Zeitschrift: Schweizer Monat : die Autorenzeitschrift für Politik, Wirtschaft und Kultur

Band: 103 (2023)

Heft: 1103

Artikel: Trauen Sie ihren Augen nicht!

Autor: Wenger, Hanna

DOI: https://doi.org/10.5169/seals-1050453

Trauen Sie Ihren Augen nicht!

Deepfakes können uns alle täuschen. Ohne Medienkompetenz und Aufklärung führt die Verbreitung von manipulierten Bildern und Videos zu einem sinkenden Vertrauen in demokratische Institutionen.

von Hanna Wenger

In unserer immer digitaler werdenden Welt spielt die künstliche Intelligenz eine immer wichtigere Rolle. Sie ermöglicht es uns, Aufgaben schneller und effizienter zu erledigen und neue, bislang ungeahnte Möglichkeiten zu erschliessen. Doch was genau ist künstliche Intelligenz und welche Auswirkungen hat sie auf unseren Alltag? Lassen Sie uns gemeinsam auf eine spannende Reise in die Welt der KI gehen und herausfinden, was sie für uns bereithält.

Hat Sie dieser Einstieg gepackt? Falls nicht, so richten Sie Ihre Kritik bitte direkt an die Tech-Firma OpenAI. Diese Einleitung wurde nämlich nicht von einem Menschen, sondern einer Maschine verfasst; wir haben «ChatGPT» beauftragt, einen spannenden Einstieg für einen Text zur künstlichen Intelligenz zu schreiben.

Maschinelle Manipulation

Künstliche Intelligenz kann heute schon vieles tun, was früher dem Menschen vorbehalten war. Und in immer mehr Bereichen wird sie auch Laien zugänglich: Inzwischen werden zahlreiche Apps für die Herstellung von künstlich erzeugten Bildern, Videos und Tonaufnahmen angeboten. Mit steigenden Fähigkeiten wächst aber auch die Gefahr des Missbrauchs. Denn ob ein Netzinhalt menschliche Realität abbildet oder ein computerbasiertes KI-Konstrukt darstellt, lässt sich mittlerweile nicht immer leicht erkennen.

Längst sind auch Bilder und Videos nicht mehr sicher vor maschineller Manipulation. Bereits Anfang 2018 machten Meldungen über sogenannte Deepfakes – ein Kofferwort aus den Begriffen «Deep Learning» und «Fake» – die Runde, nachdem die App Zao bekannt geworden war, mit der sich Gesichter auf Videos täuschend echt manipulieren liessen. Weltweit sorgte das zunächst für Vergnügen: Wer findet es schon nicht amüsant, sein eigenes Antlitz in einer Rede von Donald Trump oder einem Musikvideo von Madonna auftauchen zu lassen? Schon bald wurde aber auch die Schattenseite der Technologie sichtbar: Die App wurde zunehmend genutzt, um mit gefälschten Videos Stimmung gegen Politiker und andere Personen des öffentlichen Lebens zu machen. So wurde zum Beispiel bei einer von Russland unterstützten Desinformationskampagne das Gesicht der jungen ukrainischen Parlamentarierin Svitlana Zalishchuk[1] auf pornografischen Bildern eingeblendet, was ihre Glaubwürdigkeit als Politikerin untergrub.

Die vielen Gesichter des Deepfakes

Schon immer war es möglich, Papierdokumente und Bilder zu fälschen. Auch Fotos können bekanntlich verändert werden. So ist eine gewisse Skepsis geboten, wenn es darum geht, Berichte oder «Beweisfotos» zu akzeptieren, die nicht mit anderen Belegen übereinstimmen und keine überzeugende Herkunft aufweisen. Die hohen Hürden für überzeugende Fälschungen haben es Fachleuten jedoch bisher möglich gemacht, solche mit hoher Zuverlässigkeit zu erkennen. Aufgrund KI-gestützter Werkzeuge werden die Hindernisse für die unbemerkte Synthese jeder Form von digitalen Medien möglicherweise aber bald verschwinden. Eine solche moderne Form der Fälschungen sind eben die Deepfakes: Die realistisch wirkenden Medieninhalte, die durch Techniken der künstlichen Intelligenz generiert, abgeändert oder verfälscht worden sind, erscheinen in verschiedenen Formen.

Die wohl bekannteste Form sind Deepfakes bei Fotos. Dabei werden Änderungen an einem Gesicht oder am gesamten Körper einer Person vorgenommen oder mit physischen Komponenten einer weiteren Person ergänzt oder überblendet. Deepfakes im Audiobereich verändern die aufgenommene Stimme im Original oder ahmen die einer anderen Person nach. Die Funktion «Text-zu-Sprache» bietet die bereits weitverbreitete Möglichkeit, durch Eingeben von Text Ton zu erzeugen oder so eine bereits existierende Tonaufnahme zu überarbeiten.

