https://doi.org/10.21248/gups.fralr.26.1.04

Öffentliches Recht

Deepfakes im Lichte der KI-Verordnung

Zugleich eine kritische Analyse von Art. 50 Abs. 4 UAbs. 1 S. 1 KI-VO unter besonderer Berücksichtigung des Einsatzes von *Deepfakes* zu Werbezwecken

Janek, Moritz*

ZUSAMMENFASSUNG

Die digitale Transformation stellt das geltende Recht vor erhebliche Herausforderungen. Das zeigt sich am Beispiel der Künstlichen Intelligenz und ihren Anwendungsformen wie *Deepfakes*. Hierbei handelt es sich um durch generative KI erzeugte Bild-, Ton- oder Videoinhalte, die biometrische Merkmale von Personen täuschend echt imitieren. Dies birgt enorme Risiken, eröffnet aber insbesondere der Werbeindustrie ein erhebliches Nutzungspotenzial. Mit Art. 50 Abs. 4 UAbs. 1 S. 1 KI-VO hat der europäische Gesetzgeber erstmals eine explizite Regelung betreffend *Deepfakes* geschaffen, die Betreiber zur Kennzeichnung verpflichtet. Wenngleich diese Regelung grundsätzlich begrüßenswert ist, droht sie aufgrund fehlender Konkretisierung, einem eingeschränkten Adressatenkreis und mangelnder Identitätsoffenlegungspflicht ins Leere zu laufen. Es bedarf zeitnaher Nachbesserung, um eine maximale Durchsetzungskraft zu erreichen und zeitgleich Innovation rechtlich zu ermöglichen.

Keywords *Deepfakes*; Transparenzpflicht; KI-Verordnung; Digitalisierung

A. Technologische Innovation als Rahmenbedingung rechtlicher Veränderung

I. Digitale Transformation

Der Begriff der digitalen Transformation bezieht sich interdisziplinär auf sämtliche durch die Digitalisierung bewirkte Veränderungsprozesse. Umfasst sind neue Handlungs- und Geschäftsmöglichkeiten sowie nachhaltige Änderungen in Wirtschaft, Gesellschaft, Politik und Recht. Aufgrund dieser Tragweite wird nicht nur von einer reinen „Transformation", sondern gar von einer „digitalen Disruption" gesprochen.[1]

Am Anfang jeglicher Transformation steht eine Innovation – ein Novum, das es bisher noch nicht gab. Hierbei werden verschiedene Instrumente miteinander vereint, die in ihrem Zusammenspiel eine Neuerung herbeiführen. Insofern lassen sich verschiedene „Bausteine"[2] der digitalen Transformation herauskristallisieren. Dazu zählen: algorithmische Systeme, (personenbezogene) Daten, *Big Data*, Internet der Dinge (IoT) und *Blockchain*. In ihren unterschiedlichen Facetten und durch ihr Zusammenspiel bilden sie die Grundlage jeglicher Innovation und Transformation. Künstliche Intelligenz, die zuvorderst auf algorithmischen Systemen basiert und auf (personenbezogene) Daten angewiesen ist, wurde durch ebendiese Bausteine ermöglicht.

II. Herausforderungen de lege lata

Die rasanten technologischen Entwicklungen stellen das geltende Recht vor erhebliche Herausforderungen. Diesen muss sich die „Rechtsordnung im Sog der Erwartungen an eine am Individual- und Gemeinwohl orientierte Gestaltung der Entwicklung auch mithilfe des Rechts stellen".[3] Neue Technologien entstehen keineswegs in einem rechtsfreien Raum. Ein erster Ansatz besteht darin, bestehende Regelungen auf neuartige und bislang unbekannte Sachverhalte anzuwenden. Dabei kommen die klassischen juristischen Auslegungsmethoden zum Einsatz. Wo diese an ihre Grenzen stoßen, ist es Aufgabe des Gesetzgebers, die Fragen in Form neuer Normen zu beantworten. Das gebietet die sog. Wesentlichkeitstheorie.[4] Daneben gibt es Bereiche, in denen nahezu jegliche staatliche Regulierung fehlt. In solchen Fällen übernehmen derzeit große (Tech-)Unternehmen eine selbstregulierende Funktion, etwa durch die Einführung eines Code of Conduct. Dies widerspricht dem klassischen Verständnis des staatlichen Gewalt- und Rechtssetzungsmonopols.

Dem Recht kommt also eine primär reaktive Rolle zu. Es reagiert auf bereits eingetretene technologische

*Der Verfasser studiert Rechtswissenschaft an der Johann Wolfgang Goethe-Universität Frankfurt am Main. Daneben ist er als Wissenschaftlicher Mitarbeiter in einer international tätigen Wirtschaftskanzlei im Bereich Technology, Media & Telecommunications (TMT) tätig. Der Beitrag geht auf eine Seminararbeit im Rahmen des Seminars *„Die Digitale Transformation als Rahmenbedingung des Rechts"* bei Prof. Dr. Roland Broemel aus dem Sommersemester 2025 zurück.

Zitieren als: Janek, *Deepfakes* im Lichte der KI-Verordnung, FraLR 2026 (01), S. 23-31. DOI: 10.21248/gups.fralr.26.1.04

[1]Etwa Hoffmann-Riem (2022), Recht im Sog der digitalen Transformation, 1. Auflage, § 1 S. 4; Vesting in Eifert, Digitale Disruption und Recht, S. 16; Daub (2025), Was das Valley denken nennt, 5. Auflage, S. 106.

[2]Der Begriff, sowie die nachfolgende Auflistung, geht zurück auf bzw. orientiert sich an Hoffmann-Riem (2022), Recht im Sog der digitalen Transformation, 1. Auflage, § 4.

[3]Hoffmann-Riem (2022), Recht im Sog der digitalen Transformation, 1. Auflage, § 1 S. 1.

[4]Etwa BVerfGE 49, 89 (126f.); BVerfGE 34, 165 (192f.).

Innovationen, anstatt diese im Voraus zu gestalten. Diese Erkenntnis ist keineswegs neu. Bereits 1964 formulierte Friedrich Karl Fromm treffend: „Das Recht muß zur Kenntnis nehmen, daß aus der Utopie von gestern die Wirklichkeit von heute geworden ist, die nach der juristischen Ordnung von morgen ruft".[5]

III. Aufgaben des Rechts im Bereich der digitalen Transformation

Letztlich dient das Recht als Instrument zur Bewältigung der durch die digitale Transformation indizierten Herausforderungen, indem es Entwicklungsprozesse begleitet, strukturiert und soweit erforderlich auch begrenzt.[6] Gesetzgeberische Reaktionen sollten darauf abzielen, Chancen der Digitalisierung zu ermöglichen, gleichzeitig aber Risiken zu minimieren oder gar auszuschließen. Kurzum: Ziel des Gesetzgebers muss es sein, das Recht zur Ermöglichung von Innovation zu konzipieren.[7]

B. Chancen und Risiken von *Deepfakes*

I. „Friedliches" Nutzungspotenzial

Die vorgenannte Aufgabe trifft staatliche Akteure auch und insbesondere mit Blick auf *Deepfakes*. Ausgehend vom Wortlaut ist der Begriff *Deepfake* („*fake*") negativ konnotiert.[8] Dieser Anschein trügt vor dem enormen („friedlichen") Nutzungspotenzial.[9] Dieses tritt namentlich etwa in der Werbeindustrie zutage, die in besonderem Maße auf audiovisuelle Inhalte angewiesen ist. Deepfakes bieten hier enorme Kostensparpotentiale, etwa durch die Produktion rein virtueller Werbekampagnen.[10] Unternehmen sind in der Lage, eigene virtuelle Influencer zu kreieren, über deren Auftreten und Verhalten sie die vollständige Kontrolle behalten.[11] Darüber hinaus versetzt der Einsatz von *Deepfakes* Unternehmen in die Lage, einen einmal produzierten Werbespot ohne gesonderte Neuproduktion in eine Vielzahl von Sprachen zu übersetzen.[12] Künftig dürften schließlich sog. „hyperpersonalisierte" Werbekampagnen besonders vielversprechend erscheinen, bei denen die Konsumentscheidung und -einstellung der Kunden gezielt beeinflusst wird.[13]

II. Besonderes Risikopotenzial von Deepfakes

Mittels KI-Technologie können im digitalen Raum alternative Realitäten geschaffen werden, die die Grenze zwischen Realität und Fiktion verschwimmen lassen.[14] Für einen unachtsamen Durchschnittsbürger lassen sich synthetisch erstellte Medieninhalte nicht mehr von ihrem Original unterscheiden. Gleichzeitig können *Deepfakes* inzwischen mittels einer simplen, einfachen Eingabe generiert werden, wozu der jeweilige Nutzer keinerlei (technisches) Know-how benötigt. Sie haben sich zu einem Alltagsphänomen entwickelt.[15] Das besondere Risikopotenzial ergibt sich aus dieser einfachen Handhabung gepaart mit den destruktiven Einsatzmöglichkeiten eigens generierter *Deepfakes*. Der europäische Gesetzgeber nennt beispielhaft: Fehlinformation und Manipulation in großem Maßstab, Betrug, Identitätsbetrug und Täuschung.[16] Insgesamt kann das fatale Folgen sowohl für das Individuum als auch für die Gesellschaft haben. Aus dem Blickwinkel des Individuums stehen dabei Persönlichkeitsrechtsverletzungen und Reputationsschäden[17] im Vordergrund.[18] Aus gesellschaftlicher Perspektive stellt authentisch wirkende Desinformation ein drängendes Problem dar, sowohl für die öffentliche Meinungsbildung als auch für den demokratischen Diskurs.[19] Erschwerend trifft diese Möglichkeit mit der bereits in Gang gesetzten Neuverteilung der (Meinungs-)Macht im digitalen Raum zusammen.[20] Diese bewegt sich zunehmend weg von den traditionellen audiovisuellen Medien hin zu digitalen Plattformen wie Instagram, TikTok und X. Es droht eine Informationsverzerrung durch Simulation von Faktizität. Von den falschen Personen genutzt, bergen *Deepfakes* die Gefahr, unser gegenseitiges Vertrauen ineinander, in staatliche Stellen und damit in die Demokratie insgesamt zu erschüttern.

