

LEITFADEN ZU DEN KI-KENNZEICHNUNGSPFLICHTEN DER EU

Dritte, aktualisierte Fassung

Stand: 20.09.2026, Autoren: Dagmar Schwelle und Marco Urban

(Übersetzungen aus den englischen Originaltexten mit Unterstützung von Claude AI)

Am 2. August 2026 trat aufgrund von europarechtlichen Vorschriften eine Pflicht zur Kennzeichnung von sogenannten Deepfakes in Kraft. Da die rechtliche Definition des Begriffs Deepfake viel weiter gefasst ist als in der Alltagssprache, kann die Kennzeichnungspflicht uns alle betreffen. Denn in der Pflicht zu kennzeichnen steht der- oder diejenige, der oder die beruflich mit einem KI-System Bilder erstellt oder bearbeitet.

Noch sind viele Detailfragen offen, wahrscheinlich wird erst die Rechtsprechung Klarheit schaffen, wie die Regelungen auszulegen sind. Wir werden diesen Leitfaden daher im Laufe der Zeit immer wieder aktualisieren, ergänzen und wo nötig korrigieren, können aber keine Haftung für die Vollständigkeit und Richtigkeit übernehmen.

Wo ist die Kennzeichnungspflicht geregelt?

Die Kennzeichnungspflicht für Deepfakes steht in Art. 50 (4) AI Act, also der KI-Verordnung der Europäischen Union (Verordnung (EU) 2024/1689). Als Verordnung gilt sie unmittelbar in allen Mitgliedstaaten und muss nicht erst in deutsches Recht umgesetzt werden. Sie legt einheitliche Regeln für das Inverkehrbringen und den Einsatz von KI fest.¹

Wie die Kennzeichnungspflicht praktisch umzusetzen ist, ergibt sich aus zwei begleitenden Texten.

- Der »Code of Practice on Transparency of AI-Generated Content« (CoP)² ist ein freiwilliges Rahmenwerk, das vor allem regelt, wie der Kennzeichnungspflicht nachzukommen ist, also wie das Label auszusehen hat und wo es platziert werden muss. Der CoP wurde in einem Stakeholder-Prozess erarbeitet, an dem auch FREELENS via AG Fortschritt im Deutschen Fotorat teilgenommen hat.
- Die »Guidelines on the implementation of the transparency obligations for certain AI systems under Article 50 of Regulation (EU) 2024/1689 (the »AI Act«)« sind Leitlinien der EU-Kommission. Sie legen den AI Act aus und erläutern unbestimmte Begriffe, insbesondere auch die Definition von Deepfake. Die finale Fassung liegt seit dem 20. Juli 2026 vor.³

Wichtig ist: Weder der CoP noch die Guidelines sind rechtlich bindend, nur der Europäische Gerichtshof kann den AI Act verbindlich auslegen. Dennoch sind sie im Moment der zentrale Anhaltspunkt, wie die Kennzeichnungspflicht umzusetzen ist.

¹ Abgesehen von der Pflicht zur für Menschen sichtbaren Kennzeichnung von Deepfakes gibt es auch eine Pflicht für Anbieter von KI-Systemen, generierte Audio-, Bild-, Video- oder Textinhalte maschinenlesbar zu kennzeichnen. Wir beschränken uns hier aber im Wesentlichen auf die Kennzeichnungspflicht für Deepfakes, da nur diese uns direkt betrifft.

² <https://ec.europa.eu/newsroom/dae/redirection/document/129555>

³ <https://ec.europa.eu/newsroom/dae/redirection/document/131215>

Was steht in Art 50 (4) AI Act?

Zunächst zum Gesetzestext: »Betreiber eines KI-Systems, das Bild-, Ton- oder Videoinhalte erzeugt oder manipuliert, die ein Deepfake sind, müssen offenlegen, dass die Inhalte künstlich erzeugt oder manipuliert wurden. (...) Ist der Inhalt Teil eines offensichtlich künstlerischen, kreativen, satirischen, fiktionalen oder analogen Werks oder Programms, so beschränken sich die in diesem Absatz festgelegten Transparenzpflichten darauf, das Vorhandensein solcher erzeugten oder manipulierten Inhalte in geeigneter Weise offenzulegen, die die Darstellung oder den Genuss des Werks nicht beeinträchtigt.«

Was ist ein kennzeichnungspflichtiger Deepfake?

Die Definition des AI Act und vor allem der erste Entwurf für die Guidelines der Kommission ließen erwarten, dass künftig sehr viele in der kommerziellen Fotografie übliche Bearbeitungsformen gekennzeichnet werden müssen. Die finalen Guidelines definieren den Begriff Deepfake nun etwas enger, der Spielraum für nicht kennzeichnungspflichtige (dezent) KI-Bearbeitungen wird also größer.

Laut AI Act handelt es sich bei einem Deepfake um »einen durch KI erzeugten oder manipulierten Bild-, Ton- oder Videoinhalt, der wirklichen Personen, Gegenständen, Orten, Einrichtungen oder Ereignissen ähnelt und einer Person fälschlicherweise als echt oder wahrheitsgemäß erscheinen würde« (Art. 3 Nr. 60 AI Act).

