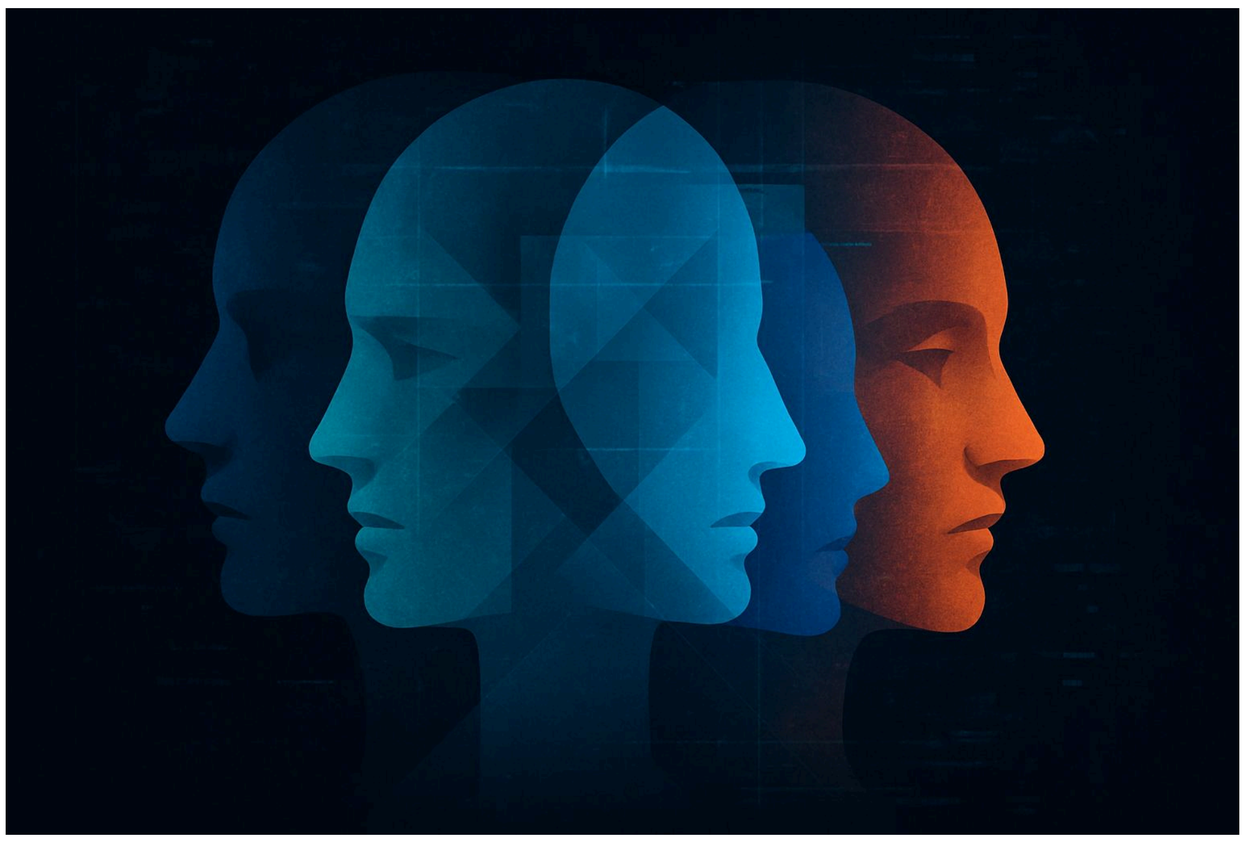
ESSAY

Ein Plädoyer für Deepfakes als Instrument der Satire

5 Februar 2024 | Anja Brauneck



Eine Art „Infokalypse“, oder auch einen „Tsunami“ an Lügen sehen viele – wohl nicht nur die üblichen Technologie-Skeptiker – durch die inzwischen alltägliche Präsenz von KI-generierten Deepfakes auf uns zukommen. Meist auch kommt ein negativer Beigeschmack auf, sobald ein Gespräch dieses Thema berührt. Das Wort Deepfake setzt sich aus den Begriffen „Deep Learning“ und „Fake“ zusammen und war ursprünglich das Pseudonym eines zweifelhaften Nutzers der Plattform Reddit. Dieser hatte 2017 erstmalig manipuliertes Videomaterial veröffentlicht. Gesichter bekannter Schauspielerinnen wurden darin – übrigens damals schon mit Hilfe einer Künstlichen Intelligenz (KI) – auf die Körper von Pornodarstellerinnen montiert.

