Amtsblatt
der Europäischen Union

DE
Reihe C

C/2026/6005

12.11.2026

MITTEILUNG DER KOMMISSION

Leitlinien für die Umsetzung der Transparenzpflichten für bestimmte KI-Systeme gemäß Artikel 50 der Verordnung (EU) 2024/1689 (KI-Verordnung)

(C/2026/6005)

Inhaltsverzeichnis

Seite

1. HINTERGRUND UND ZIELE ... 3
2. ÜBERBLICK ÜBER DIE TRANSPARENZPFLICHTEN UND HORIZONTALEN THEMEN ... 3
   - 2.1. Überblick über die Transparenzpflichten nach Artikel 50 der KI-Verordnung ... 3
   - 2.2. Begründung und Ziele ... 4
   - 2.3. Zuständige Akteure nach Artikel 50 der KI-Verordnung ... 5
   - 2.4. Ausschlüsse vom Anwendungsbereich der KI-Verordnung ... 7
     - 2.4.1. Betreiber, die im Rahmen einer ausschließlich persönlichen und nicht beruflichen Tätigkeit handeln ... 7
     - 2.4.2. Forschung & Entwicklung ... 8
     - 2.4.3. KI, die im Rahmen freier und quelloffener Lizenzen freigegeben wird ... 8
   - 2.5. Zusammenspiel mit verbotenen Praktiken, Anforderungen an Hochrisiko-KI-Systeme und KI-Kompetenz ... 9
   - 2.6. Zusammenspiel mit Vorschriften für KI-Modelle und/oder -Systeme mit allgemeinem Verwendungszweck ... 9
3. ARTIKEL 50 ABSATZ 1 DER KI-VERORDNUNG: TRANSPARENZ FÜR INTERAKTIVE KI-SYSTEME ... 9
   - 3.1. Hauptkomponenten, Konzepte und damit verbundene Transparenzpflicht(en) ... 10
     - 3.1.1. KI-Systeme, die für die direkte Interaktion mit natürlichen Personen bestimmt sind ... 10
     - 3.1.2. Informationspflicht nach Artikel 50 Absatz 1 der KI-Verordnung ... 12
   - 3.2. Ausnahmen von der Informationspflicht nach Artikel 50 Absatz 1 der KI-Verordnung ... 14
     - 3.2.1. Ausnahme für offensichtliche Interaktion mit einem KI-System ... 14
     - 3.2.2. Ausnahme für gesetzlich zu Strafverfolgungszwecken zugelassene KI-Systeme ... 16
   - 3.3. Zusammenspiel mit anderen Rechtsakten der Union ... 17
4. ARTIKEL 50 ABSATZ 2 DER KI-VERORDNUNG: KENNZEICHNUNG UND FESTSTELLUNG VON DURCH KI ERZEUGTEN ODER MANIPULIERTEN INHALTEN ... 18
   - 4.1. Hauptkomponenten und Konzepte von Artikel 50 Absatz 2 der KI-Verordnung ... 18
     - 4.1.1. KI-Systeme, die synthetische Inhalte erzeugen oder manipulieren ... 18
     - 4.1.2. Modalitäten synthetischer Inhalte im Anwendungsbereich ... 19
     - 4.1.3. Inhalte, die nicht in den Anwendungsbereich von Artikel 50 Absatz 2 der KI-Verordnung fallen. 20
   - 4.2. Die Kennzeichnungs- und Feststellungspflicht nach Artikel 50 Absatz 2 der KI-Verordnung ... 21
     - 4.2.1. Die Kennzeichnungspflicht ... 21
     - 4.2.2. Die Feststellungspflicht ... 22
     - 4.2.3. Einhaltung der Anforderungen an technische Lösungen: wirksam, interoperabel, belastbar und zuverlässig ... 22
   - 4.3. Ausnahmen von den Pflichten nach Artikel 50 Absatz 2 der KI-Verordnung ... 24
   - 4.4. Zusammenspiel mit anderen Rechtsakten der Union ... 25

5. ARTIKEL 50 ABSATZ 3 DER KI-VERORDNUNG: EMOTIONSERKENNUNGSSYSTEME UND SYSTEME ZUR BIOMETRISCHEN KATEGORISIERUNG 26

5.1. Hauptkomponenten, Konzepte und damit verbundene Transparenzpflichten nach Artikel 50 Absatz 3 der KI-Verordnung 27

5.1.1. Der Begriff „Emotionserkennungssystem“ 27

5.1.2. Der Begriff „System zur biometrischen Kategorisierung“ 27

5.1.3. Informationspflicht nach Artikel 50 Absatz 3 der KI-Verordnung 27

5.2. Ausnahme für Strafverfolgungszwecke 28

5.3. Zusammenspiel mit anderen Rechtsakten der Union 28

6. ARTIKEL 50 ABSATZ 4: KENNZEICHNUNG VON DEEPFAKES UND BESTIMMTEN TEXTVERÖFFENTLICHUNGEN 28

6.1. Hauptkomponenten, Konzepte und damit verbundene Transparenzpflichten für Deepfakes nach Artikel 50 Absatz 4 der KI-Verordnung 28

6.1.1. Der Begriff „Deepfake“ 29

6.1.2. Die Offenlegungspflicht nach Artikel 50 Absatz 4 Unterabsatz 1 der KI-Verordnung 31

6.1.3. Transparenz von künstlerischen, kreativen, satirischen, fiktionalen oder analogen Deepfake-Inhalten 32

6.1.4. Ausnahme für Strafverfolgung 34

6.1.5. Zusammenspiel mit anderen Rechtsakten der Union 34

6.2. Hauptkomponenten, Konzepte und damit verbundene Transparenzpflichten für durch KI erzeugten oder manipulierten Text nach Artikel 50 Absatz 4 der KI-Verordnung 35

6.2.1. Text, der veröffentlicht wird, um die Öffentlichkeit über Angelegenheiten von öffentlichem Interesse zu informieren 36

6.2.2. Die Offenlegungspflicht nach Artikel 50 Absatz 4 Unterabsatz 2 der KI-Verordnung 37

6.2.3. Ausnahme von der Transparenzpflicht für Text unter menschlicher Überprüfung oder redaktioneller Kontrolle und redaktioneller Verantwortung 37

6.2.4. Ausnahme für Strafverfolgung 38

6.2.5. Zusammenspiel mit anderen Rechtsakten der Union 39

7. HORIZONTALE ANFORDERUNGEN FÜR DIE NACH ARTIKEL 50 ABSATZ 5 DER KI-VERORDNUNG BEREITGESTELLTEN INFORMATIONEN 39

7.1. Klare und eindeutige Informationsbereitstellung 39

7.2. Erste Interaktion oder Aussetzung 40

7.3. Einhaltung von geltenden Barrierefreiheitsanforderungen 40

8. DURCHSETZUNG VON ARTIKEL 50 DER KI-VERORDNUNG 40

8.1. Auswirkungen der Befolgung eines als angemessen bewerteten Praxisleitfadens 41

8.2. Marktüberwachungsbehörden 42

8.3. Sanktionen 42

8.4. Anwendungsbeginn 42

9. ÜBERPRÜFUNG UND AKTUALISIERUNG DER LEITLINIEN 43

1. HINTERGRUND UND ZIELE

1. Die Verordnung (EU) 2024/1689 des Europäischen Parlaments und des Rates vom 13. Juni 2024 zur Festlegung harmonisierter Vorschriften für künstliche Intelligenz und zur Änderung bestimmter Verordnungen (im Folgenden „KI-Verordnung“) trat am 1. August 2024 in Kraft ([1]). Die KI-Verordnung enthält harmonisierte Vorschriften für das Inverkehrbringen, die Inbetriebnahme und die Verwendung künstlicher Intelligenz (im Folgenden „KI“) in der Union ([2]). Ihr Ziel besteht darin, Innovationen und die Einführung vertrauenswürdiger KI zu fördern und gleichzeitig ein hohes Schutzniveau in der Union in Bezug auf Gesundheit, Sicherheit und Grundrechte, einschließlich Demokratie und Rechtsstaatlichkeit, zu unterstützen.

2. Die KI-Verordnung verfolgt einen risikobasierten Ansatz und unterteilt KI-Systeme in vier verschiedene Risikokategorien, von denen eine KI-Systeme sind, die Transparenzrisiken bergen und den in Artikel 50 der KI-Verordnung festgelegten Pflichten unterliegen. Diese Transparenzpflichten gelten zwei Jahre nach Inkrafttreten der KI-Verordnung, d. h. ab dem 2. August 2026.

3. Gemäß Artikel 96 Absatz 1 Buchstabe d der KI-Verordnung werden diese Leitlinien der Kommission als praktische Leitlinien herausgegeben, um die zuständigen Behörden sowie die Anbieter und Betreiber von KI-Systemen dabei zu unterstützen, die Einhaltung der Transparenzpflichten nach Artikel 50 der KI-Verordnung in kohärenter, wirksamer, verhältnismäßiger und einheitlicher Weise sicherzustellen.

4. Die Ausarbeitung dieser Leitlinien stützte sich auf Beiträge einer Vielzahl von Interessenträgern, die im Rahmen einer von der Kommission organisierten breit angelegten Konsultation eingeholt wurden, sowie auf Beiträge der Mitgliedstaaten im KI-Gremium und auf Beiträge der Interessenträger zu einem Entwurf der Leitlinien, der zur Konsultation veröffentlicht wurde.

5. Diese Leitlinien sind nicht verbindlich. Eine verbindliche Auslegung der KI-Verordnung kann letztlich nur vom Gerichtshof der Europäischen Union (EuGH) vorgenommen werden.

2. ÜBERBLICK ÜBER DIE TRANSPARENZPFLICHTEN UND HORIZONTALEN THEMEN

2.1. Überblick über die Transparenzpflichten nach Artikel 50 der KI-Verordnung

6. Artikel 50 der KI-Verordnung enthält vier Transparenzpflichten, die jeweils für verschiedene Arten von KI-Systemen oder deren Ausgaben gelten.

Bestimmung	Art des KI-Systems/ Ausgabe	Transparenzpflicht	Ausnahmen oder Sonderregelungen
Artikel 50 Absatz 1	KI-Systeme, die direkt mit natürlichen Personen interagieren	Die Anbieter müssen das KI-System so entwickeln und konzipieren, dass die betreffenden natürlichen Personen informiert werden, dass sie mit einem KI-System interagieren.	Diese Pflicht gilt nicht, wenn 1) offensichtlich ist, dass sie mit einem KI-System interagieren, oder 2) das System gesetzlich zur Aufdeckung, Verhütung, Ermittlung oder Verfolgung von Straftaten zugelassen ist, es sei denn, das System steht der Öffentlichkeit zur Anzeige einer Straftat zur Verfügung.

([1]) Verordnung (EU) 2024/1689 des Europäischen Parlaments und des Rates vom 13. Juni 2024 zur Festlegung harmonisierter Vorschriften für künstliche Intelligenz (Verordnung über künstliche Intelligenz) (ABl. L, 2024/1689, 12.7.2024, ELI: http://data.europa.eu/eli/reg/2024/1689/oj).

([2]) Artikel 1 der KI-Verordnung.

Bestimmung	Art des KI-Systems/ Ausgabe	Transparenzpflicht	Ausnahmen oder Sonderregelungen
Artikel 50 Absatz 2	KI-Systeme, die synthetische Bild-, Video-, Audio- oder Textinhalte erzeugen oder manipulieren	Die Anbieter müssen sicherstellen, dass die Ausgaben des KI-Systems in einem maschinenlesbaren Format gekennzeichnet und als künstlich erzeugt oder manipuliert erkennbar sind. Dazu setzen sie technische Lösungen ein, die wirksam, interoperabel, belastbar und zuverlässig sind.	Diese Pflicht gilt nicht, wenn 1) das KI-System eine unterstützende Funktion für die Standardbearbeitung ausführt oder die Eingabedaten oder deren Semantik nicht wesentlich verändert oder 2) das KI-System zur Aufdeckung, Verhütung, Ermittlung oder Verfolgung von Straftaten gesetzlich zugelassen ist.
Artikel 50 Absatz 3	KI-Systeme zur Emotionserkennung oder biometrischen Kategorisierung	Die Betreiber müssen die von dem System betroffenen natürlichen Personen über den Betrieb des KI-Systems informieren.	Diese Pflicht gilt nicht, wenn das KI-System gesetzlich zur Aufdeckung, Verhütung oder Ermittlung von Straftaten zugelassen ist.
Artikel 50 Absatz 4	KI-Systeme, die Deepfakes oder Texte, die veröffentlicht werden, um die Öffentlichkeit über Angelegenheiten von öffentlichem Interesse zu informieren, erzeugen oder manipulieren	Die Betreiber müssen offenlegen, dass die Inhalte künstlich erzeugt oder manipuliert wurden.	Diese Pflicht gilt nicht, wenn 1) das KI-System zur Aufdeckung, Verhütung, Ermittlung oder Verfolgung von Straftaten gesetzlich zugelassen ist oder 2) die Textveröffentlichung einer menschlichen Überprüfung oder redaktionellen Kontrolle unterzogen wurde und der redaktionellen Verantwortung unterliegt. Besondere Offenlegungsregelungen gelten für Deepfakes, die Teil künstlerischer, kreativer, fiktionaler, satirischer oder analoger Werke oder Programme sind.

7. Jede dieser vier Pflichten wird in den Abschnitten 3 bis 6 dieser Leitlinien untersucht. Zusätzlich sieht Artikel 50 Absatz 5 der KI-Verordnung vor, dass die in den Absätzen 1 bis 4 genannten Informationen den betreffenden natürlichen Personen spätestens zum Zeitpunkt der ersten Interaktion oder Aussetzung in klarer und eindeutiger Weise bereitgestellt werden und den geltenden Barrierefreiheitsanforderungen entsprechen müssen. In Abschnitt 7 werden diese Anforderungen präzisiert.

8. Die verschiedenen Transparenzpflichten nach Artikel 50 der KI-Verordnung können kumulativ für ein einzelnes KI-System (bzw. dessen Ausgabe) gelten, wobei möglicherweise verschiedene Akteure (Anbieter oder Betreiber) verantwortlich sind. Wenn beispielsweise ein KI-System im Rahmen einer direkten Interaktion mit einer natürlichen Person Bilder erzeugt, muss der Anbieter des Systems möglicherweise sowohl Artikel 50 Absatz 1 als auch Artikel 50 Absatz 2 der KI-Verordnung einhalten. Wenn dieses System auch in der Lage ist, Bilder, bei denen es sich um Deepfakes handelt, oder Texte, die zu Angelegenheiten von öffentlichem Interesse veröffentlicht wurden, zu erzeugen, muss der Betreiber möglicherweise zusätzlich Artikel 50 Absatz 4 der KI-Verordnung einhalten.

2.2. **Begründung und Ziele**

9. Wie in den Erwägungsgründen 132 bis 136 der KI-Verordnung erläutert, besteht der Zweck der Transparenzpflichten darin, das Risiko von Identitätsbetrug, Täuschung, Fehlinformation, Manipulation in großem Maßstab und Betrug zu verringern und potenzielle negative Auswirkungen auf demokratische Prozesse und das gesellschaftliche Vertrauen, die von durch KI erzeugten oder manipulierten Inhalten oder Interaktionen verursacht werden, zu mindern. Die Information der Menschen über die KI-Herkunft der Interaktion und der Inhalte wird ihnen dabei helfen, fundierte Entscheidungen zu treffen, und gleichzeitig dazu beitragen, das Vertrauen in das Informationsökosystem und dessen Integrität zu wahren. Transparenz ist nicht nur eine Anforderung an die technische Konformität, sondern auch eine Voraussetzung dafür, dass natürliche Personen KI-gestützte Interaktionen und Inhalte oder die Verwendung von Emotionserkennungssystemen oder Systemen zur biometrischen Kategorisierung erkennen können. Die Gewährleistung einer transparenten und vertrauenswürdigen Nutzung interaktiver und generativer KI-Systeme wird es auch ermöglichen, die persönliche Autonomie und die europäischen Werte, die Menschenwürde und andere Grundrechte (z. B. Gedankenfreiheit, Privatsphäre, Gleichheit und Nichtdiskriminierung, Verbraucherschutz, geistiges Eigentum usw.) zu wahren und gleichzeitig die Meinungsfreiheit und die Glaubwürdigkeit von Medieninhalten aufrechtzuerhalten.

2.3. **Zuständige Akteure nach Artikel 50 der KI-Verordnung**

10. Nach Artikel 3 Nummer 3 der KI-Verordnung sind **Anbieter** natürliche oder juristische Personen, Behörden, Einrichtungen oder sonstige Stellen, die ein KI-System entwickeln oder entwickeln lassen und unter ihrem eigenen Namen oder ihrer Handelsmarke in der Union in Verkehr bringen bzw. unter ihrem eigenen Namen oder ihrer Handelsmarke in Betrieb nehmen, sei es entgeltlich oder unentgeltlich. Anbieter müssen die einschlägigen Transparenzpflichten nach Artikel 50 Absätze 1 und 2 der KI-Verordnung einhalten, die für sie gelten, unabhängig davon, ob sie in der Union oder in einem Drittland niedergelassen sind ([3]). Anbieter, die ihren Sitz außerhalb der Union haben oder sich außerhalb der Union befinden, unterliegen auch der KI-Verordnung, wenn die Ausgabe ihres KI-Systems in der Union verwendet wird ([4]). Unbeschadet der Befugnisse der zuständigen Marktüberwachungsbehörden, ihre Ausgaben als nicht mit der KI-Verordnung vereinbar zu betrachten, sollte jedoch eine zufällige, unvorhersehbare oder nicht genehmigte nachgelagerte Verwendung allein nicht die Anwendung der Pflichten auf solche Anbieter interaktiver oder generativer KI-Systeme aus Drittländern auslösen, die nicht in der Union in Verkehr gebracht oder in Betrieb genommen werden.

11. Die Anbieter müssen sicherstellen, dass ihre interaktiven oder generativen KI-Systeme durch Transparenz gekennzeichnet sind, die in die Konzeption und den Betrieb des Systems eingebettet ist, und sicherstellen, dass alle einschlägigen Transparenzpflichten gemäß Artikel 50 Absätze 1, 2 und 5 der KI-Verordnung spätestens beim Inverkehrbringen oder bei der Inbetriebnahme dieser Systeme erfüllt werden. Alle natürlichen Personen, die mit dem KI-System interagieren oder dessen in den Anwendungsbereich fallenden Ausgaben ausgesetzt sind, müssen während des gesamten Lebenszyklus des KI-Systems, wie es von diesem Anbieter in Verkehr gebracht oder in Betrieb genommen wird, wirksam und konsequent informiert werden (unabhängig davon, wer die Betreiber oder anderen Nutzer des Systems sind).

> Beispielsweise bietet ein Unternehmen Verbrauchern, beruflichen Nutzern oder anderen juristischen Personen, die das KI-System direkt für die Interaktion oder die Erzeugung von Inhalten nutzen können, eine generative oder interaktive KI-Anwendung (z. B. einen Chatbot, einen Bildgenerator oder einen KI-Agenten) auf dem Unionsmarkt unter seinem eigenen Namen oder seiner eigenen Marke an (auch durch Integration des Systems in ihre bestehenden Prozesse und seine entsprechende Verwendung ohne Änderungen oder zusätzliche Komponenten). Das Unternehmen ist ein Anbieter, der für die Einhaltung der Transparenzpflichten nach Artikel 50 Absatz 1 und/oder 2 der KI-Verordnung verantwortlich ist, unabhängig davon, ob das KI-System unentgeltlich oder entgeltlich bereitgestellt wird, und unabhängig davon, ob der Anbieter in der Union oder in einem Drittland niedergelassen ist.
>
> Ein weiteres Beispiel für einen Anbieter ist ein Unternehmen oder eine andere Organisation (z. B. eine Behörde), das bzw. die intern ein interaktives KI-System (z. B. einen Chatbot) entwickelt hat und es in der Union für den Eigengebrauch und unter seinem bzw. ihrem Namen oder seiner bzw. ihrer Marke in Betrieb nimmt.
>
> Wenn ein Unternehmen ein bereits bestehendes generatives KI-System nimmt, das von einem anderen Anbieter in Verkehr gebracht wurde, dieses System ändert (z. B. mit neuen Trainingsdaten) und es anschließend unter seinem eigenen Namen oder seiner Handelsmarke in Betrieb nimmt, wird dieses Unternehmen zum Anbieter des neuen Systems, unbeschadet der Verantwortung des Anbieters des ursprünglichen KI-Systems für Letzteres.

12. Betreiber sind natürliche oder juristische Personen, Behörden, Einrichtungen oder sonstige Stellen, die KI-Systeme in eigener Verantwortung verwenden, es sei denn, die Verwendung erfolgt im Rahmen einer persönlichen und nicht beruflichen Tätigkeit ([5]). Die „Verantwortung" für ein KI-System sollte so verstanden werden, dass die Verantwortung für die Entscheidung über den Einsatz des Systems und die Art und Weise der tatsächlichen Nutzung des Systems (einschließlich seiner Ausgaben) übernommen wird. Dies erfordert nicht unbedingt eine technische Kontrolle des Betriebs des KI-Systems, solange der Betreiber entscheidet, für welche Zwecke und wie das KI-System verwendet wird (auch in dezentralen Arbeitsabläufen und gruppeninternen Strukturen). Betreiber, die an komplexen Wertschöpfungsketten für die Produktion und Verbreitung von Inhalten beteiligt sind, sollten verhältnismäßige Maßnahmen ergreifen, um sicherzustellen, dass die Kennzeichnung der Inhalte, die sie gemäß Artikel 50 Absatz 4 der KI-Verordnung umgesetzt haben, für die Zielgruppe und die vorhersehbare Zielgruppe zum Zeitpunkt der ersten Aussetzung gemäß Artikel 50 Absatz 5 der KI-Verordnung in klarer und eindeutiger Weise angezeigt wird (z. B. durch Vertragsbedingungen mit den Vertriebspartnern, Einstellungen für die Nutzererfahrung (UX) und anzuzeigende Schnittstellen).

([3]) Artikel 2 Absatz 1 Buchstabe a, Artikel 3 Nummern 3, 9 und 11 der KI-Verordnung. Weitere Erläuterungen zu den Konzepten des Inverkehrbringens, der Inbetriebnahme und der Verwendung finden sich in den Leitlinien der Kommission zu verbotenen Praktiken der künstlichen Intelligenz, C(2025) 5052, und in der Bekanntmachung der Kommission – Leitfaden für die Umsetzung der Produktvorschriften der EU 2022, Abschnitt 2 (ABl. C 247 vom 29.6.2022, S. 1).

([4]) Artikel 2 Absatz 1 Buchstabe c und Erwägungsgrund 22 der KI-Verordnung.

([5]) Artikel 3 Nummer 4 der KI-Verordnung.

13. Betreiber fallen in den Anwendungsbereich der Transparenzpflichten nach Artikel 50 Absätze 3 und 4 der KI-Verordnung, wenn sie ihren Sitz in der Union haben oder sich in der Union befinden oder wenn sie ihren Sitz in einem Drittland haben oder sich in einem Drittland befinden, wenn die Ausgabe des KI-Systems in der Union verwendet wird ([6]). Im letzteren Fall gelten die Transparenzpflichten für Einrichtungen, die ihren Sitz außerhalb der Union haben oder sich außerhalb der Union befinden, wenn der Betreiber selbst die Verbreitung und Nutzung der KI-Ausgaben in der Union plant (d. h. durch Steuerung oder Genehmigung des Vertriebs innerhalb der Union, einschließlich durch das Einstellen von Deepfakes in das weltweit zugängliche Internet). Betreiber aus Drittländern sind jedoch nicht an die Transparenzpflichten gebunden, wenn die Inhalte des KI-Systems ein Publikum in der Union über Kanäle erreichen, die unvorhersehbar sind und sich ihrer Kontrolle entziehen.

14. Handelt es sich bei dem Betreiber eines KI-Systems um eine juristische Person, unter deren Verantwortung das System verwendet wird (z. B. ein Werbeunternehmen), sind die einzelnen Beschäftigten, die auf Anweisung und unter der Kontrolle dieser juristischen Person handeln (z. B. digitale Animatoren, Webdesigner, Ersteller von Inhalten, Journalisten), nicht als separate Betreiber für dieses System zu betrachten. Eine juristische Person bleibt auch dann Betreiberin, wenn sie in ihrem Namen und unter ihrer Verantwortung und Kontrolle Dritte (z. B. Auftragnehmer, Freiberufler) am Betrieb des Systems beteiligt.

> So ist beispielsweise ein in der Union niedergelassenes Medienunternehmen, dessen Redakteure ein KI-System verwenden, um ihre schriftliche Berichterstattung über aktuelle Ereignisse mit durch KI erzeugten Textveröffentlichungen zu Angelegenheiten von öffentlichem Interesse zu unterstützen, die auf einer weltweit zugänglichen Website veröffentlicht werden, ein Betreiber, der in den Anwendungsbereich der KI-Verordnung fällt. Ebenso ist ein in einem Drittland niedergelassenes Unternehmen, das ein KI-System verwendet, um ein Deepfake eines Prominenten zu erzeugen, das in einer in der Union gezeigten Werbung verwendet wird, auch ein Betreiber, der in den Anwendungsbereich der KI-Verordnung fällt.
>
> Dagegen ist ein Unternehmen, das sich darauf beschränkt, eine Werbeagentur mit der Produktion einer Werbung zu beauftragen, ohne Entscheidungen zu treffen und Kontrolle darüber auszuüben, ob und wie die Werbeagentur KI im Produktionsprozess einsetzt, kein Betreiber.