Die Guidelines konkretisieren diese Definition des Deepfakes anhand von vier Kriterien, die gleichzeitig erfüllt sein müssen:

1. Ähnlichkeit

Der KI-generierte oder manipulierte Inhalt muss einem realen Vorbild erkennbar ähneln. Erforderlich ist ein hoher Grad an wahrnehmbarer Ähnlichkeit, der Inhalt muss aber nicht mit dem Vorbild identisch sein. Ob eine ausreichende Ähnlichkeit vorliegt, ist im Einzelfall vom Betreiber (also von uns) zu beurteilen.

2. Bezug zu realistischen Motiven

Das dargestellte Motiv muss einer Person, einem Gegenstand, einem Ort, einer Einrichtung oder einem Ereignis ähneln, das tatsächlich existiert, plausibel existieren könnte oder in der Realität plausibel hätte existieren können. Damit können auch fiktive, aber realistisch wirkende Darstellungen erfasst sein.

Nicht erfasst sind dagegen Darstellungen, die offensichtlich unrealistisch sind und kein nennenswertes Täuschungspotenzial besitzen, etwa Menschen, die ohne technische Hilfsmittel fliegen, oder fantastische Wesen wie Drachen.

3. Erfasste Motive

- Personen: einschließlich digitaler Nachbildungen realer Personen, realistischer KI-Avatare und menschlicher Merkmale wie Aussehen, Stimme oder Verhalten.
- Gegenstände: realistische nicht-belebte Dinge wie Gebäude, Kunstwerke, Maschinen oder Konsumgüter.
- Orte: realistische Schauplätze.
- Wesen (im englischen Original ‚entities‘, im deutschen Verordnungstext irreführend mit »Einrichtungen« übersetzt): nicht-menschliche, aber belebte Wesen wie Tiere und andere Lebensformen.
- Ereignisse: realistische Szenen oder Situationen, beispielsweise historische Ereignisse oder Darstellungen beruflicher oder kommerzieller Dienstleistungen.

4. Falscher Eindruck von Echtheit oder Wahrhaftigkeit

Der Inhalt muss geeignet sein, bei einer Person den falschen Eindruck zu erwecken, er sei authentisch oder gebe die Wirklichkeit wahrheitsgemäß wieder. Dieses Kriterium betrifft das eigentliche Täuschungspotenzial des Deepfakes. Authentizität/Echtheit betrifft die Frage, ob der Inhalt hinsichtlich seiner Quelle und seines Entstehungsprozesses tatsächlich das ist, was er zu sein vorgibt. Dazu gehört etwa, ob wirklich reale Menschen oder Tiere beteiligt waren und sich so verhalten haben wie dargestellt oder auch ob es sich tatsächlich um ein Foto handelt. Wahrhaftigkeit betrifft die inhaltliche Richtigkeit des Dargestellten und seine Übereinstimmung mit der Realität.

Eine Täuschungsabsicht der Bildersteller*innen ist für das Zustandekommen eines Deepfakes nicht erforderlich. Entscheidend ist vielmehr, ob der Inhalt objektiv geeignet ist, eine Person über seine Echtheit oder Wahrhaftigkeit in die Irre zu führen.

Maßstab ist nicht ein durchschnittlicher Betrachter, sondern das im konkreten Fall »vernünftigerweise vorhersehbare Publikum«. Besondere Sorgfalt ist laut den Guidelines dann zu üben, wenn es »vernünftigerweise vorhersehbar ist, dass Deepfake-Inhalte irgendwann von Kindern, älteren Menschen oder anderen Personengruppen mit geringerer Digital- und KI-Kompetenz oder geringerem Allgemeinwissen wahrgenommen werden könnten, da diese hinsichtlich der Echtheit oder Wahrhaftigkeit der Inhalte leichter getäuscht oder in die Irre geführt werden können«.

Immerhin: Eine Weiterverbreitung durch Dritte über das vernünftigerweise vorhersehbare Publikum hinaus müssen wir nicht einkalkulieren. Wird ein Bild etwa nur in einem Abo-Bereich oder einem Unternehmens-Newsletter eingesetzt, ist nicht automatisch von einer breiten Öffentlichkeit auszugehen.

Interessant ist eine in die finalen Guidelines neu eingefügte Bemerkung: Wenn das »beabsichtigte und vernünftigerweise vorhersehbare« Publikum im konkreten Kontext gar nicht erwartet, dass der Inhalt authentisch ist, dann kann dieser definitionsgemäß gar kein Deepfake sein, selbst wenn er KI-generiert oder KI-manipuliert ist. Dies scheint die

Kennzeichnungspflicht im Bereich der Werbung zwar etwas aufzuweichen. Zwei ebenfalls in den Guidelines genannte Beispiele für Deepfakes oder Nicht-Deepfakes setzen dem aber gewisse Grenzen:

Deepfake ist demnach: »Ein KI-generiertes Bild eines Produkts in Werbung oder auf Verpackungen, das die Wahrnehmung des Publikums beeinflussen und über das tatsächliche Aussehen, die Eigenschaften oder die Verwendung des Produkts irreführen kann (z. B. indem das Produkt nicht identisch mit dem realen Produkt, ansprechender oder in besserer Qualität als in Wirklichkeit erscheint).«

Kein Deepfake ist hingegen: »Ein reales Produkt (z.B. ein Auto), das in einer Werbung vor einem KI-generierten Hintergrund und Umfeld gezeigt wird, sofern die Werbung nicht geeignet ist, das Publikum über die tatsächliche Darstellung des Produkts sowie seine Eigenschaften und Verwendung irrezuführen.«

