SocietyByte

BFH-Magazin für die Humane Digitale Transformation

# Von Deepfakes und manipulierter Realität – eine Studie von TA-Swiss

Von Anne-Careen Stoltze (BFH Wirtschaft) | 0 Kommentare

**Künstliche Intelligenz (KI) kann Bilder, Videos und Tonaufnahmen erzeugen und verändern. Das ist schon jetzt nicht mehr aus der digitalen Welt wegzudenken. Aber dies kann missbraucht und zu Manipulationen genutzt werden. Um dies einzudämmen, braucht es einen Mix aus unterschiedlichen Massnahmen. Dies zeigt die aktuelle Studie von *TA-SWISS [https://www.ta-swiss.ch/]* über Deepfakes, die BFH-Digitalexperte Prof. Dr. Reinhard Riedl begleitet hat.**

**Societybyte: Das Thema der Studie sind Deepfakes. Warum habt ihr dieses Thema aufgenommen?**

*Prof. Dr. Reinhard Riedl:* Die kurze Antwort lautet: Neu ist es möglich, mit KI Multimediamaterial zu produzieren, das täuschend echt wirkt. Das schafft bislang unbekannte Möglichkeiten für Vieles: Film, Museum, Schule, Musik, Kunst, Kriminalistik, Justiz et cetera, aber auch für Mobbing, Betrug und politische Angriffe auf die Demokratie. Es untergräbt damit auch die Glaubwürdigkeit von Multimediamaterial an sich: «Ist diese Tonaufnahme, ist dieses Video echt??» Konkret kann man sogar das Live-Manipulieren von Überwachungsvideos online einkaufen: es wird zum Teil offen angeboten. Zudem ist eine reale Gefahr, dass Staaten die demokratischen Wahlen in anderen Staaten mit Deepfakes angreifen. Wenn solche neuen Möglichkeiten sich auftun – Chancen wie Risiken – ist es Zeit für eine TA-Swiss Studie.

*Prof. Dr. Reinhard Riedl*

**Welche Fragen habt ihr untersucht?**

Aber da wir nur ein beschränktes Budget haben, geht es nicht um die Frage «Ist das Thema wichtig?», sondern um die Frage «Ist es wichtiger als andere?». Deshalb erfolgt die Auswahl der Studienthemen in einem methodischen Prozess – in Zusammenarbeit zwischen Geschäftsstelle und Leitungsausschuss. Hauptadressat der TA-Swiss Studien sind das Parlament und die Regierung. Die Hautfrage lautet also: Welche Technologietrends sind für die Legislative und die Exekutive besonders wichtig? Wo haben die Informationsbedarf? Allerdings sind wir autonom in unserer Themenwahl. Unsere Aufgabe ist beides: Monitoring aller Technologie-Entwicklungen und Detailanalyse kritischer Entwicklungen. Wichtig ist, dass bei den von uns ausgesprochenen Empfehlungen wir der wissenschaftlichen Objektivität verpflichtet bleibe. Studien, wie sie beispielsweise deutsche Minister*innen in Auftrag geben, bei denen die beteiligten Wissenschafter*innen Aktivist*innen sind oder klare politische Präferenzen haben, wären bei TA-Swiss unvorstellbar.

**Wie war die Zusammenarbeit aufgeteilt?**

Der Leitungsausschuss entschied, dass das Thema wichtig genug ist, um eine Studie zu finanzieren. Allerdings führt TA-Swiss Studien nicht selbst durch. Es schreibt sie aus, wählt die beste Offerte, organisiert eine Begleitgruppe, entscheidet über die Veröffentlichung, erstellt eine Kurzfassung und sorgt für die Verbreitung, neudeutsch die Dissemination. Meine konkrete Rolle dabei war, die Begleitgruppe zu präsidieren. Dank der guten Arbeit des Studienteams und der Kompetenz und des Engagements der Begleitgruppe war das eine spannende Aufgabe.

**Welches sind die grössten Risiken von Deep Fakes?**

Das grösste Risiko besteht darin, dass wir glauben, Deepfakes erkennen zu können. Dabei ist in Experimenten, in denen Menschen wissen, dass es sich um Deepfakes handeln könnte, die Zuordnung weitgehend zufällig. Das Erkennen durch genau Hinschauen funktioniert nicht. Das hat unter anderem auch eine Studierendengruppe bei mir in einem Experiment gezeigt. Notwendig ist deshalb – und das ist ganz schwierig – in Wahrscheinlichkeiten bezüglich der Echtheit zu denken. Natürlich werden die Technologien zur Deepfake-Erkennung immer besser, aber eben auch die Technologien zur Erstellung der Deepfakes.