15. Akteure (z. B. Anbieter und Betreiber) können in Bezug auf ein KI-System mehr als eine Rolle gleichzeitig erfüllen. Beispielsweise könnte ein Unternehmen ein internes generatives KI-System entwickeln und es für die Erzeugung von Deepfakes verwenden. In einem solchen Fall gilt es sowohl als Anbieter als auch als Betreiber, der die Einhaltung der maschinenlesbaren Kennzeichnungspflichten gemäß Artikel 50 Absatz 2 und der Anforderungen an eine klare und wahrnehmbare Kennzeichnung der Deepfakes gemäß Artikel 50 Absatz 4 der KI-Verordnung sicherstellen muss.

16. Andere Akteure (z. B. Hosting-Diensteanbieter, einschließlich Online-Plattformen, oder Fernsehveranstalter), deren Rolle sich auf die Verbreitung oder Übermittlung von durch KI erzeugten oder manipulierten Inhalten beschränkt, die von Dritten erstellt wurden, oder die durch KI erzeugte oder manipulierte Inhalte empfangen oder ihnen ausgesetzt sind, ohne die Verfügungsgewalt über die Nutzung des KI-Systems zum Zwecke der Erzeugung oder Manipulation von KI-Inhalten zu haben, sind hingegen keine Betreiber im Sinne der KI-Verordnung. Ohne nach der KI-Verordnung verantwortlich und haftbar zu sein, können diese Akteure dennoch eine wichtige Rolle in der Wertschöpfungskette spielen, um das Vertrauen in das Informationsökosystem und seine Integrität zu erhalten. Sie werden daher nachdrücklich aufgefordert, die gemäß Artikel 50 der KI-Verordnung eingeführte Markierung und Kennzeichnung beizubehalten und verhältnismäßige und technisch durchführbare Maßnahmen zu ergreifen, damit natürliche Personen die durch KI erzeugten oder manipulierten Inhalte erkennen können, unbeschadet der Verantwortung der Anbieter und Betreiber generativer KI-Systeme oder etwaiger Verpflichtungen oder Haftungen gemäß der Verordnung (EU) 2022/2065 (Gesetz über digitale Dienste) ([7]) oder anderen geltenden Rechtsvorschriften der Union oder der Mitgliedstaaten.

([6]) Artikel 2 Absatz 1 Buchstaben b und c der KI-Verordnung.

([7]) Verordnung (EU) 2022/2065 des Europäischen Parlaments und des Rates vom 19. Oktober 2022 über einen Binnenmarkt für digitale Dienste (ABl. L 277 vom 27.10.2022, S. 1).

17. Andere Akteure, die in ihrer beruflichen Eigenschaft handeln und Inhalte verbreiten (auch wenn sie nicht direkt gemäß Artikel 50 der KI-Verordnung beteiligt sind), werden ebenfalls aufgefordert, verhältnismäßige und geeignete Maßnahmen zu ergreifen, damit die natürlichen Personen, die den Inhalten letztlich ausgesetzt sind, wirksam informiert werden und in der Lage sind, die künstlich erzeugte oder manipulierte Herkunft zu erkennen. Darüber hinaus können diese Parteien verpflichtet werden, die jeweiligen Deepfakes oder Textveröffentlichungen zu Angelegenheiten von öffentlichem Interesse nach anderen geltenden rechtlichen oder berufsethischen Vorschriften zu kennzeichnen (z. B. ein Fernsehveranstalter, der ein von einer Privatperson erstelltes Deepfake im Rahmen eines Fernsehprogramms kennzeichnet und ausstrahlt). Die bloße Tatsache, dass ein Akteur Inhalte kennzeichnet oder gekennzeichnete Inhalte verwendet, bedeutet jedoch nicht, dass Letzterer als „Betreiber" im Sinne der KI-Verordnung einzustufen ist. Die KI-Verordnung erweitert oder berührt nicht die rechtliche Verantwortung dieser Akteure nach den geltenden Rechtsrahmen, wenn sie nicht gleichzeitig als Anbieter oder Betreiber von KI-Systemen gelten.

2.4. **Ausschlüsse vom Anwendungsbereich der KI-Verordnung**

18. Artikel 2 der KI-Verordnung enthält eine Reihe allgemeiner Ausschlüsse vom Anwendungsbereich der KI-Verordnung, die bereits in den Leitlinien der Kommission zu verbotenen Praktiken der künstlichen Intelligenz präzisiert wurden ([8]). In diesem Abschnitt werden daher nur die Aspekte behandelt, die für die Transparenzpflichten nach Artikel 50 der KI-Verordnung relevant sind.

2.4.1. *Betreiber, die im Rahmen einer ausschließlich persönlichen und nicht beruflichen Tätigkeit handeln*

19. Gemäß Artikel 2 Absatz 10 der KI-Verordnung gilt die Verordnung nicht für die Pflichten von Betreibern, die natürliche Personen sind und die KI-Systeme im Rahmen einer ausschließlich persönlichen und nicht beruflichen Tätigkeit verwenden ([9]). Jede Tätigkeit, durch die natürliche Personen regelmäßig einen wirtschaftlichen Vorteil erlangen oder anderweitig einer professionellen, geschäftlichen, gewerblichen, beruflichen oder freiberuflichen Tätigkeit nachgehen, sollte als „Berufstätigkeit" betrachtet werden. Der Begriff „ausschließlich persönlich" ist eine Spezifizierung des Begriffs „nicht beruflich" und bedeutet, dass die Person sowohl persönlich als auch nicht beruflich handeln muss, um in den Anwendungsbereich der Ausnahme zu fallen. Jede Nutzung durch natürliche Personen, die im Namen oder unter der Verantwortung eines im Rahmen seiner beruflichen Tätigkeit handelnden Betreibers handeln, fällt in den Anwendungsbereich der Transparenzpflichten des professionellen Betreibers gemäß Artikel 50 Absätze 3 und 4 der KI-Verordnung.

Beispiele für die persönliche und nicht berufliche Nutzung von KI-Systemen, die aus dem Anwendungsbereich ausgenommen sind:

— Einzelpersonen, die ein KI-System nutzen, um Deepfakes von Mitgliedern ihres Haushalts zu erzeugen

— Einzelpersonen, die ein KI-System persönlich nutzen, um Texte zu verschiedenen Themen zu erzeugen (z. B. Studierende, die KI nutzen, um ihre Hausaufgaben zu schreiben)

— Einzelpersonen, die ein KI-System persönlich nutzen, um Deepfakes zu erzeugen, die in den sozialen Medien verbreitet werden

20. Der Ausschluss nach Artikel 2 Absatz 10 der KI-Verordnung gilt nur in Bezug auf die Pflichten der Betreiber bei der Verwendung des Systems für ausschließlich persönliche und nicht berufliche Tätigkeiten. Das System als solches fällt weiterhin in den Anwendungsbereich der KI-Verordnung in Bezug auf die Pflichten der Anbieter, die das System in Verkehr bringen oder in Betrieb nehmen, und der Betreiber, die das System für berufliche Zwecke nutzen. Die Ausnahmeregelung berührt auch nicht die Anwendung anderer einschlägiger Rechtsvorschriften der Union und der Mitgliedstaaten.

([8]) Siehe Abschnitt 2.5 der Leitlinien der Kommission zu verbotenen Praktiken der künstlichen Intelligenz gemäß der Verordnung (EU) 2024/1689 (KI-Verordnung), Brüssel, 29.7.2025, C(2025) 5052 final.

([9]) Die Definition des Begriffs „Betreiber" schließt ebenfalls Nutzer aus, die Tätigkeiten dieser Art ausüben (siehe Abschnitt 2.3).

So sollte beispielsweise der Anbieter des KI-Systems, das zur Erzeugung von Deepfakes verwendet wird, die auf einer persönlichen Weihnachtskarte erscheinen sollen, diese Deepfakes im Einklang mit der Transparenzpflicht nach Artikel 50 Absatz 2 der KI-Verordnung in einem maschinenlesbaren Format kennzeichnen. Gleichzeitig fällt die Nutzung des Systems durch natürliche Personen für ausschließlich persönliche und nicht berufliche Tätigkeiten nicht unter die Pflicht des Betreibers gemäß Artikel 50 Absatz 4 der KI-Verordnung.

2.4.2. *Forschung & Entwicklung*

21. In Artikel 2 Absatz 6 der KI-Verordnung ist ein Ausschluss für KI-Systeme oder KI-Modelle, einschließlich ihrer Ausgabe, die eigens für den alleinigen Zweck der wissenschaftlichen Forschung und Entwicklung entwickelt und in Betrieb genommen werden, vorgesehen. Dieser Ausschluss gilt unbeschadet geltender ethischer Forschungsstandards sowie der EU-Datenschutzvorschriften und anderer Rechtsvorschriften. Im Zusammenhang mit Artikel 50 der KI-Verordnung gilt dieser Ausschluss nicht nur für interaktive KI-Systeme, sondern auch für die Ausgaben generativer KI-Systeme (einschließlich Deepfake-Inhalten), die im Rahmen der wissenschaftlichen Forschung verwendet werden. Wenn diese Systeme jedoch in Betrieb genommen werden oder ihre Ausgaben für andere Zwecke als den alleinigen Zweck der wissenschaftlichen Forschung und Entwicklung verwendet werden, müssten die einschlägigen Transparenzpflichten nach Artikel 50 der KI-Verordnung dennoch erfüllt werden.

Wenn Forschende beispielsweise ein interaktives KI-System entwickeln und in Verkehr bringen oder in Betrieb nehmen und untersuchen möchten, ob natürliche Personen in der Lage sind, bei der mündlichen Interaktion mit diesem KI-System zwischen KI-Systemen und Menschen zu unterscheiden, müssen sie die jeweiligen Transparenzpflichten nach Artikel 50 Absätze 1 und 2 der KI-Verordnung nicht umsetzen. Wird dasselbe interaktive KI-System jedoch auch für andere Zwecke als Forschung und Entwicklung in Betrieb genommen, müssten die einschlägigen Transparenzpflichten nach Artikel 50 der KI-Verordnung eingehalten werden.

22. Artikel 2 Absatz 8 der KI-Verordnung schließt zudem Forschungs-, Test- und Entwicklungstätigkeiten zu KI-Systemen oder KI-Modellen aus, bevor diese in Verkehr gebracht oder in Betrieb genommen werden. Tests unter Realbedingungen (innerhalb oder außerhalb von KI-Reallaboren) fallen nicht unter diesen Ausschluss. Darüber hinaus sollte ein KI-System, sobald es infolge solcher Forschungs- und Testtätigkeiten in Verkehr gebracht oder in Betrieb genommen wird, dennoch die einschlägigen Transparenzpflichten nach Artikel 50 der KI-Verordnung erfüllen ([10]).

2.4.3. *KI, die im Rahmen freier und quelloffener Lizenzen freigegeben wird*

23. Nach Artikel 2 Absatz 12 der KI-Verordnung fallen KI-Systeme, die unter freien und quelloffenen Lizenzen bereitgestellt werden, nicht in den Anwendungsbereich der KI-Verordnung, sofern sie nicht unter die Verbote nach Artikel 5 der KI-Verordnung oder die Transparenzpflichten nach Artikel 50 der KI-Verordnung fallen oder nicht als hochriskant eingestuft sind ([11]). Dies bedeutet, dass Anbieter und Betreiber quelloffener KI-Systeme, die in den Anwendungsbereich von Artikel 50 der KI-Verordnung fallen, weiterhin die Einhaltung ihrer jeweiligen Transparenzpflichten sicherstellen müssen.

24. Dagegen unterliegen Anbieter freier und quelloffener KI-Komponenten, die Software, Daten, Modelle, Instrumente, Dienste oder Verfahren zur Anpassung und Integration in ein KI-System umfassen und für sich genommen (teilweise oder ganz) kein KI-System darstellen, nicht den Transparenzpflichten nach Artikel 50 der KI-Verordnung. Die Anbieter solcher Systeme werden dennoch ermutigt, diese Modelle, Komponenten und Dienste in der unter Randnummer 27 beschriebenen Weise zu gestalten, soweit dies verhältnismäßig und machbar ist.

([10]) Erwägungsgrund 25 der KI-Verordnung.

([11]) In Erwägungsgrund 102 der KI-Verordnung heißt es, dass die Freigabe von Software und Daten im Rahmen einer freien und quelloffenen Lizenz „ihre offene Weitergabe erlaubt und die Nutzer [diese] kostenlos abrufen, nutzen, verändern und weiter verteilen können, auch in veränderter Form“.

2.5. Zusammenspiel mit verbotenen Praktiken, Anforderungen an Hochrisiko-KI-Systeme und KI-Kompetenz

25. In Erwägungsgrund 137 der KI-Verordnung wird präzisiert, dass die Einhaltung der Transparenzpflichten nach Artikel 50 der KI-Verordnung für ein bestimmtes KI-System nicht als Hinweis darauf ausgelegt werden kann, dass die Verwendung dieses KI-Systems oder seiner Ausgabe nach der KI-Verordnung rechtmäßig ist. Insbesondere KI-Systeme, die den Pflichten nach Artikel 50 der KI-Verordnung unterliegen, können unter Artikel 5 der KI-Verordnung fallen und somit verboten sein (z. B. die Verwendung von Emotionserkennungssystemen am Arbeitsplatz oder in Bildungseinrichtungen) ([12]). Darüber hinaus können KI-Systeme, die in den Anwendungsbereich von Artikel 50 der KI-Verordnung fallen, auch als hochriskant eingestuft werden ([13]) und den einschlägigen Anforderungen und Pflichten für solche Systeme unterliegen. Die Einhaltung von Artikel 50 der KI-Verordnung ersetzt, begrenzt oder beeinträchtigt daher nicht die Pflichten, die sich aus anderen Bestimmungen der KI-Verordnung ergeben ([14]). Gleichzeitig können KI-Systeme auch in den Anwendungsbereich von Artikel 50 der KI-Verordnung fallen, ohne gemäß Artikel 6 der KI-Verordnung als hochriskant eingestuft zu werden. Die Anforderungen an die KI-Kompetenz nach Artikel 4 der KI-Verordnung gelten auch für Anbieter und Betreiber von KI-Systemen, die in den Anwendungsbereich von Artikel 50 der KI-Verordnung fallen.

2.6. Zusammenspiel mit Vorschriften für KI-Modelle und/oder -Systeme mit allgemeinem Verwendungszweck

26. Artikel 50 der KI-Verordnung gilt für bestimmte KI-Systeme, zu denen KI-Systeme mit allgemeinem Verwendungszweck ([15]) gehören, wie Chatbots, die die direkte Interaktion mit natürlichen Personen unterstützen, oder KI-Systeme, die in der Lage sind, synthetische Inhalte zu erzeugen oder zu manipulieren (einschließlich Deepfakes) ([16]).

27. Artikel 50 der KI-Verordnung gilt nicht ausdrücklich für KI-Modelle mit allgemeinem Verwendungszweck. Wenn jedoch ein interaktives oder generatives KI-System, das in den Anwendungsbereich von Artikel 50 Absatz 1 oder 2 der KI-Verordnung fällt, auf einem von demselben Anbieter bereitgestellten KI-Modell mit allgemeinem Verwendungszweck beruht, könnten die in diesen Leitlinien dargelegten Transparenzmaßnahmen auch auf Modellebene umgesetzt werden (siehe auch Abschnitt 4). Darüber hinaus werden Anbieter anderer KI-Modelle mit allgemeinem Verwendungszweck und anderer generativer KI-Modelle, auch wenn sie nicht in den Anwendungsbereich von Artikel 50 der KI-Verordnung fallen, aufgefordert, geeignete Transparenzmaßnahmen auf Modellebene zur Identifizierung interaktiver KI-Systeme und zur Kennzeichnung durch KI erzeugter und manipulierter Inhalte umzusetzen. Dies ist besonders vorteilhaft, um nachgelagerten KI-Systemanbietern die Erfüllung ihrer Pflichten nach Artikel 50 Absätze 1 und 2 der KI-Verordnung zu erleichtern ([17]) und zur Wahrung des Vertrauens in das Informationsökosystem und seiner Integrität beizutragen. Bei KI-Modellen mit allgemeinem Verwendungszweck, die systemische Risiken bergen, können Sicherheitsmaßnahmen auf Modellebene zur Identifizierung von KI-Interaktionen und zur Kennzeichnung durch KI erzeugter oder manipulierter Inhalte gegebenenfalls Teil eines umfassenderen Pakets von Risikominderungsmaßnahmen sein, die Anbieter im Rahmen ihrer Pflicht zur Minderung systemischer Risiken gemäß Artikel 55 Absatz 1 Buchstabe b der KI-Verordnung ergreifen können.

3. ARTIKEL 50 ABSATZ 1 DER KI-VERORDNUNG: TRANSPARENZ FÜR INTERAKTIVE KI-SYSTEME

28. Artikel 50 Absatz 1 der KI-Verordnung richtet sich an Anbieter von KI-Systemen, die direkt mit natürlichen Personen interagieren und ihre Systeme so konzipieren und entwickeln müssen, dass die betreffenden natürlichen Personen darüber informiert werden, dass sie mit einem KI-System interagieren. Wie in Erwägungsgrund 132 der KI-Verordnung erläutert, besteht der Zweck dieser Informationspflicht darin, diese natürlichen Personen in die Lage zu versetzen, fundierte Entscheidungen über die Ausgaben des Systems zu treffen, eine übermäßige Abhängigkeit von solchen Systemen zu vermeiden und es diesen natürlichen Personen zu ermöglichen, ihr Vertrauen in die Inhalte und Interaktionen entsprechend zu kalibrieren.

([12]) Weitere Informationen zu Artikel 5 der KI-Verordnung sind den Leitlinien der Kommission zu verbotenen Praktiken der künstlichen Intelligenz gemäß der Verordnung (EU) 2024/1689 (KI-Verordnung) (C(2025) 5052 final) zu entnehmen.

([13]) Siehe Artikel 6 der KI-Verordnung.

([14]) Insbesondere sieht Artikel 26 Absatz 11 der KI-Verordnung eine Transparenzpflicht für Betreiber von Hochrisiko-KI-Systemen vor, die im Zusammenhang mit der Unterrichtung natürlicher Personen darüber, dass sie der Verwendung des Hochrisiko-KI-Systems unterliegen, Entscheidungen treffen oder dabei helfen, Entscheidungen zu treffen. Diese Transparenzpflichten ergänzen die in Artikel 50 der KI-Verordnung festgelegten Pflichten, da sie unterschiedliche Informationen betreffen, die natürlichen Personen zur Verfügung gestellt werden müssen, können jedoch in einzelnen Mitteilungen kombiniert werden.

([15]) Laut Artikel 3 Nummer 66 ist ein „KI-System mit allgemeinem Verwendungszweck“ ein KI-System, das auf einem KI-Modell mit allgemeinem Verwendungszweck beruht und in der Lage ist, einer Vielzahl von Zwecken sowohl für die direkte Verwendung als auch für die Integration in andere KI-Systeme zu dienen.

([16]) Artikel 50 Absatz 2 der KI-Verordnung enthält einen ausdrücklichen Verweis auf KI-Systeme mit allgemeinem Verwendungszweck.

([17]) Erwägungsgrund 133 der KI-Verordnung.

3.1. **Hauptkomponenten, Konzepte und damit verbundene Transparenzpflicht(en)**

29. Die Transparenzpflicht nach Artikel 50 Absatz 1 der KI-Verordnung gilt für Anbieter von KI-Systemen, wenn das System direkt mit natürlichen Personen interagieren soll und diese Interaktion unter keine der folgenden Ausnahmen fällt: 1) sie ist aus Sicht einer angemessen informierten, aufmerksamen und verständigen natürlichen Person offensichtlich oder 2) das System ist gesetzlich zur Aufdeckung, Verhütung, Ermittlung oder Verfolgung von Straftaten zugelassen, es sei denn, es steht der Öffentlichkeit zur Anzeige einer Straftat zur Verfügung.

3.1.1. *KI-Systeme, die für die direkte Interaktion mit natürlichen Personen bestimmt sind*

30. Damit ein System in den Anwendungsbereich von Artikel 50 Absatz 1 der KI-Verordnung fällt, müssen vier Elemente vorliegen: i) das System muss ein KI-System sein, das für ii) die Interaktion iii) auf direktem Wege iv) mit natürlichen Personen bestimmt ist.

i) **Ein KI-System**: Das System muss die Elemente der Definition von KI-Systemen ([18]) erfüllen, die einfache, nicht KI-gestützte automatisierte Reaktionsmechanismen (z. B. herkömmliche Abwesenheits-E-Mails, regelbasierte schnelle Nachrichtenantworten) ausschließt.

ii) **Zur Interaktion bestimmt:** Das KI-System muss dazu bestimmt sein, mit natürlichen Personen zu interagieren. Interaktion umfasst die Fähigkeit eines bidirektionalen Informationsaustauschs oder von Maßnahmen zwischen natürlichen Personen und dem KI-System mit echtem Gesprächs- oder Reaktionscharakter. Sie kann in verschiedenen Formen auftreten (z. B. auditiv, visuell und physisch). Die natürlichen Personen sollten in der Lage sein, Eingaben (z. B. schriftliche Texte oder andere Formen von Inhalten, Sprache oder physische Handlungen) in das KI-System einzuspeisen (z. B. über technische Kommunikationsmittel, einschließlich Online-Schnittstellen, Telefon oder E-Mail, oder direkt), während das KI-System in der Lage sein sollte, mit kontextbezogenen Ausgaben (einschließlich aller Arten von Inhalten oder Handlungen) zu reagieren und umgekehrt. Die Interaktion kann alle Arten von Ausgaben oder Formaten umfassen, die von Menschen wahrgenommen und verstanden werden können. Sie muss nicht von einem Menschen initiiert werden; sie kann auch KI-Inhalte oder Handlungen umfassen, die eine menschliche Reaktion hervorrufen. Bei der Interaktion kann es sich um einen einmaligen Austausch handeln (z. B. ein einmaliger Prompt mit anschließender einmaliger Antwort) oder sie kann über einen bestimmten Zeitraum erfolgen (Multiturn). Bei Systemen, die nur passiv Daten erheben und nicht zu einem Austausch mit natürlichen Personen in der Lage sind (z. B. automatische Zugangskontrollen mit Gesichtserkennung) oder lediglich ein einmaliges Feedback sammeln (z. B. Spam-Filter), wird nicht davon ausgegangen, dass sie mit natürlichen Personen interagieren.

iii) **Direkte Interaktion:** Die Interaktion zwischen dem KI-System und den natürlichen Personen muss direkt sein. Dies umfasst in der Regel Interaktionen in Echtzeit oder nahezu in Echtzeit (u. a. durch die Fähigkeit des Systems, Nachrichten an natürliche Personen zu verfassen und zu senden oder anderweitig mit der physischen oder virtuellen Umgebung zu interagieren, die von natürlichen Personen wahrgenommen werden kann). Direkte Interaktion schließt indirekte oder vermittelte menschliche Interaktion aus, wenn eine natürliche Person KI-Ausgaben ausgesetzt ist, ohne direkt mit dem System zu interagieren (z. B. wenn Vertreter des Kundendienstes KI-Hilfsinstrumente nutzen, um sie bei der Kommunikation mit natürlichen Personen zu unterstützen, oder wenn die KI-Ausgabe der Person nicht vom KI-System selbst, sondern von einer anderen Person, die die Inhalte verbreitet, zur Verfügung gestellt wird). Die bloße Möglichkeit für Menschen, einzugreifen oder die Ausgaben des KI-Systems zu überprüfen, sollte jedoch nicht genutzt werden, um die Anwendung der Offenlegungspflicht nach Artikel 50 Absatz 1 der KI-Verordnung zu umgehen. Fälle von Produkten und Dienstleistungen, bei denen durch KI erzeugte Reaktionen mit vom Menschen kuratierten Inhalten kombiniert werden, fallen in den Anwendungsbereich von Artikel 50 Absatz 1 der KI-Verordnung und erfordern eine Offenlegung für die spezifischen Ausgaben der Interaktion, die durch KI erzeugt wurden, es sei denn, diese KI-Ausgaben wurden von Menschen als Hauptansprechpartner der natürlichen Personen ordnungsgemäß überprüft und gesendet.

iv) **Mit natürlichen Personen:** Nach Artikel 50 Absatz 1 der KI-Verordnung muss das KI-System mit natürlichen Personen interagieren. Bei diesen Personen kann es sich um professionelle Betreiber, andere Nutzer (die das System auch für ausschließlich persönliche Tätigkeiten verwenden) oder andere Personen handeln, die das System in ihrem Namen nutzen. KI-Systeme, die in einer geschlossenen physischen Umgebung betrieben werden (z. B. als Teil einer Industriemaschine), Backend-Anrufe von Maschine zu Maschine zwischen KI-Systemen (einschließlich KI-Agenten), deren Ausgaben nicht dazu bestimmt sind, natürliche Personen zu erreichen, oder virtuelle Umgebungen ohne direkten Kontakt zu natürlichen Personen sind vom Anwendungsbereich der Pflicht ausgenommen.

([18]) Siehe Artikel 3 Nummer 1 der KI-Verordnung. Weitere Leitlinien zum Begriff „KI-System“ sind den Leitlinien der Kommission zur Definition eines Systems der künstlichen Intelligenz vom 29.7.2025 (C(2025) 5053 final) zu entnehmen.

