



Deepfakes

Deepfakes Definition¹

Der **Begriff Deepfake**² umfasst KI-generierte oder manipulierte Audio-, Bild- oder Videodateien, die vorgeben real zu sein. Es handelt sich um synthetische Medien, die durch Manipulation authentisch wirken.

Deepnudes sind eine Sonderform von Deepfakes, bei denen auf Fotos abgebildete Personen durch KI entkleidet werden. Bei **Deepfake-Pornos**³ werden die Gesichter von Betroffenen in pornografische Medien eingefügt.

Deepfakes können zu **Unterhaltungszwecken** oder als **Satire** verwendet werden. In der Regel sind sie dann von der **Meinungsfreiheit** gedeckt. Sofern eine **täuschende Darstellung von realen Personen**, Rechtsgüter wie die Ehre, den Ruf oder die Intimsphäre verletzt, können aber **unterschiedliche Normen verletzt** sein.

Deepfakes sind als eine neue Form der **Desinformation** zu verstehen. Die strafrechtlich relevante Verwendung von Deepfakes ist als **Cybercrime im weiteren Sinne** zu verstehen.

Deepfakes: Formen audiovisueller (medialer) Manipulation mittels Videos, Bilder und/oder Audioaufnahme. Erkennen der Manipulation ist mit bloßem Auge zumeist nicht möglich.

Rechtliche Beurteilung

AI-Act der Europäischen Union: Durch den AI-Act wurde für Anbieter*innen und Entwickler*innen von KI-Anwendungen eine **Kennzeichnungspflicht** (z.B. durch Wasserzeichen, Hinweise) festgelegt. Diese Pflicht gilt für alle, die KI-Tools beruflich verwenden. Ziel ist es Desinformation zu verhindern oder einzudämmen. Im Zusammenhang mit Verstößen gegen diese Kennzeichnungspflicht sind Bußgelder vorgesehen.

Eine Ausnahme besteht für die Nutzung für persönlich Zwecke durch Privatpersonen und ebenso für Werke, die offensichtlich künstlerischen, kreativen oder satirischen Zwecken dienen. Für den Fall, dass die Manipulation aber für eine*n durchschnittliche*n Betrachter*in nicht offensichtlich ist oder Satire genutzt wird, um Desinformation zu verbreiten, gilt die Ausnahme nicht.

Im Juni 2026 wurde die **Regierung** zur Schaffung (**verschärfter**) **strafrechtlicher Konsequenzen gegen Deepfakes** aufgefordert. Diese Novellierung gilt es abzuwarten. Österreich muss mit 2027 aber die **Richtlinie zur Bekämpfung von Gewalt gegen Frauen** umsetzen, in welcher auch die Verpflichtung enthalten ist, die nicht einvernehmliche Herstellung und Weitergabe sexualisierter Deepfakes unter Strafe zu stellen.

¹ Saferinternet.at, Factsheet Deepfakes und sexualisierte Gewalt:

[Factsheet_Deepfakes_und_sexualisierte_Gewalt.pdf](#), siehe u.a. zu weiteren Formen manipulierter Inhalte.

² Begriff setzt sich zusammen aus den Worten „Deep Learning“ und „Fake“.

³ Zum Fall von Collien Ulmen-Fernandes, siehe bspw. hier: <https://www.spiegel.de/thema/collien-fernandes/>

Derzeit gibt es **keinen** eigenen **Straftatbestand**, durch Deepfakes können unterschiedliche Rechtsnormen verletzt sein, bspw.:

- § 107 StGB (gefährliche Drohung)
- § 107c StGB (Cybermobbing)
- § 111 StGB (Üble Nachrede)
- § 115 StGB (Beleidigung)
- § 144 StGB (Erpressung)
- § 207a StGB (Sexuelle Missbrauchsdarstellungen)
- § 208 a StGB (Anbahnung von Sexualkontakten zu Unmündigen)
- § 78 Urhebergesetz (Recht am eigenen Bild)
- § 1330 ABGB (Ehrbeleidigung)

Gefahren und Auswirkungen für Opfer und die Gesellschaft

Weder für Lai*innen noch (mittlerweile leider auch häufig) für Expert*innen sind KI-generierte Inhalte als solche erkennbar.

Für Betroffene kann die Verbreitung von KI-generierten Inhalten verheerende berufliche wie private Konsequenzen nach sich ziehen. Bei realitätsnahen Täuschungen wird Betroffenen die Chance genommen, sich glaubwürdig zu verteidigen.

Auch Unternehmen können durch KI-generiertes Material Opfer von bspw. Erpressungen oder Identitätsdiebstählen sein.

Bedeutung im Zusammenhang mit Hate Crime

Deepfakes können auch als Tatmittel im Zusammenhang mit vorurteilsmotivierten Straftaten dienen. Bspw. können sexualisierte, abwertende Deepfakes gegen Frauen, rassistische/antisemitische Deepfakes, manipulierte Darstellungen von Angehörigen bestimmter Gruppen zur besonderen vorurteilsgeleiteten Abwertung führen. Weltanschaulich problematisch können KI-generierte Inhalte sein die bspw. von islamistischen und terroristischen Organisationen zur Produktion von Propaganda genutzt werden.⁴

Vorgaben für die Polizei

Appell an den Hausverstand: Inhalte und deren Plausibilität sind zu hinterfragen. Sofern auch technische Möglichkeiten bestehen, können z.B. Meta-Daten ausgelesen werden, um Veränderungen von Inhalten festzustellen oder Hinweise auf Urheber*innen zu finden. Die Nutzung von Plattformen wie www.mimikama.at wird empfohlen.⁵

Im PAD wurde mit 1. Jänner 2024 die Begehungsform „Deepfake“ eingeführt.⁶ Eine diesbezügliche Erfassung ist notwendig, um ein umfassendes Lagebild zu diesem Phänomen zu erhalten.

