweckt werden. Sie können auch als Instrument der Medienpädagogik zur Sensibilisierung für Falschinformationen eingesetzt werden – und damit ihr eigenes Desinformationspotenzial demontieren. Im Gesundheitsbereich gibt es ebenfalls erste Versuche. Zum Beispiel das Projekt «Deep Memory» der Universität Amsterdam. Dort geht es darum, Angststörungen mittels Deepfakes zu therapieren, indem die Betroffenen virtuell in Situationen versetzt werden, die sie normalerweise vermeiden.

*Und in der Politik?*

Da konnten wir bislang ausschliesslich Risiken ausmachen.

*Dr. des. Murat Karaboga ist wissenschaftlicher Mitarbeiter am Fraunhofer-Institut für System und Innovationsforschung (ISI) in Karlsruhe und Leiter der Deepfake-Projektgruppe.*

## Das Katz-und-Maus-Spiel muss enden

*TA: Deepfakes werden immer raffinierter. Muss uns das Sorgen machen?*

Touradj Ebrahimi: Diese auf Künstlicher Intelligenz (KI) basierende Technik der Bildsynthese bringt ein Phänomen ans Licht, das so alt ist wie die Zivilisation: das Phänomen der Fälschungen. Neu an Deepfakes ist, dass das Wissen und das Handwerk, das für die Herstellung solcher Manipulationen erforderlich ist, heute jedermann zugänglich ist. Was früher nur erfahrene Fälscher konnten, wird nun an die KI delegiert.

*Bilder zu fälschen, hat sich also sozusagen demokratisiert?*

Ich spreche lieber von Bildmanipulation. Und die ist, das möchte ich betonen, nicht an sich schlecht. Nehmen Sie die Fotos, die wir mit unseren Smartphones schießen. Niemand wird sich über die Möglichkeit beklagen, sie zu korrigieren, wenn sie unscharf oder unterbelichtet sind. Mithilfe von leicht zugänglicher Software können wir sogar Bildelemente hinzufügen oder löschen, oder völlig neue Inhalte kreieren. Ein beeindruckendes Beispiel ist der Bildgenerator DALL-E, der in wenigen Sekunden auf Grund einer Reihe vorgegebener Begriffe synthetische

Bilder erschafft. Problematisch wird das alles erst, wenn es mit böser Absicht geschieht.

*Die KI kann inzwischen so viel: Ist es überhaupt noch möglich, zwischen echt und manipuliert zu unterscheiden?*

Früher wurde die Echtheit eines Bildes mithilfe der forensischen Analyse beurteilt. Heute verwenden wir Techniken aus der digitalen Welt. Wir entwickeln Erkennungssoftware, die auf denselben Verfahren beruht wie die Deepfakes selber. Es ist ein Katz-und-Maus-Spiel: Je besser die Detektoren Bildbearbeitungen aufspüren können, desto mehr werden jene, welche Deepfakes erstellen, ihre Methoden verfeinern. Über kurz oder lang ist es aber meist die Maus, die verliert.

*Was meinen Sie damit?*

Mithilfe von «Machine Learning» (d.h. statistischem Lernen) wird die KI darauf trainiert, Anomalien zu erkennen: Die ersten Deepfakes haben zum Beispiel nie geblinzelt. Also taten unsere ersten Detektoren nichts anderes, als Jagd auf diese allzu starren Augen zu machen. Doch sobald eine Maus verliert, kommen neue hinzu. Das Rennen geht weiter, und da die Manipulationen immer komplexer

werden, werden auf lange Sicht selbst Maschinen den Unterschied zwischen echt und manipuliert nicht mehr erkennen können.

*Der Kampf gegen die Deepfakes ist zum Voraus verloren?*

Nein, aber wir müssen das Katz-und-Maus-Paradigma überwinden und einen neuen Ansatz entwickeln. Das JPEG-Gremium zum Beispiel arbeitet an einem universellen Standard: In alle digitalen medialen Inhalte sollen Metadaten eingebettet werden, die genau angeben, wann dieser Inhalt erstellt wurde, woher er stammt, und ob, wann und wie er bearbeitet wurde.

*Das Problem hat also bald eine Lösung!*

Im Prinzip ja: Aber Gesellschaft und Industrie müssen auch gewillt sein, sie anzuwenden. Voraussetzung dafür

ist einerseits, dass die Öffentlichkeit informiert und auch entsprechend geschult wird. Zudem müssen die nötigen gesetzlichen Voraussetzungen geschaffen werden. Beides ist wichtig – denn das Risiko, das Deepfakes für unsere Demokratien darstellen, ist nicht zu unterschätzen. Das hat mit der «Lügendividende» zu tun: Wenn alles manipuliert werden kann, dann lässt sich auch die Wahrhaftigkeit von allem und jedem bezweifeln. Für Lügner wird es dann ein Leichtes, selbst an sich unwiderlegbare Tatsachen in Frage zu stellen.