Deepfake-Systeme können Medien wie Bilder, Audios sowie Videos manipulieren und sind Teil der sogenannten „Meme“-Kultur. Geht es um die synthetische Kopie der Stimme, ist der Fachbegriff „Voice Cloning“. Die Künstliche Intelligenz kann das, weil sie zuvor anhand von digitalisiertem Material die Verhaltensmuster etwa einer bestimmten Person (Gesicht, Körper, Gestik, Stimme u. ä.) mittels Deep-Learning-Methoden analysiert und gelernt hat. Deepfakes sollen glaubwürdig, d. h. „echt“ erscheinen, und das funktioniert mit den neuesten Programmen immer besser. Ein besonderes Fachwissen ist – anders noch als 2017 – für die Nutzung solcher Tools nicht mehr erforderlich. Dass solche Falschmeldungen anschließend in sozialen Netzwerken mitunter massenhaft verbreitet werden und oftmals in einem politischen

Kontext stehen, schafft generell Misstrauen gegenüber in den Netzwerken veröffentlichtem Bild- und Tonmaterial. Die Tatsache, dass sich Deepfakes mitunter nur mit technischer Unterstützung aufdecken lassen, schafft dieses Misstrauen nicht aus der Welt. Allerdings ist das Phänomen sogenannter Echtheitszertifikate, die mehr oder weniger offizielle Bestätigung, dass eine Fälschung ausgeschlossen sei, wiederum ein „alter Hut“.

Fast alle technischen Entwicklungen beinhalten Risiken, das versteht sich. Vor allem aber führen diese zu Neuem. So auch die erfinderischen Deepfake-Tools von heute, und das geht in der aktuellen Deepfake-Diskussion oftmals leider unter. Von ihnen profitieren nicht nur etwa medizinische Forschungsprojekte oder Bildungseinrichtungen, sondern auch viele professionelle als auch laienhafte Kreative, die mit digitalen Inhalten arbeiten. Die künstlerische Fiktion kann „real“ erscheinen – ein ewiger Traum innerhalb mancher Richtungen der Kunst- und Filmgeschichte. Für die Filmindustrie gilt dies im Besonderen.

Vor wenigen Monaten gab es in einer Ausgabe der „ZEIT“ einen Aufruf von Stand-Up-Comedians und anderen ausübenden Künstlern. Sie glauben, ihre Befürchtungen, durch KI ersetzt zu werden, mit einem Verbot von Deepfakes auflösen zu können. „Satire braucht keine Deepfakes!“, hieß es dort u. a. Wäre das aber nicht schade? Vor allem satirisch gemeinte audio-visuelle Deepfakes, die Persönlichkeiten aus Politik und Gesellschaft durch Bild-/Video-Manipulationen „auf die Schippe“ nehmen, dabei dennoch eine ernsthaft-kritische Mission verfolgen („echte“ Satiren fordern ein solches Anliegen), erleben durch die neuen gestalterischen Möglichkeiten geradezu goldene Zeiten. Die der Satire wesenseigene gestalterische Verfremdung/Übertreibung, etwa wenn einem ihrer „Opfer“ befremdliche Worte in den Mund gelegt werden, gelingt mit Hilfe dieser neuen Technik lebensechter denn je. Wenn Deepfake-Satiren ein unterstelltes Fehlverhalten der Mächtigen anprangern, machen sie das außerdem besser als andere, nämlich mit Humor, auf den Punkt gebracht und teilweise mit großer Reichweite. Satirische Deepfakes – die Rede ist nicht von gezielter Desinformation – fördern demokratische Meinungsbildungsprozesse daher effektiver und raffinierter als viele anderen Äußerungsformen. Sie stehen deshalb unter dem Schutz der Meinungsfreiheit, wie jede andere legitime Form der Meinungsäußerung auch.

Die Reflektionsfähigkeit des Rezipienten wird allerdings massiv auf die Probe gestellt, wenn es darum geht, die nahezu perfekt „versteckte“ Botschaft eines satirischen Deepfakes zu entschlüsseln. Ihre größte Hürde ist die Erkennbarkeit. Gerade die Deepfake-Satire will absolut „wahr“ erscheinen und dennoch will sie als „unwahr“ erkannt werden. Schließlich wird etwas derart Absurdes behauptet, was nicht wahr sein „kann“. Kompliziert und höchst widersprüchlich klingt das schon, aber rechtfertigt dies gleich ein Verbot?

Machbar aber ist das. Auch satirische Deepfakes können in aller Regel als solche erkannt werden, nicht nur intuitiv. Ein Faktencheck hilft zumeist und ist heutzutage ohnehin eine viel genutzte, fast automatische Methode beim Konsum von Informationen aller Art. Gut zu wissen ist auch, dass in vielen Fällen noch immer der eigene Verstand ausreicht, um die „Fälschung“

zu erkennen. Mal ehrlich, dass etwa die Politikerin Alice Weidel im Deutschen Bundestag wirklich arabisch gesprochen haben soll, ist schwer vorstellbar. Ein ziemlich echt aussehendes Video auf TikTok hatte das behauptet. Für die besonders schweren Fälle – und die gibt es ohne Zweifel – stehen die Methoden einer KI-basierten Deepfake-Detektion zur Verfügung. An solchen Systemen jedenfalls wird längst weiter geforscht und gearbeitet.

Das Missbrauchspotential der neuen Deepfake-Programme ist dennoch nicht zu unterschätzen, insbesondere wenn diese für politische Propaganda genutzt werden, ist es erheblich. Doch die bisherigen Glaubwürdigkeitskatastrophen durch etwa „Fake News“ beschränken sich bisher – auch mit Blick auf die Vereinigten Staaten – auf einzelne Vorkommnisse, und an denen ist bis jetzt noch keine Demokratie zerbrochen. Dies ist wohl eine der größten Sorgen in der Diskussion um Deepfakes. Vorsorglich aber ist das europäische wie deutsche Rechtssystem auf den Missbrauch von Deepfakes recht gut vorbereitet. Ein generelles Deepfake-Verbot hingegen ist an keiner Stelle formuliert – und das ist gut so.