31. KI-Agenten fallen unter Artikel 50 Absatz 1 der KI-Verordnung, wenn sie in der Lage sind, bei der Ausführung der Aufgaben mit den sie anweisenden Personen oder mit anderen natürlichen Personen zu interagieren (z. B. Buchungen vorzunehmen, Korrespondenz zu verwalten, Verträge auszuhandeln oder abzuschließen, Käufe auszuführen usw.). KI-Agenten müssen so konzipiert und entwickelt werden, dass sowohl ihr künstlicher Charakter als auch die Person, in deren Auftrag sie handeln, offengelegt werden, wobei die Notwendigkeit der Transparenz in Bezug auf die Herkunft und die Übertragung von Befugnissen und die Rechenschaftspflicht für die Folgen ihres Handelns zu berücksichtigen ist. Dazu gehören Fälle, in denen KI-Agenten in der Lage sind, in komplexen Multi-Agenten-Architekturen mit anderen KI-Agenten zu interagieren, die bei der Ausführung der Aufgabe direkt mit den natürlichen Personen interagieren. Dennoch ist es dem Anbieter des KI-Agenten möglicherweise nicht immer möglich, einzelne Fälle der direkten Interaktion eines KI-Agenten mit anderen natürlichen Personen im Voraus zu ermitteln, wenn der KI-Agent in verschiedenen Kontexten eingesetzt wird. Kann der Anbieter nicht zuverlässig feststellen, ob der KI-Agent direkt mit einer natürlichen Person interagieren wird, bevor er den KI-Agenten in Verkehr bringt oder in Betrieb nimmt, ist der Agent auf Architekturebene zu konzipieren und anzuweisen, sich in jeder Situation, in der es hinreichend wahrscheinlich ist, dass er mit einer natürlichen Person interagieren wird (auch wenn eine natürliche Person eine juristische Person vertreten oder für diese handeln kann), als KI-Agent kenntlich zu machen. Darüber hinaus müssen sich KI-Agenten gegenüber den Personen, von denen sie angewiesen werden, bei wichtigen Schritten (z. B. zum Zeitpunkt der Zulassung, Meldung, Validierung usw., auch wenn der eingesetzte KI-Agent Ausgaben erhält, Ausgaben verarbeitet oder sich auf Ausgaben stützt, die von anderen KI-Systemen und nicht direkt von einer natürlichen Person erzeugt wurden) und bei jeder neuen Interaktion kenntlich machen.

Beispiele für KI-Systeme, die für die direkte Interaktion mit natürlichen Personen bestimmt sind und in den Anwendungsbereich von Artikel 50 Absatz 1 der KI-Verordnung fallen:

— KI-gestützte Sprachassistenten, Chatbots/Dialogsysteme in verschiedenen Kontexten (z. B. öffentlicher Dienst, Kundenbetreuung, Beschwerdemanagement, elektronischer Handel, Finanzen, Gesundheitsversorgung, Bildung usw.),

— KI-Hotline für Vorfälle, Betrug oder sonstige Meldungen, (humanoide) Roboter/Cobots, KI-Begleiter; robotergesteuerte Haustiere; KI-Avatare (z. B. in Virtual-Reality-Umgebungen),

— KI-Bots in sozialen Netzwerken und Medien, Codierungsagenten und andere KI-Agenten, die in der Lage sind, direkt mit natürlichen Personen zu interagieren (auch in komplexen Multi-Agenten-Architekturen).

3.2.1 **Beispiele für KI-Systeme, die nicht direkt mit natürlichen Personen interagieren und nicht in den Anwendungsbereich von Artikel 50 Absatz 1 der KI-Verordnung fallen:**

— KI-gestützte traditionelle Industrieroboter, die in einem geschlossenen industriellen Umfeld arbeiten und nicht dazu bestimmt sind, mit Menschen zu interagieren (z. B. elektronische Montageroboter),

— KI-gestützte Empfehlungssysteme, Spam-Filter, automatische Übersetzungs- oder Transkriptionstools,

— KI-gestütztes Suchen und Abrufen von Informationen ohne Erzeugung oder Änderung von Inhalten,

— KI-gestützte Techniken zur automatischen Vervollständigung von Text und Code, Systeme zur Authentifizierung/biometrischen Erkennung, Backend-Entscheidungsunterstützungssysteme, bei denen der Nutzer nur die KI-Ausgabe sieht, ohne die Möglichkeit einer direkten Interaktion mit dem System,

— KI-gestützte Systeme für vorausschauende Wartung und Optimierung in Fabriken und industriellen Anwendungen, KI-gestützte Produktvisualisierung (z. B. virtuelle Anprobe, Rauminszenierung, Produkt-Rendering), KI-gestützte Werkzeuge zur einmaligen Routenplanung, Transkription von Kundenanrufen, die für Verhaltensanalysen verarbeitet werden, usw.

3.1.2. *Informationspflicht nach Artikel 50 Absatz 1 der KI-Verordnung*

32. Nach Artikel 50 Absatz 1 der KI-Verordnung müssen Anbieter von KI-Systemen, die direkt mit natürlichen Personen interagieren, sicherstellen, dass ihre KI-Systeme (einschließlich KI-Agenten und KI-Systeme mit allgemeinem Verwendungszweck) so konzipiert und entwickelt werden, dass die betreffenden natürlichen Personen informiert werden, dass sie mit einem KI-System interagieren.

33. Während der Informationsmechanismus in die Konzeption und Entwicklung des Systems und die Art und Weise, wie das System mit den natürlichen Personen interagiert, eingebettet sein sollte, hat die Mitteilung während des Betriebs des KI-Systems und spätestens zum Zeitpunkt der ersten Interaktion mit der natürlichen Person zu erfolgen (wie in Artikel 50 Absatz 5 der KI-Verordnung vorgeschrieben, siehe Abschnitt 7). Die Anbieter sind dafür verantwortlich, sicherzustellen, dass natürliche Personen, die mit dem KI-System interagieren, nach dem Inverkehrbringen oder der Inbetriebnahme des KI-Systems über den gesamten Lebenszyklus des KI-Systems, wie es von diesem Anbieter in Verkehr gebracht oder in Betrieb genommen wurde, wirksam und konsequent informiert werden.2.2

34. **Format der Offenlegung:** In Artikel 50 Absatz 1 der KI-Verordnung ist keine bestimmte Offenlegungstechnik festgelegt. Die Anbieter können daher jede geeignete Technik wählen, solange zwei Voraussetzungen erfüllt sind: i) die in Artikel 50 Absatz 5 der KI-Verordnung festgelegten Anforderungen in Bezug auf die Bereitstellung klarer und eindeutiger Informationen spätestens zum Zeitpunkt der ersten Interaktion werden erfüllt (siehe Abschnitt 7) und ii) die Merkmale von natürlichen Personen, die aufgrund ihres Alters (z. B. Kinder oder ältere Menschen) oder einer Behinderung schutzbedürftigen Gruppen angehören, werden insoweit berücksichtigt, als das KI-System dazu bestimmt ist, mit natürlichen Personen, die diesen Gruppen angehören, zu interagieren oder dies mit hinreichender Wahrscheinlichkeit tun wird (z. B. auch dann, wenn das KI-System mit hinreichender Wahrscheinlichkeit mit der breiten Öffentlichkeit interagiert) ([19]). Gemäß Artikel 50 Absatz 5 und Erwägungsgrund 132 der KI-Verordnung müssen die Informationen auch in einer für natürliche Personen mit Behinderungen zugänglichen Form bereitgestellt werden (siehe Abschnitt 7). Wenn das System mit Kindern interagieren kann, müssen diese Mitteilungen kindgerecht, altersgerecht, leicht verständlich und für alle Kinder, auch für Kinder mit Behinderungen oder zusätzlichen Barrierefreiheitsbedürfnissen, leicht zugänglich sein. Sofern verfügbar und angemessen, wird den Anbietern empfohlen, standardisierte Offenlegungspraktiken und Kennzeichnungen zu verwenden, um die Wirksamkeit und Kohärenz der Mitteilungen und eine bessere Sensibilisierung der Nutzer zu gewährleisten.

35. **Inhalt der Mitteilung:** Nach Artikel 50 Absatz 1 der KI-Verordnung müssen Anbieter interaktiver KI-Systeme diese so konzipieren und entwickeln, dass alle natürlichen Personen, die mit solchen KI-Systemen interagieren, ausdrücklich über den künstlichen, nicht menschlichen Charakter des interagierenden Gegenübers informiert werden. Ist für den Anbieter vernünftigerweise vorhersehbar, dass das KI-System mit natürlichen Personen interagieren kann, die zu den oben genannten schutzbedürftigen Gruppen gehören, sind die Informationen entsprechend anzupassen und in einem barrierefreien Format bereitzustellen, wie es nach geltendem Unionsrecht erforderlich ist, um auch das umfassende Verständnis dieser natürlichen Personen zu gewährleisten ([20]).

36. Die Mitteilungen sollten Informationen über den künstlichen Ursprung der Interaktion in einer Weise enthalten, die dem Kontext, in dem die Interaktion stattfindet, angemessen ist. Dies sollte in einer Weise geschehen, die sicherstellt, dass die betroffenen Personen ordnungsgemäß und wirksam informiert werden, wobei jegliches Risiko einer Täuschung oder Manipulation vor und während der Interaktion zu vermeiden und gleichzeitig eine praktikable und verhältnismäßige Offenlegung sicherzustellen ist. Beispiele für die Bereitstellung der Informationen sind schriftliche Benachrichtigungen (z. B. ein Chatbot, der ein Gespräch mit dem Hinweis einleitet, dass er auf KI-Technologie beruht), visuelle Mittel (z. B. eine von einem KI-Agenten erzeugte E-Mail an eine natürliche Person, die oben ein KI-Label aufweist), akustische Mittel (z. B. ein Sprachassistent, der zu Beginn einer Sitzung erklärt, dass er KI-gestützt ist), die Offenlegung von KI-Kennungen und Berechtigungsnachweisen (z. B. KI-Agenten, die ihre KI-Identität offenlegen, soweit dies auf überprüfbare Weise möglich ist ([21])).

([19]) Siehe Erwägungsgrund 132 der KI-Verordnung. Dies setzt nicht voraus, dass das System auf diese schutzbedürftigen Gruppen ausgerichtet ist. Soll ein System mit einem Mitglied der Öffentlichkeit interagieren, so werden auch Mitglieder der schutzbedürftigen Gruppen einbezogen und sollten von den Anbietern bei der Umsetzung der Mitteilungsmaßnahmen berücksichtigt werden.

([20]) Zu den geltenden Barrierefreiheitsanforderungen siehe Abschnitt 7.3.

([21]) So können beispielsweise elektronische Attributsbescheinigungen, die gemäß der Verordnung (EU) Nr. 910/2014 eingeführt und in den gemäß derselben Verordnung eingerichteten europäischen Brieftaschen für die Digitale Identität sowie den vorgeschlagenen europäischen Unternehmensbrieftaschen zur Verfügung gestellt werden, sichere und effiziente Mittel zur Identifizierung von KI-Agenten bieten. Die Brieftaschen können elektronische Bescheinigungen speichern und verwalten, mit denen die Identität, die Attribute und die Autorisierungen des KI-Agenten überprüft werden, wodurch eine nahtlose und vertrauenswürdige Offenlegung gegenüber natürlichen Personen ermöglicht wird.

37. **Format der Mitteilungen**: Die Konzeption des KI-Systems mit multimodalen und zugänglichen Offenlegungsmethoden wird als bewährtes Verfahren empfohlen. Solche Mitteilungen können je nach Kontext, Zielgruppe und vorhersehbaren Publikumsgruppen sowie den verfügbaren Ressourcen der Anbieter, insbesondere kleiner und mittlerer Unternehmen (KMU), gegebenenfalls verschiedene Techniken kombinieren, um das Verständnis der Nutzer zu verbessern. Beispiele für geeignete Techniken können sein:

— **Offenlegung in Textform** (auf Grundlage der Benutzerschnittstelle): Hervorgehobene, leicht verständliche Kennzeichnungen oder Banner (z. B. „Sie interagieren mit einem KI-System“) und erste Grüße in Chatbots, die gegebenenfalls mit dauerhaften Abzeichen oder Kennzeichnungen kombiniert werden können, die während der gesamten Interaktion sichtbar sind und in einem angemessenen Verhältnis zum spezifischen Kontext und zu den spezifischen Risiken stehen. Darüber hinaus wird empfohlen, begleitende Offenlegungen in der Nähe der Interaktionsschnittstelle (z. B. in der Nähe des Ein-/Ausgabefelds) zu positionieren und vereinfachte Formulierungen zu verwenden, insbesondere für Nutzer mit geringerer digitaler Kompetenz und Kinder.

— **Akustische Offenlegung:** In sprachbasierten oder telefonischen Kontexten ausdrückliche mündliche Aussagen zu Beginn der Interaktion (z. B. „Dies ist ein KI-gestützter Assistent“), gegebenenfalls kombiniert mit regelmäßigen Erinnerungen bei längeren Interaktionen, insbesondere im Falle von Unterbrechungen oder einer Änderung der Rolle des KI-Systems während einer Nutzerreise. Unterschiedliche Audiosignale (z. B. Töne oder Earcons) können die Erkennung unterstützen, insbesondere für sehbehinderte Nutzer, werden jedoch für sich genommen nicht als ausreichend angesehen.

— **Visuelle/grafische Hinweise:** Konzeption des KI-Systems mit dauerhaften Symbolen, Wasserzeichen, farbigen Rahmen oder erkennbaren „KI“-Symbolen zur Ergänzung der Offenlegungen in Textform. Die Verwendung standardisierter visueller Indikatoren, auch für alle Dienste eines Anbieters, wird empfohlen, da solche Indikatoren den kognitiven Aufwand erheblich verringern und die Erkennbarkeit erhöhen können.

— **Multimodale Kombinationen:** Kombination von Text-, Audio- und visuellen Hinweisen (z. B. ein Chatbot mit einer Kennzeichnung, der gleichzeitig bei der ersten Interaktion eine Offenlegung in Textform bietet), um die Barrierefreiheit zu gewährleisten und die Klarheit für verschiedene Nutzergruppen zu erhöhen.

38. Bestimmte Techniken sind nicht erforderlich und reichen für sich genommen nicht aus, um die Transparenzpflicht nach Artikel 50 Absätze 1 und 5 der KI-Verordnung wirksam zu erfüllen, insbesondere:

— Offenlegungen, die nur in allgemeinen Geschäftsbedingungen, URLs oder Dokumentationen enthalten sind (diese Offenlegungen können die Offenlegungen im Kontext ergänzen, aber nicht ersetzen),

— maschinenlesbare Kennzeichnungen (z. B. Metadaten oder Wasserzeichen), die von den Nutzern am Interaktionspunkt nicht wahrnehmbar sind. Dies berührt nicht die Angemessenheit solcher Kennzeichnungen im Hinblick auf die Einhaltung von Artikel 50 Absatz 2 der KI-Verordnung (siehe Abschnitt 4),

— unklare oder mehrdeutige Signale (z. B. allgemeine Verweise auf einen „Assistenten“) oder menschenähnliche Darstellungen, die die Nutzer irreführen können,

— allgemeine Offenlegungen, die nicht ausreichend spezifisch für die Ausgaben und Interaktionen des KI-Systems sind (z. B. auf einer Plattform, die eine Vielzahl von Diensten oder einzelnen Eingaben bereitstellt, ist eine allgemeine Offenlegung wie „Dienste auf dieser Website nutzen KI“ unzureichend),

— technische oder fähigkeitsbasierte Beschreibungen: Aussagen, die sich ausschließlich auf zugrunde liegende Technologien beziehen (z. B. „dieses System verwendet LLM“), ohne die Funktion oder die Auswirkungen des Systems auf den Nutzer und seine künstliche Herkunft zu erläutern.

39. Offenlegungen müssen verhältnismäßig und auf die verschiedenen Kontexte und Zielgruppen des KI-Systems abgestimmt sein, um wirksam und benutzerfreundlich zu bleiben und, soweit möglich, mögliche Störungen des Nutzererlebnisses oder Gewöhnungseffekte (sogenannte „Bannerblindheit“) zu vermeiden.

40. Eine einzige, hervorgehobene Mitteilung vor der ersten Interaktion des KI-Systems mit einer bestimmten natürlichen Person dürfte in den meisten Fällen ausreichen, um die Pflicht nach Artikel 50 Absatz 1 der KI-Verordnung zu erfüllen. In bestimmten risikoreicheren Kontexten dürften jedoch regelmäßige Erinnerungen und kontextbewusste Offenlegungen erforderlich sein, um ein kontinuierliches Nutzerbewusstsein zu gewährleisten. Dies gilt insbesondere für KI-Systeme, die mit schutzbedürftigen Personen (z. B. Kindern, älteren Menschen, Menschen mit Behinderungen und Menschen mit geringerer Lese- und Schreibkompetenz) interagieren, für anhaltende und sich entwickelnde Interaktionen in sensiblen oder immersiven Kontexten (z. B. wenn Nutzer emotionale Belastung oder Schutzbedürftigkeit ausdrücken oder erleben oder der Gefahr suchtähnlicher Verhaltensweisen ausgesetzt sind), wenn ein erhöhtes Risiko besteht, dass Nutzer irregeführt werden (z. B. Finanzberatung, Versicherung, Rechtsbeistand, Gesundheitsberatung und Bearbeitung von Beschwerden) oder emotionale Bindungen oder Abhängigkeiten aufbauen (z. B. KI-Begleiter), sowie für die Notwendigkeit, das Bewusstsein und die Kontrolle über die Handlungen von KI-Agenten sicherzustellen. Die Anbieter müssen das KI-System auch so konzipieren und entwickeln, dass die Offenlegung in allen Situationen gewährleistet ist, in denen dem KI-System Fragen zu seiner Art oder zum Ursprung der Interaktion gestellt werden oder in denen aufgrund des Austauschs mit der natürlichen Person vernünftigerweise davon ausgegangen werden kann, dass diese Person wahrscheinlich über den KI-Ursprung der Interaktion irregeführt oder verwirrt wird.

3.2. **Ausnahmen von der Informationspflicht nach Artikel 50 Absatz 1 der KI-Verordnung**

41. Artikel 50 Absatz 1 der KI-Verordnung sieht zwei Ausnahmen von der Informationspflicht in Bezug auf interaktive KI-Systeme vor, die im Folgenden untersucht werden.

3.2.1. *Ausnahme für offensichtliche Interaktion mit einem KI-System*

42. Um sich auf die erste Ausnahme berufen zu können, müssen die Anbieter i) den offensichtlich künstlichen Charakter der Interaktion mit ii) einer natürlichen Person, die unter Berücksichtigung der Umstände und des Kontexts der Nutzung angemessen informiert, aufmerksam und verständig sein muss, bewerten und nachweisen.

43. Dieser Standard stützt sich auf das bestehende EU-Verbraucherschutzrecht in Bezug auf den Begriff „Durchschnittsverbraucher“ und sollte einheitlich ausgelegt werden, wobei die spezifischen und unterschiedlichen Kontexte und Verwendungen von KI-Systemen, die direkt mit natürlichen Personen interagieren, und die mit Artikel 50 Absatz 1 der KI-Verordnung verfolgten Ziele zu berücksichtigen sind.

44. Die Konzeption einer solchen „durchschnittlichen“ natürlichen Person umfasst zwei Schritte. Erstens müssen die Anbieter das Publikum berücksichtigen, mit dem das KI-System bestimmungsgemäß und mit hinreichender Wahrscheinlichkeit interagieren wird. Zweitens muss der Anbieter prüfen, wie angemessen informiert, aufmerksam und verständig ein durchschnittliches Mitglied dieses Publikums ist, wobei die Umstände und der Kontext der Nutzung zu berücksichtigen sind. Bei dieser Bewertung sind mehrere Faktoren zu berücksichtigen:

— Die Anbieter sollten die Zielgruppe des KI-Systems sowie ein breiteres potenzielles Publikum berücksichtigen, das vernünftigerweise vorhersehbar ist. Dies ist insbesondere dann der Fall, wenn die breite Öffentlichkeit leicht auf das KI-System zugreifen kann oder wenn das System in virtuellen oder physischen öffentlichen Räumen betrieben wird und in der Lage ist, mit natürlichen Personen in diesen Räumen zu interagieren. Das Publikum, das einem KI-System ausgesetzt sein kann, entspricht nicht immer der Zielgruppe.

— Die Anbieter sollten die potenziell vielfältige Zusammensetzung der Zielgruppe und des vernünftigerweise vorhersehbaren Publikums im Einklang mit Erwägungsgrund 132 der KI-Verordnung berücksichtigen. Wenn Menschen mit Behinderungen, ältere Menschen oder Minderjährige wahrscheinlich Teil dieses Publikums sind, ist das erwartete Maß an Information, Aufmerksamkeit und Verständigkeit eines durchschnittlichen Mitglieds eines solchen Publikums geringer als bei einem durchschnittlichen Mitglied eines Publikums, das diese Personengruppen nicht umfasst. Wenn das interaktive KI-System hingegen nur einer professionellen oder spezialisierten Zielgruppe zur Verfügung steht, könnte sich dies positiv auf das erwartete Maß an Wissen, Aufmerksamkeit und Verständigkeit auswirken, und ein durchschnittliches Mitglied dieser Zielgruppe könnte als besser informiert, aufmerksamer und verständiger angesehen werden.

— Im Rahmen der umfassenderen Umstände und des Kontexts der Nutzung sollten die Anbieter das Niveau der digitalen Kompetenz und der KI-Kompetenz des Zielpublikums und des vernünftigerweise vorhersehbaren Publikums sowie relevante soziale, kulturelle und sprachliche Faktoren berücksichtigen.

45. **Offensichtlichkeit**: Bei der Ermittlung der hypothetischen angemessen informierten, aufmerksamen und verständigen natürlichen Person sollten die Anbieter bewerten, ob es für eine solche Person offensichtlich sein wird, dass sie direkt mit einem KI-System interagiert. Die Ausnahme ist restriktiv auszulegen, da sie natürlichen Personen den Schutz und das Recht nimmt, durch klare und unterscheidbare Offenlegungen darüber informiert zu werden, dass sie mit KI-Systemen interagieren. Das allgemeine Bewusstsein der Verbraucher und anderer natürlicher Personen, dass es KI-Systeme (einschließlich Chatbots und Agenten) gibt, bedeutet nicht, dass sie sie bei Interaktionen erkennen. Mehrere Faktoren können relevant sein und bei der Beurteilung, ob die Interaktion offensichtlich ist, berücksichtigt werden:

— Voraussichtliche Art der Interaktion: In Fällen, in denen das KI-System physisch mit natürlichen Personen interagiert, können relevante Elemente die Sichtbarkeit mechanischer Komponenten (was die Offensichtlichkeit erhöhen würde) oder den Grad, in dem das KI-System sein nicht künstliches Äquivalent authentisch repliziert (was die Offensichtlichkeit verringern würde), umfassen. Bei visuellen oder akustischen Interaktionen können Elemente wie Schreib- oder Sprechmuster, Stimmklang (Roboterstimme oder echte menschlich klingende Stimme), die Gestaltung der Benutzerschnittstelle (z. B. ein Profilbild im Zusammenhang mit einem Chatbot, das einen Menschen zeigt) und die Fähigkeit zu einer fortgeschrittenen personalisierten Interaktion diese Bewertung beeinflussen.

— Zusammensetzung der Zielgruppe und des voraussichtlichen Publikums: Wenn das interaktive KI-System nur einem professionellen oder spezialisierten Publikum zur Verfügung steht, ohne dass potenziell schutzbedürftige Gruppen ihm ausgesetzt sind, könnte dies die mögliche Offensichtlichkeit des künstlichen Charakters der Interaktion für ein durchschnittliches Mitglied des jeweiligen Publikums erhöhen. Wenn das KI-System hingegen für die breite Öffentlichkeit oder für Verbraucher zugänglich ist, zu denen auch schutzbedürftige Personen gehören (z. B. Kinder, ältere Menschen, Menschen mit kognitiven, körperlichen oder geistigen Behinderungen oder Menschen mit geringeren KI-Kompetenzen), für die die Interaktion weniger offensichtlich ist, kann die Ausnahme nicht geltend gemacht werden, um diese schutzbedürftigen Nutzer wirksam zu schützen.

— Es sollte auch berücksichtigt werden, dass es für natürliche Personen immer schwieriger wird, zu wissen, ob eine Interaktion mit einem KI-System oder einem realen Menschen stattfindet. Aufgrund des Einsatzes von Avataren und Agenten, der allgemeinen Anonymität von Online-Interaktionen und der Fähigkeit von KI-Systemen, auf menschenähnliche Weise zu schreiben und zu interagieren, sollte die Ausnahme zur „Offensichtlichkeit“ auf Fälle beschränkt werden, in denen für eine durchschnittliche Person aus der Zielgruppe und dem vernünftigerweise vorhersehbaren Publikum fast kein Zweifel an der Art der Interaktion mehr besteht.

Beispiele für Offensichtlichkeit, für die die Transparenzpflicht nach Artikel 50 Absatz 1 der KI-Verordnung nicht gilt:

— KI-gestützte Chatbots zur Unterstützung beim Programmieren und zur Codeüberprüfung, die nur professionellen Entwicklern zur Verfügung stehen, die aufgrund ihres Fachwissens und des Kontexts der Nutzung keine begründete Erwartung haben, mit einem Menschen zu interagieren, und leicht erkennen können, dass Vorschläge von einem KI-System erzeugt werden

— Ein interner Mitarbeiterassistent für ordnungsgemäß geschultes Personal, das über KI-Kompetenz verfügt und weiß, dass es KI-Systeme für interne organisatorische Zwecke verwendet (z. B. Personalverwaltung, Rechtliches, Beschaffung, Compliance oder IT-Unterstützung).

— Interaktive KI-Systeme, die bestimmungsgemäß nur von entsprechend geschulten Angehörigen der Gesundheitsberufe verwendet werden sollen, um die medizinische Diagnose zu unterstützen und entsprechende Behandlungen vorzuschlagen.

— In Haushaltsgeräte eingebettete Umgebungs-KI mit Funktionen, die auf die Unterstützung und den Betrieb von Haushaltsgeräten beschränkt sind.

— Interaktionen mit KI-gestützten nicht-spielbaren Charakteren (Non-playable Characters, NPC) in einem Einzelspieler-Videospiel, bei dem die Art des Videospiels deutlich macht, dass keine andere natürliche Person teilnehmen und mit dem Nutzer interagieren kann, und der KI-Charakter der Interaktion auch für alle Nutzer (einschließlich schutzbedürftiger Personen) offensichtlich ist.