*Prof. Touradj Ebrahimi leitet die Gruppe für multimediale Signalverarbeitung an der EPFL und präsidiert das internationale JPEG-Standardisierungsgremium.*

*Interviews aufgezeichnet von cdh*

## Bioelektronik – Der direkte Draht zum Körper

(FS) – Direkt mit dem Körper verbundene elektronische Geräte kennen wir aus der Medizin – so regt das Cochlea-Implantat den Hörnerv an und ermöglicht gehörlosen Menschen, Töne wahrzunehmen. Doch auch ausserhalb der Medizin kommt die Bioelektronik zunehmend zum Einsatz, zum Beispiel zum intuitiven Steuern von Geräten. In einer soeben publizierten Studie kommt TA-SWISS zum Schluss, dass Anwendungen, die Körper und Technik miteinander verbinden, spezifisch zu regulieren sind. Dem Schutz von Kindern und Jugendlichen sollte dabei besondere Beachtung geschenkt werden. Denn Bioelektronik ist kein Kinderspiel: Vorsicht ist insbesondere bei zunehmend bioelektronisch unterstützten Anwendungen angebracht, zum Beispiel bei den Computergames, die bei Minderjährigen hoch im Kurs stehen. Denn darüber, wie sich Bioelektronik auf das Gehirn Heranwachsender auswirkt, ist bisher wenig bekannt.

### Nervenkitzel und Hirndoping

Auch bei erwachsenen Personen sollte bei Bioelektronik, die direkt auf das Nervensystem wirkt, genau hingesehen werden. Konkret geht es um Anwendungen, welche die Konzentration oder die Lernfähigkeit einer Person steigern oder aber im Gegenteil beruhigend und entspannend wirken sollen.

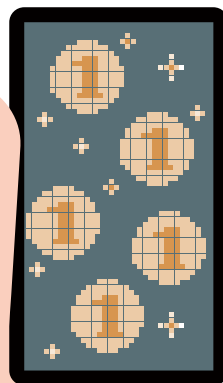
Die Studie kommt auch zum Schluss, dass die Schweiz den in der Bioelektronik tätigen Unternehmen gute Rahmenbedingungen bietet. Die intensive Hochschulforschung wird ergänzt durch eine vitale Start-up-Szene. Um diesen internationalen Spitzenplatz zu behaupten, empfehlen die Autorinnen und Autoren, die Forschung national stärker zu fördern und den Wissenstransfer von der Wissenschaft zur Wirtschaft zu unterstützen.

*Mehr zur Studie hier: [www.ta-swiss.ch/bioelektronik](http://www.ta-swiss.ch/bioelektronik)*

## Digitaler Franken – unser Geld der Zukunft?

### Ausschreibung einer Studie zur Technologiefolgen-Abschätzung

(cdh) – «Bargeld ist geprägte Freiheit», schrieb Fjodor Dostojewski im 19. Jahrhundert – lange bevor die Digitalisierung auch den Geldmarkt erfasste. Heute haben sich digitale Formen der Bezahlung in der Gesellschaft etabliert, und Bargeld wird als Zahlungsmittel immer unbedeutender. Kryptowährungen – vor nicht allzulanger Zeit noch exotische Spielerei irgendwelcher Nerds – scheinen unaufhaltsam auf dem Vormarsch. Nicht nur Techgiganten wie Google, Amazon, Meta, Apple und Microsoft, auch Zentralbanken erwägen die Einführung von digitalen Währungen.



Was bedeutet diese Entwicklung aus gesamtgesellschaftlicher Sicht? Wie wird sie sich auf die Stabilität der Zahlungs- und Verrechnungssysteme auswirken? Wie auf die Geldpolitik? TA-SWISS schreibt eine interdisziplinäre Studie zum Thema aus: Sie soll die Chancen und Risiken neuer virtueller Geldformen abschätzen.

