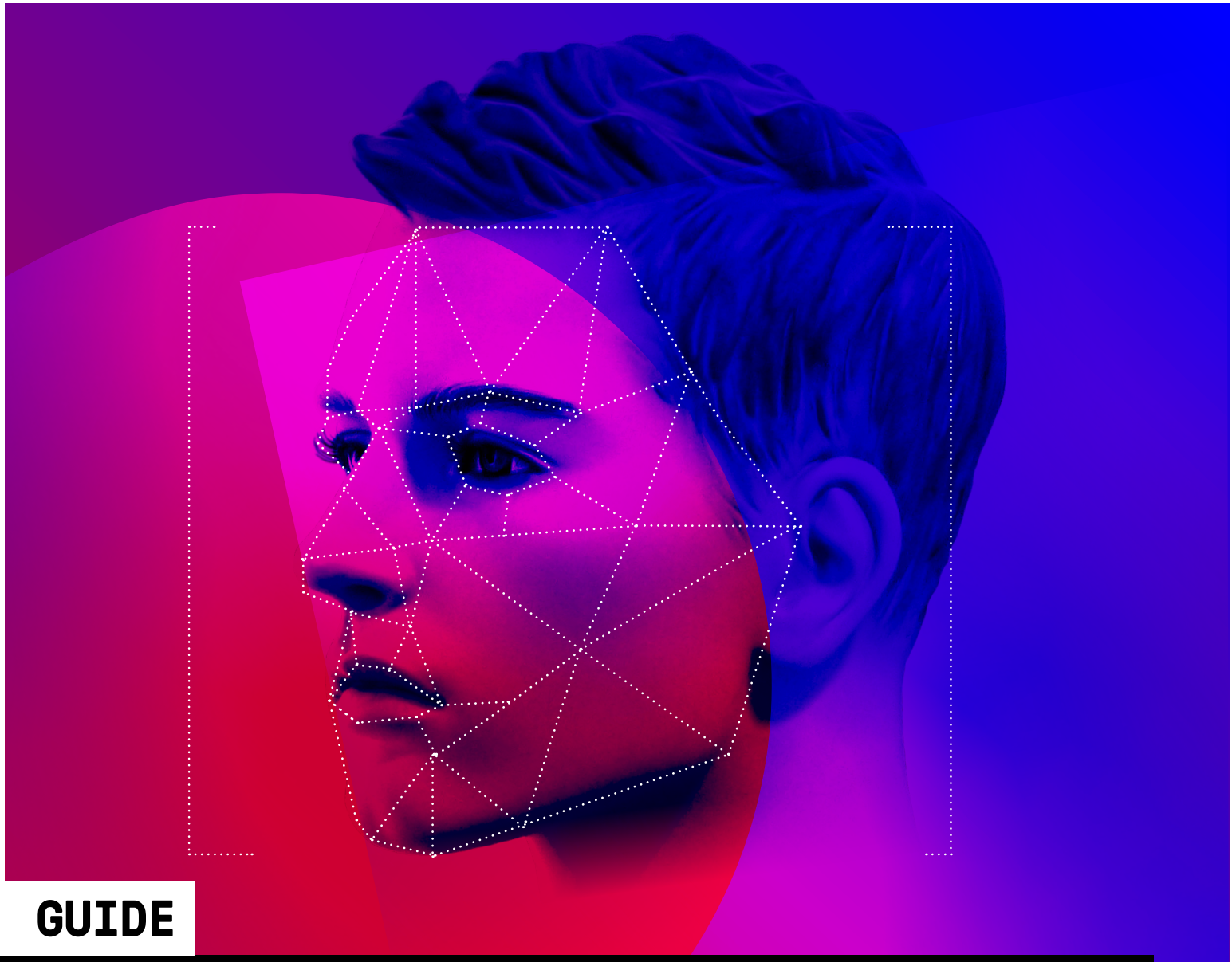


GUIDE

Deepfakes – eine rechtliche Einordnung durch die Kanzlei SKW Schwarz

Veranstalter
koelnmesse

Fachlicher und ideeller Träger der DMEXCO sowie Markeninhaber
BVDW

Unter besonderer Mitwirkung
OVK ONLINE-VERMARKTERKREIS IM BVDW

DIGITAL MARKETING EXPO & CONFERENCE

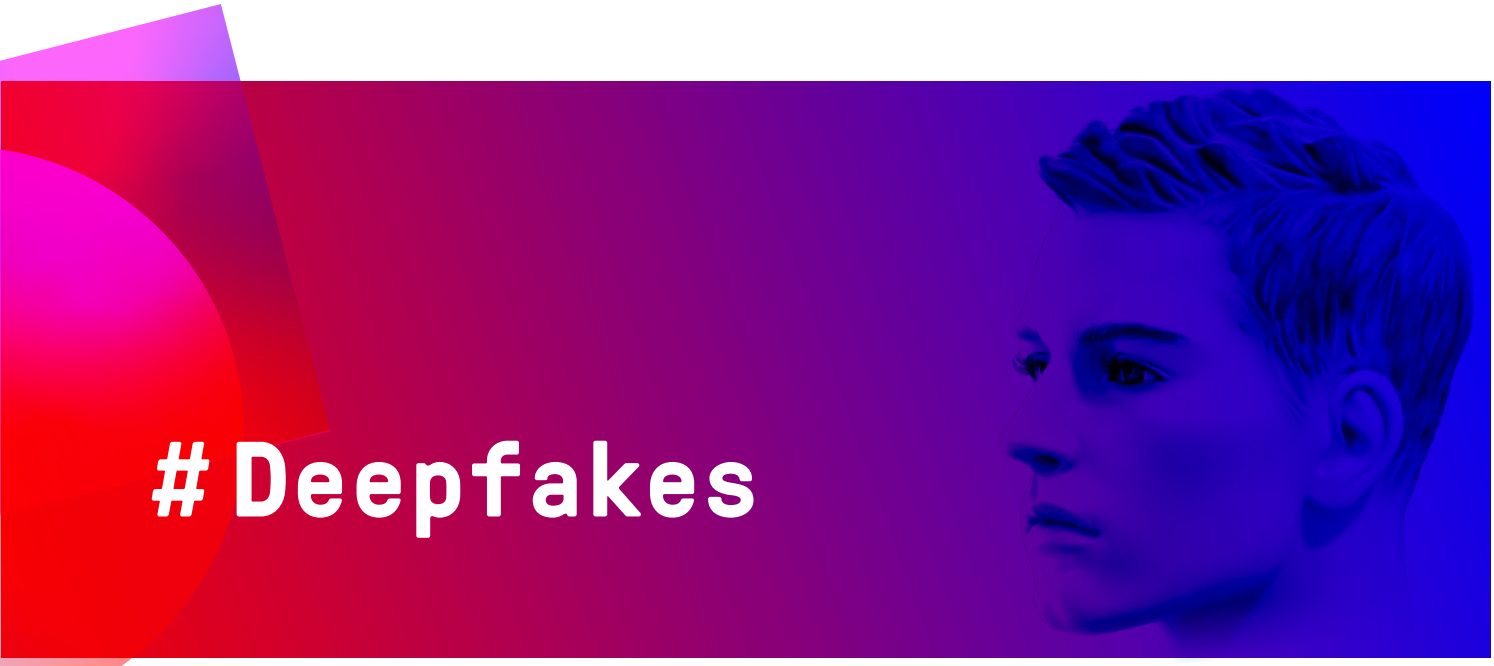
Grundlagen, Möglichkeiten und rechtliche Einordnung: Deepfakes im Marketing

Durch den technologischen Fortschritt haben sich Deepfakes rasant verbreitet und werden mittlerweile bereits von einigen Unternehmen für Werbemaßnahmen genutzt. Höchste Zeit also, dass sich Marketer:innen gründlich mit der neuen Technologie auseinandersetzen. Im DMEXCO Whitepaper befassen sich die Expert:innen der Kanzlei SKW Schwarz für dich ausführlich mit den Grundlagen, Einsatzmöglichkeiten und rechtlichen Aspekte von Deepfakes.

Inhalt

Deepfakes im Marketing und die Rolle von KI

Kann das sein? Angela Merkel wirbt für einen Stromanbieter? Skurril, aber dank moderner Deepfake-Technologie möglich. Die Berliner Agentur Try No Agency schaltete erstmals eine mittels Deepfake-Technologie erstellte Werbung in Deutschland. Darin rief ein Deepfake der damals noch amtierenden Bundeskanzlerin Merkel aufgrund der steigenden Strompreise dazu auf, den Stromanbieter zu wechseln. Immer mehr Unternehmen ziehen seither nach und setzen Deepfakes in der Werbung und in den Medien ein. Die kreativen Möglichkeiten scheinen grenzenlos. Aber was sind Deepfakes überhaupt? Wie können Unternehmen sie gewinnbringend nutzen? Und was muss ich dabei rechtlich beachten? Diese und viele weitere Fragen beantworten wir in diesem Whitepaper zu dem spannenden Thema „Deepfakes / Use of AI“.



Was sind Deepfakes?