Beispiele, die die Ausnahme der Offensichtlichkeit nicht erfüllen und für die die Transparenzpflicht nach Artikel 50 Absatz 1 der KI-Verordnung gilt:

— Ein KI-gestütztes robotergesteuertes Haustier, das seinem natürlichen Äquivalent sehr ähnlich und so konzipiert ist, dass es eine typische Interaktion zwischen Mensch und Haustier nachahmt, wodurch Unklarheit darüber entsteht, ob die Interaktion mit einem KI-System erfolgt oder nicht.

— KI-Systeme, die in immersive Umgebungen (z. B. virtuelle oder erweiterte Realität) eingebettet sind, in denen realistische Avatare oder Stimmen verwendet werden, die Menschen ähneln, bei denen Nutzer und insbesondere Kinder, ältere Menschen oder Menschen mit Behinderungen möglicherweise nicht leicht zwischen menschlicher und KI-Interaktion unterscheiden können.

— In Online-Plattformen eingebettete KI-Chatbots oder Hilfsinstrumente (Helpdesks), bei denen Nutzer direkt interagieren und KI-Ausgaben (z. B. Antworten auf Anfragen oder andere durch KI erzeugte Inhalte) empfangen, bei denen sie möglicherweise davon ausgehen, dass sie von einem Menschen erzeugt wurden.

3.2.2. *Ausnahme für gesetzlich zu Strafverfolgungszwecken zugelassene KI-Systeme*

46. Anbieter interaktiver KI-Systeme sind von der Transparenzpflicht nach Artikel 50 Absatz 1 der KI-Verordnung ausgenommen, wenn sie gesetzlich zur Aufdeckung, Verhütung, Ermittlung oder Verfolgung von Straftaten zugelassen sind, wenn geeignete Schutzvorkehrungen für die Rechte und Freiheiten Dritter bestehen.

47. Unter „gesetzlich zugelassen“ sind das Unionsrecht sowie die im Einklang mit dem Unionsrecht erlassenen nationalen Rechtsvorschriften der Mitgliedstaaten zu verstehen. In dem Gesetz, mit dem die Verwendung des KI-Systems genehmigt wird, müssen die Zwecke und Umstände der zulässigen Verwendung eindeutig festgelegt und geeignete Schutzvorkehrungen für die Rechte und Freiheiten Dritter vorgesehen werden. Angesichts des technologieneutralen Ansatzes der KI-Verordnung ist es nicht erforderlich, bestimmte KI-Systeme, bei denen die Verwendung des KI-Systems im Allgemeinen im Rahmen der gesetzlich zugelassenen Tätigkeit vorgesehen ist und dem geltenden Rechtsrahmen entspricht, gesetzlich zu identifizieren oder zu genehmigen.

48. Um unter diese Ausnahme zu fallen, muss der Zweck des KI-Systems darin bestehen, Straftaten aufzudecken, zu verhüten, zu ermitteln oder zu verfolgen (z. B. verdeckte KI-Ermittler, KI-Systeme, die unter der menschlichen Kontrolle von Ermittlern für Online-Ermittlungen oder die Kommunikation mit Personen, die an kriminellen Tätigkeiten beteiligt sein könnten, verwendet werden). Die Ausnahme ist nicht auf die Verwendung solcher KI-Systeme durch Strafverfolgungsbehörden im Sinne von Artikel 3 Nummer 48 der KI-Verordnung beschränkt, sondern kann auch interaktive (oder generative) KI-Systeme abdecken, die von anderen Behörden der Union oder der Mitgliedstaaten verwendet werden, solange ihre Verwendung zur Aufdeckung, Verhütung, Ermittlung oder Verfolgung von Straftaten gesetzlich zugelassen ist und wenn geeignete Schutzvorkehrungen für die Rechte und Freiheiten Dritter bestehen. Die Ausnahme setzt nicht voraus, dass das KI-System in der Lage ist, die Ermittlungen unabhängig durchzuführen. Wenn ein jeweiliges System sowohl für Strafverfolgungszwecke als auch für andere reguläre Zwecke genutzt werden kann, ist die Anwendbarkeit der Ausnahme im Hinblick auf den jeweiligen Einsatz zu bewerten. Die Transparenzpflichten gelten weiterhin, wenn dasselbe KI-System für andere Zwecke als die gesetzlich zugelassene Aufdeckung, Verhütung, Ermittlung oder Verfolgung von Straftaten eingesetzt wird. Für natürliche Personen muss nach wie vor Transparenz geschaffen werden, wenn sie im Rahmen seiner regulären Zwecke mit dem System interagieren.

49. Die Ausnahme gilt nicht, wenn i) das KI-System der Öffentlichkeit zugänglich ist und ii) das System Einzelpersonen eine Funktion zur Meldung von Straftaten bietet. Solche Systeme können verwendet werden, unterliegen jedoch weiterhin der Transparenzpflicht nach Artikel 50 Absatz 1 der KI-Verordnung, sodass natürliche Personen informiert werden müssen, dass sie mit solchen Systemen interagieren.

Zur Klarstellung: Diese Ausnahme betrifft ausschließlich die Pflicht nach Artikel 50 Absatz 1, natürliche Personen darüber zu informieren, dass sie mit einem KI-System interagieren ([22]). Sie gilt unbeschadet des geltenden Unionsrechts und nationalen Rechts, einschließlich der Verfahrensvorschriften über die Dokumentation, Offenlegung, Zulässigkeit und Bewertung von Beweismitteln.

([22]) Dies gilt auch für die ähnlichen Ausnahmen zugunsten von Strafverfolgungsmaßnahmen gemäß Artikel 50 Absätze 2 bis 4 der KI-Verordnung.

Beispiele für KI-Systeme, die der Öffentlichkeit zur Meldung von Straftaten zur Verfügung gestellt werden, die nicht ausgenommen sind und in den Anwendungsbereich von Artikel 50 Absatz 1 der KI-Verordnung fallen:

— KI-gestützte Polizei-Chatbots, die auf offiziellen Polizei-Websites oder in mobilen Anwendungen eingesetzt werden, die es Einzelpersonen ermöglichen, Straftaten oder Beschwerden zu melden. Diese Systeme führen die Nutzer durch strukturierte Fragen, sammeln relevante Informationen und sichten und priorisieren Berichte, bevor sie sie zur Bewertung an menschliche Beamte weiterleiten.

— KI-gestützte Telefon-Hotlines, die von der Polizei oder einer anderen Strafverfolgungsbehörde für allgemeine öffentliche Untersuchungen betrieben werden.

— KI-gestützte Hotlines für Betrug oder digitale Portale, die von Finanzinstituten oder Behörden betrieben werden und über die Nutzer mutmaßliche Finanzstraftaten melden können. Das KI-System interagiert mit dem Nutzer, um Einzelheiten zu sammeln, die Meldung zu kategorisieren und Fälle zu priorisieren, während menschliche Ermittler die Informationen validieren, bevor Maßnahmen ergriffen werden.

— Virtuelle Assistenten, die für die Sammlung von Zeugenaussagen verwendet werden, bei denen Opfer oder Zeugen Informationen über eine KI-gestützte Schnittstelle übermitteln können. Solche KI-Systeme können mehrsprachige Eingaben, Barrierefreiheitsmerkmale und eine strukturierte Vorlage von Nachweisen unterstützen, die abschließende Bewertung und die Ermittlungsschritte liegen jedoch weiterhin bei den Menschen in den Behörden.

3.3. Zusammenspiel mit anderen Rechtsakten der Union

50. Die Transparenzpflicht nach Artikel 50 Absatz 1 der KI-Verordnung gilt unbeschadet bestehender Rechtsvorschriften der Union in den Bereichen Verbraucherschutz, Datenschutz, Digitales oder Nichtdiskriminierung, die weitere Offenlegungen von Informationen vorschreiben können, die nicht speziell auf den KI-Ursprung der Interaktion ausgerichtet sind. Insbesondere sind nach der Richtlinie 2005/29/EG (im Folgenden „Richtlinie über unlautere Geschäftspraktiken“) ([23]) irreführende Handlungen oder Unterlassungen in Bezug auf die wesentlichen Merkmale eines Produkts oder einer Dienstleistung verboten. Darüber hinaus sind Unternehmer gemäß der Richtlinie 2011/83/EG (im Folgenden „Verbraucherrechte-Richtlinie“) ([24]) verpflichtet, klare und verständliche vorvertragliche Informationen über die wesentlichen Merkmale der Waren oder Dienstleistungen bereitzustellen. Wenn ein Dienst KI-gesteuert ist (z. B. ein abonnierter Chatbot, eine KI-Begleitanwendung oder ein virtueller Coaching-Dienst), kann die KI-Funktionalität als wesentliches Merkmal eingestuft werden, das offengelegt werden muss, bevor der Verbraucher vertraglich gebunden wird. Die Informationspflichten nach dem Verbraucherschutzrecht der Union gelten unabhängig davon, ob die Interaktion nach Artikel 50 Absatz 1 der KI-Verordnung als „offensichtlich“ gilt. Diese Pflichten ergänzen sich auch als allgemeine Informationspflichten im Bereich des Verbraucherschutzes, während die Meldepflicht nach der KI-Verordnung kontextbezogener ist und darauf abzielt, das kontinuierliche Bewusstsein aller natürlichen Personen im Hinblick auf eine direkte Interaktion mit einem KI-System mindestens vor deren Beginn und während der gesamten Interaktion sicherzustellen.

51. Das Gesetz über digitale Dienste enthält auch Transparenzanforderungen für Empfehlungssysteme von Online-Plattformen, einschließlich der Verpflichtung, die wichtigsten Ranking-Parameter zu erläutern. Diese Verpflichtungen ergänzen die in Artikel 50 Absatz 1 der KI-Verordnung festgelegten Transparenzpflichten, da sie einen anderen Anwendungsbereich und Zweck haben und Systeme betreffen, die grundsätzlich nicht direkt mit Nutzern im Sinne des Artikels 50 Absatz 1 der KI-Verordnung interagieren.

52. Was die Informationspflichten gegenüber betroffenen Personen nach dem Datenschutzrecht der Union betrifft, so dient die Pflicht nach Artikel 50 Absatz 1 der KI-Verordnung einem anderen Ziel als diese Pflichten und lässt diese Informationspflichten als solche unberührt.

([23]) Richtlinie 2005/29/EG des Europäischen Parlaments und des Rates vom 11. Mai 2005 über unlautere Geschäftspraktiken im binnenmarktinternen Geschäftsverkehr zwischen Unternehmen und Verbrauchern (ABl. L 149 vom 11.6.2005, S. 2).

([24]) Richtlinie 2011/83/EU des Europäischen Parlaments und des Rates vom 25. Oktober 2011 über die Rechte der Verbraucher (ABl. L 304 vom 22.11.2011, S. 64).

53. Ebenso gilt die Transparenzpflicht nach Artikel 50 Absatz 1 der KI-Verordnung unbeschadet des Unionsrechts und des nationalen Rechts über die Bereitstellung von Informationen an Personen im Rahmen von Strafverfahren (z. B. der Richtlinie 2012/13/EU oder der Richtlinie 2012/29/EU in der Fassung der Richtlinie (EU) 2026/1472) ([25]).

4. **ARTIKEL 50 ABSATZ 2 DER KI-VERORDNUNG: KENNZEICHNUNG UND FESTSTELLUNG VON DURCH KI ERZEUGTEN ODER MANIPULIERTEN INHALTEN**

54. Nach Artikel 50 Absatz 2 der KI-Verordnung müssen Anbieter von KI-Systemen, die synthetische Inhalte erzeugen, technische Lösungen umsetzen, die bestimmte Qualitätsanforderungen an die maschinenlesbare Kennzeichnung und Feststellung der Ausgaben ihrer KI-Systeme erfüllen. Ziel ist es, natürliche Personen in die Lage zu versetzen, durch KI erzeugte oder manipulierte Inhalte von anderen Inhalten (z. B. von Menschen erstellten Inhalten) zu unterscheiden und ihre Herkunft zu überprüfen und so auch zu mehr Integrität und Vertrauen in das Informationsökosystem beizutragen.

55. Die Begünstigten dieser Transparenzpflicht sind daher alle natürlichen Personen, die den durch KI erzeugten oder manipulierten Inhalten wahrscheinlich ausgesetzt sind, sowie wichtige Akteure, die eine wichtige Rolle für das Vertrauen in das Informationsökosystem und für dessen Integrität spielen, wie die zuständigen Marktüberwachungsbehörden und andere zuständige Behörden, unabhängige Forscher, Organisationen der Zivilgesellschaft, Medien, vertrauenswürdige Hinweisgeber, Grundrechtsverteidiger, Anbieter von Online-Plattformen und Online-Suchmaschinen, Rechteinhaber und andere Ersteller und Vertreiber von Inhalten usw.

4.1. **Hauptkomponenten und Konzepte von Artikel 50 Absatz 2 der KI-Verordnung**

56. Artikel 50 Absatz 2 der KI-Verordnung findet Anwendung, wenn mehrere kumulative Bedingungen erfüllt sind:

i) Das System muss als KI-System gelten.

ii) Das KI-System muss in der Lage sein, synthetische Inhalte zu erzeugen oder zu manipulieren.

iii) Der Inhalt muss in einer oder mehreren der folgenden Modalitäten vorliegen: Audio, Bild, Video oder Text.

iv) Das KI-System darf unter keine der folgenden Ausnahmen fallen: 1) das KI-System übt eine unterstützende Funktion für die Standardbearbeitung aus, 2) das KI-System verändert nicht wesentlich die vom Betreiber bereitgestellten Eingabedaten oder deren Semantik oder 3) das KI-System ist zur Aufdeckung, Verhütung, Ermittlung oder Verfolgung von Straftaten gesetzlich zugelassen.

4.1.1. *KI-Systeme, die synthetische Inhalte erzeugen oder manipulieren*

57. Artikel 50 Absatz 2 der KI-Verordnung gilt für KI-Systeme ([26]), einschließlich KI-Systemen mit allgemeinem Verwendungszweck, die synthetische Audio-, Bild-, Video- oder Textinhalte erzeugen. Artikel 50 Absatz 2 der KI-Verordnung deckt eine Untergruppe von Systemen ab, die der Definition des Begriffs „KI-System“ in Artikel 3 Nummer 1 der KI-Verordnung entsprechen und solche Ausgaben erzeugen. Artikel 50 Absatz 2 der KI-Verordnung gilt auch für KI-Systeme, die zur Manipulation von Inhalten verwendet werden.

58. Die Erzeugung von Inhalten bezieht sich auf die Erzeugung von synthetischem Material durch das KI-System (z. B. auf der Grundlage einer menschlichen Aufforderung), z. B. die Erzeugung eines synthetischen Bildes oder eines Lieds. In der Praxis bedeutet dies, dass Artikel 50 Absatz 2 der KI-Verordnung für generative KI-Systeme gelten kann, die mit einer engen Zweckbestimmung zur Erzeugung bestimmter Ausgaben konzipiert wurden. Gleichzeitig gilt Artikel 50 Absatz 2 der KI-Verordnung auch für KI-Systeme, die mehreren Zwecken dienen können, einschließlich KI-Systemen mit allgemeinem Verwendungszweck, und die eine Vielzahl von Arten von Inhalten erzeugen können, sowie für agentengestützte KI-Systeme, sofern sie synthetische Audio-, Bild-, Video- oder Textinhalte erzeugen.

([25]) Richtlinie 2012/13/EU des Europäischen Parlaments und des Rates vom 22. Mai 2012 über das Recht auf Belehrung und Unterrichtung in Strafverfahren (ABl. L 142 vom 1.6.2012, S. 1); Richtlinie 2012/29/EU des Europäischen Parlaments und des Rates vom 25. Oktober 2012 über Mindeststandards für die Rechte, die Unterstützung und den Schutz von Opfern von Straftaten (ABl. L 315 vom 14.11.2012, S. 57).

([26]) Gemäß Definition in Artikel 3 Nummer 1 der KI-Verordnung.

59. Im Gegensatz dazu bezieht sich die Manipulation von Inhalten auf bereits bestehende (synthetische oder nicht synthetische) Inhalte, die von einem KI-System verändert werden (z. B. ein bestehendes Bild oder eine Sprachaufzeichnung, die von einem KI-System gemäß menschlichen Anweisungen manipuliert wird). Damit Inhalte unter Artikel 50 Absatz 2 der KI-Verordnung fallen, reicht es aus, dass sie im Rahmen der in Artikel 50 Absatz 2 der KI-Verordnung festgelegten Ausnahmen (siehe Abschnitt 4.3) durch KI erzeugt oder über die Standardbearbeitung hinaus manipuliert wurden. Darüber hinaus schreibt Artikel 50 Absatz 2 der KI-Verordnung nicht vor, dass die Inhalte ausschließlich durch KI erzeugt oder manipuliert sein müssen. Inhalte, die mit vom Menschen geschaffenem Material gemischt werden, gelten auch als synthetische Inhalte, wenn sie nach einer der in Artikel 50 Absatz 2 der KI-Verordnung genannten Modalitäten manipuliert oder erzeugt werden.

4.1.2. *Modalitäten synthetischer Inhalte im Anwendungsbereich*

60. Artikel 50 Absatz 2 der KI-Verordnung enthält eine erschöpfende Liste der Modalitäten der unter die Transparenzpflicht fallenden synthetischen Inhalte: Audio, Bild, Video oder Text. Dazu gehören auch multimodale Inhalte, d. h. Inhalte, die aus einer Mischung dieser Modalitäten bestehen. Für die Zwecke des Artikels 50 Absatz 2 der KI-Verordnung sind die verschiedenen Modalitäten im Einklang mit den praktischen und technologischen Entwicklungen auszulegen, die sich im Laufe der Zeit entwickeln können. Vorerst sind sie wie folgt zu verstehen:

— Text bezieht sich auf diskrete symbolische Inhalte, die aus Zeichen oder Zahlen bestehen, die in einer bestimmten Reihenfolge angeordnet sind und von Menschen semantisch gelesen und interpretiert werden können.

— Bilder beziehen sich auf statische räumliche Darstellungen, die visuelle Informationen codieren und vom Menschen gesehen werden können.

— Audio bezeichnet zeitvariierende Signale, die Schall darstellen und vom Menschen gehört werden können. Dies kann Sprache, instrumentale Musik oder andere Audiosignale umfassen.

— Video bezieht sich auf zeitbasierte Sequenzen von Bildern, die möglicherweise mit Audio synchronisiert sind und vom Menschen gesehen werden können. In diesem Fall müssen sowohl Video- als auch Audioinhalte gekennzeichnet werden, und die Erkennbarkeit muss gewährleistet sein, wie nachstehend dargelegt.

61. KI-Systeme, die 3D-Bilder, 3D-Videos und 3D-Audio sowie virtuelle, erweiterte oder gemischte Realität erzeugen oder manipulieren, fallen in den Anwendungsbereich von Artikel 50 Absatz 2 der KI-Verordnung. „Virtuelle Realität“ umfasst KI-Technologien, die es den Nutzern ermöglichen, eine computergenerierte Umgebung zu erleben und mit ihr zu interagieren, in der physische Präsenz simuliert wird, und die es den Nutzern ermöglichen, in Echtzeit mit dieser Umgebung zu interagieren. Da die virtuelle und erweiterte Realität aus einer zeitbasierten Sequenz von Bildern besteht, die optional mit Audio synchronisiert ist und die visuelle Veränderung im Laufe der Zeit darstellt, kann sie als „Video“-Inhalt angesehen werden. Die zu kennzeichnende relevante „Ausgabe“ ist die endgültige Form des einzelnen durch KI erzeugten Vermögenswerts, mit besonderen Ausnahmen für die kurzlebigen interaktiven Sitzungen in Echtzeit unter Randnummer 88. Digitale Artefakte, die zur Erzeugung von 3D-Bildern, 3D-Videos und 3D-Audioinhalten verwendet werden, wie z. B. eine Punktewolke oder ein 3D-Netz, müssen nicht gekennzeichnet werden.

62. Digitale Zwillinge, die eine virtuelle Nachbildung von Personen, physischen Objekten oder Systemen bieten und mehrere Datenarten integrieren, fallen ebenfalls in den Anwendungsbereich von Artikel 50 Absatz 2 der KI-Verordnung, wenn sie durch KI erzeugt oder manipuliert werden und in Form von Audio, Bild, Video oder Text vorliegen, es sei denn, sie fallen unter die Ausnahmen und gelten als „industrielle Anwendungen“ oder „B2B-Anwendungen“ gemäß Randnummer 87 oder als kurzlebige Echtzeitinhalte gemäß Randnummer 88.

63. Artikel 50 Absatz 2 der KI-Verordnung kann auch auf KI-Agenten Anwendung finden, wenn der KI-Agent eine Handlung vornimmt, deren Ausgabe durch KI erzeugte oder manipulierte Inhalte sind, die von natürlichen Personen in Form von Audio, Bild, Video oder Text wahrnehmbar sind. Diese Ausgaben müssen daher wie nachstehend beschrieben gekennzeichnet und erkennbar sein. Im Gegensatz dazu gelten Handlungen von KI-Agenten, die keine durch KI erzeugten oder manipulierten Inhalte umfassen (z. B. Zwischenverarbeitungsschritte wie Argumentation und Denkkette), oder durch KI erzeugte Handlungen, die nicht dazu bestimmt sind, von natürlichen Personen direkt wahrgenommen zu werden (z. B. eine Webanfrage oder eine Browseraktion), nicht als synthetische Inhalte im Sinne von Artikel 50 Absatz 2 der KI-Verordnung und müssen die Kennzeichnungs- und Feststellungspflichten nicht erfüllen.

4.1.3. *Inhalte, die nicht in den Anwendungsbereich von Artikel 50 Absatz 2 der KI-Verordnung fallen*

64. Inhalte, die von einem KI-System auf der Grundlage einer einfachen Datenverarbeitung erzeugt werden und die nicht speziell durch KI erzeugt oder manipuliert sind, z. B. ein gerenderter Frame, fallen nicht in den Anwendungsbereich von Artikel 50 Absatz 2 der KI-Verordnung.

65. Darüber hinaus fällt die Ausgabe eines KI-Systems, das lediglich bestehende Inhalte wiedergibt oder die Darstellung oder Anordnung bereits bestehender Inhalte ermöglicht, nicht unter Artikel 50 Absatz 2 der KI-Verordnung. Dies kann z. B. bei Musikspiellisten oder anderen Empfehlungssystemen der Fall sein, die nur bestehende Inhalte auf der Grundlage von Nutzerpräferenzen oder -aktivitäten auswählen oder einstufen, ohne neue Inhalte zu erstellen oder die Inhalte selbst zu manipulieren. Gleiches gilt für KI-Systeme, die in internen Analyseprozessen verwendet werden, die Daten extrahieren und strukturieren, diese aber nicht zusammenfassen.

66. Gleiches trifft auch auf die Ein- und Ausgabe eines KI-Systems, bei der es sich um bloße Beobachtungen und Aufzeichnungen von Daten aus physischen oder virtuellen Umgebungen handelt, die nicht durch KI erzeugt oder manipuliert (z. B. durch Roboter oder andere KI-gestützte Sensoren) oder Datenübertragungen durch KI-Systeme ohne Veränderung sind. Dies umfasst eine Vielzahl von Ein- und Ausgaben von KI-Systemen, insbesondere in industriellen Umgebungen wie Fertigungsdaten für Roboter, die Daten beobachten, der von KI-gestützten intelligenten Zählern aufgezeichnete Verbrauch, Netzfrequenz- und Spannungsmessungen und die Aufzeichnung von GPS-Standortdaten von Fahrzeugen usw.

67. Da das Ziel der Transparenzpflicht in Artikel 50 Absatz 2 der KI-Verordnung darin besteht, Menschen in die Lage zu versetzen, durch KI erzeugte oder manipulierte Inhalte zu unterscheiden, um den Risiken der Täuschung und Manipulation zu begegnen und die Integrität und das Vertrauen in das Informationsökosystem zu gewährleisten, fallen darüber hinaus alle Inhalte, die nicht mit diesem Ziel in Zusammenhang stehen und die nicht wahrnehmbar sind oder nicht dazu bestimmt sind, von natürlichen Personen ausgelegt, überprüft oder als Handlungsgrundlage verwendet zu werden, nicht unter diese Bestimmung.

68. Insbesondere ist davon auszugehen, dass die folgenden Inhalte nicht in den Anwendungsbereich von Artikel 50 Absatz 2 der KI-Verordnung fallen:

— Ausgaben, die in Form einer kurzen Sequenz von Zahlen, Symbolen oder Buchstaben erzeugt werden (z. B. einzelne Wörter, Bildunterschriften, Alt-Text, UI-Labels und kleinformatige Grafiken, Bildunterschriften und andere Datenkennzeichnungen);

— Quellcode, worunter Inhalte zu verstehen sind, die in einer Programmier-, Skript-, Auszeichnungs-, Abfrage- oder Konfigurationssprache geschrieben sind und dazu bestimmt sind, von einem Rechnersystem interpretiert, zusammengestellt oder ausgeführt zu werden, unabhängig davon, ob er aus Textzeichen besteht. Dies kann auch Kommentare in natürlicher Sprache und Kontextinformationen umfassen, die integraler Bestandteil des Quellcodes sind. Ebenfalls ausgenommen sind Software Development Kits (SDK), Structured Query Language (SQL), Infrastructure-as-Code und „YAML Ain't Markup Language" (YAML), JavaScript Object Notation (JSON), Konfigurationsschemas, Skripte, maschinenlesbare Spezifikationen, Anwendungsprogrammierschnittstellen (API und Softwarebibliotheken);

— Ausgaben eines KI-Systems, die bestimmungsgemäß ausschließlich von Maschine zu Maschine übertragen und automatisch ohne Aussetzung gegenüber Menschen verarbeitet wird, z. B. Kommunikation zwischen Agenten, Spam-Abwehrsignale, solange sie von natürlichen Personen nicht wahrgenommen werden;

— Ausgaben, die nur in geschlossenen Umgebungen in Arbeitsabläufen der Industrie und der Produktentwicklung verwendet werden (z. B. für die Produktion von Filmen, Animationen, Spielen oder Werbung), es sei denn, sie sind die Endausgabe des KI-Systems in Form von durch KI erzeugten oder manipulierten Text-, Audio-, Bild- oder Videoinhalten. In diesen Fällen muss nur die Endausgabe der Arbeitsablaufstufen gekennzeichnet und erkennbar sein.