---

[5] Fromm, GRUR 1964, 304 (306).

[6] Paal/Kumkar, ZfDR 2021, 97 (99).

[7] Hoffmann-Riem, JuS 2023, 617 (618); Hoffmann-Riem (2016), Innovation und Recht, 1. Auflage, S. 33.

[8] Die wohl erste namentliche Bezeichnung eines *Deepfakes* als ebensolchen entstammte einer Reddit-Gruppe aus dem Jahr 2017. Seinerzeit wurde ein KI-Tool dazu genutzt, Prominente Gesichter auf existierende pornographische Video-Clips zu projizieren. Wenngleich der Begriff umstritten war und auch immer noch ist, hat er sich in der Folgezeit durchgesetzt. Vertiefend hierzu Vincent, Why we need a better definition of 'deepfake', https://www.theverge.com/2018/5/22/17380306/deepfake-definition-ai-manipulation-fake-news (zuletzt aufgerufen am 26.04.2026).

[9] Thiel, ZRP 2021, 202 (203); Lantwin, MMR 2019, 574 (574f.).

[10] Während der Covid-Pandemie drehte der US-Sender „Hulu" einen Deepfake-Werbespot. Mithilfe von KI und sog. „Face Swap" entstand so ein Werbespot mit drei Sportlern. Siehe hierzu etwa https://www.ispot.tv/ad/n3Gt/hulu-live-sports-the-deepfake-featuring-saquon-barkley-baker-mayfield (zuletzt aufgerufen am 26.04.2026).

[11] Ein prominentes Beispiel ist die virtuelle Influencerin „Lil Miquela", die allein auf Instagram über 2,3 Millionen Follower hat (Stand April 2026).

[12] In einem Werbespot der „Malaria Must Die"-Initiative sieht man David Beckham insgesamt neun Sprachen flüssig sprechen. Zu dem Werbespot: https://www.youtube.com/watch?v=QiiSAvKJIHo (zuletzt aufgerufen am 26.04.2026).

[13] Pawelec/Bieß (2021), Deepfakes, 1. Auflage S. 62.

[14] Kraetzig, CR 2024, 207 (208) Rn. 3.

[15] Kumkar/Griesel, KIR 2024, 117 (118).

[16] Erwg. 133 S. 2 KI-VO.

[17] Hinderks, ZUM 2022, 110 (113); Kumkar/Griesel, KIR 2024, 117 (118).

[18] Erneut befeuert wurde die rechtspolitische Debatte um die Strafbarkeit rufschädigender Deepfakes durch die Berichterstattung des Spiegels, nach der Christian Ulmen über viele Jahre hinweg ua eine Vielzahl pornographischer Deepfakes seiner Ex-Frau Colien Fernandes erstellt und verbreitet haben soll (https://www.spiegel.de/netzwelt/netzpolitik/collien-fernandes-erstattet-anzeige-gegen-ex-mann-christian-ulmen-a-6abfb991-1665-4469-9c8e-3cc5a2cb4f29; zuletzt aufgerufen am 26.04.2026); Zum jüngsten gesetzgeberischen Reformvorhaben siehe nunmehr einen Referentenentwurf des BMJV vom 16.04.2026 zur Stärkung des zivil- und strafrechtlichen Schutzes vor digitaler Gewalt (online abrufbar unter https://www.bmjv.de/SharedDocs/Downloads/DE/Gesetzgebung/RefE/RefE_GgdG.pdf?__blob=publicationFile&v=2; zuletzt aufgerufen am 26.04.2026).

[19] Kumkar/Rapp, ZfDR 2022, 199 (201); Kraetzig, CR 2024, 207 (208) Rn. 3.

[20] Kumkar/Griesel, KIR 2024, 117 (118); Dörr, WRP 2021, 168-173; Ingold, MMR 2020, 82-86.

C. Regulierung von *Deepfakes* nach Art. 50 Abs. 4 UAbs. 1 S. 1 KI-VO

Diese Problematik vor Augen wurde der europäische Gesetzgeber mit Art. 50 Abs. 4 UAbs. 1 S. 1 der KI-VO[21] tätig und betritt insoweit Neuland. Vor Erlass dieser Regelung existierte keine explizite *Deepfake*-Regulierung.

I. Regulatorische Grundlagen der KI-VO im Überblick

Mit der KI-VO, dem ersten verbindlichen Regelwerk für Künstliche Intelligenz auf europäischer Ebene, verfolgt der europäische Gesetzgeber einen horizontalen, risikobasierten Ansatz.[22] Die verschiedenen KI-Systeme werden in vier Risikostufen eingeteilt und entsprechenden Pflichten unterworfen. Je höher das Risiko, das mit der Nutzung oder dem Inverkehrbringen eines KI-Systems verbunden ist, desto strenger die Anforderungen an diese Systeme.[23] Zu unterscheiden sind dabei verbotene KI-Praktiken (Art. 5 KI-VO), Hochrisiko-KI-Systeme (zB bei erheblichen Risiken für Gesundheit, Sicherheit oder Grundrechte, Art. 6–49 KI-VO), sowie KI-Systeme mit begrenzten (zB Chatbots) oder minimalen Risiken (zB KI-gestützte Spamfilter).

II. Inhalt und Regelungsbereich

Das besondere Risikopotenzial, das aus der einfachen Handhabung und Erstellung von *Deepfakes* bei gleichzeitig potenziell schwerwiegenden Auswirkungen erwächst, hat den Unionsgesetzgeber dazu veranlasst, eine ausdrückliche Regelung zu *Deepfakes* in die KI-Verordnung aufzunehmen.[24] Schutzzweck ist die Vermeidung von Täuschung über die vermeintliche Faktizität durch Schaffung von Transparenz.[25]

1. Erfasste KI-Systeme Mit Art. 3 Nr. 60 KI-VO enthält die KI-VO bereits eine Definition des Begriffes *Deepfake*. Demnach handelt es sich 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*“ lässt. Kurzum: *Deepfakes* sind jegliche täuschend echt wirkende Video-, Bild- oder Tonaufnahmen, die einen Realitätsbezug aufweisen.

a) KI-generierter Inhalt Für das Vorliegen eines *Deepfakes* bedarf es zunächst eines „*durch KI erzeugten oder manipulierten Bild-, Ton- oder Videoinhalt* (Art. 3 Nr. 60 KI-VO). Denkbar sind somit zweierlei Fälle: der *Output* kann entweder von Grund auf neu generiert sein oder einen ursprünglichen Inhalt manipulieren. Künstlich generierte Texte sind nicht umfasst, sondern unterliegen allein den in Art. 50 Abs. 4 UAbs. 2 S. 1 KI-VO niedergelegten Transparenzanforderungen.