Der Begriff „Deepfake“ setzt sich zusammen aus „Deep Learning“ und „Fake“. Deepfakes sind täuschend echt wirkende, mittels künstlicher Intelligenz (KI) erzeugte oder manipulierte Text-, Bild-, Audio- oder Videodateien. So können etwa synthetisch erzeugte Audio- und Videoaufnahmen von berühmten Persönlichkeiten in Werbespots und in den Medien eingesetzt werden, ohne dass die jeweilige Person tatsächlich in

den Aufnahmeprozess involviert war. Dabei wirken Bilder, Gestik und Mimik so authentisch, dass die Fälschung für Laien und zunehmend auch für technische Expert:innen nicht erkennbar ist.

Wie werden Deepfakes erzeugt?

Deepfakes werden mittels Deep Learning erzeugt, einer Form der künstlichen Intelligenz. Dabei werden künstliche **neuronale Netzwerke** eingesetzt, also Netzwerke, in denen künstliche Neuronen durch Algorithmen die Nervenzellen im Gehirn nachahmen. Diese KI-Systeme werden durch die Zufuhr großer Datenmengen in Form von Text-, Bild-, Audio- oder Videodateien trainiert. Das System analysiert die Daten und erkennt die darin enthaltenen Muster und Eigenschaften. Aus diesen bei der Analyse herausgefilterten Mustern und Eigenschaften kann die KI dann neue Inhalte generieren. Der künstlich generierte Inhalt, zum Beispiel das täuschend echt wirkende Bild einer Person, ist dann der sogenannte Deepfake. Durch den Prozess werden die KI-Systeme kontinuierlich leistungsfähiger, da sie mit immer größeren Datenmengen trainiert werden. Manche Programme können Deepfakes sogar in Echtzeit erstellen. Auch Apps zur Erzeugung von Deepfakes sind mittlerweile für jeden zugänglich, sodass Deepfakes heute mittels Laptop oder Smartphone autonom und ohne jede Expertise erstellt werden können. Nutzer:innen müssen lediglich die entsprechenden Dateien in Form von Text-, Bild-, Audio- oder Videomaterial zur Verfügung stellen.