4.2. **Die Kennzeichnungs- und Feststellungspflicht nach Artikel 50 Absatz 2 der KI-Verordnung**

69. Artikel 50 Absatz 2 der KI-Verordnung enthält eine Transparenzpflicht für Anbieter generativer KI-Systeme, die in den Anwendungsbereich dieser Bestimmung fallen, die zwei unterschiedliche, aber inhärent miteinander verknüpfte Elemente umfasst. Erstens müssen die Anbieter sicherstellen, dass die Ausgaben des KI-Systems in einem maschinenlesbaren Format gekennzeichnet werden. Zweitens müssen die Anbieter sicherstellen, dass die Ausgaben als künstlich erzeugte oder manipulierte Inhalte erkennbar sind. Für jede eingesetzte Kennzeichnungslösung müssen die Anbieter sicherstellen, dass entsprechende Mittel für die Feststellung zur Verfügung stehen (wie nachstehend dargelegt), damit natürliche Personen, die den Inhalten ausgesetzt sind, und andere relevante Akteure die künstlich erzeugten Inhalte des KI-Systems erkennen und von anderen Inhalten unterscheiden können. Eine technische Lösung ist als eine Kombination von Kennzeichnungstechniken und Mitteln zur Feststellung zu verstehen, die der Anbieter eingeführt hat, um der Transparenzpflicht nach Artikel 50 Absatz 2 der KI-Verordnung nachzukommen.

70. Jedes der beiden oben genannten Elemente muss erfüllt sein, um die Ziele der Transparenzpflichten nach Artikel 50 Absatz 2 der KI-Verordnung wirksam zu erreichen. Die Erfüllung nur eines Elements (z. B. die maschinenlesbare Kennzeichnung von Ausgaben, ohne dass die Mittel zu ihrer Feststellung zur Verfügung stehen) reicht nicht aus, um dieser Bestimmung nachzukommen ([27]).

4.2.1. *Die Kennzeichnungspflicht*

71. Der Umfang der Kennzeichnungspflicht beschränkt sich auf die Anbringung von Kennzeichnungen in einem maschinenlesbaren Format. Ein maschinenlesbares Format bedeutet, dass Kennzeichnungen so strukturiert sind, dass Softwareanwendungen sie ohne menschliches Eingreifen leicht identifizieren, erkennen und extrahieren können. Wahrnehmbare Markierungen und Kennzeichnungen sind gegebenenfalls nicht als ergänzende Maßnahme ausgeschlossen, um den Betreibern die Einhaltung ihrer Pflicht zur Kennzeichnung von Deepfakes gemäß Artikel 50 Absatz 4 der KI-Verordnung zu erleichtern.

72. Um der Kennzeichnungspflicht nachzukommen, können Anbieter auf eine einzige Kennzeichnungstechnik oder eine Kombination von Techniken zurückgreifen, sofern ihre gesamte technische Lösung maschinenlesbar ist und die Anforderungen an Wirksamkeit, Interoperabilität, Belastbarkeit und Zuverlässigkeit bis zu dem gesetzlich vorgeschriebenen Grad erfüllt, wie in Artikel 50 Absatz 2 Satz 2 der KI-Verordnung vorgesehen (siehe Abschnitt 4.2.3).

73. In Erwägungsgrund 133 der KI-Verordnung sind Beispiele für solche Techniken aufgeführt, darunter Wasserzeichen, Metadatenidentifizierungen, kryptografische Methoden zum Nachweis der Herkunft und Authentizität des Inhalts, Protokollierungsmethoden, Fingerabdrücke oder andere Techniken sowie eine Kombination solcher Techniken. Zwar werden Methoden zum Nachweis der Herkunft und Authentizität in Erwägungsgrund 133 der KI-Verordnung erwähnt, doch sind Anbieter nicht verpflichtet, eine vollständige Herkunftskette mit Informationen über den Ursprung des Inhalts und Änderungen oder andere relevante Aussagen über den Verlauf des Inhalts zu erfassen oder aufzubewahren. Solche Herkunftsmethoden können jedoch auch zur Einhaltung von Artikel 50 Absatz 2 der KI-Verordnung verwendet werden und dazu beitragen, dass natürliche Personen durch KI erzeugte oder manipulierte Inhalte von authentischen Inhalten unterscheiden können.

74. Anbieter können die Kennzeichnungslösung für die Ausgabe auf verschiedenen Stufen der Wertschöpfungskette umsetzen (z. B. nach der Erzeugung von Inhalten auf der Ebene des KI-Systems, auf der Ebene des zugrunde liegenden KI-Modells oder integriert in den Inferenzprozess des KI-Systems). Sie können auf die Kennzeichnungslösung zurückgreifen, die von einem vorgelagerten Modellanbieter oder einem Dritten, der die Lösung bereitstellt, umgesetzt wird (z. B. als offener Standard oder spezialisierter Dienst), sofern die Kennzeichnungslösung mit Artikel 50 Absatz 2 der KI-Verordnung im Einklang steht. Dieser Rückgriff lässt die Verantwortung des Anbieters des KI-Systems für den Nachweis der Einhaltung von Artikel 50 Absatz 2 der KI-Verordnung unberührt.

([27]) In Erwägungsgrund 133 der KI-Verordnung wird präzisiert, dass es angemessen ist, die Anbieter dieser Systeme zu verpflichten, technische Lösungen zu integrieren, die die Kennzeichnung in einem maschinenlesbaren Format und die Feststellung ermöglichen, dass die Ausgabe von einem KI-System und nicht von einem Menschen erzeugt oder manipuliert wurde. In Erwägungsgrund 135 der KI-Verordnung wird auf die Pflicht in Bezug auf die Feststellung verwiesen.

4.2.2. *Die Feststellungspflicht*

75. Anbieter von KI-Systemen, die in den Anwendungsbereich von Artikel 50 Absatz 2 der KI-Verordnung fallen, müssen sicherstellen, dass die Ausgaben ihrer Systeme als durch KI erzeugte oder manipulierte Inhalte erkennbar sind. Dies bedeutet, dass der Anbieter verpflichtet ist, dafür zu sorgen, dass die Mittel zur Feststellung den Personen zur Verfügung stehen, die den Inhalten potenziell ausgesetzt sind. Darüber hinaus müssen solche Feststellungslösungen gemäß Artikel 50 Absatz 5 der KI-Verordnung in der Lage sein, vom Menschen lesbare Ergebnisse zu erzielen, aus denen hervorgeht, ob die Inhalte durch KI erzeugt oder manipuliert wurden.

76. Ein Feststellungsinstrument ist ein Mechanismus, der erkennt, ob Inhalte durch KI erzeugt oder manipuliert wurden, und in der Regel technische Marker oder Signaturen identifiziert, die ihre Herkunft bestätigen. Um vollständige Interoperabilität zu gewährleisten, müssen sich die Anbieter auf öffentlich zugängliche branchenübliche Feststellungslösungen stützen, die es Dritten ermöglichen, Feststellungsmechanismen einzuführen, die idealerweise lokal auf dem digitalen Gerät ausführbar sind und eine nahtlose Integration in eigenständige, isolierte Netze ermöglichen. Stehen solche Normen nicht zur Verfügung, insbesondere in der Anfangsphase der Umsetzung von Artikel 50 Absatz 2 der KI-Verordnung für Wasserzeichentechnologien, kann sich der Anbieter auf seine eigene Feststellungslösung oder auf eine Feststellungslösung Dritter oder eine gemeinsame Feststellungslösung stützen, sofern diese Lösungen die Interoperabilität mit den von anderen Anbietern verwendeten Feststellungslösungen gewährleisten (siehe Abschnitt 4.2.3). Diese Möglichkeit sollte zeitlich begrenzt sein, bis harmonisierte Normen und eine standardisierte anbieterunabhängige interoperable Feststellungslösung, die sicher, datenschutzfreundlich und lokal ausführbar ist, entstehen.

77. In Artikel 50 Absatz 5 der KI-Verordnung ist ferner festgelegt, dass die in Artikel 50 Absatz 2 der KI-Verordnung genannten Informationen den natürlichen Personen, die den Inhalten ausgesetzt sind, spätestens zum Zeitpunkt der ersten Interaktion oder Aussetzung in klarer und eindeutiger Weise bereitzustellen sind. Diese Informationen sind insbesondere das Ergebnis der Feststellung, das zeigt, ob es sich um durch KI erzeugte oder manipulierte Inhalte handelt, und nicht die durch KI erzeugten oder manipulierten Inhalte selbst. Unter „zum Zeitpunkt der ersten Aussetzung" ist der Zeitpunkt zu verstehen, zu dem natürliche Personen bereit sind, die Herkunft eines bestimmten Inhalts zu überprüfen und auf die Feststellungslösung zuzugreifen, um die Ergebnisse der Feststellung zu erhalten.

78. Wenn der Anbieter des KI-Systems bei der Bereitstellung einer Feststellungslösung oder dem Zugang zu einer Feststellungslösung von einem anderen Akteur abhängig ist, muss er sicherstellen, dass die Feststellungslösung mit Artikel 50 Absätze 2 und 5 der KI-Verordnung im Einklang steht und dass diese Akteure die Ergebnisse der Feststellung im Einklang mit Artikel 50 Absatz 5 der KI-Verordnung anzeigen (siehe Abschnitt 7).

4.2.3. *Einhaltung der Anforderungen an technische Lösungen: wirksam, interoperabel, belastbar und zuverlässig*

79. Nach Artikel 50 Absatz 2 Satz 2 der KI-Verordnung müssen die technischen Lösungen für die Kennzeichnung und Feststellung wirksam, interoperabel, belastbar und zuverlässig sein. Diese Anforderungen sind wie folgt zu verstehen:

— **Wirksamkeit** bezieht sich auf die Fähigkeit der von den Anbietern implementierten technischen Lösung, ihre Markierungen zu erkennen und natürliche Personen in die Lage zu versetzen, von ihrem KI-System künstlich erzeugte oder manipulierte Inhalte zu unterscheiden und so zum Vertrauen in das Informationsökosystem und zu dessen Integrität beizutragen.

— **Zuverlässigkeit** bezieht sich auf die Fähigkeit der technischen Lösung, durch KI erzeugte oder manipulierte Inhalte unter Nennbedingungen in der Vielfalt der vom KI-System des Anbieters erzeugten oder manipulierten Inhalte genau zu erkennen und von anderen Inhalten zu unterscheiden.

— **Belastbarkeit** bezieht sich auf die Fähigkeit der technischen Lösung, durch KI erzeugte oder manipulierte Inhalte, die vom KI-System des Anbieters erzeugt werden, unter unterschiedlichen Bedingungen genau zu erkennen und von anderen zu unterscheiden, was sowohl häufige Veränderungen als auch feindliche Angriffe umfasst.

— **Interoperabilität** bezieht sich auf die Fähigkeit der verschiedenen technischen Lösungen für die Kennzeichnung und Feststellung, nahtlos über mehrere Systeme, Akteure, Kontexte und technische Implementierungen hinweg zu funktionieren, um die Feststellung von durch KI erzeugten oder manipulierten Inhalten zu ermöglichen, und zwar unabhängig von den Kennzeichnungstechniken, die von Anbietern verschiedener KI-Systeme eingesetzt werden.

80. Die Anbieter müssen sicherstellen, dass die Kombination der technischen Lösungen für die Kennzeichnung und Feststellung alle oben genannten Anforderungen in dem gesetzlich vorgeschriebenen Umfang ganzheitlich erfüllt, wobei die Grenzen und die Komplementarität der verschiedenen eingesetzten Lösungen zu berücksichtigen sind. Die betreffenden technischen Lösungen müssen – soweit technisch möglich – alle Anforderungen erfüllen und die Besonderheiten und Beschränkungen der verschiedenen Arten von Inhalten, die Umsetzungskosten und den allgemein anerkannten Stand der Technik, wie er in den einschlägigen technischen Normen zum Ausdruck kommen kann, berücksichtigen.

81. „Technisch mögliche" Lösungen im Sinne des Artikels 50 Absatz 2 der KI-Verordnung sind Lösungen, die für die in den Anwendungsbereich der Bestimmung fallenden Modalitäten unter Verwendung der derzeit verfügbaren Technologien, Methoden und technischen Verfahren innerhalb der spezifischen technischen Architektur und des betreffenden operativen Umfelds umgesetzt werden können. Der Anbieter ist nicht verpflichtet, eine technische Lösung zu nutzen, die noch nicht entwickelt oder auf dem Markt verfügbar oder deren Umsetzung technisch nicht möglich ist. Die technische Möglichkeit ist ein objektiver Begriff, der nicht von den spezifischen Ressourcen und Fähigkeiten der einzelnen Anbieter abhängt.

82. Nach Artikel 50 Absatz 2 der KI-Verordnung müssen Anbieter technisch mögliche Lösungen umsetzen, die dem allgemein anerkannten Stand der Technik entsprechen, um die vier Qualitätsanforderungen an Wirksamkeit, Belastbarkeit, Zuverlässigkeit und Interoperabilität zu erfüllen.

83. Der „Stand der Technik" ist zu verstehen als ein zu einem bestimmten Zeitpunkt entwickelter Stand des technischen Könnens in Bezug auf Produkte, Verfahren und Dienstleistungen, der auf den einschlägigen gesicherten Erkenntnissen von Wissenschaft, Technik und Erfahrung beruht und als bewährtes Verfahren in der Technik anerkannt ist ([28]). Stand der Technik bedeutet nicht unbedingt, dass es sich um Ergebnisse der neuesten wissenschaftlichen Forschung handelt, die sich noch im Versuchsstadium befinden oder technologisch noch nicht ausgereift sind. Die Anbieter müssen ihre Kennzeichnungs- und Feststellungslösungen kontinuierlich, zeitnah und verhältnismäßig an die Weiterentwicklung der Technologie und des Stands der Technik anpassen.

84. Nach Artikel 50 Absatz 7 der KI-Verordnung können sich Anbieter auf einen Praxisleitfaden stützen, der gemäß dieser Bestimmung als angemessen bewertet wird, um die Einhaltung der Kennzeichnungs- und Feststellungspflichten nach Artikel 50 Absätze 2 und 5 der KI-Verordnung nachzuweisen, und der an den Stand der Technik angepasst ist. Weitere Informationen über die Auswirkungen der Einhaltung eines solchen Leitfadens und die Überwachung anderer Anbieter, die sich nicht an einen Leitfaden halten, sind Abschnitt 8.1 zu entnehmen.

85. Die Kosten für die Umsetzung bestimmter technischer Lösungen für die Kennzeichnung und Feststellung können ebenfalls berücksichtigt werden. In bestimmten Fällen, wie in den folgenden Punkten beschrieben, können die Kosten in keinem Verhältnis zu marginalen Gewinnen stehen, die einen geringeren Wert haben, wenn es darum geht, Menschen in die Lage zu versetzen, durch KI erzeugte oder manipulierte Inhalte zu unterscheiden, und die Integrität des Informationsökosystems und das Vertrauen in dieses Ökosystem zu gewährleisten.

86. Es kann daher eng definierte Fälle geben, in denen eine technische Lösung auf der Grundlage weniger robuster Metadatenkennzeichnungen ausreicht, um Artikel 50 Absatz 2 der KI-Verordnung zu entsprechen, wobei die Besonderheiten der Ausgaben des KI-Systems zu berücksichtigen sind, wenn sie von Natur aus ein geringeres Risiko der Täuschung, Manipulation oder negativer Auswirkungen auf das Informationsökosystem bergen. Dies ist insbesondere der Fall, wenn ein generatives KI-System in physische Produkte eingebettet ist, die Ausgaben in einer technisch kontrollierten und geschlossenen Umgebung erzeugen, die hauptsächlich instruktiven Charakter hat (z. B. ein KI-System, das in Navigationssysteme in Fahrzeugen eingebettet ist), soweit wirksame technische Maßnahmen in das Produkt eingebettet sind, um zu verhindern, dass die Ausgaben die Produktumgebung verlassen, z. B. indem sie erfasst oder exportiert und extern verbreitet werden.

([28]) Siehe Anhang II des Durchführungsbeschlusses C(2025) 3871 final der Kommission über einen Normungsauftrag an das Europäische Komitee für Normung und das Europäische Komitee für elektrotechnische Normung in Bezug auf Hochrisiko-KI-Systeme zur Unterstützung der Verordnung (EU) 2024/1689 des Europäischen Parlaments und des Rates und zur Aufhebung des Durchführungsbeschlusses C(2023) 3215 final.

87. Es kann auch begrenzte Fälle von „industriellen KI-Anwendungen“ oder „Business-to-Business-Anwendungen“ (mit Ausnahme öffentlicher und verbraucherorientierter KI-Systeme) geben, in denen der Grundsatz der Verhältnismäßigkeit und der begrenzte Nutzen für die Transparenzziele bedeuten, dass keine Kennzeichnung und Feststellung von durch KI erzeugten oder manipulierten Ausgaben erforderlich ist. Dies gilt nur, wenn die folgenden Voraussetzungen kumulativ erfüllt sind:

i) Die vom KI-System erzeugten Ausgaben sind rein technischer Natur, z. B. Konstruktionsentwürfe, Arbeitsabläufe in der industriellen Produktion, technische Anweisungen, Ausgaben, die infolge prädiktiver Systemwartungsprozesse in industriellen Umgebungen generiert werden, interne Dokumentationsprozesse oder Produktionsschritte und Arbeitsabläufe, bevor die Ausgaben fertiggestellt und anderen externen Personen oder der Öffentlichkeit zur Verfügung gestellt werden.

ii) Die vom KI-System erzeugten Ausgaben sollen nur von einer begrenzten, vorab festgelegten Zahl natürlicher Personen wahrgenommen und verarbeitet werden, die in beruflicher Funktion innerhalb der Organisation des Anbieters und des Betreibers handeln.

iii) Die vom KI-System erzeugten Ausgaben dürfen nicht dazu bestimmt sein, außerhalb des Unternehmens weitergegeben zu werden oder von externen Personen verwendet zu werden, wobei geeignete Schutzvorkehrungen getroffen werden müssen, um eine vernünftigerweise vorhersehbare Fehlanwendung zu verhindern (z. B. Cloud-Isolation, rollenbasierte Kontrollen).

88. Die Erzeugung von Inhalten in Echtzeit, die kurzlebig sind und sofort konsumiert werden, ohne dass sie aufgezeichnet, gespeichert oder weiter verbreitet werden (z. B. in Videospielen, Virtual-Reality-Anwendungen), kann ebenfalls ausgenommen werden, wenn die Kennzeichnung technisch nicht möglich ist und die Personen, die den Inhalten ausgesetzt sind, darauf hingewiesen werden, dass die Inhalte durch KI erzeugt oder manipuliert sind (z. B. Offenlegung innerhalb der Nutzungserfahrung, Benachrichtigungen auf Sitzungsebene).

4.3. **Ausnahmen von den Pflichten nach Artikel 50 Absatz 2 der KI-Verordnung**

89. Artikel 50 Absatz 2 der KI-Verordnung sieht drei ausdrückliche Ausnahmen von den in dieser Bestimmung festgelegten Transparenzpflichten vor.

90. Die erste Ausnahme betrifft KI-Systeme, soweit sie eine unterstützende Funktion für die Standardbearbeitung ausführen. Unter Standardbearbeitung ist der Prozess der Vorbereitung vorhandener Inhalte für die Veröffentlichung oder Verbreitung (z. B. kleine Bearbeitungen zur Verbesserung der Lesbarkeit und der Grammatik, der Qualität und des Formats) zu verstehen, bei dem keine neuen Inhalte erzeugt werden. Durch Standardbearbeitung soll sichergestellt werden, dass die Inhalte unter anderem frei von offensichtlichen technischen oder grammatikalischen sprachlichen Fehlern sind, mit den geltenden Anforderungen an Layout, Präsentation, Formatierung oder Barrierefreiheit in Einklang stehen oder sektorspezifischen Gepflogenheiten entsprechen. Die Bearbeitung geht über die Standardbearbeitung hinaus, wenn der Inhalt erheblich geändert wird (wesentliche Änderungen, strukturelle Änderungen usw.), was sich auf seine Bedeutung, seinen Stil oder seine Absicht auswirkt.

91. Die zweite Ausnahme gilt, wenn ein KI-System die vom Betreiber bereitgestellten Eingabedaten oder deren Semantik nicht wesentlich verändert. Eine Veränderung sollte als wesentlich gelten, wenn die Eingabedaten oder ihre Semantik während des Prozesses der Ausgabeerzeugung durch das KI-System erheblich manipuliert wurden. Dies muss auf einer Bewertung relevanter Faktoren beruhen, wie Format, Art der Medieninhalte, Stil und Änderungen der Inhalte, die sich auf ihre Bedeutung, ihren Stil oder ihre Absicht auswirken. Festzustellen, ob dies der Fall ist, bedarf einer Einzelfallprüfung.

92. Wenn ein KI-System sowohl für die Erzeugung oder Manipulation von Inhalten als auch für nicht wesentliche geringfügige Änderungen von Eingabedaten verwendet werden kann, gilt die Transparenzpflicht nach Artikel 50 Absatz 2 der KI-Verordnung nicht für Inhalte, die geringfügig und nicht wesentlich verändert wurden.

Beispiele für Standardbearbeitung und geringfügige Änderungen, für die die Ausnahmeregelung nach Artikel 50 Absatz 2 der KI-Verordnung gilt:

— Grammatikkorrektur und Rechtschreibprüfung, sprachliche und geringfügige stilistische Überarbeitungen, die den Inhalt, die Bedeutung, den Stil oder die Botschaft von Texten nicht verändern, durch KI erzeugte Übersetzungen von Texten,

— Formatierung, Formatumwandlung, technische Komprimierung, Reduzierung irrelevanter Inhalte oder Entfernung im Interesse größerer Klarheit, ohne die Bedeutung des Inhalts oder den wesentlichen Inhalt zu verändern,

— Geringfügiges Zuschneiden, geringfügige Farbanpassungen oder -korrekturen, Aufhellung oder Verdunkelung, Schärfung zur Verbesserung der Klarheit oder sonstige Standardbearbeitung und technische Korrekturen,

— Entfernung von Staubstellen durch eine schmutzige Linse oder einen schmutzigen Sensor, Entfernung von durch Blitzlicht verursachten roten Augen; Löschung und Verschleierung von Hintergründen, die in der Originaldatei sichtbar sind, Pixelierung oder Verwischung von Gesichtern,

— Neuskalierung eines Videoclips, Anpassung von Lautstärke und Klang; begrenzte Videostabilisierung,

— Standardanpassungen von Farbe und Kontrast, geringfügige Anpassungen der Wiedergabegeschwindigkeit, geringfügige Korrekturen des Horizonts eines Bildes, Anwendung von Pixelfiltern zur Verstärkung bestimmter Teile eines Bildes oder Anwendung von Farbkarten auf Graustufenbilder, Fertigstellung von Randbildern, Umwandlung eines Schwarz-Weiß-Bildes in Farbbilder oder Videos und umgekehrt,

— Pixelfüllung zur Anpassung des Seitenverhältnisses oder zur Verbesserung der Klarheit, automatische Übergangsclips, andere nicht substanzielle Bearbeitungen zur optischen und technischen Verbesserung,

— Medizinische technische Bildverarbeitung oder -rekonstruktion innerhalb von medizinischen Geräten (z. B. Anmerkungen, Kästen, Konturen, Kreise, Heatmaps),

— Transkriptionen von Gesprächen; durch KI erzeugte Inhalte, die authentische menschliche Eingaben nur durch assistive Technologien transformieren, die es Menschen mit Behinderungen ermöglichen, zu kommunizieren (z. B. unterstützte Kommunikation (Augmentative and Alternative Communication, AAC) oder benutzerdefinierte neurale Stimmen (Custom Neural Voices, CNV)), da sie die Bedeutung der Inhalte nicht semantisch verändern.

Beispiele für semantische Änderungen, die eine Kennzeichnung nach Artikel 50 Absatz 2 der KI-Verordnung erfordern:

— Durch KI erzeugte Zusammenfassungen von Texten; Umformulierung oder Neuformulierung von Texten, wodurch Stil, Struktur und Bedeutung über eine rein grammatikalische und geringfügige stilistische Korrektur hinaus verändert werden,

— Entfernung, Austausch oder Einfügung von Gegenständen oder Personen in bestehende Bilder und Videos, wodurch die Bedeutung des Inhalts oder der wesentliche Inhalt verändert wird; Austausch des Gesichts oder wesentliche Änderung des Gesichts,

— Synthese realistischer Äußerungen in der Stimme einer bestimmten Person oder Erstellung realistischer Videos, in denen Ereignisse dargestellt werden, die nicht stattgefunden haben; Veränderung der Körperform oder der Hautfarbe einer Person; extreme Aufhellung, Verdunkelung, Farb- und Kontraständerungen, die die Bedeutung, die Absicht und die Botschaft des Inhalts verändern,

— Erstellung von zusammengesetzten Bildern oder Videoclips, wodurch die Darstellung von Personen, Gegenständen, Ereignissen oder Tatsachen verändert wird, sowie jede andere wesentliche Änderung des Inhalts.