b) Realitätsbezug Damit der *Output* als *Deepfake* zu kategorisieren ist, muss er jedenfalls „*wirklichen Personen, Gegenständen, Orten, Einrichtungen oder Ereignissen ähneln*“ (Art. 3 Nr. 60 KI-VO). Nicht erfasst sind insoweit solche Inhalte, die erkennbar fiktional sind.[26] Anknüpfend hieran stellt sich jedoch die Frage, welche Anforderungen an das „Ähneln“ zu stellen sind. Teilweise wird verlangt, dass die porträtierten Inhalte einen Anknüpfungspunkt in der Realität haben müssen (mithin ein Foto oder Video bereits real existierte, bevor es mittels KI verändert wurde).[27] Von anderen Autoren wird es als ausreichend erachtet, dass die generierten Inhalte frei erfunden sind, aber ihrer Art nach real existieren könnten bzw. real existierenden Personen bzw. Gegenständen „merklich“ ähneln (sog. potenzielle Realität).[28] Richtig ist zunächst, dass der sehr weit geratene Wortlaut des Art. 3 Nr. 60 KI-VO einer Einschränkung bedarf,[29] was für erstgenannte Ansicht spricht.[30] Der Begriff „*ähneln*“ steht sprachlich im Zusammenhang mit den zuvor genannten „*wirklichen Personen, Gegenständen, Orten, Einrichtungen oder Ereignissen*“. Ebenso ist in der englischen (Original-)Fassung die Rede von einem generierten Inhalt, „*that resembles existing persons, objects, places, entities or events*“. Hieraus könnte man den Schluss ziehen, dass der Gesetzgeber mit dem Wort „*wirklich*“ (engl. „*existing*“) bewusst einen tatsächlichen und nicht nur potenziellen Realitätsbezug zum Ausdruck bringen wollte. Einer solchen Interpretation steht jedoch der in ErwG 120 KI-VO zum Ausdruck kommende gesetzgeberische Wille entgegen.[31] Mit ErwG 120 S. 1 KI-VO stellt der Gesetzgeber klar, dass die im Rahmen der KI-VO normierten Erkennungs- und Offenlegungspflichten, mithin auch diejenige des Art. 50 Abs. 4 KI-VO, ebenso der wirksamen Durchsetzung des DSA[32] dienen sollen. Im Rahmen des ErwG 120 S. 2 KI-VO wird explizit Bezug auf die Pflichten der Anbieter sehr großer Online-Plattformen und sehr großer Online-Suchmaschinen (iSd Art. 33 Abs. 1 DSA) genommen, „*systematische Risiken zu ermitteln und zu mindern, die sich aus der Verbreitung von künstlich erzeugten oder manipulierten Inhalten*“ ergeben können.[33] Dabei hatte der Gesetzgeber insbesondere etwaige negative „*Auswirkungen auf demokratische Prozesse, den gesellschaftlichen Diskurs und Wahlprozesse, unter anderem durch Desinformation*“ vor Augen.[34] Das zeigt, dass Sinn und Zweck der Regulierung von *Deepfakes* insbesondere der Schutz

[21] Verordnung (EU) 2024/1689 des Europäischen Parlaments und des Rates vom 13. Juni 2024 zur Festlegung harmonisierter Vorschriften für künstliche Intelligenz.

[22] Buchalik/Gehrmann, CR 2024, 145 (146) Rn. 5.

[23] Wendt/Wendt (2024), Künstliche Intelligenz, 1. Auflage, § 3 Rn. 41; für eine graphische Darstellung vgl. etwa https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai (zuletzt aufgerufen am 14.01.2026).

[24] European Parliamentary Research Service (EPRS), Generative AI and watermarking, https://www.europarl.europa.eu/RegData/etudes/BRIE/2023/757583/EPRS_BRI(2023)757583_EN.pdf, S. 2 (zuletzt aufgerufen am 14.01.2026).

[25] Schwartmann/Keber/Zenner (2024), Praxisleitfaden KI-VO, 1. Auflage, Teil 2 Kap. 1 Rn. 469.

[26] Lauber-Rönsberg in BeckOK KI-Recht, Art. 50 Rn. 62.

[27] So wohl Martini in Martini/Wendehorst KI-VO, Art. 50 Rn. 102.

[28] So etwa Becker, CR 2024, 353 (361) Rn. 70; Lauber-Rönsberg in BeckOK KI-Recht, Art. 50 Rn. 62; Engel-Bunsas, RDi 2025, 292 (298) Rn. 37; Kumkar/Griesel, KIR 2024, 117 (120).

[29] Becker, CR 2024, 353 (361) Rn. 68.

[30] So iE auch Lauber-Rönsberg in BeckOK IT-Recht, Art. 50 Rn. 62; Engel-Bunsas, RDi 2025, 292 (298) Rn. 37.

[31] Lauber-Rönsberg in BeckOK IT-Recht, Art. 50 Rn. 62; Engel-Bunsas, RDi 2025, 292 (298f.) Rn. 37.

[32] Verordnung (EU) 2022/2065 des Europäischen Parlaments und des Rates vom 19. Oktober 2022 über einen Binnenmarkt für digitale Dienste.

[33] ErwG 120 S. 2 KI-VO.

[34] ErwG 120 S. 2 KI-VO.

der Allgemeinheit vor Desinformation ist.[35] Hierfür macht es keinen Unterschied, ob der generierte Inhalt einen exakten Anknüpfungspunkt in der Realität hat oder nicht. Angesichts des stetigen technologischen Fortschrittes werden *Deepfakes* zudem zunehmend realistischer. Es wird immer schwieriger, KI-generierte Inhalte von der Realität zu unterscheiden. Daraus folgt, dass eine weite Auslegung des Begriffs „*ähneln*“ geboten ist. Selbstredend kann dies zu Abgrenzungsschwierigkeiten führen. Mit Blick auf den Sinn und Zweck der Vorschrift, sind diese aber hinzunehmen, um Einzelfälle richtig zu erfassen. Ebenso hat das zwar zur Folge, dass nahezu alle KI-generierten Inhalte von Art. 3 Nr. 60 KI-VO erfasst und mithin der Transparenzpflicht nach Art. 50 Abs. 4 UAbs. 1 S. 1 KI-VO unterliegen. Eine so weite Auslegung des Schutzbereiches von Art. 50 Abs. 4 UAbs. 1 S. 1 KI-VO ist angesichts dessen gewichtigen Schutzzweckes gerechtfertigt. Im Übrigen sorgen Ausnahmen von der Transparenzpflicht (C.II.3) für eine angemessene Lösung im Einzelfall. Nur so kann ein umfassender und hinreichender Schutz der Gesellschaft erreicht werden.

Nach alledem genügt somit grundsätzlich eine potenzielle Realität. Richtigerweise bedarf aber auch dieses Ergebnis einer Einschränkung. Andernfalls droht der Anwendungsbereich des Art. 50 Abs. 4 UAbs. 1 S. 1 KI-VO auszuufern und sämtliche KI-generierte Inhalte zu erfassen. In Ergänzung zur potenziellen Realität ist deshalb zu verlangen, dass diese sich in den Grenzen des realistisch Erwartbaren hält. Als Maßstab dürfte sich die öffentliche Wahrnehmung unter Zugrundelegung eines durchschnittlichen Adressatenkreises anbieten.[36]

Mit Blick auf die Werbeindustrie könnte sich hier jedoch ein anderes ergeben. Hier wird der Inhalt in der Absicht generiert, das KI-Erzeugnis real erscheinen zu lassen, regelmäßig jedoch nicht mit einer Täuschungsabsicht bzgl. des Inhalts. Deshalb bergen sie geringere und andersgeartete Gefahren als (zB politisch motivierte) *Deepfakes*.[37] Indes vermag dies an der Auslegung des Begriffs „ähneln“ nichts zu ändern. Dieser muss einheitlich ausgelegt werden. Der vom Gesetzgeber angestrebte Schutz vor Desinformation überwiegt gegenüber dem Interesse an der Werbung mit *Deepfakes*. Insofern erübrigen sich auch Überlegungen über eine Bereichsausnahme für Werbezwecke durch teleologische Reduktion der Vorschrift. Hätte der Gesetzgeber das gewollt, hätte er eine solche Regelung positivrechtlich normiert. Aus Art. 50 Abs. 4 UAbs. 1 S. 3 KI-VO folgt, dass unter anderem auch künstlerische und satirische Inhalte erfasst sind. Daraus ergibt sich, dass der Gesetzgeber den Schutzbereich umfassend versteht, sodass für eine teleologische Reduktion bzw. eine Bereichsausnahme kein Platz besteht.

c) Eindruck der Echtheit oder Wahrheitsgemäßheit Schließlich muss der generierte *Output* „einer Person fälschlicherweise als echt oder wahrheitsgemäß erscheinen" (Art. 3 Nr. 60 KI-VO). Es muss jedenfalls die reale Möglichkeit einer Täuschung hinsichtlich der Faktizität des *Outputs* gegeben sein.[38] Angesichts des Normzweckes sind hieran niedrige Anforderungen zu stellen. Mithin genügt es, wenn ein durchschnittlich versierter Nutzer bei flüchtiger Betrachtung davon ausgehen durfte, dass es sich um einen authentischen Inhalt handelt.[39] Dafür spricht im Übrigen auch die Suggestivkraft von Bildern.[40] Daher muss es sich dem Adressaten geradezu aufdrängen, dass es sich um eine Fälschung handelt, damit der generierte Inhalt einer Person nicht als fälschlicherweise echt erscheint. Das tragende Argument ist auch hier der Schutz der Gesellschaft vor Desinformation und Täuschungen.

2. Normadressat Normadressat des Art. 50 Abs. 4 UAbs. 1 S. 1 KI-VO ist der „Betreiber" des genutzten KI-Systems. Gemäß Art. 3 Nr. 4 KI-VO ist das zunächst jede „natürliche oder juristische Person, Behörde, Einrichtung oder sonstige Stelle, die ein KI-System in eigener Verantwortung verwendet". Ausgenommen werden sodann private Verwendungen „im Rahmen einer persönlichen und nicht beruflichen Tätigkeit" (Art. 3 Nr. 4 KI-VO).

