

Deepfakes

ECHT ODER NICHT?

Ein Bild von Papst Franziskus im weißen Daunenmantel, ein Foto von Donald Trumps angeblicher Verhaftung oder Nacktbilder von Taylor Swift auf der Plattform X: Diese Bilder gingen um die Welt und sorgten für Aufsehen. Alle Beispiele sind „Deepfakes“ und haben gemeinsam, dass künstliche Intelligenz (KI) bei der Fälschung der Bilder beteiligt war <https://t1p.de/faktencheck-papst>. Solche realistischen Fälschungen tauchen in Bild- oder Videoform inzwischen immer häufiger auf. Ein Grund ist, dass die Technologie hinter Deepfakes immer ausgefeilter wird.

Nicht nur für Kinder und Jugendliche ist schwer zu erkennen, ob ein Bild oder Video echt ist oder nicht. Es wird immer wichtiger, die Fähigkeit zu entwickeln, echte von falschen Informationen zu unterscheiden und Bilder sowie Videos in ihrem Kontext zu verstehen und zu überprüfen.

AKTUELLES

Neue KI erstellt fotorealistische Videos

Die Firma Open AI hat ein neues KI-Modell entwickelt, das Textbefehle in fotorealistische Videos umwandelt. Das nun vorgestellte Programm heißt Sora und ist in der Lage, Videos von einer Länge bis zu 60 Sekunden zu generieren. Obwohl es noch nicht für die breite Öffentlichkeit verfügbar ist, sorgt es bereits jetzt für großes Aufsehen im Netz. Auf der Webseite von Sora <https://openai.com/sora> kann man sich ein Bild davon machen, wie täuschend echt manche der Videos aussehen.

www.klicksafe.de/news/kuenstliche-intelligenz-erstellt-fotorealistische-videos



MANIPULIERTE MEDIENINHALTE DURCH KI

Deepfakes sind durch künstliche Intelligenz (KI) erstellte oder manipulierte Foto-, Video- oder Sprachaufzeichnungen, die täuschend echt wirken. Der Begriff leitet sich von den englischen Worten Deep Learning (eine Form des maschinellen Lernens) und Fake (Fälschung) ab. Viele Deepfakes werden aus Spaß oder zur Unterhaltung erstellt und veröffentlicht. Allerdings gibt es auch immer mehr Fälle, in denen solche Fälschungen mit bösen Absichten produziert werden, um das Ansehen von öffentlichen oder privaten Personen zu schädigen. Die einfach zu bedienende Technologie hinter den Deepfakes ermöglicht die Erstellung überzeugender gefälschter Inhalte wie Fake-Pornos oder gefälschter Social-Media-Profile. Die

Reportage des Y-Kollektives zeigt, wie einfach das geht: <https://t1p.de/virtual-influencers>.

KI-Systeme sind leicht zugänglich, oft reicht eine einfache Registrierung. Sogar Kinder können sie problemlos nutzen. Unternehmen wie Microsoft und Google integrieren KI-Technologien vermehrt in ihre Produkte und Plattformen. Das bedeutet, dass Kinder und Jugendliche in Zukunft nahezu unausweichlich mit KI-Inhalten in Berührung kommen werden <https://t1p.de/klicksafe-ki-JMS>.

Es ist die Aufgabe von Erwachsenen, bereits bei Kindern ein Verständnis für die Chancen und Risiken der neuen technologischen Möglichkeiten zu schaffen. Die

THEMEN IN DIESER AUSGABE

- Echt oder nicht?
- Manipulierte Medieninhalte durch KI
- Eigene Daten schützen
- Informieren und aufklären

TERMINE

Digitales Fortbildungsangebot für die SPFH: Update „Familien mit Migrationshintergrund“

Online am 18.04.2024

<https://t1p.de/spfh-fortbildung>



AKTION JUGENDSCHUTZ

Landesarbeitsstelle
Baden-Württemberg
0711-23737-0
info@ajs-bw.de
www.ajs-bw.de
Redaktion: Petra Wolf

EIGENE DATEN SCHÜTZEN

Sobald ausreichend Video- oder Bildmaterial einer Person online verfügbar ist, können neue manipulierte Inhalte erstellt werden. So können Heranwachsende selbst Opfer von Deepfakes werden, denn sie teilen gerne Selfies und Videos in sozialen Medien. Wird man selbst Opfer von Deepfakes, ist es wichtig, die Fotos/Videos als Beweismaterial abzuspeichern und Screenshots zu erstellen. Danach sollten die sozialen Netzwerke kontaktiert werden, auf denen die manipulierten Inhalte geteilt wurden. Die Plattformen sind verpflichtet, rechts-

verletzende Inhalte schnell zu prüfen und zu löschen. Zusätzlich kann Anzeige bei der Polizei erstattet werden, denn rechtlich gesehen verletzen Deepfakes das Persönlichkeitsrecht und können zu strafrechtlichen Konsequenzen führen <https://tip.de/handysektor-tipps>.