Welche Arten von Deepfakes gibt es?

Mittels Deepfakes können insbesondere Gesichter, Stimmen und Texte gefälscht werden.

Beim sogenannten **Face Swapping** werden Gesichter in einem Video ausgetauscht. Ziel ist es, das Gesichtsbild einer anderen Person mit derselben Mimik, Gesichtsbeleuchtung und Blickrichtung in ein Video zu übertragen. Das Ganze funktioniert so: Neuronale Netze extrahieren kodiert die relevanten Mimik- und Beleuchtungsinformationen aus einem Gesichtsbild und erzeugen daraus ein entsprechendes neues, gefälschtes Gesichtsbild. So gibt es etwa ganze Filmszenen von „Kevin allein

zu Haus“, in denen das Gesicht der Hauptfigur Kevin, gespielt von Macaulay Culkin, durch das von Sylvester Stallone ausgetauscht wurde.

Beim sogenannten **Face Reenactment** können täuschend echte Videos einer Person durch die Manipulation von Mimik-, Kopf - und Lippenbewegungen erstellt werden. Die KI produziert ein 3D-Modell des Gesichts der Zielperson, das von der manipulierenden Person anhand eines Videostreams kontrolliert werden kann. Diese kann so täuschend echte Gesichtsausdrücke erzeugen. So ist es möglich, gefälschte Videos von Personen zu erstellen. Die US-Organisation „RepresentUS“ hat beispielsweise aus Anlass der US-Wahlen 2020 im Rahmen einer Antikorruptionskampagne ein mittels Deepfake-Technologie gefälschtes Video von Kim Jong-Un erzeugt, in dem dieser die US-Bürger vor dem Untergang der US-Demokratie warnt.

Durch das **Synthetisieren von Gesichtsbildern** ist es außerdem möglich, neue (Pseudo-)Identitäten zu erstellen. Die Verfahren beschränken sich bisher auf einzelne Bilder. So gibt es auf Instagram beispielsweise virtuelle Influencer, die vollständig KI-generiert sind, was mit bloßem Auge nicht erkennbar ist.

Zur Fälschung von Stimmen können mittels sogenannter **Text-to-Speech-Verfahren** Texte in eine Audio-Datei umgewandelt werden. Hierbei wird ein Text vorgegeben, der durch die KI verarbeitet und in ein Audio-Signal umgewandelt wird, das sich wie die Zielperson anhört. Weiterhin ist es möglich, in einer Audiodatei die Stimme durch das sogenannte **Voice-Conversion-Verfahren** zu verändern. Hier gibt die anwendende Person ein Audiosignal vor, indem sie zum Beispiel einen Text einspricht. Dieses Audiosignal wird dann zu einer Zielstimme konvertiert.

Bei der **Fälschung von Texten** werden KI-basierte Texte erstellt, bei denen auf den ersten Blick nicht unterschieden werden kann, ob sie von einem Menschen oder

von einer Maschine verfasst worden sind. Ein derzeit prominentes Beispiel dafür ist ChatGPT, aber auch andere Textgeneratoren, Chatbots oder Social Bots. Meist reichen wenige Wörter aus, anhand derer die KI einen Text vervollständigen oder neu erstellen kann, sodass Nachrichten, Blog-Einträge, Chat-Antworten oder ganze Texte generiert werden können.

Wozu können Deepfakes verwendet werden?

Deepfakes sind vielseitig einsetzbar – sowohl im positiven als auch im negativen Sinne. Die Anwendungsmöglichkeiten reichen von Satire und Unterhaltung über bewusste Desinformation und Propaganda bis hin zur gezielten Diskreditierung einzelner Personen. Deepfakes können sogar im Gesundheitswesen eingesetzt werden. So entwickelte das kanadische Start-up Lyrebird Produkte mithilfe der Deepfake-Technologie, die beispielsweise ALS-Patienten bei Verlust der Sprachfähigkeit wieder eine Stimme geben. Auch im Bereich Kultur sind Deepfakes kreativ einsetzbar. Im Dalí-Museum in Florida werden die Besucher:innen zum Beispiel von dem bereits verstorbenen Künstler Salvador Dalí mittels Deepfake-Technologie in Lebensgröße begrüßt. Um den Deepfake-Dalí zu erstellen wurde eine KI mit etwa 6.000 Bildern aus Dalí-Videoaufnahmen trainiert.

Welche Möglichkeiten gibt es, Deepfakes in den Medien einzusetzen?

Auch in den Medien gibt es zahlreiche Möglichkeiten, Deepfakes gewinnbringend einzusetzen. So können durch den Einsatz von Deepfakes Produktionsprozesse in allen relevanten Branchen wie der Film- und Musikindustrie vereinfacht und beschleunigt werden.

In der **Filmproduktion** eröffnet die Deepfake-Technologie vielseitige neue Erzählmöglichkeiten. Deepfakes werden beispielsweise zur Nachbildung verstorbener Schauspieler:innen eingesetzt. Dadurch ist es möglich, die Dreharbeiten mit dem Deepfake bereits verstorbener Schauspieler:innen abzuschließen, sodass die Filmfigur nicht ebenfalls sterben oder das Drehbuch angepasst werden muss. Außerdem können

Schauspieler:innen mittels Deepfake-Technologie verjüngt werden, wie es beispielsweise mit Will Smith in „Gemini Man“ geschehen ist, als der von ihm gespielte Henry Brogan seinen 25 Jahre jüngeren Klon trifft. Dies erspart insbesondere eine aufwendige Maske bei der Filmproduktion. Durch Deepfakes können Filme zudem so einfach wie noch nie in verschiedene Sprachen synchronisiert werden. Ein großer Vorteil daran ist, dass die Qualität der Synchronisation durch die neue Technologie sogar steigt. Denn die Lippenbewegungen von Schauspieler:innen können automatisch an die jeweilige Synchronsprache angepasst werden, wodurch ein realistischeres Filmerlebnis entsteht. Außerdem kann die Originalstimme von Schauspieler:innen in allen Sprachen beibehalten werden, sodass es nicht mehr nötig ist, Synchronstimmen zu casten. Der polnische Film „The Champion of Auschwitz“ (2021) von Regisseur Maciej Barczewski ist der erste Film, der mittels KI synchronisiert wurde.