Besonders wichtig ist für uns Fotografinnen und Fotografen natürlich, welche Formen der KI-gestützten Bildbearbeitung dazu führen könnten, dass aus einem Foto ein Deepfake wird. Einen wichtigen Anhaltspunkt, was möglich ist und was nicht, kann Artikel 50 (2) AI Act liefern, der eigentlich nicht direkt uns, sondern die Anbieter von KI-Systemen betrifft. Diese müssen den Output ihrer Systeme maschinenlesbar kennzeichnen, außer wenn die Systeme »eine unterstützende Funktion für die Standardbearbeitung ausführen oder die vom Betreiber bereitgestellten Eingabedaten oder deren Semantik nicht wesentlich verändern«. Daraus kann man den Schluss ziehen, dass diese Modifikationen in der Folge auch keine Deepfakes erzeugen. Laut den Guidelines handelt es sich dabei z.B. um:

- Rauschreduzierung oder -entfernung zur Verbesserung der Klarheit, ohne die Bedeutung oder die Substanz des Inhalts zu verändern;
- geringfügige Farbanpassungen oder -korrekturen, Aufhellen oder Abdunkeln, Schärfen zur Verbesserung der Klarheit oder sonstige Standardbearbeitungen und technische Korrekturen;
- Entfernung von Staubflecken, die durch ein verschmutztes Objektiv oder einen verschmutzten Sensor verursacht wurden; Entfernung des Rote-Augen-Effekts; Löschen und Unkenntlichmachen von Hintergründen, die in der Originaldatei sichtbar sind; Verpixelung oder Unscharfmachen von Gesichtern;
- Standardanpassungen von Farbe und Kontrast, geringfügige Anpassungen der Wiedergabegeschwindigkeit, geringfügige Korrekturen zur Begradigung des Horizonts eines Bildes, Anwendung von Pixelfiltern zur Hervorhebung bestimmter Bildbereiche oder Anwendung von Farbpaletten auf Graustufenbilder, Ergänzung von Bildrändern (edge image completion), Umwandlung eines Schwarz-Weiß-Bildes oder -Videos in ein Farbbild bzw. -video und umgekehrt;
- Pixel Filling zur Anpassung des Seitenverhältnisses oder zur Verbesserung der Klarheit, ..., sonstige nicht-substanzielle Bearbeitungen zu kosmetischen und technischen Zwecken.

»Semantische Änderungen« hingegen müssen von den Anbietern der KI-Systeme maschinenlesbar gekennzeichnet werden. Das bedeutet nicht automatisch, dass damit auch

Deepfakes entstehen, da dafür alle oben genannten Kriterien erfüllt sein müssen. Dennoch lässt sich an diesen Beispielen ablesen, welche Bearbeitungen mit KI zu Deepfakes führen könnten:

- Entfernen, Ersetzen oder Einfügen von Objekten oder Personen in bestehenden Bildern und Videos, wodurch sich die Bedeutung und Substanz des Inhalts verändern; Gesichtsaustausch oder substanzielle Veränderung des Gesichts;
- Erstellung realistischer Videos, die Ereignisse darstellen, die nicht stattgefunden haben; Veränderung der Körperform oder der Hautfarbe einer Person; extreme Aufhellung, Abdunkelung sowie Farb- und Kontrastveränderungen, die die Bedeutung, die Absicht und die Aussage des Inhalts verändern;
- Erstellung von zusammengesetzten Bildern oder zusammengesetzten Videoclips, die die Darstellung von Personen, Objekten, Ereignissen oder Fakten verändert; jede sonstige substanzielle Veränderung des Inhalts.

Auffällig ist hier, dass gemäß den finalen Guidelines Entfernen, Hinzufügen oder Verändern von Elementen nicht schon per se Kennzeichnungspflicht auslösen, sondern erst dann, wenn sie eine substanzielle Veränderung des Inhalts bewirken. Das war im Entwurf noch deutlich strenger geregelt.

Diese Tendenz setzt sich auch im Kapitel betreffend Deepfakes fort. In Absatz 116 der Guidelines heißt es, dass je nach Kontext und Publikum eine »KI-gestützte Manipulation unbedeutender inhaltlicher oder technischer Aspekte bereits bestehender Inhalte für die Beurteilung der Authentizität oder Wahrhaftigkeit des Inhalts durch eine Person von geringer Relevanz sein kann«, womit kein Deepfake entsteht. Aufgezählt werden hier: die Bearbeitung von Hintergrunddetails (z. B. das Entfernen von Passanten), Anpassungen der Beleuchtung, Farbkorrektur, Rauschreduzierung oder -entfernung, kosmetische Anpassungen und Optimierungen, Hintergrunderweiterungen, Anpassungen oder Ersetzungen von Hintergründen zu eindeutig ästhetischen Zwecken, Kompositionen und Arrangements bestehender Produkte. Als zulässiger Kontext werden hier beispielhaft Produktwerbung oder Verpackungen genannt.