Mit Blick auf die Werbeindustrie wird sich regelmäßig folgende Rollenverteilung ergeben: Unternehmen A entwickelt ein KI-System, mit dem *Deepfakes* generiert werden können. Unternehmen B erwirbt das System von Unternehmen A (und implementiert es ggf. im eigenen IT-System). Unternehmen B stellt das KI-System nunmehr seinen Angestellten (zB der Marketing-Abteilung) zur Verfügung, damit diese Werbeinhalte erstellen oder modifizieren können. Letztere führen eine lediglich untergeordnete Funktion aus. Die übergeordnete Kontrolle der Rahmenbedingungen liegt bei Unternehmen B, das sich im Übrigen sämtliche Bedienungshandlungen zurechnen lassen muss.[41] Insofern sind nicht die einzelnen Angestellten, sondern Unternehmen B als „Betreiber“ im Sinne des Art. 3 Nr. 4 KI-VO zu klassifizieren. In dem Beispiel würde Unternehmen A die Rolle des „Anbieters“ gemäß Art. 3 Nr. 3 KI-VO einnehmen. Diese Rollenverteilung scheint der Regelfall zu sein, sodass die meisten Unternehmen in der Praxis als „Betreiber“ und mithin als Normadressat einzustufen sein werden.[42] Mit Blick auf die Rollenverteilung kann es in der Praxis zu Abgrenzungsschwierigkeiten kommen. Solche Abgrenzungsprobleme stellen sich etwa bei *Software-as-a-Service* (SaaS) oder *Embedded Software*, da hier übergeordnete Kontrolle und die konkrete Bedienungshandlung auseinanderfallen.[43]

Nach Art. 3 Nr. 4 Hs. 2 KI-VO gilt nicht als Betreiber, wer ein KI-System lediglich „*im Rahmen einer persönlichen und nicht beruflichen Tätigkeit verwendet*“. Spätestens mit der Veröffentlichung von *Deepfakes* gegenüber größeren Gruppen oder mittels öffentlich zugänglicher (*Social Media*-)Accounts dürfte dieser privilegierte Bereich indes verlassen sein.[44] Zudem muss der Betreiber grundsätzlich über eine Niederlassung innerhalb der EU verfügen (vgl.

[35] Engel-Bunsas, RDi 2025, 292 (299) Rn. 37.
[36] Vgl. Kumkar/Griesel, KIR 2024, 117 (120).
[37] Becker, CR 2024, 353 (361) Rn. 70.
[38] Martini in Martini/Wendehorst KI-VO, Art. 50 Rn. 103.
[39] Kumkar/Griesel, KIR 2024, 117 (120).
[40] Lauber-Rönsberg in BeckOK KI-Recht, Art. 50 Rn. 63.
[41] Kirschke-Biller/Füllsack in BeckOK KI-Recht, Art. 3 Rn. 125.
[42] Vgl. Remmertz (2025), Legal Tech-Strategien für die Rechtsanwaltschaft, 2. Auflage, § 2 Rn. 385.
[43] Kirschke-Biller/Füllsack in BeckOK KI-Recht, Art. 3 Rn. 126. Ebenso kann es zu Abgrenzungsschwierigkeiten zwischen Betreiber und (End-)Nutzern kommen. Vgl. hierzu weitergehend Schuh/Witt, EuDIR 2025, 142 (145–148) Rn. 13-23.
[44] Becker, CR 2024, 353, (361) Rn. 67.

Art. 2 Abs. 1 lit. b KI-VO) um Adressat zu sein.[45] Jedoch genügt es gemäß Art. 2 Abs. 1 lit. c KI-VO, dass der Betreiber seinen Sitz in einem Drittland hat, sofern der von dem KI-System generierte *Output* in der EU verwendet wird oder EU-Bürger hiervon betroffen sind. Hierdurch wird sichergestellt, dass auch solche Medieninhalte erfasst sind, die zwar außerhalb der Europäischen Union generiert wurden, jedoch potenziell ihren disruptiven Effekt innerhalb der EU entfalten könnten.[46]

3. Ausnahmen

a) Straftaten, Art. 50 Abs. 4 UAbs. 1 S. 2 KI-VO Die Transparenzpflicht entfällt gemäß Art. 50 Abs. 4 UAbs. 1 S. 2 KI-VO, sofern und soweit die Verwendung des *Deepfakes* zur „*Aufdeckung, Verhütung, Ermittlung oder Verfolgung von Straftaten*" gesetzlich zugelassen ist. Normiert wird damit eine Öffnungsklausel sowohl für den europäischen als auch den nationalen Gesetzgeber.

b) Erleichterung des Art. 50 Abs. 4 UAbs. 1 S. 3 KI-VO Keine Ausnahme im eigentlichen Sinne, dennoch eine Erleichterung sieht Art. 50 Abs. 4 UAbs. 1 S. 3 KI-VO für offensichtlich künstlerische, kreative, satirische, fiktionale oder analoge Werke oder Programme vor. Hier beschränkt sich die Kennzeichnungspflicht darauf, den generierten Inhalt in einer Weise offenzulegen, die die Darstellung oder den Genuss des Werks nicht beeinträchtigt. Hiervon sollen aber nur solche Werke erfasst sein, bei denen von Anfang an ausgeschlossen ist, dass sie real sein könnten.[47] Mit dieser Regelung möchte der Gesetzgeber das Spannungsverhältnis zur Kunstfreiheit (Art. 13 Abs. 1 Var. 1 GRCh) auflösen.[48] Bereits aus diesem Grund und unter Berücksichtigung allgemeiner juristischer Grundsätze sind Art. 50 Abs. 4 UAbs. 1 S. 3 KI-VO sowie die dort verwendeten Begriffe im Lichte von Art. 13 Abs. 1 Var. 1 GRCh auszulegen.[49]

c) (Analoge) Anwendung des Art. 50 Abs. 4 UAbs. 2 S. 2 Alt. 2 KI-VO In Bezug auf KI-generierte Texte besteht nach Art. 50 Abs. 4 UAbs. 2 S. 2 Alt. 2 KI-VO die Möglichkeit, sich durch die Übernahme redaktioneller Verantwortung von der Kennzeichnungspflicht zu befreien. Für *Deepfakes* ist eine vergleichbare Privilegierung nicht vorgesehen. Da insbesondere Unternehmen ein erhebliches praktisches Interesse daran haben dürften, die Entstehungsumstände KI-generierter Bild-, Ton- oder Videoinhalte nicht offenlegen zu müssen, stellt sich die Frage, ob die Ausnahme des Art. 50 Abs. 4 UAbs. 2 S. 2 Alt. 2 KI-VO auch auf *Deepfakes* Anwendung finden kann. Das hätte den Vorteil, dass sich Unternehmen durch Etablierung eines internen Kontrollprozesses und Übernahme redaktioneller Verantwortung von der Kennzeichnungspflicht des Art. 50 Abs. 4 UAbs. 1 S.1 KI-VO „befreien" könnten. Im Ergebnis scheidet jedoch sowohl eine direkte als auch eine analoge Anwendung aus.

Eine direkte Anwendung scheitert bereits am eindeutigen Wortlaut des Art. 50 Abs. 4 UAbs. 2 S. 2 KI-VO. Dieser spricht von „*Diese Pflicht gilt nicht*", was lediglich Bezug auf die in S. 1 statuierte Kennzeichnungspflicht hinsichtlich KI-erzeugter Textinhalte nimmt. Ebenso ist es nicht fernliegend, dass Unternehmen in der Praxis diesen Weg gehen würden, um die Entstehungsumstände des Werbeinhaltes nicht offenlegen zu müssen. Dadurch würde die Kennzeichnungspflicht des Art. 50 Abs. 4 UAbs. 1 S. 1 KI-VO de facto ins Leere laufen. Somit stehen einer direkten Anwendung auch teleologische Gesichtspunkte entgegen.