Die wohl trügerischste und damit auch einflussreichste Manipulation ist das Video-Deepfake. Denn auch bei bewegten Bildern können Gesichter heute einfach ausgetauscht werden. Zudem ist Gesichtsverwandlung möglich, bei der das Gesicht einer Person durch einen nahtlosen Übergang in ein anderes Gesicht übergeht. Technisch machbar ist auch das Ganzkörperpuppenspiel, bei dem die

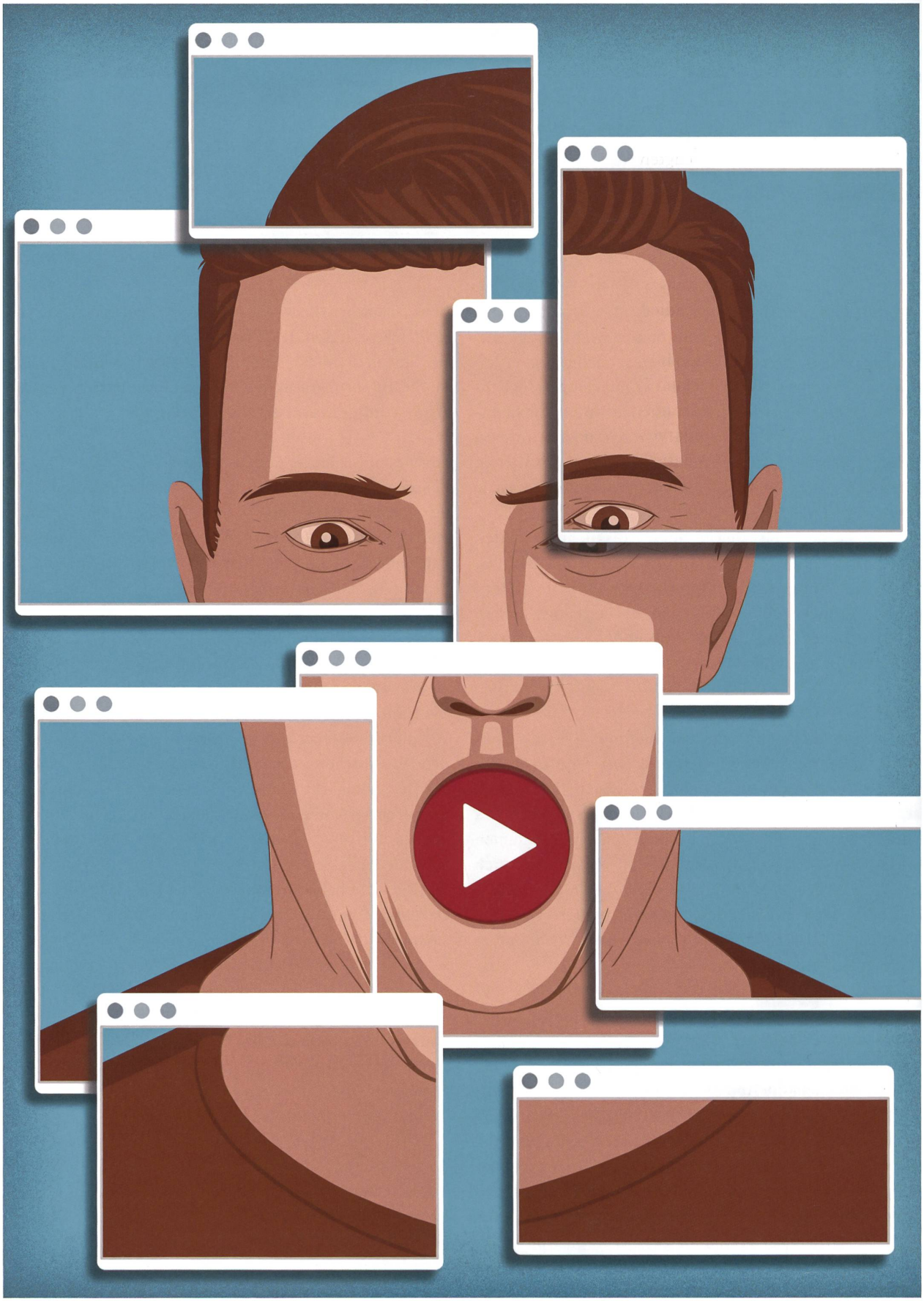

Illustration von Stephan Schmitz.

Bewegungen einer Person auf den Körper einer anderen Person übertragen werden – ähnlich wie Schauspieler ihre Bewegungen auf animierte Figuren im Film übertragen. Bei hochwertigen Video-Deepfakes ist es zudem möglich, die Lippensynchronisation zu verändern und mit einer passenden Audiospur zu unterlegen.

Hinter den künstlich erzeugten Bildern, Audios und Videos steckt weit mehr als nur eine ins Programm eingespeiste Quelldatei. Die Erstellung von glaubwürdigen Deepfakes erfordert derzeit noch Fachwissen, umfangreiche Schulungen, teure Hardware und spezielle Software. Darüber hinaus ist die Bearbeitung qualitativ hochwertiger digitaler Bild-, Video- und Audiodateien grundsätzlich zeit- und arbeitsaufwendig.