Weiter heißt es: »Umgekehrt kann eine substanzielle KI-gestützte Bearbeitung von Hintergrunddetails journalistischer Bilder, die über übliche technische, redaktionelle Praktiken hinausgeht, die Wahrnehmung der Authentizität und Wahrhaftigkeit des Inhalts durch eine Person negativ beeinflussen.«

Gemäß Guidelines wären kleinere, nicht-substanzielle Veränderungen nun wohl ohne Kennzeichnung zulässig, wenn sie die Echtheitserwartung nicht beeinträchtigen. Allerdings könnten sich solche Eingriffe im redaktionellen Bereich aus Gründen journalistischer Ethik verbieten.

Leider liefern die Guidelines (bis auf die oben genannten Beispiele mit generierten Produktbildern oder Hintergründen) keine weiteren praxisnahen Beispiele für unsere Arbeitsfelder.

Wir haben daher versucht, vom Service Desk des AI Office Hinweise zu konkreten Anwendungsfällen aus unserer Praxis zu bekommen. Die Antworten blieben eher vage.

Der AI Act Service Desk hat der Wiedergabe seiner Antworten in zusammengefasster Form nicht widersprochen, kann für die Richtigkeit dieser Zusammenfassung aber nicht einstehen. Maßgeblich sind allein der AI Act (Verordnung (EU) 2024/1689) einschließlich seiner Erwägungsgründe sowie die Guidelines der Kommission (C(2026) 5054 final vom 20. Juli 2026). Die Aussagen des Service Desk sind jedenfalls nur allgemein und unverbindlich.

Wir haben folgende Fälle abgefragt:

- Portrait eines CEO für Unternehmens-PR: folgende KI-gestützte Bearbeitungen: Entfernung von Hautunreinheiten, Aufhellung von Augenringen, Entfernung von Falten im Hemd, Abschwächung eines Doppelkinns, leichte Korrektur abstehender Ohren, Entfernung von Steckdosen und Mülleimern im Hintergrund, Entfernung eines Stativs, das in das Bild ragt.
- Foto einer Stadtlandschaft für Tourismus-PR: folgende KI-gestützte Bearbeitungen: Entfernung von Müll, Entfernung von temporären Zäunen oder Kränen und von Graffiti-Tags, Entfernung einer obdachlosen Person.
- Architekturfoto für das Marketing eines Architekturbüros: umfangreiche KI-gestützte Bearbeitung: Entfernung von Stromleitungen, geparkten Autos, Verkehrszeichen, Mülleimern, Austausch des Himmels.

Der Service Desk hat darauf vor allem mit Verweis auf die entsprechenden Punkte in den Guidelines geantwortet. Er hat aber auch angefügt, dass - je nach den Besonderheiten des individuellen Falles - die Entfernung von Kleiderfalten, Mülleimern, Steckdosen oder Stativen in einem Corporate-Portrait als irrelevant eingeschätzt werden könnte. Ebenso die Entfernung von Stromleitungen, Mülleimern oder geparkten Autos in einem Architekturfoto. Hier ist nicht nur interessant, was der Service Desk in seine Aufzählung aufgenommen hat, sondern auch, was er von unseren oben genannten Beispielen für KI-gestützte Bearbeitungen nicht erwähnt hat. Ob und welche Schlüsse daraus gezogen werden können, ist freilich unklar.

Wichtig sind auf jeden Fall der Einsatzkontext und die Erwartungen des Publikums, also ob ein Bild der Unternehmenskommunikation oder der Werbung dient oder ob es sich um ein Pressefoto handelt. Was in offensichtlich werblichem Kontext vielleicht nicht gekennzeichnet werden muss, müsste in der Unternehmenskommunikation möglicherweise doch und bei Verwendung in redaktionellem Kontext auf jeden Fall gekennzeichnet werden.

Etwas aussagekräftiger war eine Antwort auf eine Detailfrage betreffend ein Foto für Tourismus-PR: wir hatten geschrieben, dass wir annehmen, dass die KI-gestützte »Renovierung« einer stark verfallenen Fassade einer Sehenswürdigkeit ein Deepfake sein könnte. Hier besteht nach Einschätzung des Service Desk tatsächlich das Risiko, dass ein Bild als Deepfake angesehen werden kann, wenn eine KI-Bearbeitung dazu führt, dass ein Ort in einem Zustand präsentiert wird, den Besucher tatsächlich nicht vorfinden werden.

Auch nach Beantwortung unserer Fragen bleibt unser Fazit: die Übergänge sind fließend. Wir müssen bei Grenzfällen selbst abwägen und entscheiden, ob wir kennzeichnen oder nicht.

Wer muss kennzeichnen?

Kennzeichnen muss der »Betreiber«, also »eine natürliche oder juristische Person, Behörde, Einrichtung oder sonstige Stelle, die ein KI-System in eigener Verantwortung verwendet, es sei denn, das KI-System wird im Rahmen einer persönlichen und nicht beruflichen Tätigkeit verwendet«.

Hier offenbart sich nun leider ein grundsätzlicher Webfehler des AI Act. Denn Normadressaten sind nicht in erster Linie die Plattformen und Verlage und Unternehmen, die KI-generierte und KI-bearbeitete Bilder veröffentlichen. Diese gelten nur dann als »Betreiber«, wenn die Bilder unter ihrer Verantwortung und Kontrolle gepromptet oder mit KI modifiziert wurden. In den meisten Fällen blicken uns unsere Kunden aber nicht über die Schulter und leiten uns bei der Bearbeitung unserer Bilder auch nicht an. Daher sind in aller Regel wir die eigenverantwortlichen »Betreiber« der KI-Systeme, und sind somit in der Pflicht, ein sichtbares Label anzubringen.