Ebenso besteht kein Raum für eine analoge Anwendung des Art. 50 Abs. 4 UAbs. 2 S. 2 Alt. 2 KI-VO auf *Deepfakes*. Hierzu bedürfte es dem Vorliegen einer Regelungslücke bei vergleichbarer Interessenlage zwischen dem geregelten und dem nicht geregelten Fall.[50] Daran fehlt es hier. Richtig ist zunächst, dass eine vergleichbare Regelung in Art. 50 Abs. 4 UAbs. 1 KI-VO nicht normiert ist. Mit Blick auf die Systematik und die klare Trennung der Kennzeichnungspflichten der beiden Unterabsätze des Art. 50 Abs. 4 KI-VO lässt sich jedoch der Umkehrschluss ziehen, dass der Gesetzgeber für *Deepfakes* eine solche Privilegierung gerade nicht treffen wollte. Auch mit Blick auf den teleologischen Regelungsgehalt ergibt sich kein anderes Ergebnis. Ziel der Kennzeichnungspflicht des Art. 50 Abs. 4 UAbs. 2 S. 1 KI-VO ist gerade nicht, offenzulegen, dass sich der Autor bei der Erstellung des Inhalts einer KI bedient hat.[51] Vielmehr wollte der Unionsgesetzgeber die Gefahr massenhafter ungeprüfter KI-generierter Falschinformationen ohne menschliche Verantwortlichkeit minimieren.[52] Insofern ist davon auszugehen, dass es sich hierbei um eine beabsichtigte Nichtanwendbarkeit der Ausnahme für Texte im Rahmen des Art. 50 Abs. 4 UAbs. 1 KI-VO handelt. Es liegt allenfalls eine Lücke de lege ferenda vor, zu deren Schließung der Rechtsanwender nicht befugt ist.[53]

4. Rechtsfolge: Kennzeichnungspflicht Die Rechtsfolge ist eine Verpflichtung des KI-Systembetreibers „*offenzulegen, dass die Inhalte künstlich erzeugt oder manipuliert wurden*" (Art. 50 Abs. 4 UAbs. 1 S. 1 KI-VO).[54] Wie genau die Offenlegung zu erfolgen hat, wird nicht näher konkretisiert. In Art. 50 Abs. 5 S. 1 KI-VO ist lediglich festgelegt, dass die Offenlegung in klarer und eindeutiger Weise zu erfolgen hat. Im Übrigen muss sie auch den geltenden Barrierefreiheitsanforderungen entsprechen (Art. 50 Abs. 5 S. 2 KI-VO). Die Offenlegung hat gegenüber dem Rezipienten des generierten Medieninhaltes zu erfolgen.[55] Hingegen muss nicht offengelegt werden, wer hinter dem *Deepfake* im Sinne einer Verantwortlichkeit steht.[56] Gemäß Art. 50 Abs. 5 S. 1 KI-VO muss die Offenlegung spätestens zum Zeitpunkt der ersten Interaktion mit dem generierten Inhalt erfolgen. Mit Blick auf Art. 50 Abs. 4 UAbs. 1

[45] ErwG 21 KI-VO.

[46] Kumkar/Griesel, KIR 2024, 117 (121).

[47] Engel-Bunsas, RDi 2025, 292 (300) Rn. 44; Martini in Martini/Wendehorst KI-VO, Art. 50 Rn. 110.

[48] ErwG 134 S. 2f. KI-VO; Martini in Martini/Wendehorst KI-VO, Art. 50 Rn. 110; Lauber-Rönsberg in BeckOK KI-Recht, Art. 50 Rn. 72; Engel-Bunsas, RDi 2025, 292 (300) Rn. 44; Rößling, ZfDR 2024, 187 (188).

[49] Vgl. etwa Lenaerts, EuR 2012, 3 (3).

[50] EuGH, Rs. 165/84, ECLI:EU:C:1985:507, Rn. 14 - Krohn; Ahmling (2012), Analogiebildung durch den EuGH im Europäischen Privatrecht, 1. Auflage, S. 147.

[51] Martini in Martini/Wendehorst KI-VO, Art. 50 Rn. 114.

[52] Martini in Martini/Wendehorst KI-VO, Art. 50 Rn. 114.

[53] Vgl. Engisch (2024), Juristisches Denken, 13. Auflage, S. 203f.; Maus, ZfPW 2023, 25 (29f.).

[54] Vgl. insoweit auch ErwG 134 S. 1 KI-VO.

[55] Engel-Bunsas, RDi 2025, 292 (299).

[56] Block, EuCML 2024, 184 (189); Hinderks, ZUM 2022, 110 (113).

S. 3 KI-VO kann sich hier für künstlerische, kreative, satirische, fiktionale oder analoge Inhalte etwas anderes ergeben. Unter Umständen kann in einem solchen Fall eine Offenlegung erst im Nachgang erfolgen.[57] Gleichwohl müssen geeignete Schutzvorkehrungen für die Rechte und Freiheiten der Rezipienten bestehen.[58]

Gemäß Art. 50 Abs. 7 S. 1 KI-VO wird das Büro für Künstliche Intelligenz (*AI-Office*) Praxisleitfäden ausarbeiten, um eine wirksame Umsetzung der Pflicht in Bezug auf die Feststellung und Kennzeichnung von *Deepfakes* zu erleichtern. So soll eine unionsweit einheitliche Auslegung und Anwendung der Transparenzvorschriften des Art. 50 KI-VO sichergestellt werden, die dieser allein mangels Detailtiefe nicht leisten kann.[59] Zur Genehmigung solcher Verhaltenskodizes kann die Kommission nach dem in Art. 98 Abs. 2 KI-VO niedergelegten Verfahren Durchführungsrechtsakte erlassen (Art. 56 Abs. 6 KI-VO).[60] Daneben können weitere Transparenzpflichten unionaler oder nationaler Vorschriften treten, die gemäß Art. 50 Abs. 6 KI-VO sodann kumulativ zu erfüllen sind.

5. Sanktionsmöglichkeiten Gemäß Art. 99 Abs. 4 lit. g KI-VO sieht der Verordnungsgeber für einen Verstoß gegen die Transparenzpflicht ein Bußgeld in Höhe von bis zu 15.000.000 Euro vor. Handelt es sich bei dem Normadressaten um ein Unternehmen, so kann das Bußgeld sogar bis zu 3 % des weltweiten Jahresumsatzes des vorangegangenen Geschäftsjahres betragen, je nachdem welcher Betrag höher ist. Privilegiert werden hierbei kleine und mittlere Unternehmen (KMU), einschließlich Start-up-Unternehmen, indem der jeweils niedrigere Betrag gilt (Art. 99 Abs. 6 KI-VO).

III. Kritische Analyse des Art. 50 Abs. 4 UAbs. 1 S. 1 KI-VO

1. Art und Weise der Kennzeichnung Konkretisierende Vorgaben seitens des *AI-Office* sind noch nicht erfolgt. Gerade eine solche konkretisierende Regelung wäre allerdings mit Blick auf die Praxis wünschenswert. Daraus ergeben sich Folgefragen betreffend die konkrete Art der Offenlegung, nutzbare und erforderliche (technische) Möglichkeiten zur Kennzeichnung, sowie des Zeitpunkts der Kennzeichnung.

a) „Offenlegen": Ausdrücklich oder konkludent Zunächst stellt sich die Frage, was unter dem Begriff „*offenlegen*" im Sinne des Art. 50 Abs. 4 UAbs. 1 S. 1 KI-VO zu verstehen ist, also ob sich hieraus eine Pflicht zur ausdrücklichen Kennzeichnung[61] ergibt, oder ob unter Umständen auch eine konkludente Offenlegung möglich ist. Letzteres wäre insbesondere dann der Fall, wenn aus einer Gesamtwürdigung sämtlicher Begleitumstände auf die Eigenschaft als *Deepfake* geschlossen werden kann. Richtigerweise bedarf es einer ausdrücklichen Kennzeichnung.

Gegen die Möglichkeit einer konkludenten Kennzeichnung spricht zunächst der Wortlaut des ErwG 134 KI-VO. Dieser spricht zwar auch nur von „*kennzeichnen*". In der englischen Version wird hier allerdings von „*labelling*" gesprochen. Das legt nahe, dass dem unionalen Gesetzgeber im Gesetzgebungsverfahren eine ausdrückliche Kennzeichnung vorschwebte.[62] Hierfür sprechen im Übrigen auch teleologische Gesichtspunkte. Durch die Kennzeichnungspflicht sollen auch und insbesondere unerfahrene Rezipienten geschützt werden.[63] Letztlich würde eine konkludente Kennzeichnung zu erheblichen Abgrenzungsschwierigkeiten in der Praxis führen. Für jeden Einzelfall müsste gesondert festgestellt werden, ob sich die Eigenschaft als *Deepfake* bereits aus dem Medieninhalt selbst ergibt. Das würde dem Schutzzweck der Norm zuwiderlaufen.

b) (Technische) Möglichkeit der Kennzeichnung Ebenso enthält die KI-VO ihrerseits keinerlei Anforderungen an die technisch notwendige, aber auch erforderliche Art der Kennzeichnung. Aus ErwG 134 KI-VO („*klar und deutlich*") lässt sich einzig ableiten, dass aus der Kennzeichnung hervorgehen muss, dass der Inhalt künstlich erzeugt oder manipuliert worden ist.[64] Betreibern stehen grundsätzlich diverse Möglichkeiten zur Verfügung. Hierbei muss zunächst zwischen visuellen (Bild- und Videoinhalten) und auditiven Inhalten (Toninhalte) differenziert werden.