Auch in der **Musikbranche** ermöglicht die Nutzung von Deepfakes mittlerweile vieles, was vor Kurzem noch undenkbar erschien. So können Songtexte und ganze Kompositionen mittels KI erstellt und mit der KI-generierten Stimme bestimmter Künstler:innen wiedergegeben werden. Dabei ist für Zuhörer:innen nicht erkennbar, dass der Song ohne den stimmlichen Einsatz menschlicher Künstler:innen produziert wurde. Die KI kann sogar so weit trainiert werden, dass sie ganze Musikstile nachahmt. Wissenschaftler:innen des SONY CSL Research Lab haben erstmals vollständige Songs durch KI schreiben lassen, darunter zum Beispiel den Song „Daddy’s Car“ im Stil der Beatles. Deepfakes ermöglichen in der Musikproduktion damit eine Prozessoptimierung in großem Stil. Auch können die Stimmen verstorbener Künstler:innen mittels Deepfake nachgeahmt und für fortlaufende Musikproduktionen eingesetzt werden. Ein weiteres Beispiel für den Einsatz von Deepfakes in der Musik lieferte der YouTuber Roberto Nickson, der mittels Deepfake den Rap-Song „Music is forever changed“ mit der Stimme des Rappers Kanye West aufgenommen hat. Den Song hat Nickson mithilfe von Voice-Conversion-Verfahren selbst eingerappt. Das Ergebnis klingt so, als würde tatsächlich Kanye West selbstreflektiert über seine verbalen Attacken gegen die jüdische Community rappen.

Im **Journalismus** kann mittels Deepfakes einfach und ohne großen Zeitaufwand Content generiert werden. Es ist möglich, den Stil bestimmter Journalist:innen täuschend echt nachzuahmen. Auch Übersetzungen können vollautomatisch vorgenommen werden. Zudem können illustrierende Bilder und Videos je nach Bedarf frei erschaffen werden, was insbesondere für satirische Darstellungen interessant sein kann. Der Kreativität sind insoweit keine Grenzen gesetzt. Dabei sind derartige Deepfake-Illustrationen immer schwieriger als solche zu erkennen. Selbst bei vermeintlich echt aussehenden Fotos ist es für die Konsument:innen nicht mehr ohne Weiteres nachvollziehbar, ob sie die Wirklichkeit abbilden oder rein KI-generiert sind.

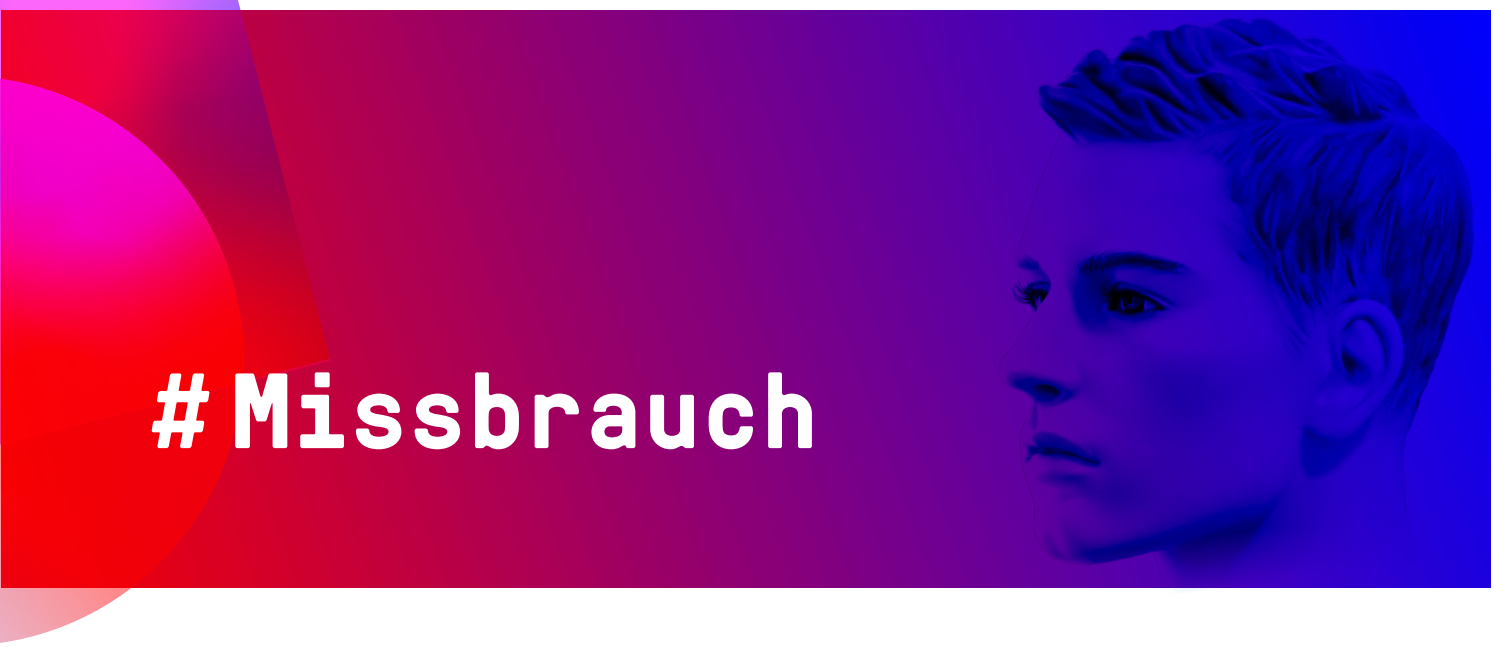
Welche Möglichkeiten gibt es, Deepfakes in der Werbung einzusetzen?