Nur in folgenden Fällen wären die Auftraggeber*innen selbst (auch) »Betreiber«:

- wenn sie den KI-Einsatz explizit anweisen oder vorgeben (z. B. durch exaktes Briefing oder Vertrag),
- wenn der Fotograf/die Fotografin in die redaktionellen Workflows eines Unternehmens eingegliedert ist, oder
- wenn der Kunde/die Kundin nach Lieferung selbst KI-Bearbeitungen vornimmt.

Plattformen, Verlage, Unternehmen, Agenturen - also alle, die unsere Bilder veröffentlichen - werden von den Guidelines zwar »nachdrücklich ermutigt«, die sichtbaren und unsichtbaren KI-Kennzeichnungen zu erhalten, sind selbst aber nicht direkt in der Pflicht. Eine Ausnahme bilden sehr große Online-Plattformen (VLOPs): Sie unterliegen über Art. 35 Digital Services Act der EU mitunter eigenen Pflichten, KI-Inhalte erkennbar zu machen. Für unsere typischen Auftraggeber - Verlage, Agenturen, Unternehmen - gilt das aber nicht.

Das ist ein wirklich großes Defizit der gesamten Regelungsarchitektur: Die Betreiber müssen zwar zwingend kennzeichnen, die Publizierenden müssen die Kennzeichnung aber nicht unbedingt beibehalten. Hier wird einmal mehr deutlich, dass sich die Autoren des AI Act nicht wirklich mit den Verwertungsketten der unterschiedlichen Werkformen beschäftigt haben. Wir haben in den Konsultationsprozessen immer wieder auf diese Diskrepanz aufmerksam gemacht.

Gekennzeichnet werden muss spätestens zum Zeitpunkt der ersten Interaktion oder Wahrnehmung durch eine natürliche Person. Nach unserer derzeitigen Einschätzung muss schon bei der Übergabe an die Kund*innen ein sichtbares KI-Label angebracht werden. Die genannten natürlichen Personen sind also wohl die Mitarbeiter*innen unserer Kund*innen, wie Artbuyer*innen, Layouter*innen oder Grafiker*innen.

Dies ist natürlich hochgradig unpraktisch, da zum Zeitpunkt der Übergabe noch nicht klar ist, ob und gegebenenfalls wie das Bild bei der tatsächlichen Veröffentlichung beschnitten werden soll (mehr dazu unter: Wie lässt sich die Kennzeichnungspflicht bei der Übergabe an Kund*innen praktisch umsetzen?)

Wie muss gekennzeichnet werden?

Die konkreten Anforderungen für die Kennzeichnung liefert der Code of Practice. Kernstück ist ein sichtbares Label, das als zentrales Element das großgeschriebene Kürzel »AI« trägt. Eine nationalsprachliche Variante (also »KI«) ist nur dort vorgesehen, wo nationales Recht die Verwendung des Englischen ausschließt.

Die Buchstaben müssen die gleiche Höhe haben und ihre Proportionen beim Skalieren behalten; Größe, Farbe, Kontrast und Typografie dürfen variieren, solange das Label klar, zugänglich und für einen Menschen unterscheidbar bleibt.

Die EU stellt dafür drei Icons zur Verfügung (herunterzuladen [hier](#)): eines für vollständig KI-generierte Inhalte (»AI GENERATED«), eines für mit KI bearbeitete Inhalte (»AI MODIFIED«) und ein Basis-Icon (»AI«), das um eine zweite (interaktive) Informationsebene ergänzt werden kann.

AI GENERATED

AI MODIFIED



Man muss das EU-Icon aber nicht verwenden, ein eigenes Label genügt, solange es die genannten Vorgaben – also eine klare, unterscheidbare und barrierefreie Offenlegung – erfüllt. Im AI Act selbst ist übrigens anders als im CoP das englische »AI« nicht vorgeschrieben. Für Nutzungen rein im deutschsprachigen Raum wäre ein Label mit den Buchstaben »KI« somit auch möglich. Wer den CoP unterzeichnet hat, sollte allerdings beim dort vorgesehenen »AI«-Kürzel bleiben.

Das Label muss ohne Zutun und ohne anhaltende Aufmerksamkeit der Betrachtenden sofort erkennbar sein, lange genug sichtbar bleiben und sich vor jedem Hintergrund deutlich abheben. Es muss direkt in den Inhalt selbst eingebettet werden, beim Bild etwa in einer freien Ecke (im CoP wird die obere rechte Ecke genannt), und zwar spätestens bei der ersten Wahrnehmung durch einen Menschen. Falls es gleichwertige Alternativen zu einer Platzierung direkt im Bild gibt, können diese angewendet werden (z.B. ein Overlay, durch das beim Betrachtenden der Eindruck erweckt wird, als sei das Label direkt in den Inhalt eingebettet). Bei Videos muss das Label am Anfang und zusätzlich soweit möglich in regelmäßigen Abständen bzw. nach Unterbrechungen (z. B. Werbeblöcken) erscheinen.