Betreffend Bild- und Videoinhalte bieten sich grundsätzlich zwei verschiedene Kennzeichnungsmöglichkeiten an. Einerseits könnte ein gut lesbarer Hinweis bei- oder vorangestellt werden, der auf den KI-generierten Ursprung hinweist (zB: „*Nachfolgender Werbespot wurde mittels KI erstellt.*").[65] Mit Blick auf Art. 50 Abs. 5 S. 1 KI-VO und obige Ausführungen muss der Hinweis spätestens mit Beginn eines Werbespots oder ähnlichem erfolgen. Demnach wäre eine Kennzeichnung im Nachgang grundsätzlich nicht mehr ausreichend.[66] Alternativ könnte man auf das sog. „watermarking" zurückgreifen, bei welchem der Hinweis in nicht zu übersehbarer Weise direkt in den visuellen Inhalt implementiert wird.[67] Dieses darf nicht mit dem Watermarking im Sinne des Art. 50 Abs. 2 S. 1 KI-VO verwechselt werden.[68] Im Gegensatz dazu muss der Hinweis hier für das menschliche Auge gerade erkennbar sein.[69] Letzteres dürfte wohl praktikabler sein, kann es doch unmittelbar durch das generierende Programm eingefügt werden. Diese Möglichkeit stößt stellenweise auf Kritik, könne es doch einfach wieder herausretuschiert werden.[70] Das vermag indes nicht zu überzeugen. Es

[57] Martiny, ZUM 2025, 200 (211); Kumkar/Griesel, KIR 2024, 117 (122).

[58] ErwG 134 S. 2 KI-VO; Martiny, ZUM 2025, 200 (211).

[59] Wendt/Wendt (2024), Künstliche Intelligenz, 1. Auflage, § 9 Rn. 17.

[60] Schwartmann/Keber/Zenner (2024), Praxisleitfaden KI-VO, 1. Auflage, Teil 2 Kap. 1 Rn. 452; Wendt/Wendt (2024), Künstliche Intelligenz, 1. Auflage, § 9 Rn. 17.

[61] So etwa Kumkar/Rapp, ZfDR 2022, 199 (226); Martini/Wendehorst KI-VO, Art. 50 Rn. 108; Lauber-Rönsberg in BeckOK KI-VO, Art. 50 Rn. 70; Kumkar/Griesel, KIR 2024, 117 (121); Merkle, RDi 2024, 414 (420) Rn. 45; Hinderks, ZUM 2022, 110 (113).

[62] Kumkar/Griesel, KIR 2024, 117 (121).

[63] Lauber-Rönsberg in BeckOK KI-VO, Art. 50 Rn. 70; Martini in Martini/Wendehorst KI-VO, Art. 50 Rn. 108.

[64] Hinderks, ZUM 2022, 110 (113).

[65] Kumkar/Rapp, ZfDR 2022, 199 (226).

[66] Martiny, ZUM 2025, 200 (211); Kumkar/Griesel, KIR 2024, 117 (122).

[67] Lauber-Rönsberg in BeckOK KI-Recht, Art. 50 Rn. 70; Kumkar/Rapp, ZfDR 2022, 199 (226); Hinderks, ZUM 2022, 110 (112).

[68] Hierzu vertiefend Vasse'i, RDi 2024, 406.

[69] ErwG 134 KI-VO.

[70] Ua Valta/Vasel, ZRP 2021, 142 (142).

kann hier keinen Unterschied machen, welcher Kennzeichnungsmethode der Pflichtige sich bedient. Ein etwaiger Vorspann eines Videoinhaltes kann ebenso entfernt werden. In solchen Fällen greift das Sanktionsregime der KI-VO.

Hinsichtlich auditiver Elemente dürfte es jedenfalls erforderlich sein, dass der Hinweis in zeitlicher Nähe zum Toninhalt angebracht wird.[71] Bei längeren Toninhalten kann es zudem erforderlich sein, die Kennzeichnung in regelmäßigen Intervallen zu wiederholen.[72]

Insgesamt erscheinen sämtliche vorgenannten Möglichkeiten gangbar. In Ermangelung einer spezifizierenden positivrechtlichen Regelung und eines Praxisleitfadens seitens des *AI-Office* kommt den Betreibern – jedenfalls bis zur Schaffung eines solchen – ein Einschätzungsspielraum zu.[73] Im Rahmen dieser Einschätzung sind sämtliche Einzelfallumstände zu berücksichtigen. Insbesondere muss der (potenziellen) Adressatenkreis und dessen Medienaffinität besondere Berücksichtigung finden.[74] Je schutzwürdiger die Adressatengruppe (zB Kinder oder ältere Personen), desto leichter verständlich und sichtbar sind die Hinweise zu gestalten. Das Regelungsproblem wurde in der Praxis erkannt. Seither gibt es privatrechtliche Bemühungen, gewisse Standards hinsichtlich der Kennzeichnung zu etablieren. Nennenswert ist insoweit die *„Coalition for Content Provenance and Authentity“*.[75] Im Rahmen ihres *„Conformance Program“* wird mittels sog. *„Content Credentials“* dargelegt, ob und in welchem Umfang KI zum Einsatz kam.[76]

c) Ausblick Zwar tritt die Transparenzpflicht erst am 2. August 2026 in Kraft (Art. 113 Abs. 2 KI-VO), dennoch wären frühzeitige Leitlinien seitens des *AI-Office* für die Praxis wünschenswert. Bisher besteht erhebliche Rechtsunsicherheit dahingehend, in welcher Art und Weise, in welcher Größe und mit welchem Text die Offenlegung zu erfolgen hat. Das ist misslich, denn Art und Weise der Offenlegung sind entscheidend für die Effektivität der Transparenzpflicht.[77] Gegenwärtig kommt jedoch Bewegung in diese Thematik. Die EU hat am 17. Dezember 2025 einen ersten Entwurf eines *„Code of Practice on Transparency of AI-Generated Content“* veröffentlicht.[78] Dieser sieht im Wesentlichen die Schaffung eines interaktiven EU-weiten Icons vor (vgl. etwa S. 24). Das *Icon* soll es ermöglichen, Täuschungsgrade zu unterscheiden, zusätzliche Informationen bei Interaktionen bereitzustellen und passende Platzierungen je nach Inhaltsformat zu wählen (S. 24f.). Die Veröffentlichung eines finalen *Code of Practice* ist für Mai–Juni 2026 geplant. Diese Lösungsmöglichkeit ist erfreulich. Gleichwohl bleibt der weitere Prozess und die näheren Konkretisierungen abzuwarten. Nur so kann die aktuell bestehende Rechtsunsicherheit beseitigt werden und sich das volle „friedliche“ Nutzungspotenzial von *Deepfakes* entfalten.

2. Reichweite Neben der praktischen Ausgestaltung der Kennzeichnungspflicht ist auch die (eingeschränkte) Reichweite des Art. 50 Abs. 4 UAbs. 1 S. 1 KI-VO nicht optimal gelungen. Dieser unterwirft nämlich ausschließlich den Betreiber des KI-Systems der Kennzeichnungspflicht.

a) Persönliche oder nicht berufliche Tätigkeit Zunächst sind aufgrund des Betreiberbegriffs des Art. 3 Nr. 4 Hs. 2 KI-VO vom Anwendungsbereich solche Personen ausgenommen, die im Rahmen einer persönlichen und nicht beruflichen Tätigkeit handeln. Diese Einschränkung ergibt sich im Übrigen bereits unmittelbar aus der Definition des Anwendungsbereichs der KI-VO (Art. 2 Abs. 10 KI-VO). Ausgenommen sind somit Verbraucher als klassische „Privatpersonen“. Diese Grundsatzentscheidung ist im Allgemeinen zu begrüßen. Mit Blick auf die fehlende Transparenzpflicht, ist das jedoch mehr als bedenklich. Es ist zu erwarten, dass (gerade) auch Privatpersonen von den destruktiven Einsatzmöglichkeiten Gebrauch machen werden und somit eine Vielzahl im Internet kursierender *Deepfakes* droht, nicht erfasst zu sein.[79] Deshalb ist die Ausnahme des Anwendungsbereiches restriktiv auszulegen.[80] Die handelnde natürliche Person ist bereits dann als Betreiber zu qualifizieren, wenn sie die eigene Privatsphäre nur teilweise überschreitet.[81] Das dürfte jedenfalls dann der Fall sein, wenn der *Deepfake* in größeren Gruppen oder offenen Accounts verbreitet wird.[82] Hierdurch entstehen zwar im Einzelfall Abgrenzungsschwierigkeiten. Angesichts der Alternative, dass ein Großteil der im Internet befindlichen *Deepfakes* keiner Transparenzpflicht unterliegen würde, ist dies aber hinzunehmen. Nur so sorgt der Ansatz des Gesetzgebers für Einzelfallgerechtigkeit.

b) Ausweitung auf Anbieter Unter Effektivitätsgesichtspunkten wäre es sinnvoll gewesen, auch den Anbieter des KI-Systems der Kennzeichnungspflicht zu unterwerfen. Er unterliegt ohnehin der Transparenzpflicht nach Art. 50 Abs. 2 S. 1 KI-VO und muss sicherstellen, dass die Ausgabe seines KI-Systems in einem maschinenlesbaren Format und als künstlich erzeugt oder manipuliert gekennzeichnet ist. Diese Norm umfasst ebenso *Deepfakes* und komplementiert das regulatorische Programm.[83] Insoweit dürfte es nur einen untergeordneten Aufwand darstellen, die Kennzeichnung auf diejenige auszuweiten, die im Sinne des Art. 50 Abs. 4 UAbs. 1 S. 1 KI-VO erforderlich wäre. Vorteil davon wäre, dass der Anbieter die technische Möglichkeit hätte, den *Deepfake* als solchen *„by Design“* zu kennzeichnen.[84] Das würde die Durchsetzung der Norm insgesamt erleichtern. Zwar kann das Wasserzeichen in der Folgezeit herausretuschiert werden.[85] Jedoch würde

[71] Engel-Bunsas, RDi 2025, 292 (299); Martini in Martini/Wendehorst KI-VO, Art. 50 Rn. 108.