Der Digital Service Act (DSA) verpflichtet seit 2024 große Plattformen und Suchmaschinen, Missbrauchsrisiken etwa durch Deepfakes insbesondere mit Hilfe von Kennzeichnungspflichten zu minimieren. Ab August 2026 tritt die KI-Verordnung (KI-VO) fast vollständig in Kraft. Ergänzend zum DSA gelten dann Kennzeichnungspflichten auch für die Anbieter und die unmittelbaren Verwender von Deepfake-Systemen. Zumal die Deepfake-Technik oftmals personenbezogene Daten automatisiert verarbeitet, greift zudem die Datenschutzgrundverordnung (DSGVO) mit ihren Regularien. Daneben schützt das nationale Recht individuelle Rechtsgüter, die spezifisch durch Deepfakes jetzt häufiger bedroht werden: Das Strafrecht etwa kriminalisiert u. a. verschiedene Formen der Beleidigung (§§ 185, 186, 187 StGB) und insbesondere die Politikerbeleidigung (§ 188 StGB, ein 2021 nachgeschärfter Straftatbestand noch von 1951). Dann gibt es das Verbot der falschen Verdächtigung (§ 164 StGB, wenn ein Deepfake als vermeintliches Beweismittel genutzt wird) sowie das Verbot der Verbreitung pornografischer Inhalte (§ 184 StGB, bei Darstellungen von Gewalt, Kindern, Jugendlichen). Ferner schützt der Bildnisschutz nach dem Kunsturhebergesetz (§ 22 KUG, auch künstlich generierte Bilder einer Person fallen darunter) sowie das verfassungsrechtlich garantierte allgemeine Persönlichkeitsrecht. Dieses gewährt dem Einzelnen einen umfassenden Schutz vor unbefugten Eingriffen in sein Selbstbestimmungsrecht und seine Stimme. Der Urheberschutz hingegen steht bei Deepfakes weniger im Vordergrund. Zwar imitieren Deepfakes mitunter urheberrechtlich geschütztes Material, sie reproduzieren dieses aber überwiegend nicht und Nachahmungen sind vom Urheberrecht nicht geschützt.

Ein rechtlicher Flickenteppich ist das, sagen die Kritiker, die noch stärker regulieren wollen. Das stimmt zwar, und in Bezug auf das Unionsrecht gibt es hinsichtlich der Um- und Durchsetzung der neuen Bestimmungen aus der KI-VO nicht unerheblich viele offene Fragen. Zudem wird das deutsche Straf-, Zivil- und Verfassungsrecht nicht jeden missbräuchlichen Einzelfall gänzlich

absichern können. In Summe sind diese Regelungen trotzdem allemal beachtlich, auch im internationalen Vergleich. Zudem schützen die Staatsanwälte wie Gerichte und greifen bei Deepfakes generell hart durch, wie erste Urteile zeigen.

Doch leider trifft dies auch – satirische – Deepfakes, und manches Urteil dazu kann durchaus Angst machen. Man denke nur an den umstrittenen Faeser-Post auf „X“ eines Chefredakteurs vom Frühjahr. Zur Erinnerung: ein Meme, auf dem die damalige Innenministerin mit einem Schild in der Hand zu sehen ist, auf dem der Satz steht „Ich hasse die Meinungsfreiheit!“. Das Urteil in erster Instanz lautete: Freiheitsstrafe von sieben (!) Monaten wegen Verleumdung gegen Personen des politischen Lebens gemäß § 188 StGB – immerhin auf Bewährung. Dabei muss man sagen, es hatte sich ganz offensichtlich um Satire gehandelt und weniger um eine bewusste Verunglimpfung einer Politikerin. Darüber war sich die Medienrechtsszene überwiegend einig und vermutlich hatten das die meisten Nutzer auf „X“ auch so gesehen.

Nur weil das Erkennen satirischer Inhalte in Deepfakes eine erhöhte Medienkompetenz bzw. Aufmerksamkeit erfordert, darf der traditionell eingeräumte rechtliche Freiraum für Satiren nicht vergessen werden. Dieser gilt in nahezu allen Judikaturen für z. B. die Karikatur – warum dann nicht auch für Deepfakes? Schließlich entwickeln sich nicht nur die Techniken weiter, sondern mit ihr zugleich die Fähigkeiten der Nutzer. In diesem Fall geht es um den Lernprozess, Inhalte immer kritischer zu reflektieren, gegebenenfalls mit Hilfsmitteln. Das sollte man den Rezipienten durchaus zutrauen. Sie werden den „Fake“ in aller Regel schon erkennen. Um Deepfakes in Schach zu halten, wird vermutlich das der größte Trumpf sein. Die vielen neuen Satiriker dürfen/sollen folglich gerne weitermachen.