Sonderfall künstlerischer, kreativer, satirischer, fiktionaler oder analoger Werke:

Ist ein Bild »Teil eines offensichtlich künstlerischen, kreativen, satirischen, fiktionalen oder analogen Werks oder Programms«, gilt eine abgeschwächte Pflicht. Das Label muss zwar auch dann gesetzt werden, darf aber so platziert werden, dass es die Darstellung und den Genuss des Werks nicht beeinträchtigt, also etwa außerhalb bzw. neben dem Bild- oder Videorahmen, in begleitenden Hinweisen oder einer Beschreibung, im Vor- oder Abspann, in einem Ausstellungsflyer, auf dem Eintrittsticket oder der Verpackung, bzw. bei digitaler Ausspielung in einem Overlay oder einer anklickbaren Schaltfläche.

Für diese Ausnahme gelten aber strenge Regeln: Erstens muss der künstlerische/kreative Charakter offensichtlich sein. Inhalte, deren Charakter für das Publikum unklar oder mehrdeutig ist, fallen aus dem Sonderregime heraus, ebenso Inhalte, die primär informieren oder kommerziell wirken sollen. Die Guidelines formulieren es ausdrücklich einzelfallabhängig: Manche Inhalte (etwa Werbung oder Dokumentationen) mit Deepfake-Anteilen können in der einen Situation als offensichtlich kreatives oder fiktionales Werk gelten, in einer anderen aber nicht. Für unseren typischen Berufsalltag heißt das: wir sollten uns auf diese Ausnahme besser nicht verlassen.

Wie lässt sich die Kennzeichnungspflicht bei der Übergabe an Kund*innen praktisch umsetzen?

Eine besondere Herausforderung entsteht, wenn Fotograf*innen kennzeichnungspflichtige Bilder an Kund*innen übergeben, die anschließend über den endgültigen Bildzuschnitt, das Format und den Veröffentlichungskontext entscheiden. Ein im Bild angebrachtes Label kann durch spätere Bearbeitungen oder den Zuschnitt verloren gehen.

Die finalen Guidelines sehen für solche komplexen Produktions- und Vertriebsketten vor, dass Betreiber »verhältnismäßige Maßnahmen« ergreifen sollen, um sicherzustellen, dass die Kennzeichnung für das vorgesehene und vernünftigerweise vorhersehbare Publikum bei der ersten Wahrnehmung klar und unterscheidbar angezeigt wird. Als Beispiele für verhältnismäßige Maßnahmen nennen die Guidelines unter anderem vertragliche Vereinbarungen mit Vertriebspartnern.

Wir haben den AI Act Service Desk gefragt, ob Fotograf*innen die Kennzeichnungspflicht vertraglich auf ihre Kund*innen übertragen können, da ja in der Regel erst diese über das endgültige Verwendungsformat entscheiden. Der Service Desk wollte sich dazu nicht festlegen. Wir verstehen seine Antwort im Kontext aber so, dass vertragliche Vereinbarungen zwar hilfreich zur Koordination in den oben genannten komplexen Produktions- und Vertriebsketten sind, dass sie uns aber nicht von der Pflicht entbinden, Deepfakes bereits bei der Übergabe an die Kund*innen sichtbar mit Label zu kennzeichnen.

Auch die Dokumentation von KI-Bearbeitungen in den Metadaten von Bildern oder durch Einbettung von C2PA reicht nicht aus, um unseren Kennzeichnungspflichten nachzukommen, da diese Informationen für natürliche Personen ohne technische Hilfsmittel nicht wahrnehmbar sind. Als flankierende Maßnahmen sind Informationen über Bearbeitungsschritte in den Metadaten aber jedenfalls sinnvoll.

Eine von FREELENS vorgeschlagene Möglichkeit, die Umsetzung der Kennzeichnungspflicht praktisch handhabbar zu machen, wäre, die Bilddatei um einen weißen Rand zu erweitern, in dem ein deutlich sichtbares Label angebracht wird, und zugleich die Kund*innen vertraglich zu verpflichten, dass sie selbst das Label bei der endgültigen Veröffentlichung korrekt platzieren müssen. Wir haben den AI Act Service Desk um eine Einschätzung dieses Vorgehens gebeten, der Service Desk wollte sich aber auch hier nicht festlegen. Wir finden jedoch, dass dieses

Vorgehen den Anforderungen des AI Act genügen müsste, da das Label bei der Öffnung der Datei sofort sichtbar wäre.

Bis zu einer endgültigen Klärung sollte davon ausgegangen werden, dass Fotograf*innen kennzeichnungspflichtige Inhalte grundsätzlich mit einer wahrnehmbaren Kennzeichnung im Bild ausliefern müssen. Zusätzlich sollten sie laut Guidelines wie gesagt »verhältnismäßige Maßnahmen« ergreifen, um sicherzustellen, dass das Label am Ende beim Zielpublikum ankommt. Das hängt natürlich von euren Möglichkeiten ab, auf das Verhalten des Kunden Einfluss zu nehmen. Ihr solltet den Kunden jedenfalls darauf hinweisen, dass das Label beibehalten werden muss. Wenn euer Kunde/eure Kundin aber das KI-Label abschneidet, dürfte euch das regelmäßig nicht anzulasten sein, sofern ihr die verhältnismäßigen Maßnahmen ergriffen habt.

Textbausteine für mögliche vertragliche Regelungen findet ihr im Anhang. Sie sind unverbindliche Formulierungshilfen und sollten im Einzelfall rechtlich geprüft werden.

Sollen Fotograf*innen nun den Code of Practice unterzeichnen?