Es existieren zwar bereits einige einfach bedienbare Apps, die für ein breites Publikum zugänglich sind. Deren Ergebnisse überzeugen jedoch noch nicht wirklich und können in der Regel rasch entlarvt werden. Aufgrund der rasch fortschreitenden exponentiellen Entwicklung sollte es aber nicht mehr allzu lange dauern, bis auch einfach anwendbare Apps täuschungsechte Deepfakes produzieren können.

Im Kampf gegen missbrauchte Technologie

Politische und gesellschaftliche Institutionen – vom Geheimdienstausschuss des US-Repräsentantenhauses bis hin zur Europäischen Kommission – äussern sich vermehrt über die potentiellen Gefahren von Deepfakes, die von gefälschten Kriegserklärungen bis hin zu sexueller Belästigung und Diskriminierung durch gefälschte Pornovideos reichen. Die Ängste vor Deepfakes werden zudem vermehrt von der Öffentlichkeit geteilt: Eine Umfrage des Pew Research Centers ergab, dass 63 Prozent der Erwachsenen in den USA befürchten, dass «veränderte Videos und Bilder grosse Verwirrung über die Fakten aktueller Ereignisse stiften».[2] Die Befürchtung, dass Deepfake-Inhalte den Journalismus und vertrauenswürdige Informationsquellen potentiell untergraben könnten, ist also verbreitet.

Zudem gibt es auch ein anderes Problem: Mit dem Aufkommen äusserst glaubwürdiger Deepfakes könnten selbst korrekte Videoinhalte oder Aufnahmen von denjenigen als Deepfakes verleumdet werden, die den Inhalt für unvorteilhaft halten. Daraus entsteht ein als «Lügendividende» bezeichnetes Phänomen. Die Verbreitung von Deepfakes könnte somit insgesamt zu einem sinkenden Vertrauen in Nachrichteninstitutionen führen, indem sie Misstrauen selbst gegenüber legitimen Formen von Nachrichten und Informationen schürt. Und wo die Grenze zwischen Wahrheit und Lüge zunehmend verworrener wird, gerät auch die öffentliche Meinungsbildung unter Druck.

So sind Medienhäuser, Plattformunternehmen und Bildungsinstitutionen gefragt, dieser Gefahr zu begegnen. Sie werden eine neue Reihe von Werkzeugen benötigen, die es ihnen ermöglichen, die Authentizität digitaler Dateien zu überprüfen und zu erklären. Dabei wird es keine Schwarz-Weiss-Lösungen geben, werden doch Deepfakes und synthetische Medien auf ihre eigene Art und Weise wertvolle historische Artefakte sein, die zwingend aufgedeckt und geklärt werden müssen.

Im Kampf um den Nachweis der Echtheit werden laufend neue Programme benötigt, die es etwa Medienschaffenden erleichtern, Quellen zu prüfen und Deepfakes zweifelsfrei als solche zu enthüllen. Bildungsinstitutionen sind gefragt, vermehrt auf die Medienkompetenzförderungen von Lernenden Wert zu legen und das kritische Hinterfragen von Nachrichten zu vermitteln, insbesondere im Umgang mit den sozialen Medien. Es ist aber auch Aufgabe der Social-Media-Plattformen, Inhalte zu prüfen und die Verbreitung von Deepfakes einzudämmen, was zunehmend ein schwieriges Unterfangen darstellen wird.

Kritischer Umgang ist nötig

Deepfake-Inhalte bringen durchaus auch Chancen mit sich: So können KI-Verfahren auch als nützliches Instrument bei der Strafverfolgung sowie der polizeilichen Ermittlung und Analyse genutzt werden. Auch in der Film- und Gamingindustrie ergeben sich zweifelsohne neue Möglichkeiten.

Weil die Herstellung von Deepfake-Videos in den letzten Jahren deutlich schneller, einfacher und billiger geworden ist, wird es zunehmend schwieriger, Fakten von Fiktion zu unterscheiden. Menschen mit bösen Absichten können innert kürzester Zeit immensen Schaden anrichten, was nicht nur das Vertrauen in die Öffentlichkeit, sondern auch in die Demokratie gefährdet. Wie oben ausgeführt sind die Medien und das Bildungswesen in der Pflicht. Letztlich liegt es aber auch an uns allen als Zuschauer, Zuhörerin oder Social-Media-User, kritisch zu bleiben und Inhalte zu hinterfragen, bevor wir sie empört oder euphorisch teilen. ‹

[1] www.codastory.com/disinformation/how-disinformation-became-a-new-threat-to-women/
[2] www.pewresearch.org/fact-tank/2019/06/14/about-three-quarters-of-americans-favor-steps-to-restrict-altered-videos-and-images/

Hanna Wenger
forscht im Rahmen ihrer Masterarbeit an der Universität Fribourg zum Thema Deepfake. Sie arbeitet im Marketing des «Schweizer Monats».