**Wie schätzt du die Lage in der Schweiz ein?**

Die aktuelle Gefahrenlage ist überschaubar, das hat unter anderem die Studie gezeigt, aber ihre Entwicklung ist nicht vorhersagbar. Nicht die Technologie verändert die Gesellschaft, sondern die Technologienutzung. Die Frage ist nicht, was KI kann. Die Frage ist, wer KI wofür nutzt. Das ist nur kurzfristig antizipierbar, mittel- bis langfristig ist es völlig unklar. Es ist ja schon die Technologieentwicklung selbst schwer prognostizierbar. Aber ich will mich nicht mit schlauen Trivialitäten vor der Antwort drücken: Die Schweiz ist nicht besonders gefährdet, sollte sich aber darauf vorbereiten, Opfern von Deepfakes zu helfen und auf Deepfake-Angriffe aus totalitären Staaten zu reagieren. Das dabei aufgebaute Knowhow wird man auch dann weiter nutzen können, wenn kriminelle Delikte und politische Angriffe mit Deepfakes die Ausnahme bleiben.

**Und im Vergleich zu anderen Ländern bzw. der EU?**

Was den Vergleich mit anderen Ländern betrifft, so spreche ich hier als Privatperson. Nicht als Wissenschafter, nicht als Mitarbeiter der BFH und nicht als Mitglied des TA-Swiss Leitungsausschusses: Die Schweiz ist zu klein, zu heterogen und geopolitisch zu wenig wichtig, um ein attraktives Ziel abzugeben. Es ist eher wahrscheinlich, das aus der Schweiz heraus EU-Wahlen angegriffen werden als das ferne Grossmächte die Schweizer Wahlen angreifen lassen. Das heisst allerdings auch: Die Polizei muss auch auf Taten aus der Schweiz im Ausland vorbereitet sein. Denn wenn es passiert, wird der Hinweis, dass das Ausländer waren, nicht reichen, um die resultierende Beziehungskrise zu managen.

**Wie ist das Studienteam vorgegangen?**

Die Forschung setzte auf fünf Methoden: eine Literaturanalyse, eigene technischen Experimenten, eine Medienanalyse, Expert*inneninterviews und eine Bevölkerungsumfrage. Die aufgearbeiteten Ergebnisse wurde jeweils der Begleitgruppe vorgelegt, die dazu Feedback gab. Vier Perspektiven wurden vom Forschungsteam besonders ausführlich behandelt: rechtliche Aspekte, Deepfakes im Journalismus, Deepfakes in der Politik und Deepfakes in der Wirtschaft. Da der Stand der Technik und die erwartbaren technischen Entwicklungen grundlegend sind, werden die technischen Grundlagen im Schlussbericht schon in der Einführung besprochen. Anschliessend folgen dann die rechtlichen Aspekte – konkret der Schutz vor Deepfakes, Deepfakes in Gerichtsverfahren, öffentlich-rechtliche Vorgaben und zukünftige Regulierungsoptionen – bevor die Bereiche Journalismus, Politik und Wirtschaft analysiert werden.

**Wie war eure Zusammenarbeit?**

Im Fazit werden die Querbezüge zwischen den einzelnen Bereich nochmals aufgezeigt. Das war nicht einfach, weil Forschende natürlich durch die Peer-Review-Logik darauf konditioniert sind, eng fokussiert zu arbeiten. Aber ich denke, es ist gelungen. Auf der Basis der Forschungsergebnisse wurden auch Empfehlungen ausgearbeitet. Natürlich wurden auch diese Empfehlungen intensiv zwischen Forschungsteam und Begleitgruppe diskutiert. Dieser Austausch war teilweise stark emotional, aber immer konstruktiv, und das Forschungsteam hat die Vorschläge gut aufgenommen. In Summe ein herausfordernder Arbeitsprozess. Wir waren als Begleitgruppe sicher nicht immer einfach für das Forschungsteam, aber das Endergebnis hat vom Vorgehen profitiert.