In der Werbung bietet der Einsatz von Deepfakes ebenfalls neue Möglichkeiten. Im Vordergrund steht das Schalten von **hyperpersonalisierter Werbung**. Solche Formen der personalisierten Werbung schaffen eine stärkere Markenbindung und können daher gewinnbringend genutzt werden. Dabei ist zum Beispiel nur die Herstellung eines einzigen Werbeclips erforderlich, mit dem verschiedene **Zielgruppen** angesprochen werden können. Durch Veränderung der Mimik und Lippenbewegungen eines Models sowie der Generierung neuer Audiodateien können individualisierte Botschaften je nach Zielgruppe erstellt werden. So wäre es beispielsweise problemlos möglich, Models mit unterschiedlicher Hautfarbe, Alter und Größe für bestimmte Zielgruppen anzupassen. Weiterhin könnten Werbespots mit der synthetisch erzeugten Stimme einer berühmten Persönlichkeit erstellt werden und das sogar in verschiedenen Sprachen.

Mit derselben Methode können Kampagnen individuell an **Einzelpersonen** gerichtet werden. Den Ideen sind keine Grenzen gesetzt: Werbemodels können potenzielle Käufer:innen mit Namen ansprechen oder der Hintergrund der Werbeaufnahme kann dahingehend verändert werden, dass das Model die Kleidung in einem Umfeld präsentiert, das den Kund:innen bekannt ist, zum Beispiel in deren Heimatstadt. So startete Zalando eine Werbekampagne mit dem Model Cara Delevingne, bei welcher mithilfe

der Deepfake-Technologie mit nur einem Videodreh rund 60.000 individualisierte Videobotschaften für Nutzer:innen in verschiedenen Kleinstädten und Dörfern erzeugt wurden.

Auch das Online-Einkauferlebnis kann durch den Einsatz von Deepfakes persönlicher, realistischer und sogar spielerischer gestaltet werden. Ein anschauliches Beispiel ist die Möglichkeit der virtuellen Anprobe. Dabei müssen potenzielle Käufer:innen Ausgangsmaterial in Form von Videos oder Fotos zur Verfügung stellen. Die KI wandelt diese in Echtzeit zu Deepfakes um und zeigt die User:innen beim Tragen der angebotenen Ware. Zudem können Social Media Influencer:innen künstlich mittels KI generiert werden oder virtuelle Nutzungsanleitungen für Produkte erstellt werden. Die Möglichkeiten, Deepfakes in der Werbung einzusetzen, sind damit wohl lange nicht ausgeschöpft.



#Missbrauch

Wo besteht Missbrauchspotenzial durch Deepfakes?

Die Zeiten, in denen Fotos und Videos als eine meist verlässliche, objektiv „richtige“ Tatsachengrundlage gegolten haben, scheinen durch die Deepfake-Technologie vorbei zu sein. Deepfakes zu erstellen ist ein Kinderspiel. Apps wie Faceswap sind für alle zugänglich und einfach zu bedienen, sodass die Erstellung von Deepfakes längst massentauglich geworden ist. Problematisch wird es dort, wo Deepfakes zur Verbreitung von Falschinformationen eingesetzt werden. Da audiovisuellen Inhalten in der Regel eine sehr hohe Authentizität beigemessen wird, stellen Deepfakes für die öffentliche Meinungsbildung eine große Gefahr dar. Innerhalb weniger Sekunden sind Falschinformationen über die sozialen Medien und das Internet verbreitet. Gleichzeitig wird es immer schwieriger, Deepfakes als solche zu erkennen. Die Nutzung von Deepfakes zur Desinformation oder sogar zu Propagandazwecken ist ein ernst zu nehmendes Problem und gefährdet zunehmend das Vertrauen in die öffentliche Berichterstattung.

Des Weiteren kann die Verwendung von Deepfakes zur Diskreditierung und Verleumdung von Personen missbraucht werden. So kann es zu üblen und nachhaltigen Rufschädigungen von Prominenten und Privatpersonen kommen, beispielsweise durch den Einsatz von Face Swapping in nicht jugendfreien Videos. Aber auch in anderen Bereichen wird der Einsatz von Deepfakes zur Gefahr. Biometrische Systeme zur

Online-Identifikation können mittels Deepfakes überwunden werden und im Bereich des Social Engineering können gezielte Phishing-Angriffe zur Information- und Datengewinnung durchgeführt werden. Das Potenzial der Nutzung der Deepfake-Technologie zu Betrugszwecken ist also immens.

All das schadet dem Image von Deepfakes und kann zu einer Ablehnung innerhalb der relevanten Zielgruppen führen, selbst dann, wenn Unternehmen Deepfakes in redlicher Weise einsetzen.

Welche Gegenmaßnahmen sind geeignet?