[72] Lauber-Rönsberg in BeckOK KI-Recht, Art. 50 Rn. 70.

[73] Merkle, RDi 2024, 414; Kumkar/Griesel, KIR 2024, 117 (121).

[74] Vgl. Kumkar/Griesel, KIR 2024, 117 (121f.).

[75] https://c2pa.org/ (zuletzt aufgerufen am 26.04.2026).

[76] https://contentcredentials.org/ (zuletzt aufgerufen am 26.04.2026).

[77] Kumkar/Rapp, ZfDR 2022, 199 (224).

[78] Online abrufbar unter https://digital-strategy.ec.europa.eu/en/library/first-draft-code-practice-transparency-ai-generated-content (zuletzt aufgerufen am 26.04.2026).

[79] Blocher, KIR 2025, 225 (226); Martini in Martini/Wendehorst KI-VO, Art. 50 Rn. 105.

[80] Hinderks, ZUM 2022, 110 (111); Lauber-Rönsberg in BeckOK KI-Recht, Art. 50 Rn. 67.

[81] Kumkar/Griesel, KIR 2024, 117 (121); Martini in Martini/Wendehorst KI-VO, Art. 50 Rn. 105.

[82] Becker, CR 2024, 353 (361) Rn. 67; Kirschke-Biller/Füllsack in BeckOK KI-Recht, Art. 3 Rn. 123.1.

[83] Martini in Martini/Wendehorst KI-VO, Art. 50 Rn. 107.

[84] Kumkar/Griesel, KIR 2024, 117 (120); Hinderks, ZUM 2022, 110 (112).

[85] Valta/Vasel, ZRP 2021, 142 (142).

so das Risiko unwissender und mithin fälschlicher Nichtkennzeichnung minimiert.[86] Ebenso würde eine böswillige Nichtkennzeichnung technisch erschwert.[87]

3. Durchsetzung Verbesserungswürdig ist zudem, dass die hinter dem *Deepfake* stehende (natürliche oder juristische) Person nicht offenbart werden muss. Das Problem ist zweischichtig. Einerseits wird die Verhängung etwaiger Sanktionen gem. Art. 99 Abs. 4 lit. g KI-VO erschwert. Andererseits entstehen erhebliche Schwierigkeiten hinsichtlich der Durchsetzung etwaiger Ansprüche der Rezipienten oder derjenigen, deren Persönlichkeitsrechte durch den generierten Inhalt verletzt wurden.[88] Unterstellt man, es würden die materiell-rechtlichen Anspruchsvoraussetzungen vorliegen, bliebe in der Praxis vielfach unklar, gegen wen diese zu richten sind.

Hiergegen könnte eine Erweiterung der Kennzeichnungspflicht hinsichtlich der Identität des Betreibers Abhilfe schaffen, dass kumulativ die Eigenschaft des Inhalts als *Deepfake* als auch dessen Identität offenbart werden muss. Dieses Problem wurde im Rahmen des Gesetzgebungsverfahrens gesehen. In einer Stellungnahme schlug das *Committee on Legal Affairs* (JURI) vor, Art. 50 Abs. 4 UAbs. 1 S. 1 KI-VO um einen weiteren Halbsatz zu ergänzen, wonach auch der Name der dahinterstehenden Entität offengelegt werden sollte.[89] Bedauerlicherweise konnte sich dieser Vorschlag nicht durchsetzen. Zwar wird der Betreiber in Bezug auf etwaige Rechtsverletzungen natürlicher Personen nach den Art. 12 ff. der DSGVO dazu verpflichtet sein, seine Identität zu offenbaren.[90] Die Vorschriften der DSGVO treten neben diejenigen der KI-VO und bleiben hinsichtlich ihrer Anwendbarkeit insoweit unberührt (Art. 2 Abs. 7 S. 1 KI-VO, ErwG 10 S. 4 KI-VO). Gleichwohl bleibt der Betroffene auf die Erfüllung der Informationspflicht durch den Betreiber angewiesen. Mit Blick auf die Praxis wird sich häufig das Problem stellen, dass dessen Identität unbekannt und der Anspruch insoweit ebenso wenig durchsetzbar sein wird. Zudem schafft dies lediglich hinsichtlich etwaiger Persönlichkeitsrechtsverletzungen Abhilfe und bietet gerade aus demokratietheoretischer Perspektive keinerlei Mehrwert.[91] Der Unionsgesetzgeber sollte also die Norm auf Grundlage des Vorschlags des *Committee on Legal Affairs* nachbessern. Nur so lässt sich ein effektiver Schutz der betroffenen Rezipienten erreichen.

4. Verbot von Deepfakes Angesichts des erheblichen „friedlichen" Nutzungspotenzial von *Deepfakes*, ist es begrüßenswert, dass sich der Gesetzgeber gegen ein vollständiges Verbot von *Deepfakes* entschieden hat. Ein solches hätte wohl ohnehin einen unverhältnismäßigen Eingriff in die Kunst- und Kommunikationsfreiheiten (Art. 11 Abs. 1, Art. 13 Abs. 1 Var. 1 GRCh) dargestellt und wäre daher nicht realisierbar gewesen.[92]

5. Deepfakes als Hochrisiko-KI-System Der europäische Gesetzgeber verfolgt hinsichtlich *Deepfakes* einen singulären Regulierungsansatz, als er diese nur einer Transparenzpflicht unterwirft. Gegenwärtig sind sie nicht als Hochrisiko-KI-System eingestuft. Die einzige Ausnahme hierbei sind solche Systeme, deren Zweck in der Manipulation von Wahlen liegt (Art. 6 Abs. 2 iVm Anhang III Nr. 8 lit. b S. 1 KI-VO). Es ist jedenfalls diskussionswürdig, ob eine solche Einstufung nicht generell hätte erfolgen sollen. Bereits im Gesetzgebungsverfahren sorgte diese Frage für erhebliche Auseinandersetzungen. Während dies einerseits unter Verweis auf das destruktive Nutzungspotenzial auf Zustimmung stieß,[93] begrüßen andere das gesetzgeberische Ergebnis.[94] Letzteres verdient Zustimmung.

Zunächst könnte für eine Einstufung als Hochrisiko-KI sprechen, dass mit den Vorschriften betreffend Hochrisiko-KI (Art. 6–49 KI-VO) eine differenzierte Regulierung von *Deepfakes* möglich wäre.[95] Dagegen spricht indes die unterschiedliche Schutzrichtung. Hochrisiko-KI-Systeme bergen erhebliche Gefahren für die Rechte des Einzelnen, namentlich unter anderem, da sie diskriminierende Entscheidungen treffen können und diese für den durchschnittlichen Nutzer nicht durchschaubar sind.[96] Zwar dient die Regulierung von *Deepfakes* auch dem Schutz des Einzelnen, jedoch hinsichtlich des Missbrauchspotenzials und den daraus resultierenden Risiken insbesondere dem Schutz der Gesellschaft.[97] Hinzu kommt, dass eine Vielzahl der Normen die Rezipienten zwar theoretisch umfassender schützen könnten, in der Praxis jedoch viele Vorschriften faktisch ins Leere liefen und den Prozess verkomplizieren würden.[98] Auch würde das drängende Problem der mangelnden Kenntnis von der Identität des Betreibers hierdurch nicht behoben. Durch die Einstufung als Hochrisiko-KI müsste der Anbieter das KI-System gemäß Art. 49 Abs. 1 KI-VO registrieren lassen. Zudem müsste der Anbieter gem. Art. 18 Abs. 1 KI-VO die dort genannten Dokumente aufbewahren. Beides zeigt, dass eine Einstufung als Hochrisiko-KI für das Individuum keinen wirklichen Mehrwert bringen würde.

Im Übrigen ist die Kommission gemäß Art. 7 Abs. 1 KI-VO ermächtigt, den hier maßgeblichen Anhang III zu ändern. Somit wäre jederzeit eine Einstufung als Hochrisiko-KI möglich, sollte in der Folgezeit hierfür ein Bedürfnis entstehen. Einzig der Verwaltungsaufwand für

[86] Hinderks, ZUM 2022, 110 (112).