93. Die dritte Ausnahme gilt für generative KI-Systeme, die gesetzlich zur Erzeugung oder Manipulation synthetischer Inhalte zur Aufdeckung, Verhütung, Ermittlung oder Verfolgung von Straftaten zugelassen sind. Die unter den Randnummern 46 bis 48 für diese Ausnahme enthaltenen Erläuterungen sind auch in diesem Fall relevant.

4.4. Zusammenspiel mit anderen Rechtsakten der Union

94. Die Kennzeichnungs- und Feststellungspflichten nach Artikel 50 Absatz 2 der KI-Verordnung konzentrieren sich darauf, wie die Inhalte erstellt wurden und welchen künstlichen Ursprung sie haben, und nicht darauf, wer die Inhalte erstellt hat. Alle von Anbietern von KI-Systemen verwendeten Kennzeichnungs- und Feststellungslösungen müssen daher mit dem geltenden Datenschutzrecht der Union im Einklang stehen, einschließlich der Datenschutzgrundsätze und -pflichten, wie der Verpflichtung zum Datenschutz durch Technikgestaltung und durch datenschutzfreundliche Voreinstellungen, der Datenminimierung und der Grundsätze der Sicherheit und Vertraulichkeit. So sollten beispielsweise zum Zwecke der Kennzeichnung und Feststellung die Informationen über den Urheber der Inhalte nicht verarbeitet und alle anderen verarbeiteten personenbezogenen Daten sollten gelöscht werden, sobald der Zweck der Feststellung erfüllt ist, die Sicherheit und Vertraulichkeit müssen gewährleistet sein usw.

95. In Erwägungsgrund 136 der KI-Verordnung wird ferner die besondere Bedeutung der Transparenzpflichten nach Artikel 50 Absätze 1, 2 und 4 der KI-Verordnung hervorgehoben, um die wirksame Umsetzung des Gesetzes über digitale Dienste zu erleichtern. Dies gilt insbesondere für die Pflicht der Anbieter sehr großer Online-Plattformen oder sehr großer Online-Suchmaschinen, systemische Risiken sorgfältig zu ermitteln, zu analysieren, zu bewerten und zu mindern, die aus der Verbreitung von künstlich erzeugten oder manipulierten Inhalten entstehen können, insbesondere das Risiko tatsächlicher oder vorhersehbarer negativer Auswirkungen auf demokratische Prozesse, den gesellschaftlichen Diskurs und Wahlprozesse, unter anderem durch Desinformation (Artikel 34 und 35 des Gesetzes über digitale Dienste). Maschinenlesbare Kennzeichnungen können es diesen Anbietern ermöglichen, durch KI erzeugte oder manipulierte Inhalte zu erkennen und eine angemessene Kennzeichnung und Offenlegung bereitzustellen, wodurch das Risiko von durch KI erzeugter Desinformation in ihren Diensten verringert wird.

96. Die Pflichten zur Markierung und Kennzeichnung von durch KI-Systeme erzeugten oder manipulierten Inhalten gemäß Artikel 50 Absätze 2 und 4 der KI-Verordnung berühren nicht die Pflicht in Artikel 16 Absatz 6 des Gesetzes über digitale Dienste für Anbieter von Hostingdiensten, gemäß Artikel 16 Absatz 1 des Gesetzes über digitale Dienste eingegangene Meldungen über illegale Inhalte zu bearbeiten ([29]). Insbesondere sollte die Kennzeichnung von durch KI erzeugten oder manipulierten Inhalten nicht die Beurteilung der Rechtswidrigkeit der spezifischen Inhalte und die Entscheidung darüber nach anderen Rechtsrahmen beeinflussen. Diese Beurteilung sollte ausschließlich anhand der Vorschriften für die Rechtmäßigkeit der Inhalte nach diesen Rechtsrahmen vorgenommen werden ([30]). Wenn beispielsweise ein gekennzeichnetes Deepfake-Bild als potenzielles Material über sexuellen Kindesmissbrauch gekennzeichnet wird oder gegen das Marken- oder Urheberrecht verstößt, hat die Anbringung einer Kennzeichnung keinen Einfluss auf die potenzielle Rechtswidrigkeit der Inhalte nach diesen Rechtsrahmen. Diese Bewertung sollte ausschließlich auf der Grundlage der jeweils geltenden Rechtsvorschriften (z. B. Strafrecht, Marken- oder Urheberrechtsvorschriften) durchgeführt werden.

97. Umgekehrt kann die maschinenlesbare Kennzeichnung, die gemäß Artikel 50 Absatz 2 der KI-Verordnung in die Inhalte eingebettet ist, Anbietern sehr großer Online-Plattformen und sehr großer Online-Suchmaschinen dabei helfen, ihren Verpflichtungen gemäß den Artikeln 34 und 35 des Gesetzes über digitale Dienste nachzukommen, systemische Risiken zu ermitteln, zu analysieren und zu bewerten, die sich aus der Verbreitung künstlich erzeugter oder manipulierter Inhalte ergeben können, insbesondere das Risiko tatsächlicher oder vorhersehbarer negativer Auswirkungen auf demokratische Prozesse, den gesellschaftlichen Diskurs und Wahlprozesse, auch durch Desinformation ([31]), und angemessene, verhältnismäßige und wirksame Maßnahmen zur Minderung dieser Risiken zu ergreifen, die auf die ermittelten besonderen systemischen Risiken zugeschnitten sind. Anbieter sehr großer Online-Plattformen und sehr großer Online-Suchmaschinen können auch als Anbieter von KI-Systemen gelten, falls sie die einschlägigen Bedingungen der KI-Verordnung erfüllen, wenn sie KI-Systeme in ihre Dienste integrieren (siehe auch Abschnitt 6.1.5).

98. Anbieter von Online-Plattformen, Online-Suchmaschinen und andere einschlägige Akteure, die Inhalte entlang der Wertschöpfungskette verbreiten, werden insbesondere ermutigt, die von den Anbietern gemäß Artikel 50 Absatz 2 der KI-Verordnung eingeführten Kennzeichnungen zu bewahren und ihre Identifizierung zu ermöglichen, damit die natürlichen Personen, die den Inhalten ausgesetzt sind, über ihre Herkunft informiert werden können.

5. **ARTIKEL 50 ABSATZ 3 DER KI-VERORDNUNG: EMOTIONSERKENNUNGSSYSTEME UND SYSTEME ZUR BIOMETRISCHEN KATEGORISIERUNG**

99. Nach Artikel 50 Absatz 3 der KI-Verordnung sind Betreiber von Emotionserkennungssystemen und Systemen zur biometrischen Kategorisierung verpflichtet, natürliche Personen, die von diesen Systemen betroffen sind, über den Betrieb des Systems zu informieren.

100. In Erwägungsgrund 132 der KI-Verordnung wird erläutert, dass der Zweck dieser Pflicht darin besteht, sicherzustellen, dass natürliche Personen (einschließlich Menschen mit Behinderungen) sich bewusst sind, dass sie Emotionserkennungssystemen und Systemen zur biometrischen Kategorisierung ausgesetzt sind, die in ihre Privatsphäre eingreifen können. Die Pflicht gilt daher unabhängig davon, ob die Personen solchen Systemen in Echtzeit ausgesetzt sind oder ob diese Systeme nachträglich betrieben werden.

([29]) Siehe auch Erwägungsgrund 11 der KI-Verordnung.
([30]) Siehe auch Erwägungsgrund 137 der KI-Verordnung.
([31]) Erwägungsgrund 136 der KI-Verordnung.

5.1. **Hauptkomponenten, Konzepte und damit verbundene Transparenzpflichten nach Artikel 50 Absatz 3 der KI-Verordnung**

5.1.1. *Der Begriff „Emotionserkennungssystem“*

101. In Artikel 3 Nummer 39 der KI-Verordnung wird ein „Emotionserkennungssystem“ definiert als „ein KI-System, das dem Zweck dient, Emotionen oder Absichten natürlicher Personen auf der Grundlage ihrer biometrischen Daten festzustellen oder daraus abzuleiten“. Die ausführliche Erläuterung dieses Begriffs in den Leitlinien der Kommission über die Einstufung von Hochrisiko-KI-Systemen, einschließlich der darin enthaltenen praktischen Beispiele, ist im Zusammenhang mit Artikel 50 Absatz 3 der KI-Verordnung gleichermaßen relevant.

102. Da alle Emotionserkennungssysteme auch als hochriskant eingestuft werden, sofern sie nicht nach Artikel 5 Absatz 1 Buchstabe f der KI-Verordnung am Arbeitsplatz und in Bildungseinrichtungen verboten sind, sollte die Transparenzpflicht nach Artikel 50 Absatz 3 der KI-Verordnung in Verbindung mit den anderen Schutzvorkehrungen und Anforderungen gelten, die für Hochrisiko-KI-Systeme gelten.

5.1.2. *Der Begriff „System zur biometrischen Kategorisierung“*

103. In Artikel 3 Nummer 40 der KI-Verordnung wird ein „System zur biometrischen Kategorisierung“ definiert als „ein KI-System, das dem Zweck dient, natürliche Personen auf der Grundlage ihrer biometrischen Daten bestimmten Kategorien zuzuordnen, sofern es sich um eine Nebenfunktion eines anderen kommerziellen Dienstes handelt und aus objektiven technischen Gründen unbedingt erforderlich ist“. Die ausführliche Erläuterung dieses Begriffs in den Leitlinien der Kommission über die Einstufung von Hochrisiko-KI-Systemen, einschließlich seiner spezifischen Elemente sowie der darin enthaltenen praktischen Beispiele, ist im Zusammenhang mit Artikel 50 Absatz 3 der KI-Verordnung gleichermaßen relevant.

104. Wichtig ist, dass Artikel 50 Absatz 3 der KI-Verordnung, sofern nicht das Verbot nach Artikel 5 Absatz 1 Buchstabe g der KI-Verordnung gilt, für alle Systeme zur biometrischen Kategorisierung (z. B. Alters- oder Geschlechtsklassifizierung auf der Grundlage biometrischer Daten) gilt, unabhängig davon, ob sie in den Anwendungsbereich von Hochrisiko-KI-Systemen gemäß der KI-Verordnung fallen oder nicht.

5.1.3. *Informationspflicht nach Artikel 50 Absatz 3 der KI-Verordnung*

105. **Umfang der bereitzustellenden Informationen**: Die Betreiber sind verpflichtet, die betroffenen natürlichen Personen darüber zu informieren, dass sie einem im Betrieb befindlichen Emotionserkennungssystem oder System zur biometrischen Kategorisierung ausgesetzt sind. Nach der KI-Verordnung ist es nicht erforderlich, Informationen über die Gründe für den Betrieb des Systems aufzunehmen oder darüber, ob das System auch Informationen für andere Zwecke als die Emotionserkennung oder die biometrische Kategorisierung verarbeitet, die nach dem Datenschutzrecht der Union offengelegt werden müssen (siehe Abschnitt 5.3).

106. **Adressaten der Informationen:** Nach Artikel 50 Absatz 3 der KI-Verordnung müssen die Betreiber alle natürlichen Personen, die dem Betrieb des KI-Systems ausgesetzt sind, einschließlich Kindern, informieren.

107. **Mittel zur Bereitstellung der Informationen:** Die KI-Verordnung schreibt keine spezifischen Mittel für die Bereitstellung der Informationen vor, außer dass die Informationen gemäß Artikel 50 Absatz 5 der KI-Verordnung in klarer und eindeutiger Weise und im Einklang mit den geltenden Barrierefreiheitsanforderungen (die den Bedürfnissen natürlicher Personen mit Behinderungen entsprechen) bereitgestellt werden müssen. Wie natürliche Personen, die dem System ausgesetzt sind, zu informieren sind, kann vom Einsatzort (z. B. virtuelle Spielplattform, stationärer Handel, Bahnhof), den möglichen Adressaten (z. B. Kinder, ältere Menschen, Menschen mit Behinderungen, Kunden), der Dauer und Art der Aussetzung (z. B. langfristig, einmalig, kontinuierlich, intermittierend) und dem möglichen Bestehen einer Beziehung zwischen dem Betreiber und den exponierten Personen (z. B. wenn es einen bestehenden Kommunikationskanal mit der Person gibt) abhängen. Je nach Anwendungsfall können die Informationen schriftlich, durch standardisierte Symbole (auch in elektronischer Form), mündlich oder durch Kombinationen dieser Mittel bereitgestellt werden.

Beispiele für die Unterrichtung natürlicher Personen:

— Ein zentral platziertes Pop-up-Fenster mit einer Onboarding-Nachricht vor dem Start eines Computerspiels, aus der hervorgeht, dass das Gesicht des Spielers aufgezeichnet wird und seine Emotionen erfasst werden.

— Ein sichtbarer Hinweis an jedem möglichen Eingang zu einem Ausstellungsraum mit der Information, dass die Gesichtsbilder der Besucher beim Betreten des Raums erfasst werden, um sie einer bestimmten Altersgruppe zuzuordnen.

108. **Zeitplan für die Bereitstellung der Informationen**: Die Informationen müssen gemäß Artikel 50 Absatz 5 der KI-Verordnung spätestens zum Zeitpunkt der ersten Interaktion mit dem KI-System oder der ersten Aussetzung gegenüber dem KI-System bereitgestellt werden. Gegebenenfalls ist eine vorherige Übermittlung der erforderlichen Informationen nicht ausgeschlossen.

5.2. Ausnahme für Strafverfolgungszwecke

109. Diese Informationspflicht gilt nicht für gesetzlich zur Aufdeckung, Verhütung oder Ermittlung von Straftaten zugelassene Emotionserkennungssysteme und Systeme zur biometrischen Kategorisierung, die im Einklang mit dem Unionsrecht verwendet werden, sofern geeignete Schutzvorkehrungen für die Rechte und Freiheiten Dritter bestehen. Die Erläuterungen zu dieser Ausnahme unter den Randnummern 46 bis 48 sind mit einem Unterschied auch in diesem Fall relevant. Während die anderen Ausnahmen in Artikel 50 der KI-Verordnung ein ausdrückliches Gesetz erfordern, das die nicht-transparente Verwendung von KI-Systemen erlaubt, ermöglicht Artikel 50 Absatz 3 der KI-Verordnung eine solche Verwendung, wenn dies nach den Rechtsvorschriften über die Befugnisse der Strafverfolgungsbehörden zulässig ist, um Emotionserkennungssysteme und Systeme zur biometrischen Kategorisierung zur Aufdeckung, Verhütung oder Ermittlung von Straftaten zu nutzen, ohne die betroffenen Personen zu informieren, vorbehaltlich geeigneter Schutzvorkehrungen.

5.3. Zusammenspiel mit anderen Rechtsakten der Union

110. Die Informationspflicht nach Artikel 50 Absatz 3 der KI-Verordnung an sich macht die Verwendung eines Emotionserkennungssystems und Systems zur biometrischen Kategorisierung nicht rechtmäßig. Ebenso wenig legitimiert sie in die Privatsphäre eingreifende oder diskriminierende Verwendungen, die nach Artikel 5 der KI-Verordnung oder anderen Rechtsvorschriften der Union rechtswidrig und verboten sein könnten. Betreiber müssen die Informationspflicht nach Artikel 50 Absatz 3 der KI-Verordnung zusätzlich zu allen geltenden Informationspflichten nach dem Datenschutzrecht der Union erfüllen. In bestimmten Fällen können Betreiber in Erwägung ziehen, den Informationen, die betroffenen Personen nach dem Datenschutzrecht der Union zur Verfügung gestellt werden, Mitteilungen gemäß Artikel 50 Absatz 3 der KI-Verordnung hinzuzufügen.

6. ARTIKEL 50 ABSATZ 4: KENNZEICHNUNG VON DEEPFAKES UND BESTIMMTEN TEXTVERÖFFENTLICHUNGEN

111. In Artikel 50 Absatz 4 der KI-Verordnung sind zwei getrennte Pflichten für Betreiber generativer KI-Systeme festgelegt, die – außer in bestimmten Fällen – eine klare und unterscheidbare Offenlegung i) von Deepfakes und ii) von durch KI erzeugten oder manipulierten Texten erfordern, die veröffentlicht werden, um die Öffentlichkeit über Angelegenheiten von öffentlichem Interesse zu informieren. Diese Kennzeichnungspflichten gelten zusätzlich und unbeschadet der Pflichten zur maschinenlesbaren Kennzeichnung und Feststellung gemäß Artikel 50 Absatz 2 der KI-Verordnung, die für Anbieter von KI-Systemen gelten, die synthetische Bilder, Video-, Audio- oder Textinhalte erzeugen oder manipulieren.

6.1. Hauptkomponenten, Konzepte und damit verbundene Transparenzpflichten für Deepfakes nach Artikel 50 Absatz 4 der KI-Verordnung

112. Artikel 50 Absatz 4 Unterabsatz 1 der KI-Verordnung findet Anwendung, wenn mehrere Bedingungen erfüllt sind:

i) Das System muss ein KI-System sein.

ii) Es muss von Betreibern für berufliche Zwecke verwendet werden (siehe Abschnitte 2.3 und 2.4.1).

iii) Es muss Bild-, Ton- oder Videoinhalte erzeugen oder manipulieren, bei denen es sich um Deepfakes handelt.

iv) Das KI-System darf nicht unter die Ausnahme für eine gesetzlich zulässige Verwendung zur Aufdeckung, Verhütung, Ermittlung oder Verfolgung von Straftaten fallen.

6.1.1. *Der Begriff „Deepfake“*

113. In Artikel 3 Nummer 60 der KI-Verordnung wird „Deepfake“ definiert als ein „durch KI erzeugter oder manipulierter Bild-, Ton- oder Videoinhalt, der wirklichen Personen, Gegenständen, Orten, Einrichtungen oder Ereignissen ähnelt und einer Person fälschlicherweise als echt oder wahrheitsgemäß erscheinen würde“. Der Unterschied zwischen Manipulation und Erzeugung wurde in Abschnitt 4.1 erläutert. Zur Klärung des Begriffs „Deepfake“ werden im Folgenden vier kumulative Kriterien aus dieser Definition untersucht: i) Ähnlichkeit mit ii) wirklichen iii) Personen, Gegenständen, Orten, Einrichtungen oder Ereignissen, iv) die einer Person fälschlicherweise als echt oder wahrheitsgemäß erscheinen würden.

i) **Ähnlichkeit:** In Artikel 3 Nummer 60 der KI-Verordnung ist festgelegt, dass durch KI erzeugte oder manipulierte Bild-, Ton- oder Videoinhalte wirklichen Subjekten „ähneln“ müssen, um ein Deepfake darzustellen. In Erwägungsgrund 134 der KI-Verordnung wird hinzugefügt, dass eine solche Ähnlichkeit „merklich“ sein sollte. Durch KI erzeugte oder manipulierte Bild-, Ton- oder Videoinhalte ähneln merklich einem Subjekt, wenn ein hoher Grad an Ähnlichkeit zwischen den Deepfake-Inhalten und dem Subjekt (einschließlich seiner erkennbaren Elemente), das durch das Deepfake simuliert wird, besteht. Der Inhalt muss nicht mit dem Subjekt identisch sein. Ob der Grad der Ähnlichkeit merklich ist, ist eine Einzelfallbewertung, die vom Betreiber auf der Grundlage eines objektiven Vergleichs zwischen dem simulierten Subjekt und den Deepfake-Inhalten vorzunehmen ist, wobei unter anderem darauf abzustellen ist, inwieweit das Deepfake charakteristische oder ausgeprägte Merkmale darstellt.

ii) **Wirklich:** In Artikel 3 Nummer 60 der KI-Verordnung ist festgelegt, dass durch KI erzeugte oder manipulierte Bild-, Ton- oder Videoinhalte „wirklichen“ Subjekten ähneln müssen, um ein Deepfake darzustellen. Um den Zweck von Artikel 50 Absatz 4 der KI-Verordnung, das Risiko von Identitätsbetrug, Täuschung, Fehlinformation, Manipulation und Betrug zu verringern, wirksam zu erfüllen, impliziert das Merkmal der Wirklichkeit, dass die durch KI erzeugten oder manipulierten Bild-, Ton- oder Videoinhalte realistischen Subjekten (z. B. Personen, Gegenständen, Orten) ähneln sollten. Daher reicht es aus, dass simulierte Personen, Gegenstände, Orte, Einrichtungen oder Ereignisse jemandem oder etwas ähneln, der bzw. das in Wirklichkeit existiert, plausibel existieren kann oder plausibel existiert haben könnte, um als Deepfake zu gelten, da solche Inhalte eine wichtige Quelle für Täuschung, Manipulation und Desinformation sein können. Simulierte Personen, Gegenstände, Orte, Einrichtungen oder Ereignisse, die beispielsweise gegen Natur- oder Physikgesetze verstoßen oder Lebensformen darstellen, die in der Biologie nicht allgemein anerkannt sind (z. B. Menschen, die ohne mechanische Hilfsmittel fliegen, Drachen oder Elefanten, die Auto fahren) und kein Irreführungspotenzial haben, gelten hingegen als unrealistisch und fallen daher nicht in den Anwendungsbereich der Transparenzpflicht.

iii) **Personen, Gegenstände, Orte, Einrichtungen oder Ereignisse:** In Artikel 3 Nummer 60 der KI-Verordnung ist festgelegt, dass durch KI erzeugte oder manipulierte Bild-, Ton- oder Videoinhalte wirklichen „Personen, Gegenständen, Orten, Einrichtungen oder Ereignissen“ ähneln sollten, um eine Deepfake darzustellen. Unter „Personen“ sind realistische Menschen zu verstehen (einschließlich digitaler Nachbildungen echter Personen, realistischer durch KI erzeugter menschlicher Avatare oder Personen sowie persönlicher Merkmale oder Ausdrucksformen wie Bild, Stimme, Verhalten, Leistungen usw.). Unter „Gegenstände“ sind realistische, unbelebte materielle Gegenstände zu verstehen, einschließlich Gebäuden, Kunstwerken, Maschinen, Verbrauchsgütern usw. Unter „Orte“ sind reale Schauplätze zu verstehen. Unter „Einrichtungen“ sind realistische, nichtmenschliche, aber belebte Wesen zu verstehen, die Tiere oder andere biologische Lebensformen umfassen. Unter „Ereignisse“ sind realistische Szenen oder Situationen zu verstehen, an denen Personen, Gegenstände, Orte und Einrichtungen beteiligt sein können (z. B. Anspielung auf historische Ereignisse oder Darstellung von professionellen Dienstleistungen oder Verbraucherdienstleistungen).

iv) **Fälschliches Erscheinen als echt oder wahrheitsgemäß gegenüber einer Person:** Viertens wird in Artikel 3 Nummer 60 der KI-Verordnung spezifiziert, dass der durch KI erzeugte oder manipulierte Bild-, Ton- oder Videoinhalt, der wirklichen Personen, Gegenständen, Orten, Einrichtungen oder Ereignissen ähnelt, „einer Person fälschlicherweise als echt oder wahrheitsgemäß erscheinen würde“. Dieses Kriterium bezieht sich auf das wesentliche Merkmal von Deepfake-Inhalten und ihre Fähigkeit, eine Person über die Echtheit oder Wahrhaftigkeit der Inhalte zu täuschen oder irrezuführen ([32]). Die Echtheit von Inhalten bezieht sich darauf, ob es sich bei den Inhalten tatsächlich um das handelt, was sie in Bezug auf ihre Quelle oder ihren Erstellungsprozess vorgeben (einschließlich z. B. der Beteiligung echter Menschen oder Tiere und ihres tatsächlichen Verhaltens oder ihrer tatsächlichen Handlungen, des tatsächlichen Erscheinungsbilds oder der tatsächlichen Verwendung von Gegenständen oder der Erbringung von Dienstleistungen oder des korrekten Ablaufs eines Ereignisses). Die Wahrhaftigkeit bezieht sich auf den Wahrheitsgehalt des Inhalts (z. B. die sachliche Richtigkeit der Darstellungen im Deepfake).

([32]) Daher ist dieses Kriterium unabhängig und getrennt von dem Begriff der Täuschung im Sinne der Richtlinie 2005/29/EG über unlautere Geschäftspraktiken im binnenmarktinternen Geschäftsverkehr zwischen Unternehmen und Verbrauchern (Richtlinie über unlautere Geschäftspraktiken), https://eur-lex.europa.eu/legal-content/DE/TXT/HTML/?uri=CELEX:32005L0029, zu verstehen.