Um der Missbrauchsgefahr durch Deepfakes entgegen zu wirken, steht zunächst die **Aufklärung** der Öffentlichkeit im Vordergrund. Diese muss Hinblick auf die Existenz von Deepfakes und die damit verbundenen Risiken sensibilisiert werden und es muss ein Bewusstsein für das bestehende Missbrauchspotenzial geschaffen werden. Nur über Wissen kann die Fähigkeit vermittelt werden, eine differenzierte Einschätzung darüber zu treffen, ob der konsumierte Inhalt in Form von Bildern, Videos, Ton oder Text „echt“ oder ein Deepfake ist. So gibt es einige Merkmale, die Betrachter:innen stutzig machen sollten: Sind bei Gesichtern sichtbare Übergänge oder verwaschene Konturen zu sehen, klingt eine Stimme metallisch, monoton oder ist die Betonung beziehungsweise die Aussprache falsch, so liegt es nahe, dass es sich um ein Deepfake handelt.

Allerdings werden Deepfakes immer besser. Sie zu erkennen wird selbst für technisch Versierte immer schwieriger. Somit sind der Gesetzgeber und die Technik gefragt, um eine einfache und zuverlässige **Identifikation** von Deepfakes zu ermöglichen. Durch kryptographische KI-Verfahren ist es heutzutage möglich, Ausgangsmaterial durch eine digitale Signatur zu schützen und identifizierbar zu machen. Dies ermöglicht eine

sichere Einordnung von vertrauenswürdigen Quellen. Außerdem werden zur Erkennung von Deepfakes neuronale Netzwerke eingesetzt, dieselbe Technologie, die zur Erstellung der Deepfakes verwendet wird. Dabei wird die KI mit Original- und Fake-Material trainiert. Der Algorithmus lernt, die wesentlichen Merkmale der Originale und der Deepfakes zu erkennen, und kann so zwischen Original und Fälschung unterscheiden. Auch der europäische Gesetzgeber ist mittlerweile tätig geworden. So sieht ein Regulierungsentwurf der EU-Kommission zu KI-Systemen vor, dass alle mittels Deepfake-Technologie erstellten Materialien als solche gekennzeichnet werden müssen. Außerdem werden die KI-Verfahren, die zur Erkennung von Deepfakes genutzt werden, stetig verbessert und weiterentwickelt.

#Rechtslage

Wie ist die Rechtslage zu Deepfakes derzeit ausgestaltet?

Derzeit sind Deepfakes in Deutschland keinem eigenen Regelungssystem unterworfen. Somit finden die geltenden zivil-, straf-, und öffentlich-rechtlichen Regelungen Anwendung. Auch auf europäischer Ebene existieren noch keine Gesetze. Jedoch liegen derzeit schon Vorschläge der EU-Kommission für einen Regulierungsentwurf zu KI-Systemen vor. Zudem ist seit November 2022 der Digital Services Act (DSA) in Kraft, mit dem illegale Inhalte schneller entfernt werden können.

Welche Rechte können durch den Einsatz von Deepfakes verletzt werden?

Bei der Verwendung von Deepfakes kommt es vor allem zu Verletzungen des allgemeinen Persönlichkeitsrechts. Weiterhin kann der Einsatz von Deepfakes zu Verletzungen des Urheberrechts, Datenschutzrechts, aber auch des Strafrechts führen.

Kann der Einsatz von Deepfakes strafrechtliche Konsequenzen nach sich ziehen?

Strafrechtlich sind bei der Verwendung von Deepfakes besonders die Delikte gegen die persönliche Ehre relevant – namentlich **Beleidigung, üble Nachrede und Verleumdung** gemäß §§ 185 ff. Strafgesetzbuch (StGB). Wer zum Beispiel eine Darstellung mittels Deepfake erstellt, die eine Person herabwürdigt oder fehlerhaft darstellt, macht sich der Beleidigung strafbar. Wer eine unwahre Tatsache über Betroffene verbreitet, begeht eine Verleumdung. Zu beachten ist, dass im Rahmen des Strafrechts auch die Verbreitung einer Beleidigung strafbar ist. Hiervon ist auch das Teilen von beleidigenden Medieninhalten im Internet umfasst.

Weiterhin ist im Zusammenhang mit Deepfakes eine Strafbarkeit wegen **Betrugs** denkbar. Beispielhaft ist hierfür der sogenannte C-Level-Fraud, bei welchem Mitarbeiter:innen der telefonischen Anweisung von höherrangigen Manager:innen folgen, deren Stimme mittels Deepfake generiert wurde, und daraufhin eine Überweisung auf das Konto der Täter:innen tätigen. Zudem können Deepfakes zur **Erpressung** oder **Nötigung** verwendet werden.

Im Zusammenhang mit Wahlen gibt es noch keine Regelungen, die mittels Deepfake erzeugte Falschinformationen strafrechtlich ahnden. In Teilen der USA finden sich dagegen schon entsprechende Regelungen. Demnach hat das Verbreiten von Deepfakes durch Politiker:innen bis zu 60 Tage vor einer Wahl strafrechtliche Konsequenzen zur Folge.

Allgemeines Persönlichkeitsrecht

Wie können Persönlichkeitsrechte durch den Einsatz von Deepfakes verletzt werden?