Einreichfrist für Projektskizzen ist der 4. November 2022. Die Ausschreibungsunterlagen sind unter folgendem Link zu finden:

[www.ta-swiss.ch/projekte#ausschreibungen](http://www.ta-swiss.ch/projekte#ausschreibungen)

## Der Digitalisierung auf der Spur – Jahresbericht 2021



(EE und cdh) – Eigentlich sollte ein Jahresbericht nach hinten schauen: Darauf, was im vergangenen Jahr geschaffen wurde. Doch wenn das Vollbrachte eine Fortsetzung verlangt, dann gehört zum Rück- eben auch der Ausblick. Im Sommer 2021 hat die Stiftung TA-SWISS der Öffentlichkeit das Resultat ihrer Auseinandersetzung mit dem Thema «Demokratie und Digitalisierung» in Form einer

dreiteiligen Studie sowie einer vielbeachteten Ausstellung im Polit-Forum Bern vorgestellt. Eine der Schlussfolgerungen dieser Arbeit lautet: Demokratie ist nie erreicht – sie braucht Pflege. Ein demokratisches Politsystem misst sich an der Inklusion und am Kompromiss. Es braucht die Ausrichtung am Kollektiv und das Recht auf politische Mitbestimmung der Einzelnen, die Pflege eines sozialen Raumes und den Schutz der individuellen Privatsphäre, Solidarität und Eigenverantwortung. Davon lebt Demokratie – von den einzelnen Bürgerinnen und Bürgern, von all den Vermittlern und Intermediären: politischen Organisationen, Vereinen, zivigesellschaftlichen Institutionen und ihren unterschiedlichen kulturellen Identitäten. Damit wäre das Stichwort gegeben: Kultur. Auch sie ist prägender Bestandteil dieses Landes. Und auch sie entwickelt sich immerzu. Welchen Einfluss die Digitalisierung auf diese Entwicklung hat, wie sie sich auf das Kulturschaffen, seine Rezeption und Verbreitung, die Teilhabe an Kulturangeboten, die Art, wie Werke gesammelt, aufbewahrt und archiviert werden, und nicht zuletzt den Status der Kunstschaffenden selber auswirkt, ist Gegenstand einer neuen grossangelegten Studie von TA-SWISS. Erste Reflexionen dazu legt der Jahresbericht 2021 vor – nicht zuletzt mit einem äusserst lesenswerten Essay von Tobi Müller, Kulturjournalist und Autor in Berlin.

### Publikationen

**Studie «Wenn Menschen ihren Körper mit Technik vernetzen – Grundlagen und Perspektiven nicht-medizinischer Bioelektronik»**, Anne Eckhardt, Andreas Abegg, Goran Seferovic, Samra Ibric, Julia Wolf, Hrsg. TA-SWISS, vdf Hochschulverlag an der ETH Zürich, 2022. Im Buchhandel (ISBN: 978-3-7281-4137-8) oder als kostenloses E-Book im Open Access ([www.vdf.ch](http://www.vdf.ch)) erhältlich.

**«Nervenkitzel durch Bioelektronik»**, Kurzfassung zur Bioelektronikstudie, Hrsg. TA-SWISS, Bern, 2022. Online verfügbar unter: [www.ta-swiss.ch/publikationen](http://www.ta-swiss.ch/publikationen)

**Jahresbericht 2021 – Kultur und Digitalisierung**, Hrsg. TA-SWISS, Bern, 2022. Online verfügbar unter: [www.ta-swiss.ch/publikationen/jahresberichte](http://www.ta-swiss.ch/publikationen/jahresberichte)

Herausgeber:  
TA-SWISS  
Stiftung für Technologiefolgen-Abschätzung  
Brunngasse 36, 3011 Bern  
[ta-swiss.ch](http://ta-swiss.ch)

Redaktion: Christine D'Anna-Huber  
Layout: Hannes Saxer  
Erscheint 3 – 4 Mal jährlich  
Texte: Christine D'Anna-Huber (cdh), Elisabeth Ehrensperger (EE), Murat Karaboga, Laetitia Ramelet, Fabian Schlupe (FS)  
Übersetzung: Alexandra de Bourbon, pro-verbial, Zürich  
Bilder / Illustrationen: Hannes Saxer  
Printauflage: deutsch 2100 / französisch 700  
Elektronisch: deutsch 2500 / französisch 600

### Bestellschein

Bitte senden Sie mir die folgenden Unterlagen (kostenlos)

- Ex. **Studie «Wenn Menschen ihren Körper mit Technik vernetzen»** (solange Vorrat)
- Ex. **Kurzfassung «Nervenkitzel durch Bioelektronik»**  
Bitte Sprache angeben: D , F , I , E
- Jahresbericht 2021 – Kultur und Digitalisierung**  
Bitte Sprache angeben: D , F
- Ich möchte den **TA-SWISS-Newsletter** lieber elektronisch erhalten.

E-Mail Adresse \_\_\_\_\_

Name / Vorname \_\_\_\_\_

Institution \_\_\_\_\_

Strasse \_\_\_\_\_

PLZ / Ort \_\_\_\_\_

Bitte retour an: TA-SWISS, Brunngasse 36, 3011 Bern  
Sie können unsere Publikationen auch per E-Mail bestellen:  
[info@ta-swiss.ch](mailto:info@ta-swiss.ch)

mitglied der  
 akademien der  
wissenschaften schweiz