114. Ob Inhalte „einer Person fälschlicherweise als echt oder wahrheitsgemäß erscheinen“ würden, sollte als Ganzes bewertet werden, wobei der Grad der Ähnlichkeit, die potenzielle inhaltliche Botschaft des Inhalts, die beabsichtigten und vorhersehbaren Einsatzkontexte, die Umgebung, in der die Inhalte präsentiert werden, sowie die beabsichtigte und vernünftigerweise vorhersehbare Publikumszusammensetzung und ihre Erwartungen zu berücksichtigen sind. Diese Bewertung ist jedoch objektiv und erfordert nicht die Absicht des Betreibers, die natürlichen Personen, die den Inhalten ausgesetzt sind, zu täuschen oder irrezuführen, damit diese Inhalte ein Deepfake darstellen. In Bezug auf den Einsatzkontext und die Erwartungen des Publikums bedeutet dies, dass, wenn die Inhalte in bestimmten Einsatzkontexten verwendet werden, in denen das bestimmungsgemäße und vernünftigerweise vorhersehbare Publikum nicht erwartet, dass die Inhalte echt oder wahrheitsgemäß sind, die durch KI erzeugten oder manipulierten Inhalte nicht *fälschlicherweise* als echt oder wahrheitsgemäß erscheinen können, auch wenn sie tatsächlich nicht echt oder unwahr sind. In solchen Fällen ist das vierte Kriterium der Begriffsbestimmung von „Deepfake“ nicht erfüllt. So ist es beispielsweise nicht wahrscheinlich, dass die KI-Erzeugung oder -Manipulation von Hintergrundszenen, Spezialeffekten oder die technische Vor- und Nachbearbeitung im Rahmen von Standard-Filmproduktionsprozessen dazu führen, dass Inhalte dem Publikum *fälschlicherweise* als echt oder wahrheitsgemäß erscheinen. Umgekehrt dürfte dies beim Einsatz von KI zur Erzeugung oder Manipulation wesentlicher Elemente, die sich auf die Wahrnehmung des Publikums auswirken, der Fall sein (z. B. vollständig durch KI erzeugte Akteure, digitale Replikationen echter oder verstorbener Akteure, Altersanpassung, simulierte Darbietungen, einschließlich Teilen davon, oder nicht echte oder unwahre Darstellung von Personen, Gegenständen, Orten oder Ereignissen in Dokumentarfilmen). In Bezug auf Bild- und Videoinhalte wird anerkannt, dass ein hoher Grad an Fotorealismus es wahrscheinlicher macht, dass solche Inhalte als Deepfake angesehen werden sollten (da sie wahrscheinlich wirklichen Subjekten ähneln), aber Fotorealismus allein ist für die Bewertung nicht entscheidend. Es sollte noch geprüft werden, ob solche Inhalte geeignet sind, eine Person hinsichtlich der Echtheit oder Wahrhaftigkeit der Inhalte zu täuschen oder irrezuführen, wobei unter anderem der beabsichtigte und vorhersehbare Einsatzkontext zu berücksichtigen ist.

115. Im Gegensatz zu Artikel 50 Absatz 1 der KI-Verordnung sollte die Bewertung des fälschlichen Erscheinens einer Person als echt oder wahrheitsgemäß nicht auf einer hypothetischen „durchschnittlichen“ Person beruhen, von der erwartet wird, dass sie den Inhalten ausgesetzt ist (siehe Abschnitt 3.2.1), sondern die mögliche unterschiedliche Zusammensetzung des vernünftigerweise vorhersehbaren Publikums, das den Deepfake-Inhalten ausgesetzt sein könnte, gebührend berücksichtigen. Eine gebührende Berücksichtigung der Publikumszusammensetzung ist besonders wichtig, wenn vernünftigerweise vorhersehbar ist, dass die Deepfake-Inhalte zu irgendeinem Zeitpunkt von Kindern, älteren Menschen oder anderen Personengruppen mit geringeren digitalen und KI-Kompetenzen oder einem geringeren allgemeinen Wissensniveau wahrgenommen werden können, da sie in Bezug auf die Echtheit oder Wahrhaftigkeit der Inhalte leichter getäuscht oder irregeführt werden können. In solchen Fällen kann ein falsches Erscheinen des Inhalts gegenüber diesem Teil des vorhersehbaren Publikums als echt oder wahrheitsgemäß ausreichen, um Inhalte als Deepfake einzustufen. Von den Betreibern wird jedoch nicht erwartet, dass sie eine mögliche weitere Verbreitung der Deepfake-Inhalte durch Dritte über das vom Betreiber vernünftigerweise vorhersehbare Publikum hinaus unter Berücksichtigung der beabsichtigten Vertriebskanäle prüfen (wenn z. B. Deepfake-Inhalte nur auf einem abonnierten Teil einer Website oder als Teil eines Unternehmens-Newsletters angezeigt werden, bedeutet dies nicht, dass der Betreiber standardmäßig eine breite öffentliche Zugänglichkeit in Betracht ziehen sollte).

116. Ebenso kann die KI-gestützte Manipulation unbedeutender materieller oder technischer Aspekte bereits bestehender Inhalte von geringer Bedeutung für die Bewertung der Echtheit oder Wahrhaftigkeit der Inhalte durch eine Person sein, sodass die daraus resultierenden Inhalte nicht zu Deepfakes werden. Dies könnte beispielsweise die Bearbeitung von Hintergrundinformationen (z. B. Entfernung von Passanten), Beleuchtungseinstellungen, die Anpassung von Audioparametern, Farbkorrekturen, Lärmminderung oder -beseitigung, die Verbesserung der Zugänglichkeit oder Dateikomprimierung, optische Anpassungen und Verbesserungen umfassen. Ob die Manipulation in diese Kategorie fällt, hängt vom Kontext und den Auswirkungen auf die Wahrnehmung der Echtheit oder Wahrhaftigkeit des Inhalts durch die Personen im konkreten Fall ab. Beispielsweise dürften KI-gestützte Farbkorrekturen, Hintergrunderweiterungen bestehender Inhalte, Anpassungen oder der Austausch von Hintergründen für eindeutig ästhetische Zwecke, Zusammensetzungen und Anordnungen bestehender Produkte oder die Neuskalierung von Bildern, die in Produktwerbung oder -verpackung verwendet werden, nur geringfügige Auswirkungen auf die Wahrnehmung einer Person in Bezug auf die Echtheit und Wahrhaftigkeit der Werbung und des Produkts haben. Umgekehrt kann sich eine umfangreiche KI-gestützte Bearbeitung von Hintergrunddetails an journalistischen Bildern, die über die üblichen technischen und redaktionellen Praktiken hinausgeht, negativ auf die Wahrnehmung einer Person in Bezug auf die Echtheit und Wahrhaftigkeit der Inhalte auswirken.

Beispiele für Deepfakes nach Artikel 3 Nummer 60 der KI-Verordnung:

— Durch KI manipuliertes Bild einer Szene mit zwei echten Profifußballern vor einem Gebäude, das einem Fußballstadion ähnelt.

— Durch KI erzeugte Audiodatei, die das Klonen von Stimmen der regelmäßigen Moderatoren eines Zeitungspodcasts und eines Gastgesprächspartners umfasst, die über Nachrichtenereignisse diskutieren.

— Durch KI erzeugtes Video einer Person, die einem Politiker ähnelt und eine Rede vor einem Publikum hält.

— Durch KI erzeugtes Video mit einer durch KI erzeugten Darstellung eines prominenten Influencers in einem Werbekontext.

— Durch KI erzeugtes Video mit einem realistischen synthetischen Avatar eines Geschäftsführers eines Unternehmens, der den Beschäftigten zu ihrer Arbeit und den Unternehmensergebnissen des vergangenen Jahres gratuliert.

— Ein durch KI erzeugtes Bild eines Produkts in der Werbung oder Verpackung, das die Wahrnehmung des Publikums beeinflussen und in Bezug auf das tatsächliche Aussehen, die Merkmale oder die Verwendung des Produkts irreführen kann (z. B. indem das Produkt nicht mit dem tatsächlichen Produkt identisch, attraktiver oder von besserer Qualität erscheint als im realen Leben).

Beispiele, die keine Deepfakes nach Artikel 3 Nummer 60 der KI-Verordnung darstellen:

— Durch KI erzeugtes Bild einer Szene mit einer über den Eiffelturm fliegenden Sphinx.

— Durch KI erzeugtes Video von Mäusen, die im Rahmen einer Werbekampagne für einen Käsehersteller in menschlicher Sprache über die beste Käsesorte diskutieren.

— KI-Sprachreplikation für fiktive Charaktere (z. B. für Hörbücher, Spiele oder Animationen), wenn keine Täuschung über die Identität der Erzähler vorliegt.

— Durch KI manipulierte Hörfunksendungen, bei denen technische Audioparameter (z. B. normalisierende Lautstärken, Lärmminderung, Audiokompression) angepasst werden, ohne die von den Sprechern tatsächlich gesprochenen Worte oder ihre Sprechweise zu ändern.

— Durch KI erzeugte Karikatur eines bereits bestehenden Bildes, das ein historisches Ereignis darstellt.

— Video eines Moderators eines wissenschaftlichen Fernsehprogramms, der eine Gebirgsregion durchquert und sich mit dem Rückzug von Gletschern befasst, was durch eine durch KI manipulierte Hintergrundanimation veranschaulicht wird, die zeigt, wie die Gletscher von ihrem historischen Niveau auf ihr derzeitiges Niveau zurückgegangen sind.

— Durch KI erzeugte fiktive Umgebungen (z. B. fiktive Wälder und Burgen) in Videospielen.

— Filme, in denen reale (nicht durch KI manipulierte) Schauspieler vor einem durch KI erzeugten Hintergrund spielen (z. B. Darstellung eines fiktiven Prärie-Abschnitts, historischer Gebäude alter Städte usw.).

— Ein echtes Produkt (z. B. ein Auto), das in einer Werbung vor einem durch KI erzeugten Hintergrund und einer durch KI erzeugten Umgebung gezeigt wird, solange die Werbung nicht geeignet ist, das Publikum über die tatsächliche Darstellung des Produkts sowie seine Merkmale und Verwendung irrezuführen.

6.1.2. *Die Offenlegungspflicht nach Artikel 50 Absatz 4 Unterabsatz 1 der KI-Verordnung*

117. Nach Artikel 50 Absatz 4 Unterabsatz 1 der KI-Verordnung sind Betreiber von KI-Systemen, die Bild-, Ton- oder Videoinhalte erzeugen oder manipulieren, die Deepfakes darstellen, verpflichtet, „offen[zu]legen, dass die Inhalte künstlich erzeugt oder manipuliert wurden“. In Erwägungsgrund 134 der KI-Verordnung wird präzisiert, dass

Betreiber, die KI-Systeme zum Erzeugen oder Manipulieren von Deepfake-Inhalten verwenden, klar und deutlich offenlegen sollten, dass solche Inhalte künstlich erzeugt oder manipuliert wurden, indem sie die Ausgaben von KI entsprechend kennzeichnen und auf ihren künstlichen Ursprung hinweisen. Die gemäß Artikel 50 Absatz 4 Unterabsatz 1 der KI-Verordnung angewandten Kennzeichnungs- oder Offenlegungsmethoden sollten für natürliche Personen verständlich und wahrnehmbar sein (z. B. mit sichtbaren oder akustischen Kennzeichnungen), ohne dass sie auf spezifische technische Instrumente angewiesen sind oder gezielte Maßnahmen durchführen müssen. Daher können sich Betreiber nicht auf die maschinenlesbare Kennzeichnung verlassen, die der Anbieter gemäß Artikel 50 Absatz 2 der KI-Verordnung in den Inhalt einbettet, da diese Kennzeichnungen für die natürlichen Personen, die den Deepfake-Inhalten ausgesetzt sind, nicht unmittelbar klar und deutlich sind.

118. Nach Artikel 50 Absatz 7 der KI-Verordnung können sich Betreiber an einen Praxisleitfaden halten, der gemäß Artikel 50 Absatz 7 der KI-Verordnung als angemessen bewertet wird, um die Einhaltung dieser Offenlegungspflicht nachzuweisen. Weitere Informationen über die Auswirkungen der Einhaltung eines solchen Leitfadens und die Überwachung anderer Betreiber, die sich nicht an einen Leitfaden halten, sind Abschnitt 8.1 zu entnehmen.

6.1.3. *Transparenz von künstlerischen, kreativen, satirischen, fiktionalen oder analogen Deepfake-Inhalten*

119. Artikel 50 Absatz 4 Unterabsatz 1 der KI-Verordnung sieht eine abgeschwächte Transparenzpflicht für Deepfakes vor, die Teil offensichtlich künstlerischer, kreativer, satirischer, fiktionaler oder analoger Werke oder Programme sind. Bei solchen Inhalten beschränkt sich die Transparenzpflicht auf die Offenlegung des Deepfakes in angemessener Weise, die die Darstellung oder den Genuss des Werks nicht beeinträchtigt.

120. **Relevante Inhaltskategorien**: Um von der abgeschwächten Transparenzpflicht zu profitieren, sollten Deepfakes (Bild, Audio oder Video) mindestens einer der folgenden Inhaltskategorien angehören: künstlerische, kreative, satirische, fiktionale oder analoge Werke oder Programme. Ob ein Deepfake unter eine oder mehrere dieser Kategorien fällt, muss vom Betreiber im Einzelfall geprüft werden. Für die Zwecke des Artikels 50 Absatz 4 der KI-Verordnung sind diese Kategorien wie folgt zu verstehen:

— Unter künstlerischen Werken sind Werke zu verstehen, die zu künstlerischen Zwecken geschaffen wurden, einschließlich Musik, filmischer Arbeit und bildender Kunst.

— Unter kreativen Werken sind Werke zu verstehen, die kreative Entscheidungen erfordern. Werke, die hauptsächlich auf funktionellen oder technischen Erwägungen beruhen, können nicht als kreativ angesehen werden.

— Unter satirischen Werken sind Werke zu verstehen, die darauf abzielen, die Gesellschaft, die Politik, die Wirtschaft oder Persönlichkeiten des öffentlichen Lebens durch den Einsatz humoristischer Techniken (einschließlich Ironie, Sarkasmus, Verhöhnung, Nachahmung usw.) zu kritisieren.

— Unter fiktiven Werken sind kulturelle Werke zu verstehen, an denen Personen, Gegenstände, Orte, Einrichtungen oder Ereignisse in einem imaginären, aber realitätsnahen Umfeld beteiligt sind.

— Unter analogen Werken sind Werke zu verstehen, die Kernmerkmale oder ähnliche ausdrucks- oder funktionsbezogene Merkmale mit den oben genannten Kategorien teilen, aber nicht genau in eine Kategorie passen.

121. Die oben genannten Inhaltskategorien können auch für Sendungen gelten, die als Einzelbestandteil eines von einem Mediendiensteanbieter erstellten Sendeplans oder Katalogs zu verstehen sind und deren Form und Inhalt mit der Form und dem Inhalt von Fernsehprogrammen vergleichbar sind ([33]). Der Verweis auf „Werke“ im Zusammenhang mit Artikel 50 Absatz 4 der KI-Verordnung hat keinen Einfluss auf die Frage, ob die Inhalte nach dem Unionsrecht über das Urheberrecht und verwandte Schutzrechte schutzfähig sind.

([33]) Siehe in ähnlicher Weise Artikel 1 Buchstabe b der AVMD-Richtlinie.

122. **Offensichtlich:** Die Tatsache, dass ein Deepfake in eine (oder mehrere) der oben genannten Inhaltskategorien fällt, sollte für die natürlichen Personen, die ihm ausgesetzt sind, klar ersichtlich sein. Dies bedeutet, dass die Kategorien eng auszulegen sind, was dadurch gerechtfertigt ist, dass aufgrund der potenziellen Auswirkungen auf die Meinungsfreiheit und die Freiheit der Kunst und der Wissenschaft eine weniger strenge Offenlegungsregelung gilt, während die Risiken von Fehlinformationen, Manipulation und Täuschung noch angemessen angegangen werden müssen. Daher sind Inhalte, deren Art für das Publikum möglicherweise unklar oder mehrdeutig ist, vom Anwendungsbereich dieser weniger strengen Transparenzregelung ausgenommen. Zu den relevanten Faktoren für die Beurteilung der Offensichtlichkeit künstlerischer, kreativer, satirischer oder fiktionaler Inhalte gehören: i) die Frage, ob der Inhalt Formate oder Stile aufweist, die für die Inhaltskategorien charakteristisch sind (z. B. Ironie oder Übertreibung bei satirischen Werken, bestimmte Kunststile, charakteristische Themen, Techniken oder Ästhetik usw.), ii) der Kontext, in dem die Inhalte präsentiert werden (z. B. ob die Plattform, das Medium oder der Ort, an dem die Inhalte präsentiert, zur Verfügung gestellt oder verbreitet werden, mit einer künstlerischen, kreativen, satirischen oder fiktiven Nutzung verbunden ist), und iii) die Erwartungen des Publikums (z. B. ein Film, eine Spielumgebung oder eine Szene virtueller Realität). Darüber hinaus schließt die Bedingung, dass Inhalte „offensichtlich" in eine der oben genannten Inhaltskategorien fallen sollten, Inhalte vom Anwendungsbereich dieser Inhaltskategorien aus, wenn sie ausschließlich informativer oder kommerzieller Art und als solche erkennbar sind (z. B. Nachrichtenberichterstattung). In dieser Hinsicht könnten einige Arten von Inhalten (z. B. Werbung oder Dokumentarfilme), die Deepfakes enthalten, in bestimmten, spezifischen Situationen als offensichtlich kreative oder fiktive Werke angesehen werden, in anderen jedoch nicht, da die Bewertung fallspezifisch ist. Wenn in dem Deepfake mehrere Charaktere kombiniert werden (z. B. informativ und kreativ), sollte der informative Charakter stets überwiegen und die Einhaltung der standardmäßigen Kennzeichnungsanforderungen erfordern.

123. **Geeignete Offenlegung, die die Darstellung oder den Genuss des Werks nicht beeinträchtigt**: Deepfakes, die Teil offensichtlich künstlerischer, kreativer, satirischer oder fiktiver Werke oder Programme sind, sind nicht von der Transparenzpflicht nach Artikel 50 Absatz 4 Unterabsatz 1 der KI-Verordnung ausgenommen. Die Betreiber müssen nach wie vor die KI-Herkunft der Inhalte oder deren Manipulation offenlegen, können dies jedoch in geeigneter Weise tun, ohne die Darstellung oder den Genuss des Werks zu beeinträchtigen. In Erwägungsgrund 134 der KI-Verordnung wird präzisiert, dass eine solche geeignete Offenlegung die normale Nutzung und Verwendung nicht beeinträchtigen und gleichzeitig den Nutzen und die Qualität des Werks aufrechterhalten sollte. Die Festlegung, welche Offenlegungsmaßnahmen als geeignet anzusehen sind, erfordert die Prüfung des Einzelfalls, bei der die Betreiber alle relevanten Faktoren (einschließlich der Art des Werks, des Publikums, des Kontexts usw.) berücksichtigen können. In jedem Fall müssen die Betreiber Artikel 50 Absatz 5 der KI-Verordnung einhalten (siehe Abschnitt 7). Nach Artikel 50 Absatz 7 der KI-Verordnung können sich Betreiber an einen Praxisleitfaden halten, der gemäß dieser Bestimmung als angemessen bewertet wird, um die Einhaltung dieser Offenlegungspflicht nachzuweisen. Weitere Informationen über die Auswirkungen der Einhaltung eines solchen Leitfadens und die Überwachung anderer Betreiber, die sich nicht an einen Leitfaden halten, sind Abschnitt 8.1 zu entnehmen.

124. **Geeignete Schutzvorkehrungen für die Rechte und Freiheiten Dritter:** In Erwägungsgrund 134 der KI-Verordnung wird präzisiert, dass für die Einhaltung der abgeschwächten Transparenzpflicht und die Verwendung künstlerischer, kreativer, satirischer oder fiktiver Deepfakes „geeignete Schutzvorkehrungen für die Rechte und Freiheiten Dritter bestehen" müssen. Daher müssen die Betreiber sicherstellen, dass die Rechte und Freiheiten Dritter (einschließlich z. B. ihres Rechts auf den Schutz personenbezogener Daten oder der Rechte des geistigen Eigentums) bei der Erstellung oder Veröffentlichung von Deepfakes angemessen geschützt und geachtet werden, auch wenn die Deepfakes als künstlerisch, kreativ, satirisch oder fiktional anzusehen sind. Die Berufung auf die abgeschwächte Transparenzpflicht kann keine Rechtfertigung für die Missachtung der Grundrechte von Einzelpersonen oder Rechteinhabern nach dem Unionsrecht im Bereich des geistigen Eigentums oder dem Datenschutzrecht der Union sein ([34]).

Beispiele für Deepfakes, die kreative, satirische, fiktionale oder analoge Werke sind:

— Künstlerisches/fiktionales Werk: Filme (oder Film-Trailer) mit durch KI erzeugten oder manipulierten Deepfakes wirklicher Schauspieler, deren Alter angepasst wurde, oder digitaler Nachbildungen toter Schauspieler oder anderer Personen, die in einem Kino oder auf einer Streaming-Plattform verfügbar sind.

— Künstlerisches/kreatives Werk: Durch KI erzeugte Musik in jeder Art von Genre, die dem individuellen Stil wirklicher Künstler ähnelt.

([34]) Dies gilt auch für veröffentlichte Texte, mit denen die Öffentlichkeit über Angelegenheiten von öffentlichem Interesse informiert wird.

— Satirisches/fiktionales Werk: Durch KI manipuliertes Bild eines wirklichen Politikers, in dem er in einer Szene zu sehen ist, die eindeutig bestimmte politische Entscheidungen, die von dieser Person getroffen oder unterstützt werden, humorvoll kritisieren soll.

— Analoges kreatives/fiktionales Werk: Durch KI erzeugte Spielbilder mit Deepfake-Simulationen echter, wirklicher Personen.

Beispiele für Deepfakes, die keine künstlerischen, kreativen, satirischen, fiktionalen oder analogen Werke darstellen:

— Durch KI manipuliertes Video im Stil eines Teleshopping-Kanals, das eine Deepfake-Simulation von Menschen umfasst, die in einer durch KI erzeugten Szene für ein Produkt werben und die funktionale Nutzung des Produkts durch die simulierten Verbraucher darstellen, um die Zuschauer zum Kauf des Produkts zu bewegen.

— Durch KI erzeugtes Bild von Prominenten, das ihre Beteiligung an Aktivitäten impliziert, die nie stattgefunden haben, ohne jeglichen fiktiven, satirischen oder analogen Zweck.

— Durch KI manipuliertes Video mit einem realistischen synthetischen Influencer, der ein gesponsertes echtes Produkt testet und sich ausschließlich auf die Darstellung der Produktfunktionen konzentriert.

— Durch KI erzeugtes Video mit realistischen Holocaust-Szenen, die auf öffentlich zugänglichen Social-Media-Plattformen geteilt werden.

6.1.4. *Ausnahme für Strafverfolgung*

125. Wenn die Verwendung eines Deepfake (z. B. eines durch KI erzeugten Videos, in dem Ereignisse simuliert werden, die nie stattgefunden haben) zur Aufdeckung, Verhütung, Ermittlung oder Verfolgung von Straftaten gesetzlich zugelassen ist, sind Betreiber von der Transparenzpflicht nach Artikel 50 Absatz 4 Unterabsatz 1 der KI-Verordnung ausgenommen. Die Erläuterungen unter den Randnummern 46 bis 48 für eine ähnliche Ausnahme in Artikel 50 Absatz 1 der KI-Verordnung sind auch für diese Bestimmung relevant.

6.1.5. *Zusammenspiel mit anderen Rechtsakten der Union*

126. Die Transparenzpflicht nach Artikel 50 Absatz 4 Unterabsatz 1 der KI-Verordnung interagiert mit den Pflichten der Anbieter sehr großer Online-Plattformen und sehr großer Online-Suchmaschinen zur Ermittlung und Minderung systemischer Risiken gemäß den Artikeln 34 und 35 des Gesetzes über digitale Dienste, insbesondere Artikel 35 Absatz 1 Buchstabe k des Gesetzes über digitale Dienste ([35]). In der letztgenannten Bestimmung werden unter den möglichen Risikominderungsmaßnahmen, die Anbieter sehr großer Online-Plattformen und sehr großer Online-Suchmaschinen ergreifen können, Maßnahmen aufgeführt, mit denen sichergestellt wird, dass „eine Einzelinformation, unabhängig davon, ob es sich um einen erzeugten oder manipulierten Bild-, Ton- oder Videoinhalt handelt, der bestehenden Personen, Gegenständen, Orten oder anderen Einrichtungen oder Ereignissen merklich ähnelt und einer Person fälschlicherweise als echt oder wahrheitsgemäß erscheint, durch eine auffällige Kennzeichnung erkennbar ist, wenn sie auf ihren Online-Schnittstellen angezeigt wird“. Darüber hinaus können Anbieter sehr großer Online-Plattformen und sehr großer Online-Suchmaschinen auch eine benutzerfreundliche Funktion bereitstellen, die es den Nutzern ermöglicht, solche Informationen anzuzeigen. Diese Bestimmung des Gesetzes über digitale Dienste scheint zwar mit Artikel 50 Absatz 4 der KI-Verordnung vergleichbar zu sein, ergänzt sie jedoch und enthält zwei wichtige Unterscheidungen:

— **Unterschiedlicher sachlicher Anwendungsbereich**: Im Vergleich zu Artikel 50 Absatz 4 der KI-Verordnung deckt Artikel 35 Absatz 1 Buchstabe k des Gesetzes über digitale Dienste ein breiteres Spektrum von Inhalten ab, die fälschlicherweise echt oder wahrheitsgemäß erscheinen können, unabhängig von den zur Erstellung der Inhalte verwendeten Technologien (d. h. er deckt Inhalte ab, die mithilfe von KI-Systemen oder anderen Instrumenten erzeugt oder manipuliert wurden).

([35]) Siehe auch Abschnitt 4.4 für weitere Informationen über das Zusammenspiel zwischen Artikel 50 der KI-Verordnung und dem Gesetz über digitale Dienste.