Beim Erstellen und Verwenden von Deepfakes wird regelmäßig das allgemeine Persönlichkeitsrecht (Art. 2 Abs. 1 GG i.V.m. Art. 1 Abs. 1 GG) verletzt, was allenfalls bei erkennbar satirischen Darstellungen gerechtfertigt sein kann. Dabei handelt es sich in der Regel um Verletzungen des **Rechts am eigenen Bild** und des **Rechts am gesprochenen Wort**. Das allgemeine Persönlichkeitsrecht schützt nämlich nach der Rechtsprechung des Bundesverfassungsgerichts „vor der Verbreitung eines technisch manipulierten Bildes, das den Anschein erweckt, ein authentisches Abbild einer Person zu sein“. Das Persönlichkeitsrecht ist immer dann verletzt, wenn **keine Einwilligung** der abgebildeten Person in die Verbreitung des eigenen Bildes vorliegt. Die Verletzung kann auch nicht über die Meinungsfreiheit (Art. 5 GG) gerechtfertigt werden, da es sich bei einem Deepfake um eine täuschend echte Fälschung, mithin um eine unwahre Information handelt. Unrichtige Informationen sind von der Meinungsfreiheit nicht geschützt. Etwas anderes gilt allenfalls bei erkennbar satirischen Darstellungen. Nur in diesem Fall kann die Verletzung des Persönlichkeitsrechts über die Meinungs-, Presse- oder Kunstfreiheit nach Art. 5 GG gerechtfertigt werden.

Welche möglichen Konsequenzen drohen im Falle einer Verletzung der Persönlichkeitsrechte?

Wer in seinem allgemeinen Persönlichkeitsrecht verletzt ist, dem steht in aller Regel ein Unterlassungs- und Beseitigungsanspruch zu. Bei einer schwerwiegenden Persönlichkeitsverletzung ist auch eine Geldentschädigung denkbar. Hinsichtlich der Verbreitung von Bildnissen ohne die Einwilligung der abgebildeten Person sehen §§ 22, 33 des Gesetzes betreffend das Urheberrecht an Werken der bildenden Künste und der Photographie (KUG) sogar eine Freiheitsstrafe von bis zu einem Jahr oder eine Geldstrafe vor.

Welche Rechte können im Bereich des Urheberrechts durch den Einsatz von Deepfakes verletzt werden?

Durch den Einsatz von Deepfakes können Urheberrechte vor allem dann verletzt werden, wenn bei der Erstellung des Deepfakes auf urheberrechtlich geschütztes Material zurückgegriffen und dieses verändert wird. Denn dies ist nur innerhalb enger Schranken erlaubt und stellt sonst eine Bearbeitung des urheberrechtlich geschützten Werks dar, zu welcher der Urheber nach dem Urheberrechtsgesetz seine Einwilligung geben muss.

Das Bildnis oder die Stimme einer Person sind jedoch nicht urheberrechtlich geschützt. Anspruchsberechtigt ist also nicht die abgebildete Person, sondern der Urheber, bzw. derjenige, der am Werk die Leistungsschutzrechte hat. Wird zum Beispiel eine bekannte Filmszene verändert, um Werbebotschaften zu überbringen, können die Inhaber:innen von Schutzrechten an dieser Filmszene, also zum Beispiel Filmproduzent:innen, Ansprüche geltend machen, sofern sie nicht in die Bearbeitung eingewilligt haben.

Der Eingriff in das Urheberrecht durch die Bearbeitung kann allerdings gerechtfertigt sein, wenn es sich bei der Bearbeitung, also dem mittels Deepfake erstellten Bild oder Video, um eine Parodie, Karikatur oder ein Pastiche handelt. Um diese Ausnahmen in Anspruch zu nehmen, ist es aber erforderlich, dass eine inhaltliche oder künstlerische Auseinandersetzung mit dem Werk stattgefunden hat. Ob die Voraussetzungen für eine der gesetzlichen Ausnahmen erfüllt sind, muss im Einzelfall geprüft werden. Gerade in einem werblichen Kontext steht jedoch immer auch ein kommerzieller Gedanke hinter einer Bearbeitung, was häufig gegen die Annahme einer der genannten Ausnahmen spricht.

Welche möglichen Konsequenzen drohen im Falle einer Verletzung des Urheberrechts?

Bei einem Verstoß gegen das Urheberrecht haben Verletzte (= Urheber oder Leistungsschutzberechtigte) einen Anspruch auf Unterlassung, Beseitigung und Schadensersatz, sowie – falls das Werk bereits vervielfältigt wurde – einen Anspruch auf Vernichtung und Rückruf des rechtswidrig hergestellten oder verbreiteten Werks. Gerade dies kann bei über das Internet geteilten Inhalten zur Herausforderung werden.

#Datenschutz

Welche Rechte können im Bereich des Datenschutzes durch den Einsatz von Deepfakes verletzt werden?

Die Herstellung und Verbreitung von Deepfakes natürlicher Personen stellt regelmäßig eine Verarbeitung personenbezogener Daten dar. Personenbezogene Daten sind solche, die sich auf eine natürliche Person beziehen und diese identifizierbar machen, zum Beispiel der Name oder die physischen und genetischen Merkmale einer Person. Im Fall von Deepfakes sind vor allem äußerliche Merkmale von Personen wie deren Gesichtszüge, Körpersprache und Stimme von der Datenverarbeitung betroffen.