## Zu welchen Ergebnissen ist das Forschungsteam gekommen?

- Erstens dass die Wahrnehmung durch das Labelling bestimmt wird. Deepfakes werden negativ wahrgenommen, mit KI erzeugte Inhalte hingegen nicht. Der wissenschaftliche Begriff «synthetische Medien» ist gar nicht bekannt. Bei Deepfakes werden Gefahren für die Gesellschaft, konkret die Schweizer Demokratie, gesehen, individuelle Risiken dagegen kaum.
- Zweitens helfen Tipps für den Umgang wenig. Vertrautheit mit den digitalen sozialen Medien ist wichtig. Das macht klar, wie wichtig das Fach Medien und Informatik ist. Realistisch betrachtet gibt es dort aber selten genügend Ressourcen für eine reflektierende Auseinandersetzung mit Deepfakes.
- Drittens stellen Deepfakes für den Journalismus eine fachliche und ökonomische Herausforderung, Soziale Plattformen können beim Verbreiten von Videos ignorieren, ob es sich um Deepfakes handelt, Journalist*innen sind der sachgerechten Darstellung verpflichtet. In der Schweizer Praxis sind Journalist*innen vorerst aber fast nur im Auslandsjournalismus konfrontiert. Grundsätzlich könnte im übrigen das Deepfake-Phänomen auch den wahrgenommenen Wert von Qualitätsjournalismus steigern.
- Viertens wurden im Bericht der Studiengruppe die verschiedensten Gefahren dargestellt. Hier ist Wirtschaftsspionage unter Zuhilfenahme von Deepfakes für die Schweiz eine relevante Bedrohung, weil viele Schweizer Unternehmen attraktive Angriffsziele darstellen. Ausserdem müssen sich die Gerichte neu mit den Fälschungsmöglichkeiten von Überwachungsvideos auseinandersetzen. Und fünftens werden im Bericht auch zahlreiche Chancen beschrieben. In der Filmindustrie sind Deepfakes bereits angekommen, viele andere Sektoren haben das theoretische vorhandene Potential jedoch noch kaum genutzt. Auf absehbare Zeit dürfte das auch nicht stattfinden.

## Und welche Massnahmen empfehlt ihr?

Die Empfehlungen lauten: Selbstverantwortung wahrnehmen und den technischen Fortschritt für die Verteidigung nutzen, Plattformen in die Pflicht nehmen und Opferschutz stärken, international bei der Verfolgung von Täter*innen zusammenzuarbeiten – und vor allem mehr Aufklärungen über Gefahren und Chancen. Wir sollten über verschiedenste Kanäle – darunter natürlich auch die Schule – das Bewusstsein stärken, dass es einfach ist auf Deepfakes hereinzufallen. Es wird voraussichtlich keine Lösung für das Deepfakes-Problem geben. Wir werden lernen müssen, damit zu leben.

**Gerne noch etwas zum Ausblick – zu welchen Regulierungen sollte sich die Schweiz durchringen?**

Es gibt zwei prioritäre Handlungsbereiche: Der Aufbau von Knowhow bei Polizei und Justiz sollte verstärkt vorangetrieben werden in der Schweiz, auch wenn dies nur durch internationale Zusammenarbeit geht. Im Gegenteil ist die internationale sogar wünschenswert. Und die Bildung sollte verstärkt werden. Es geht dabei nicht darum, dass Lehrer*innen zu KI-Expert*innen werden, sondern dass sie dazu befähigt werden, den kritischen Umgang mit Information aus dem Internet zu unterrichten. Dazu benötigen sie Unterrichtsmaterial und Unterrichtsstunden. Wenn wir in diesen beiden Handlungsbereichen vorwärts kommen, haben wir schon viel erreicht. Daneben sollte man sich schon jetzt Gedanken über die Opferhilfe machen und allenfalls durch Wettbewerbe positive Nutzungen von Deepfakes fördern, beispielsweise in der Wissensvermittlung. Grossen zusätzlichen Regulierungsbedarf sehe ich dagegen derzeit nicht.

## Über die Studie

Die Stiftung TA-SWISS [https://www.ta-swiss.ch/] , ein Kompetenzzentrum der Akademien der Wissenschaften Schweiz, setzt sich mit den Chancen und Risiken neuer Technologien auseinander. Sie hat die Studie bei einem interdisziplinären Team in Auftrag gegeben. Diese wurde unter der Leitung von Murat Karaboga (Fraunhofer-Institut für System- und Innovationsforschung ISI in Karlsruhe) durchgeführt. Das Forscherteam wurde von einer Expert*innengruppe beratend begleitet, deren Präsident Prof. Dr. Reinhard Riedl ist.

In der Studie «Deepfakes und manipulierte Realitäten» wird eine Auslegeordnung gemacht. Sie zeigt auf, welche Rahmenbedingungen bereits jetzt für Deepfakes gelten und wo noch Regulierungsbedarf besteht. Zudem wird in der Studie untersucht, inwieweit sich Bürgerinnen und Bürger von gefälschten Inhalten aufs Glatteis führen lassen. Im Hinblick auf die Chancen von Deepfakes wird exemplarisch aufgezeigt, in welchen Bereichen synthetisch erzeugte Inhalte einen Mehrwert aufweisen.

Zur Studie [https://zenodo.org/records/11643644] und zur Kurzfassung [https://zenodo.org/records/11643843] .

AUTHOR: ANNE-CAREEN STOLTZE

Anne-Careen Stoltze ist Redaktorin des Wissenschaftsmagazins SocietyByte und Host des Podcasts "Let's Talk Business". Sie arbeitet in der Kommunikation der BFH Wirtschaft, sie ist Journalistin und Geologin.

Posts from Anne-Careen Stoltze | Website

Create PDF

## Ähnliche Beiträge

Es wurden leider keine ähnlichen Beiträge gefunden.

0

COMMENTS