Das AI Office ermuntert alle Anbieter und Betreiber von KI-Systemen, den Code of Practice on Transparency of AI-Generated Content zu unterzeichnen. Da alle von uns, die beruflich KI-Tools verwenden, wie oben geschildert »Betreiber« sind, fallen auch wir in den Kreis der potenziellen Unterzeichner*innen.

Die im CoP aufgeführten Maßnahmen sollen bei der praktischen Erfüllung der Verpflichtungen gemäß AI Act unterstützen. Wer den CoP unterschreibt, bekennt sich dazu, die betreffenden Maßnahmen zu ergreifen, das soll die Nachweisbarkeit eines AI-Act-konformen Verhaltens erleichtern.

Aber lohnt sich eine Unterschrift tatsächlich? In den Guidelines wird angeführt, dass den Unterzeichner*innen von den Aufsichtsbehörden ein höheres Vertrauen entgegengebracht werden soll, während Nicht-Unterzeichner*innen im Prüffall intensivere Nachfragen zu erwarten haben und einen höheren Aufwand treiben müssen, um ihr regelkonformes Verhalten zu belegen.

Wir nehmen an, dass eine Unterzeichnung wohl eher für jene von uns sinnvoll ist, die tatsächlich vor allem KI-Bilder generieren und nicht bloß Fotos mit KI-Tools bearbeiten. Wer sich gegen eine Unterzeichnung entscheidet, sollte jedenfalls seine eigene Compliance-Logik gründlich dokumentieren.

Hier findet ihr das Formular für die Unterzeichnung des Code of Practice: <https://digital-strategy.ec.europa.eu/en/library/how-sign-code-practice-transparency-ai-generated-content>. Eventuell wird künftig auch eine Online-Lösung statt des derzeitigen Word-Formulars zur Verfügung gestellt.

Grundsätzlich gilt: Wir sind — ähnlich wie bei der DSGVO — nicht im Fokus dieser Verordnung. Wir müssen sie beachten, aber es steht nicht zu befürchten, dass demnächst ein Beamter der Aufsichtsbehörde an der Türe klingelt, um die Einhaltung des AI Act zu kontrollieren.

Ab wann muss gekennzeichnet werden?

Alle ab dem 2. August 2026 geschaffenen »Deepfakes« müssen spätestens zum Zeitpunkt der ersten Interaktion oder Wahrnehmung gekennzeichnet werden. Vor dem 2. August generierte Bilder oder mit KI bearbeitete Fotos müssen nachträglich nicht gekennzeichnet werden. Das AI Office »ermuntert« die »Betreiber« aber, dies dennoch zu tun. Im Interesse der Transparenz gegenüber den Kund*innen sollten wir das auch tun.

Die Anforderungen von redaktionellen Kund*innen werden wohl noch über den AI Act hinausgehen. Denn in aller Regel ist die Manipulation von Fotografien hier nicht zulässig, egal ob mit oder ohne KI. Hier sollte schon heute gelten: Keine Pixel verschieben, keine Pixel entfernen, keine Pixel hinzufügen.

Apropos Inkrafttreten: Art. 50 (2) AI Act enthält ja auch eine Pflicht für »Anbieter« von KI-Systemen (also Midjourney, Adobe, Gemini etc.), synthetische Audio-, Bild-, Video- oder Textinhalte maschinenlesbar zu kennzeichnen (also unsichtbar per Wasserzeichen, Metadaten, mit kryptografischen Methoden etc.) und als künstlich erzeugt oder manipuliert erkennbar zu machen. Diese Pflicht, die ursprünglich auch schon am 2. August in Kraft treten sollte, wurde nun vom sogenannten »Digital-Omnibus« auf 2. Dezember 2026 verschoben für KI-Systeme, die bereits vor dem 2. August auf dem Markt waren. Sinnvollerweise hätten diese unsichtbare Markierung und das sichtbare Label von Beginn an Hand in Hand gehen sollen. Die Kennzeichnung hätte dann automatisiert werden können, was uns die Arbeit sicherlich erleichtert hätte.

Welche Konsequenzen kann ein Verstoß haben?

Theoretisch drohen empfindliche Bußgelder: Bei Verstößen gegen die Transparenzpflichten sieht der AI Act Strafen von bis zu 15 Millionen € vor oder - wenn dieser Betrag höher ist - bis zu 3 % des gesamten weltweiten Jahresumsatzes des vorangegangenen Geschäftsjahres. Für kleine und mittlere Unternehmen einschließlich Soloselbstständiger gilt dagegen jeweils der niedrigere der beiden Beträge als Obergrenze. Zudem müssen die Behörden bei der Bemessung die wirtschaftliche Tragfähigkeit kleiner Anbieter und Betreiber berücksichtigen.

Wie streng und wie schnell die nationalen Aufsichtsbehörden die Transparenzpflichten gegenüber einzelnen Solo-Selbstständigen tatsächlich durchsetzen werden, ist noch offen. Zur Sicherheit solltet ihr immer dokumentieren, welche KI-Tools ihr in welchem Umfang einsetzt, wie ihr die Kennzeichnungspflicht prüft und wie ihr die Kennzeichnung praktisch sicherstellt. Eine Archivierung der Originaldateien sowie der an die Kund*innen übergebenen Dateien ist ebenfalls sinnvoll.