Die Verarbeitung personenbezogener Daten **ohne Einwilligung** der betroffenen Person oder andere Rechtsgrundlage verstößt gegen das Datenschutzrecht (Art. 6 Datenschutz-Grundverordnung - DSGVO). Als Verarbeitung gilt insbesondere die im Zusammenhang mit den jeweiligen Deepfakes stehende Speicherung, Veränderung und Offenlegung von Daten. Die Verarbeitung dieser Daten ist nur zulässig, sofern die Einwilligung der betroffenen Person vorliegt, oder eine andere Rechtsgrundlage für die Verarbeitung einschlägig ist. Zu beachten ist jedoch, dass die Einwilligung immer nur für die konkrete Verarbeitung gilt, für die sie erteilt wurde. Eine Videoaufnahme, mit der sich die betroffene Person ursprünglich einverstanden erklärt hat, die aber anschließend ohne ihr Wissen für die Herstellung eines Deepfakes verwendet

wird, ist damit unzulässig. Auf sogenannte Haushaltsaufnahmen, also Verarbeitung personenbezogener Daten durch natürliche Personen zur Ausübung ausschließlich persönlicher oder familiärer Tätigkeiten, finden die Bestimmungen der DSGVO keine Anwendung. Ihre Grenze findet die Privilegierung zu Haushaltsaufnahmen aber dort, wo die Daten beziehungsweise Deepfakes einem unbegrenzten Personenkreis zur Verfügung gestellt werden, etwa in sozialen Netzwerken.

Welche möglichen Konsequenzen drohen im Falle einer Rechtsverletzung im Bereich des Datenschutzes?

Im Falle der unrechtmäßigen Verarbeitung personenbezogener Daten sieht die seit 2018 geltende DSGVO einen Löschungsanspruch (Art. 17 DSGVO) sowie einen Schadensersatzanspruch (Art. 82 DSGVO) der Betroffenen vor, wobei im Falle von Deepfakes vor allem der Ersatz für immaterielle Schäden interessant ist. Nach jüngster EuGH Rechtsprechung braucht man für einen Schadenersatz jedoch den Nachweis eines konkreten Schadens. Das heißt, ein bloßes Störgefühl oder Unwohlsein, weil die personenbezogenen Daten verarbeitet wurden, reicht für einen Schadenersatz noch nicht aus. Verantwortlich im Sinne des Datenschutzrechts ist, wer die Deepfakes **herstellt oder verbreitet**.

#Anspruchs-durchsetzung

Wie können die Ansprüche durchgesetzt werden?

Das deutsche Recht bietet vielfältige Vorschriften, die zum Schutz gegen widerrechtlich erstellte Deepfakes dienen können. Jedoch ist der effektive Rechtsschutz in der Praxis schwierig. Problematisch erweist sich vor allem die prozessuale Durchsetzung der Ansprüche, da Plattformbetreiber:innen in der Regel nicht ohne Weiteres in die Pflicht genommen werden können. Auch die Identität der Ersteller:innen von Deepfakes ist nur schwer zu ermitteln. Angesichts der hohen Verbreitungsgeschwindigkeit von Inhalten im Internet ist es daher für eine effektive Anspruchsdurchsetzung meist schon zu spät, sobald ein Deepfake erst einmal online ist.

#Ausblick

Wie entwickelt sich die Rechtslage aktuell?

Durch die schnelle Weiterentwicklung der Deepfake-Technologie wird die Qualität der Deepfakes immer besser, was zur Folge hat, dass auch die Identifizierung eines Deepfakes immer schwieriger wird. Die KI-Programme benötigen zunehmend weniger Ausgangsmaterial für die Erstellung von Deepfakes mit guter Qualität. Dadurch wird das Erstellen von qualitativen Deepfakes immer einfacher, auch für Laien. Die Weiterentwicklung von Gegenmaßnahmen ist daher unerlässlich. Das Bundesinnenministerium hat sich bereits für einen „klaren Rechtsrahmen“ beim Einsatz von KI ausgesprochen. Die EU-Kommission schlug im April 2021 den Entwurf einer Verordnung für Künstliche Intelligenz (**Artificial Intelligence Act**) zur Schaffung eines legalen Rahmens für die Entwicklung, den Einsatz und die Nutzung von KI-Systemen in der EU vor. Über den Entwurf hat das Parlament im Juni 2023 bereits final abgestimmt, sodass durchaus möglich ist, dass die Verordnung noch 2023 in Kraft tritt.

Der Entwurf sieht insbesondere die Klassifizierung von KI-Systemen vor. Diese werden nach ihrem Risiko für Betroffene eingeteilt. Je risikoreicher das System ist, desto höher sind die zu beachtenden Pflichten, bis hin zum strikten Verbot. Die Verordnung sieht zudem eine **Transparenzpflicht** vor. Danach müssen Nutzer:innen eines KI-Systems offenlegen, wenn Inhalte wie Deepfakes künstlich erzeugt oder manipuliert wurden.

Bei einem Verstoß sieht der Entwurf eine Haftung von bis zu 20 Millionen Euro oder – im Falle von Unternehmen – von bis zu 4 Prozent des gesamten weltweiten Jahresumsatzes des vorangegangenen Geschäftsjahres vor, je nachdem, welcher Betrag höher ist. Eine effektive Umsetzung ist jedoch nur gewährleistet, wenn Deepfakes überhaupt erkannt werden können. Daher müssen die Methoden zur technischen Erkennbarkeit von Deepfakes stetig verbessert werden.

Koelnmesse GmbH

Messeplatz 1

50679 Cologne

T. +49 1806 145 514*

info@dmexco.com

Bilder: DKunert/pixabay.com, jessica45/pixabay.com

Autoren: Cynthia Smponias, Vivian Pérez, Margret Knitter, Helena Kasper, Marguerita von Bechtolsheim

*(0,20 EUR/Anruf aus dem dt. Festnetz; max. 0,60 EUR/Anruf aus dem Mobilfunknetz)

Bleib auf dem Laufenden!

Weitere spannende Stories, Whitepaper, Podcasts & Co. findest du in unserem Newsbereich!

Zu den News