[87] Hinderks, ZUM 2022, 110 (112); Martini in Martini/Wendehorst KI-VO, Art. 50 Rn. 106.

[88] Block, EuCML 2024, 184 (199); Lauber-Rönsberg in BeckOK KI-Recht, Art. 50 Rn. 58; Kumkar/Rapp, ZfDR 2022, 199 (211f.).

[89] Committee on Legal Affairs, Opinion AI Act, https://artificialintelligenceact.eu/wp-content/uploads/2022/09/AIA-JURI-Rule-57-Opinion-Adopted-12-September.pdf, S. 45 f. (zuletzt aufgerufen am 26.04.2026).

[90] Vertiefend hierzu: Hinderks, ZUM 2022, 110 (116f.).

[91] Hinderks, ZUM 2022, 110 (116).

[92] Rostalski/Weiss, ZfDR 2021, 329 (352f.); Lauber-Rönsberg in BeckOK KI-Recht, Art. 50 Rn. 56; Kumkar/Rapp, ZfDR 2022, 199 (225); für ein Verbot etwa Linardatos, LTO, https://www.lto.de/recht/hintergruende/h/deepfakes-regulierung-europa-eu-schaden-demokratie-manipulation (zuletzt aufgerufen am 26.04.2026).

[93] Etwa Rostalski/Weiss, ZfDR 2021, 329 (352); Hinderks, ZUM 2022, 110 (117f.); Ebert/Spiecker gen. Döhmann, NVwZ 2021, 1188 (1192).

[94] Etwa Kumkar/Griesel, KIR 2024, 117 (126); Block, EuCML 2024, 184 (191).

[95] Hinderks, ZUM 2022, 110 (117); Valta/Vasel, ZRP 2021, 142 (142f.).

[96] ErwG 64 S. 1, 73 S. 1 KI-VO.

[97] Kraetzig, CR 2024, 207 (208) Rn. 3; Kumkar/Rapp, ZfDR 2022, 199 (201); Kumkar/Griesel, KIR 2024, 117 (118); Hinderks, ZUM 2022, 110 (113).

[98] Vgl. auch: Block, EuCML 2024, 184 (191).

den jeweiligen Anbieter würde steigen. Im Sinne einer Innovationsförderung ist dies unangemessen.

6. Rückwirkung Die Transparenzpflicht des Art. 50 Abs. 4 UAbs. 1 S. 1 KI-VO wird am 02. August in Kraft treten (Art. 113 Abs. 2 KI-VO). Klar ist, sie wird alle ab diesem Zeitpunkt generierten *Deepfakes* der Offenlegung unterwerfen. Es drängt sich deshalb die Frage auf, ob und inwieweit *Deepfakes* erfasst werden, die vor diesem Zeitpunkt generiert wurden, mithin ob Art. 50 Abs. 4 UAbs. 1 S. 1 KI-VO auch eine rückwirkende Wirkung entfaltet. Das hätte zur Folge, dass sämtliche *Deepfakes* nachträglich als solche gekennzeichnet werden müssten. Eine positivrechtliche Klärung dieser Frage oder ein klarstellender Leitfaden seitens des *AI-Office* liegen nicht vor. Aus Gründen der Rechtssicherheit wäre dies wünschenswert gewesen. Im Ergebnis ist eine Rückwirkung des Art. 50 Abs. 4 UAbs. 1 S. 1 KI-VO abzulehnen. *Deepfakes* unterfallen demnach erst ab dem 02. August 2026 einer Kennzeichnungspflicht.

Zunächst ist festzuhalten, dass der Wortlaut des Art. 50 Abs. 4 UAbs. 1 S. 1 KI-VO hier weder in die eine noch in die andere Richtung ausgelegt werden kann und insoweit offen formuliert ist. Insbesondere unterscheidet er nicht zwischen solchen *Deepfakes*, die vor oder nach dem 02. August 2026 erstellt wurden. Für eine Rückwirkung sprechen teleologische Gesichtspunkte. Der Rezipient des generierten Inhaltes kann nicht erkennen, wann ein Medium veröffentlicht wurde. Gleichwohl erscheint er ebenso schutzwürdig.

Dem begegnen mit Blick auf grundlegende Rechtsprinzipien erhebliche Bedenken. Sollten auch vor dem 02. August 2026 generierte *Deepfakes* erfasst sein, handelt es sich de facto um einen Fall der (sog. echten) Rückwirkung. Für solche *Deepfakes*, die vor Verabschiedung der KI-VO generiert wurden, war eine etwaige Kennzeichnungspflicht nicht erkennbar. Ebenso wie im deutschen Recht besteht auch auf europäischer Ebene ein grundsätzliches Rückwirkungsverbot.[99] Mitnichten gilt dieses absolut.[100] Erforderlich ist jedenfalls, dass sich die Rückwirkung (mittelbar) aus der Verordnung selbst ergibt. Ein Blick auf Art. 111 KI-VO zeigt, dass die KI-VO ihrerseits sehr wohl Vorschriften kennt, die eine etwaige Rückwirkung regeln. Gemäß Art. 111 Abs. 2 S. 1 KI-VO sind Betreiber eines KI-Systems im Falle wesentlicher Veränderungen hieran verpflichtet, die Vorschriften der KI-VO auch dann einzuhalten, wenn das KI-System vor dem 02. August 2026 in den Verkehr gebracht oder in Betrieb genommen wurde. Nach Art. 111 Abs. 2 S. 2 KI-VO gelten für Anbieter und Betreiber von Hochrisiko-KI-Systemen, die bestimmungsgemäß von Behörden verwendet werden, die Anforderungen der KI-VO unbeschadet des Inkrafttretens dieser. Im Gegenzug wird eine Umsetzungsfrist von vier Jahren gewährt. Letztlich regelt Art. 111 Abs. 3 KI-VO, dass Anbieter von GPAI-Modellen die erforderlichen Maßnahmen treffen müssen, damit auch vor Inkrafttreten der KI-VO in den Verkehr gebrachte Modelle die Anforderungen der KI-VO erfüllen. Daraus lässt sich im Sinne eines argumentum e contrario schließen, dass der Gesetzgeber eine Rückwirkung positivrechtlich normiert hätte, sofern er eine solche gewollt hätte.

Gleichwohl kann es gerade für Unternehmen empfehlenswert sein, bereits vorher eine Kennzeichnung eigener synthetisch erstellter Werbung vorzunehmen. So könnten Unternehmen potenziellen Kunden proaktiv suggerieren, dass man die Risiken Künstlicher Intelligenz ernst nimmt und menschliche Werte schätzt (*Ethical AI*).

D. Fazit und Überlegungen de lege ferenda

Das „friedliche“ Nutzungspotenzial von *Deepfakes* ist enorm. Gleichwohl bergen sie erhebliche Gefahren sowohl für die Gesellschaft als auch für das Individuum. Der gesetzgeberische Ansatz, dem Gefahrenpotenzial von *Deepfakes* mittels einer Kennzeichnungspflicht zu begegnen, ist begrüßenswert. Bereits bestehende Normen treten ergänzend hinzu und komplementieren den Regulierungsansatz. Es besteht erheblicher Nachbesserungsbedarf. Wünschenswert wäre, dass auch die Identität des Betreibers des zur Erstellung eines *Deepfakes* verwendeten KI-Systems von der Transparenzpflicht umfasst wird. Andernfalls ergeben sich erhebliche Durchsetzungsschwierigkeiten. Zudem ist bis dato unklar, wie die Kennzeichnung zu erfolgen hat. Hier bedarf es einer baldigen Klarstellung durch das *AI-Office*. Letztlich wäre es besser, den Anbieter als Normadressat des Art. 50 Abs. 4 UAbs. 1 S. 1 KI-VO einzubeziehen, um der Norm im Wege einer *„by Design“*-Lösung maximale Durchsetzungskraft zu verschaffen. Es bleibt abzuwarten, ob der Unionsgesetzgeber hier nachjustiert. Aktuelle politische Spannungen lassen die Regulierungen zunehmend sanfter ausfallen, um die (technologische) Innovation in das eigene Land zu locken. Eine solche Entwicklung verkennt jedoch das erhebliche Gefahrenpotenzial, das *Deepfakes* entfalten können. Nur eine rechtssichere, lückenlos vollziehbare Kennzeichnungspflicht vermag den Schutz vor Desinformation und weiteren Risiken effektiv zu gewährleisten. Zugleich kann auf diese Weise das erhebliche wirtschaftliche und gesellschaftliche Nutzungspotenzial von *Deepfakes* rechtlich abgesichert werden.

[99] Schmahl in Schulze/Janssen/Kadelbach (2020), Europarecht, 4. Auflage, § 6 Rn. 44; Günther, EuZW 2000, 329 (329).

[100] EuGH, Rs. C-414/04, ECLI:EU:C:2006:742, Rn. 52; EuGH, Rs. C-4/10, ECLI:EU:C:2011:484, Rn. 25.