Jedenfalls solltet ihr besser nicht darauf hoffen, dass man nicht erkennen kann, dass ihr generative KI eingesetzt habt. Denn die Anbieter von KI-Systemen müssen KI-generierte Inhalte künftig nicht nur maschinenlesbar/unsichtbar kennzeichnen, sondern auch Erkennungsverfahren öffentlich verfügbar machen, mit denen KI-Bilder oder mit KI manipulierte Fotos identifiziert werden können. Vermutlich wird es künftig Tools geben, mit denen Bilder auf Markierungen aller KI-Systeme überprüft werden können.

Wie geht man nun am besten mit der Situation um?

Die finalen Guidelines lassen, wie oben beschrieben, einen gewissen Spielraum an KI-Gebrauch zu, ohne in die Kennzeichnungspflicht zu rutschen. Im Zweifel gilt aber: besser einmal zu viel gekennzeichnet als einmal zu wenig.

Wer auf die neuen Tools nicht verzichten will, sollte jedenfalls im Vorfeld abklären, ob das aus Kund*innensicht überhaupt erwünscht oder zulässig ist. Manche Auftraggeber*innen - vor allem aus dem redaktionellen Bereich - haben auch strengere Anforderungen als der AI Act. Komplette Transparenz über das eigene Vorgehen, in schriftlicher Konversation und auch exakt dokumentiert in den Metadaten, ist wichtiger denn je. Idealerweise entwickelt ihr einen Workflow, der von der Kamera bis zur Ausgabe alle Bearbeitungsschritte dokumentiert. Am besten sowohl in den Metadaten als auch - etwas robuster - mittels C2PA/Content Credentials.

Unterliegen auch Selbständige einer Schulungspflicht?

Die bisherige Pflicht zum Erwerb von KI-Kompetenz aus Artikel 4 AI Act wird durch den sogenannten »Digital-Omnibus« etwas aufgeweicht (der Erwerb von KI-Kompetenz muss nun nicht mehr »sichergestellt«, sondern nur noch »unterstützt« werden). Zugleich werden nun die Unterzeichner*innen des CoP angehalten, Maßnahmen zur Förderung der Kenntnisse betreffend Kennzeichnungspflichten zu ergreifen. Bei selbständigen KI-Betreibern wie euch bedeutet das, dass ihr selbst für eure Fortbildung sorgen müsst, immer in angemessenem Verhältnis zur Größe und den verfügbaren Ressourcen eures Unternehmens.

FREELENS bietet regelmäßig Schulungen an, die diese Kenntnisse vermitteln.

ANHANG - TEXTBAUSTEINE FÜR VERTRAGLICHE REGELUNGEN MIT KUND*INNEN

Nach derzeitigem Stand können wir euch folgende Textbausteine für vertragliche Regelungen mit euren Kund*innen bereitstellen. Sie sind unverbindliche Formulierungshilfen, kein Ersatz für eine anwaltliche Prüfung im Einzelfall. Sie entfalten ihre volle Wirkung eher als individuell ausgehandelte Vereinbarung denn als AGB-Klauseln:

(1) Soweit gelieferte Bild- oder Videoinhalte unter Einsatz generativer KI-Werkzeuge so erzeugt oder bearbeitet wurden, dass sie als Deepfake im Sinne von Art. 3 Nr. 60 KI-Verordnung einzustufen sind, liefert der Fotograf/die Fotografin diese mit einem sichtbaren AI-Label entsprechend dem »Code of Practice on Transparency of AI-Generated Content« (CoP) aus. Form und Position legen die Parteien im Rahmen der Anforderungen des CoP einvernehmlich fest.

(2) Der Auftraggeber/die Auftraggeberin verpflichtet sich, das vom Fotografen/von der Fotografin gesetzte AI-Label und die Metadaten des Bildes bei jeder Veröffentlichung und Weitergabe zu erhalten. Wird der Inhalt zugeschnitten, ins Layout integriert oder anderweitig bearbeitet, sodass das Kennzeichen entfällt, verdeckt oder unkenntlich wird, bringt der Auftraggeber/die Auftraggeberin es spätestens zum Zeitpunkt der ersten Wahrnehmung durch eine natürliche Person an geeigneter, klar und deutlich erkennbarer Stelle des veröffentlichten Inhalts erneut an.

(3) Sind die Veröffentlichungsformate zu divers für eine Platzierung des AI-Labels an einer bestimmten Stelle im Bild, platziert der Fotograf/die Fotografin das AI-Label auf einem dem eigentlichen Bild angefügten weißen Rand. Der Auftraggeber/die Auftraggeberin verpflichtet sich, die Kennzeichnung in der finalen Veröffentlichung gemäß den Vorgaben des AI Act und des Code of Practice selbst neu anzubringen und die Kennzeichnung so aufrechtzuerhalten.

(4) Der Auftraggeber/die Auftraggeberin stellt den Fotografen/die Fotografin von allen Ansprüchen, Bußgeldern und Kosten (einschließlich angemessener Rechtsverteidigungskosten) frei, die daraus entstehen, dass er/sie eine ihm/ihr nach dieser Vereinbarung obliegende Kennzeichnung unterlässt, entfernt oder fehlerhaft anbringt.

(5) Nimmt der Auftraggeber/die Auftraggeberin selbst KI-Bearbeitungen vor, ist er/sie für deren Kennzeichnung allein verantwortlich.