Es ist ratsam, möglichst wenig Material von sich oder den eigenen Kindern im Netz online zu teilen und auch Heranwachsende dafür zu sensibilisieren, Inhalte in sozialen Netzwerken nur für einen ausgewählten Freundeskreis sichtbar zu machen.

Es ist die Aufgabe von Erwachsenen, bereits bei Kindern ein Verständnis für die Chancen und Risiken der neuen technologischen Möglichkeiten zu schaffen.

INFORMIEREN UND AUFKLÄREN

Kinder und Jugendliche müssen verstehen, dass mithilfe von künstlicher Intelligenz täuschend echte Medieninhalte von jeder beliebigen Person erstellt werden können.

Eine gute Möglichkeit, um über die Technologie hinter KI-generierten Bildern zu sprechen, ist die Seite www.thispersondoesnotexist.com. Bei jedem Neuaufruf ist das Bild einer Person zu sehen, die es gar nicht gibt. Die Seite www.whichfacesreal.com zeigt immer zwei Gesichter, eines davon ist echt, eines davon mit einer KI konstruiert. Wer in der Familie kann erkennen, welches Gesicht real ist? Gemeinsam kann mit Heranwachsenden mit Fotoanwendungen wie www.midjourney.com oder <https://stablediffusionweb.com/#demo> experimentiert werden. Beide erstellen Bilder anhand von schriftlichen Anweisungen. Die Ergebnisse können genutzt werden, um die Frage zu beantworten: Woran erkennen wir, dass die Bilder nicht echt sind? Anzeichen können optische Details wie Bildflackern, seltsame Formen von Gliedma-

ßen oder unwirkliche Proportionen und Perspektiven sein. Ganz grundsätzlich helfen bei der Frage, ob es sich um Deepfake-Bilder oder Videos handelt, alle Methoden, die allgemein zur Überprüfung von Falschinformationen nützlich sind: Ergibt das Gesehene überhaupt einen Sinn? Wurde der Inhalt auf einer vertrauenswürdigen Seite veröffentlicht oder (nur) auf Social Media? Haben Faktencheck-Portale wie correctiv.org oder mimi-kama.org die Inhalte schon überprüft? <https://www.fluter.de/deepfakes-erkennen>

Es ist wichtig, dass Eltern Kindern und Jugendlichen vermitteln, dass KI-Anwendungen verantwortungsvoll genutzt werden müssen. Es ist inakzeptabel, damit absichtlich schädliche Inhalte zu erstellen. Selbst scheinbar harmlose Scherze wie das Erstellen peinlicher Fotos von Mitschüler:innen durch KI können ernsthafte negative Auswirkungen für die betroffene Person haben und rechtliche Konsequenzen nach sich ziehen.

LINKS, TIPPS

- Deep Fakes & Co. — Wie umgehen mit Risiken von Künstlicher Intelligenz? **Klicksafe-Expert:innentalk** <https://kurzelinks.de/klicksafe-deepfakes>
- **For Your (Dis-)Information** Die YouTube-Quiz-Show geht Desinformationen im Netz auf den Grund <https://kurzelinks.de/FYI>

FÜR KINDER

- **Logo!** Die Kindernachrichten des ZDF erklären Deepfakes: Was ist das und wie entstehen sie? <https://kurzelinks.de/logo-deepfakes> (ab 8 Jahren)

FÜR JUGENDLICHE

- Deepfake-Videos: Erkennst du die Fälschung? **Reportage** der Funk-Reporter <https://kurzelinks.de/funk-deepfakes>
- **BAIT**—der Faktencheck auf TikTok gibt jungen Menschen das Handwerkszeug, um Fakes und Verschwörungsmymen zu enttarnen @bait.faktencheck