⁴ Vgl. dazu BMI, Verfassungsschutzbericht 2025, 115.

⁵ Siehe auch: <https://www.mimikama.org/ki-und-deepfakes/deepfakes-erkennen/>.

⁶ BMI-Erlass vom 29.12.2023, GZ: 2023-0.924.244.

Judikatur OLG Wien, 2 R 178/25y

29. April 2026;

Verstoß gegen § 78 UrhG (KI-generierte Inhalte, Deepfakes, Bildnisschutz, Persönlichkeitsrecht)

Sachverhalt:

Der Kläger ist ein in Österreich bekannter Publizist, der sich gegen die beklagte Partei, einen Medienakteur wegen deren Verbreitung von KI-generierten Bildnissen zur Wehr setzte. Die Bildnisse illustrierten kritische Beiträge zum Vorgehen des Klägers (gegen Hass im Netz) und stellten diesen u.a. mit verzerrtem, hämischen oder manischem Lachen dar. Der Kläger hatte der Veröffentlichung der Bildnisse, die er als Hasspostings bezeichnete, nicht zugestimmt, weswegen er die Unterlassung, Urteilsveröffentlichung und Zahlung begehrte. Die beklagte Partei bekämpfte die Entscheidungen des Erstgerichts mit Berufung auf die Meinungsäußerungsfreiheit, Pressefreiheit, Informationsfreiheit und Kunstfreiheit.

Kernaussagen des Urteils des OLG Wien⁷:

- Bei den Veröffentlichungen handelt es sich **nicht um Satire**, welche **durch charakterisierende Übertreibung und Verzerrung der Realität** darauf abzielt zu **provozieren** und zu **bewegen**, da für den **durchschnittlichen Betrachter nicht erkennbar** ist, dass es sich bei den Veröffentlichungen **nicht um authentische Darstellungen** des Klägers handle. Durch den durchschnittlichen Betrachter ist nicht erkennbar, dass die Gesichtszüge des Klägers künstlich erzeugt sind und bearbeitet wurden.
- Daher handelt es sich bei den Bildnissen um **Deepfakes**, bei denen ein **durchschnittlicher Betrachter davon ausgeht**, dass es sich um **reale Bildnisse** handelt und bei welchen der Anschein einer Echtheit vermittelt werden soll.
- Eine **Kennzeichnung** der Veröffentlichungen als KI-generierte Inhalte **fehlte**. Mit dem im Sommer 2024 in Kraft getretenen EU Artificial Intelligence Act (VO (EU) 2024/1689; AI-Act) wurde festgelegt, dass **KI-generierte oder manipulierte Inhalt**, die real wirken („Deepfakes“) **gekennzeichnet** werden müssen. Eine **Kennzeichnungspflicht**, zur Vermeidung einer Täuschung oder Manipulation durch KI-generierte Bildbearbeitungen, besteht seit August 2026.
- Die **Grenze der Meinungsfreiheit** besteht dann, wenn diese auf der **Unwahrheit** basiert bzw. **kein sachlicher Anknüpfungspunkt** vorhanden ist. **Deepfakes**, die eine **unwahre Darstellung der Realität** sind, sind daher **unter keinen Umständen gerechtfertigt**. Eine Interessenabwägung ist entbehrlich, würde in Ermangelung eines künstlerischen oder informativen Mehrwertes der Veröffentlichungen aber zu Lasten der beklagten Partei ausfallen.
- Im Generieren und Verbreiten der Bildnisse ist ein **Verstoß gem. § 78 Urhebergesetz (UrhG)** begründet.

FAZIT: Verstoß gegen § 78 UrhG.

Das Urteil des OLG Wien finden Sie unter dem folgenden [LINK](#).

⁷ Anm.: Der Weg zum OGH steht noch offen.

STOP HATE CRIME

Allgemeine Informationen



Rückblick:

- Mit Inkrafttreten der Prozessbegleitungs-Regulierungsverordnung wurde die Grundausbildung für **psychosoziale Prozessbegleiter*innen** neu geregelt. Die Abt. III/S/1 hat für die Zusatzausbildung „Modul für volljährige Opfer von Hass-im-Netz-Delikten“ die **Lehrgangsleitung** übernommen. Ein erster Termin fand im **März 2026** statt.
- Der **Leitfaden** zur Unterstützung nationaler Behörden bei der **Planung, Durchführung und Evaluierung öffentlicher Kampagnen gegen Hass und Hetze** wurde auf **Deutsch** übersetzt.
- Der **Lagebericht Hate Crime 2025** wurde im **Juli 2026** veröffentlicht.



Zukunft

- Erarbeitung der **englischen Version** zum **Lagebericht Hate Crime 2025**.
- Die **Abt. III/S/1** begann in enger Abstimmung mit der **Justiz** die inhaltliche Konzeption des im Regierungsprogramm **beschlossenen Nationalen Aktionsplans zur Bekämpfung von Hate Crime**.



- ✓ **Lagebericht 2025**
- ✓ **Lehrgang Hass im Netz für psychosoziale Prozessbegleiter*innen**
- ✓ **Neuer Leitfaden für Kampagnen auf Deutsch**