— **Unterschiedlicher personenbezogener Anwendungsbereich**: Nach Artikel 50 Absatz 4 der KI-Verordnung sind Betreiber von KI-Systemen, die Deepfakes erzeugen oder manipulieren, verpflichtet, solche Inhalte zu kennzeichnen, während Artikel 35 Absatz 1 Buchstabe k des Gesetzes über digitale Dienste für Anbieter sehr großer Online-Plattformen und sehr großer Online-Suchmaschinen gilt, die die Inhalte verbreiten oder solche Inhalte durch Suche anzeigen lassen. So können beispielsweise Betreiber von KI-Systemen, die gemäß der KI-Verordnung in eine sehr große Online-Plattform integriert sind, als „Nutzer" im Sinne von Artikel 3 Buchstabe b des Gesetzes über digitale Dienste ([36]) angesehen werden, wenn sie die sehr große Online-Plattform nutzen, um mit dem jeweiligen KI-System erzeugte Inhalte zu verbreiten, während sehr große Online-Plattformen als Anbieter des KI-Systems gelten können, wenn die Bedingungen des Artikels 3 Nummer 3 der KI-Verordnung erfüllt sind. Stellen Anbieter sehr großer Online-Plattformen oder sehr großer Online-Suchmaschinen solchen Betreibern Kennzeichnungsinstrumente zur Verfügung, die es ihnen ermöglichen, ihre Deepfake-Inhalte im Einklang mit Artikel 50 Absatz 4 der KI-Verordnung zu kennzeichnen (d. h. eine klare und unterscheidbare Offenlegung des KI-Ursprungs), können sich diese Betreiber auf solche Instrumente stützen, um ihrer Transparenzpflicht gemäß dieser Bestimmung im Zusammenhang mit den verwendeten sehr großen Online-Plattformen oder sehr großen Online-Suchmaschinen nachzukommen. Die Bereitstellung einer solchen Funktion berührt nicht die Verantwortung der Betreiber gemäß der KI-Verordnung für die Erfüllung ihrer Kennzeichnungspflichten nach Artikel 50 Absatz 4 der KI-Verordnung. Darüber hinaus gelten Anbieter sehr großer Online-Plattformen und sehr großer Online-Suchmaschinen auch als Betreiber eines KI-Systems im Sinne der KI-Verordnung, wenn sie ein KI-System unter ihrer Verantwortung für ihre eigenen beruflichen Zwecke verwenden (z. B. indem sie visuelle Elemente erstellen, die als Teil des Marketingmaterials verwendet werden können).

127. Wenn das Deepfake Informationen über eine direkt oder indirekt identifizierbare (lebende) natürliche Person enthält, müssen Betreiber als Verantwortliche nach dem Datenschutzrecht der Union auch die entsprechenden Verpflichtungen nach dem Datenschutzrecht der Union erfüllen ([37]). Dazu gehören beispielsweise die Berufung auf eine geeignete Rechtsgrundlage für die Verarbeitung der personenbezogenen Daten, Informationen über die Verarbeitung von Daten über die dargestellten betroffenen Personen usw.

128. Betreiber von KI-Systemen, die Deepfakes erzeugen, müssen darüber hinaus sicherstellen, dass sie die einschlägigen Rechtsvorschriften der Union über geistiges Eigentum (wie das Markenrecht der Union und das Unionsrecht über das Urheberrecht oder verwandte Rechte) einhalten, wenn ihre Deepfakes Gegenstände enthalten, die nach diesen Rechtsvorschriften geschützt sind. Die Anbringung einer Kennzeichnung oder eines Symbols auf einem Deepfake hat als solche keine Auswirkungen auf dessen Anspruch auf urheberrechtlichen Schutz, der im Einklang mit dem geltenden Urheberrecht der Union zu bewerten ist.

129. Wenn wirkliche Personen in Deepfakes abgebildet werden, müssen die Betreiber auch die einschlägigen rechtlichen Regelungen zum Schutz der Persönlichkeitsrechte gegenüber persönlichen Bildern oder Stimmen einhalten. Die Transparenzpflicht nach Artikel 50 Absatz 4 Unterabsatz 1 der KI-Verordnung bedeutet nicht, dass durch KI erzeugte oder manipulierte Deepfakes, die nach dem geltenden Unionsrecht oder nationalen Recht schädlich und rechtswidrig sind (z. B. irreführende Werbung oder Strafrecht wie Material über sexuellen Kindesmissbrauch und nicht einvernehmliche intime Bilder), erzeugt und verbreitet werden können (siehe auch Abschnitt 4.4).

6.2. Hauptkomponenten, Konzepte und damit verbundene Transparenzpflichten für durch KI erzeugten oder manipulierten Text nach Artikel 50 Absatz 4 der KI-Verordnung

130. Artikel 50 Absatz 4 Unterabsatz 2 der KI-Verordnung findet Anwendung, wenn mehrere Bedingungen erfüllt sind:

i) Das System muss ein KI-System sein.

ii) Es muss zur Erzeugung oder Manipulation von Text (siehe Abschnitt 4.2.1) verwendet werden, der veröffentlicht wird, um die Öffentlichkeit über Angelegenheiten von öffentlichem Interesse zu informieren.

iii) Die Verwendung des KI-Systems fällt unter keine der Ausnahmen, wenn 1) die durch KI erzeugten Inhalte einem Prozess der menschlichen Überprüfung oder redaktionellen Kontrolle unterzogen wurden und eine natürliche oder juristische Person die redaktionelle Verantwortung für die Veröffentlichung der Inhalte trägt oder 2) die Verwendung zur Aufdeckung, Verhütung, Ermittlung oder Verfolgung von Straftaten gesetzlich zugelassen ist.

([36]) Erwägungsgrund 2 Gesetz über digitale Dienste.

([37]) Weitere Einzelheiten sind den gemeinsamen Leitlinien der Kommission und des EDSA zum Zusammenspiel zwischen der KI-Verordnung und dem EU-Datenschutzrecht (in Vorbereitung) zu entnehmen.

6.2.1. *Text, der veröffentlicht wird, um die Öffentlichkeit über Angelegenheiten von öffentlichem Interesse zu informieren*

131. Zur Klärung des Anwendungsbereichs von Artikel 50 Absatz 4 Unterabsatz 2 der KI-Verordnung werden im Folgenden drei Elemente untersucht: i) veröffentlichter Text ii) zur Information der Öffentlichkeit iii) über Angelegenheiten von öffentlichem Interesse.

i) **Veröffentlichter Text:** Der durch KI erzeugte oder manipulierte Text sollte veröffentlicht werden. Dies bedeutet, dass der Text für eine unbestimmte, relativ große Zahl unabhängiger, potenzieller Leser gleichzeitig und/oder nacheinander zugänglich sein sollte, unabhängig davon, ob dies gegen Bezahlung (z. B. Abonnements) geschieht oder nicht ([38]). Text gilt hingegen nicht als veröffentlicht, wenn der Zugang auf bestimmte Personen beschränkt ist, die einer geschlossenen, privaten Gruppe angehören (z. B. eine kleine, geschlossene Gruppe in einer Instant-Messaging-App oder wenn die Gruppe zu klein oder unbedeutend ist). Beispiele für Texte, die nicht als veröffentlicht gelten, sind z. B. private, interpersonelle Korrespondenz (für berufliche Zwecke) oder organisationsinterne Texte oder Mitteilungen (z. B. Veröffentlichungen in internen Unternehmensnetzwerken).

ii) **Zur Information der Öffentlichkeit:** Der durch KI erzeugte oder manipulierte veröffentlichte Text muss darauf abzielen, die Öffentlichkeit zu informieren. Dies bedeutet, dass mit dem Text Kenntnisse, Meinungen oder Fakten vermittelt werden sollten. Dagegen können kurze Texte, in denen Kenntnisse, Meinungen oder Fakten nicht wesentlich wiedergegeben werden, nicht als Information der Öffentlichkeit angesehen werden.

iii) **Über Angelegenheiten von öffentlichem Interesse:** Der Text muss die Öffentlichkeit „über Angelegenheiten von öffentlichem Interesse“ informieren. Im Allgemeinen sollten solche Angelegenheiten für die Gesellschaft insgesamt relevant sein, sei es auf lokaler, nationaler, Unions- oder internationaler Ebene, und einer öffentlichen Debatte oder Kontrolle würdig sein. In diesem Zusammenhang sollte davon ausgegangen werden, dass in den Texten Angelegenheiten von öffentlichem Interesses behandelt werden, wenn sie Themen aus den Bereichen Politik und demokratische Prozesse, öffentliche Verwaltung und Dienste, Justizverwaltung und Strafverfolgung, Schutz der Grundrechte, öffentliche Sicherheit, öffentliche Gesundheit, Umweltschutz, Verbrauchersicherheit sowie wirtschaftliche, finanzielle, politische, wissenschaftliche oder kulturelle Entwicklung abdecken, die für die öffentliche Debatte relevant sein können. Angelegenheiten, die als von öffentlichem Interesse angesehen werden können, können sich im Laufe der Zeit und über Kontexte hinweg weiterentwickeln.

Beispiele für Text, der veröffentlicht wird, um die Öffentlichkeit gemäß Artikel 50 Absatz 4 der KI-Verordnung über Angelegenheiten von öffentlichem Interesse zu informieren, und der als künstlich erzeugt oder manipuliert gekennzeichnet werden muss, es sei denn, er unterliegt einer menschlichen Überprüfung/redaktionellen Kontrolle (siehe Abschnitt 6.2.3):

— Durch KI erzeugte Zusammenfassung eines von Menschen verfassten Artikels auf der Website einer Zeitung, in dem über eine kürzlich ergangene Entscheidung eines Stadtrats diskutiert wird.

— Durch KI manipulierte Teile eines Lifestyle-Website-Artikels, in dem die Auswirkungen verschiedener Ernährungsweisen auf eine bestimmte Krankheit bei Frauen mittleren Alters verglichen werden.

— Durch KI manipulierte Unternehmensberichte, die auf der Website eines börsennotierten Unternehmens veröffentlicht werden und Informationen für Anleger enthalten.

— Durch KI erzeugte Mitteilung auf dem Profil eines meteorologischen Instituts in den sozialen Medien, mit der die Bürgerinnen und Bürger vor Sturm gewarnt und damit verbundene Vorsichtsmaßnahmen genannt werden.

Beispiele für Text, der nicht veröffentlicht wird, um die Öffentlichkeit über Angelegenheiten von öffentlichem Interesse zu informieren, der nicht in den Anwendungsbereich von Artikel 50 Absatz 4 der KI-Verordnung fällt:

— Durch KI erzeugte Fantasy-Romane.

— Durch KI manipulierter Text, der Teil der Werbung oder Produktbeschreibungen eines Unternehmens ist (ohne Angaben z. B. zu Gesundheit, Verbrauchersicherheit oder Nachhaltigkeit).

— Nachrichtenzusammenfassung, die von einem Chatbot erzeugt wird, der nur dem Nutzer zur Verfügung steht, der den Chatbot gepromptet hat.

— Durch KI manipulierter Text eines Beraters für eine Kundenberatung in Bezug auf Maßnahmen, die zur Einhaltung der geltenden Rechtsvorschriften zu ergreifen sind.

([38]) Vgl. entsprechend Mediakabel, C-89/04, EU:C:2005:348, Rn. 33; Lagardère, C-192/04, EU:C:2005:475, Rn. 31; SGAE, C-306/05, EU:C:2006:764, Rn. 37-40; Stichting Brein, C-527/15, EU:C:2017:300, Rn. 43-44 sowie die dort angeführte Rechtsprechung; GEMA, Rechtssache C-135/23, EU:C:2024:526, Rn. 38 sowie die dort angeführte Rechtsprechung.

6.2.2. *Die Offenlegungspflicht nach Artikel 50 Absatz 4 Unterabsatz 2 der KI-Verordnung*

132. Laut Artikel 50 Absatz 4 Unterabsatz 2 der KI-Verordnung müssen Betreiber eines KI-Systems, das Text erzeugt oder manipuliert, der veröffentlicht wird, um die Öffentlichkeit über Angelegenheiten von öffentlichem Interesse zu informieren, offenlegen, dass der Text künstlich erzeugt oder manipuliert wurde. In Erwägungsgrund 134 der KI-Verordnung wird präzisiert, dass Betreiber von Textveröffentlichungen, die in den Anwendungsbereich der Bestimmung fallen, ähnlich wie bei Deepfakes klar und eindeutig offenlegen sollten, dass solche Inhalte künstlich erzeugt oder manipuliert wurden, indem sie die KI-Ausgabe entsprechend kennzeichnen und deren künstlichen Ursprung offenlegen. Wie bei Deepfakes müssen die gemäß Artikel 50 Absatz 4 Unterabsatz 2 der KI-Verordnung angewandten Kennzeichnungs- oder Offenlegungsmethoden (einschließlich Haftungsausschlüssen) für natürliche Personen klar und wahrnehmbar sein (z. B. durch sichtbare oder akustische Maßnahmen), ohne dass sie auf spezifische technische Instrumente angewiesen sind oder gezielte Maßnahmen durchführen müssen. Nach Artikel 50 Absatz 7 der KI-Verordnung können sich Betreiber an einen Praxisleitfaden halten, der gemäß dieser Bestimmung als angemessen bewertet wird, um die Einhaltung dieser Offenlegungspflicht nachzuweisen. Weitere Informationen über die Auswirkungen der Einhaltung eines solchen Leitfadens und die Überwachung von Betreibern, die sich nicht an einen solchen Leitfaden halten, sind Abschnitt 8.1 zu entnehmen.

6.2.3. *Ausnahme von der Transparenzpflicht für Text unter menschlicher Überprüfung oder redaktioneller Kontrolle und redaktioneller Verantwortung*

133. Artikel 50 Absatz 4 Unterabsatz 2 der KI-Verordnung sieht eine Ausnahme von der in dieser Bestimmung festgelegten Transparenzpflicht vor, wenn zwei kumulative Bedingungen erfüllt sind: i) der durch KI erzeugte oder manipulierte Text muss einer menschlichen Überprüfung oder redaktionellen Kontrolle unterzogen worden sein, und ii) eine juristische oder natürliche Person muss die redaktionelle Verantwortung für die Veröffentlichung tragen. Die Betreiber können sich gegebenenfalls auf einschlägige geltende berufliche oder berufsethische Normen stützen, um die Einhaltung dieser Anforderungen wie nachstehend dargelegt nachzuweisen.

i. Text unter menschlicher Überprüfung oder redaktioneller Kontrolle

134. Die erste Voraussetzung für die Anwendung der Ausnahme von Artikel 50 Absatz 4 Unterabsatz 2 der KI-Verordnung erfordert, dass der durch KI erzeugte oder manipulierte Text einer menschlichen Überprüfung oder redaktionellen Kontrolle unterzogen wurde. Menschliche Überprüfung bezieht sich auf die bewusste Prüfung des wesentlichen Inhalts durch eine oder mehrere natürliche Personen, die über einschlägige Kenntnisse und fachliches Urteilsvermögen in Bezug auf den zu prüfenden Gegenstand verfügen (z. B. akademische gegenseitige Begutachtung oder professionelle Validierungsketten). Die Faktenprüfung im Hinblick auf die inhaltliche Richtigkeit ist eine Mindestanforderung, die Teil dieser Überprüfung sein sollte. Redaktionelle Kontrolle bezieht sich auf die Kontrolle, die in der Praxis von einer verantwortlichen redaktionellen Einrichtung (z. B. einem Chefredakteur) über den Inhalt ausgeübt wird, die befugt ist, den Text aus inhaltlichen Gründen zu genehmigen, zu ändern oder abzulehnen (einschließlich der Faktenprüfung von Informationen und der Gewährleistung der Vertrauenswürdigkeit von Quellen). Dies gilt unbeschadet bestehender Überprüfungs- und redaktioneller Verfahren und professioneller Standards, die für Mediendiensteanbieter gelten.

135. Oberflächliche, rein formale oder verfahrenstechnische Prüfungen (z. B. Rechtschreibprüfung oder grammatikalische Korrektur), das bloße Bestehen einer redaktionellen Strategie, automatisierte Überprüfungsverfahren oder flüchtige redaktionelle Genehmigung ohne substanzielle Auseinandersetzung seitens des menschlichen Überprüfers oder der redaktionellen Einrichtung können die Voraussetzungen für eine menschliche Überprüfung oder redaktionelle Kontrolle für die Zwecke dieser Ausnahme nicht erfüllen.

136. Werden KI-Systeme verwendet, um Inhalte nach einer redaktionellen Abzeichnung zu ändern, zu ergänzen oder umzuformulieren, müssen die daraus resultierenden Inhalte für die Zwecke des Artikels 50 Absatz 4 der KI-Verordnung als durch KI erzeugt oder manipuliert behandelt werden. Jede substanzielle KI-Intervention, die nach dem Prozess der menschlichen Überprüfung oder redaktionellen Kontrolle erfolgt, führt daher dazu, dass die Ausnahme hinfällig wird.

137. Nach Artikel 50 Absatz 7 der KI-Verordnung können sich Betreiber an einen Praxisleitfaden halten, der gemäß dieser Bestimmung als angemessen bewertet wird, um die Einhaltung der in dieser Bestimmung festgelegten Bedingung der menschlichen Überprüfung oder redaktionellen Kontrolle (insbesondere durch Betreiber, die keine Mediendiensteanbieter sind, die bestehenden redaktionellen beruflichen oder berufsethischen Normen unterliegen) nachzuweisen. Weitere Informationen über die Auswirkungen der Einhaltung eines solchen Leitfadens und die Überwachung von Anbietern, die sich nicht an einen solchen Leitfaden halten, sind Abschnitt 8.1 zu entnehmen.

ii. *Redaktionelle Verantwortung*

138. Die zweite Voraussetzung für die Anwendung der Ausnahme von Artikel 50 Absatz 4 Unterabsatz 2 der KI-Verordnung erfordert, dass eine juristische oder natürliche Person die redaktionelle Verantwortung für die Veröffentlichung der Inhalte trägt. Dies bedeutet, dass diese Person die letzte rechtliche Verantwortung für die Veröffentlichung der Inhalte tragen muss, einschließlich der menschlichen Überprüfung oder redaktionellen Kontrolle (z. B. eine Einzelperson, eine Redaktionsleitung oder das Verlagsunternehmen). Um die öffentliche Rechenschaftspflicht und das Vertrauen der Öffentlichkeit zu gewährleisten, sollten im Einklang mit den bestehenden professionellen Medienstandards die Identität und die Kontaktdaten der juristischen Person, der natürlichen Person oder der Funktion mit redaktioneller Verantwortung an einem leicht auffindbaren Ort öffentlich zugänglich gemacht werden (sofern noch nicht anderweitig verfügbar). Dies kann online erfolgen, z. B. über die allgemeinen Geschäftsbedingungen einer Website oder andere nutzerorientierte Rechtsinformationen. Offline können solche Informationen in das Kolophon oder Impressum einer Veröffentlichung aufgenommen werden.

Beispiele für Text, der einer menschlichen Überprüfung oder redaktionellen Kontrolle unterliegt und für den eine juristische oder natürliche Person die redaktionelle Verantwortung trägt:

— Ein durch KI manipulierter Artikel in einer Zeitung oder eine durch KI erzeugte Zusammenfassung eines Artikels, der der redaktionellen Kontrolle des jeweiligen Chefredakteurs mit redaktioneller Verantwortung durch die juristische Person, die die Zeitung veröffentlicht, unterliegt.

— Ein durch KI manipulierter akademischer Blog, der einer internen gegenseitigen Begutachtung unterzogen wurde und bei dem das jeweilige Forschungszentrum, das den Blog verwaltet, die redaktionelle Verantwortung trägt.

— Durch KI erzeugte öffentliche Sicherheitswarnungen, die von einem öffentlichen Bediensteten genehmigt wurden, bevor sie unter der Verantwortung der zuständigen öffentlichen Agentur für Bevölkerungsschutz an die Bürgerinnen und Bürger verteilt werden.

— Durch KI erzeugte Nachhaltigkeitsberichte, die auf der Website eines börsennotierten Unternehmens veröffentlicht und von Fachleuten in einschlägigen Funktionen (z. B. Compliance) einer menschlichen Überprüfung unterzogen wurden.

— Durch KI unterstützte Übersetzung eines vom Menschen verfassten Artikels, bei der die Übersetzung einer menschlichen Überprüfung unterzogen wurde.

Beispiele für Text, der nicht der erforderlichen menschlichen Überprüfung oder redaktionellen Kontrolle unterzogen wird und daher nicht in den Anwendungsbereich der Ausnahme fällt:

— Eine Website, auf der durch KI erzeugte Artikel zur Unionspolitik ohne bewusste menschliche Überprüfung oder redaktionelle Kontrolle veröffentlicht werden.

— Durch KI erzeugte Artikel, die von einem anderen KI-System überprüft und bearbeitet werden und bei denen ein menschlicher Redakteur vor der Veröffentlichung lediglich eine oberflächliche grammatikalische Prüfung durchführt.

— Ein durch KI erzeugtes, selbstveröffentlichtes Buch über den Klimawandel, das auf einer E-Commerce-Plattform zur Verfügung gestellt wird und keiner Überprüfung durch eine kompetente natürliche oder juristische Person (und auch nicht durch die Plattform) unterzogen wurde.

6.2.4. *Ausnahme für Strafverfolgung*

139. Wenn die Verwendung von KI-Systemen zur Erzeugung von Text, der ohne menschliche Überprüfung veröffentlicht wird, um die Öffentlichkeit über Angelegenheiten von öffentlichem Interesse zu informieren (z. B. eine durch KI erzeugte polizeiliche Warnung vor einer terroristischen Bedrohung, die in sozialen Medien veröffentlicht wird), gesetzlich zur Aufdeckung, Verhütung, Ermittlung oder Verfolgung von Straftaten zugelassen ist, sind Betreiber von der Transparenzpflicht nach Artikel 50 Absatz 4 der KI-Verordnung ausgenommen. Die Erläuterungen unter den Randnummern 46 bis 48 für eine ähnliche Ausnahme in Artikel 50 Absatz 1 der KI-Verordnung sind auch für diese Bestimmung relevant.

6.2.5. *Zusammenspiel mit anderen Rechtsakten der Union*

140. Im Medienrecht der Union ist die redaktionelle Verantwortung ein entscheidendes Konzept, das in Artikel 2 Nummer 8 der Verordnung (EU) 2024/1083 (Europäisches Medienfreiheitsgesetz oder EMFA) ([39]) definiert ist. In diesem Kontext bezeichnet „redaktionelle Verantwortung" „die Ausübung einer wirksamen Kontrolle sowohl hinsichtlich der Auswahl der Sendungen oder Presseveröffentlichungen als auch hinsichtlich ihrer Zusammenstellung für die Zwecke der Bereitstellung eines Mediendienstes, unabhängig von dem Bestehen einer Haftung für den bereitgestellten Dienst nach nationalem Recht". In bestimmten Situationen kann sich diese Definition mit dem in Artikel 50 Absatz 4 Unterabsatz 2 der KI-Verordnung verwendeten Begriff der redaktionellen Verantwortung überschneiden, insbesondere wenn die Betreiber des KI-Systems auch als Mediendiensteanbieter im Sinne des EMFA gelten. Die redaktionelle Verantwortung im Sinne von Artikel 50 Absatz 4 der KI-Verordnung sollte daher im Einklang mit dem bestehenden Besitzstand im Medienbereich ausgelegt werden, ungeachtet der Tatsache, dass es sich nach wie vor um ein eigenständiges Konzept handelt, das auch in einem breiteren Kontext und für andere Betreiber gelten kann, sofern sie die redaktionelle Verantwortung für die Textveröffentlichung übernehmen. Mediendiensteanbieter können sich auf ihre bestehenden redaktionellen Prozesse und Standards stützen, um die Ausnahme nach Artikel 50 Absatz 4 Unterabsatz 2 der KI-Verordnung in Anspruch nehmen zu können.

7. **HORIZONTALE ANFORDERUNGEN FÜR DIE NACH ARTIKEL 50 ABSATZ 5 DER KI-VERORDNUNG BEREITGESTELLTEN INFORMATIONEN**

141. Laut Artikel 50 Absatz 5 der KI-Verordnung sollten die gemäß Artikel 50 Absatz 1 bis 4 der KI-Verordnung bereitzustellenden Informationen den betreffenden natürlichen Personen spätestens zum Zeitpunkt der ersten Interaktion oder Aussetzung in klarer und eindeutiger Weise bereitgestellt werden. Die Informationen sollten zudem den geltenden Barrierefreiheitsanforderungen entsprechen. Zu den in diesem Zusammenhang betroffenen natürlichen Personen gehören die Personen, die direkt mit dem KI-System interagieren (Artikel 50 Absatz 1 der KI-Verordnung), aktive oder passive Nutzer und andere Personen, die den durch KI erzeugten oder manipulierten synthetischen Inhalten ausgesetzt sind (mit oder ohne direkte Interaktion mit dem KI-System) (Artikel 50 Absätze 2 und 4 der KI-Verordnung), sowie Personen, die der Verwendung von Emotionserkennungssystemen oder Systemen zur biometrischen Kategorisierung unterliegen (Artikel 50 Absatz 3 der KI-Verordnung).

7.1. **Klare und eindeutige Informationsbereitstellung**

142. Informationen gelten als in klarer Weise bereitgestellt, wenn sie für die betreffende natürliche Person (einschließlich Menschen mit Behinderungen) wahrnehmbar, leicht verständlich und zugänglich sind. Informationen gelten als auf eindeutige Weise bereitgestellt, wenn leicht erkennbar ist, dass sie von anderen Informationen und der Umgebung, in der der Inhalt dargestellt wird, getrennt sind. Sie müssen auch für das Publikum leicht verständlich sein, auch für bestimmte Gruppen wie Kinder oder Menschen mit Behinderungen, wenn sie dem Publikum angehören, das den Inhalten voraussichtlich ausgesetzt ist oder mit dem System interagiert ([40]). Um die Informationen klar und eindeutig bereitzustellen, können die einschlägigen Informationen unter anderem im Rahmen der Interaktion (gemäß Artikel 50 Absatz 1 der KI-Verordnung) oder als Teil der einschlägigen Inhalte (gemäß Artikel 50 Absätze 2 und 4 der KI-Verordnung) dargestellt werden. Informationen gelten nicht als klar und eindeutig bereitgestellt, wenn sie von natürlichen Personen unter normalen Aussetzungs- oder Interaktionsbedingungen leicht übersehen werden können (z. B. nur als Teil eines Handbuchs oder versteckt unter Menüoptionen in einer Online-Schnittstelle, Teil der Nutzungsbedingungen, die von den Nutzern häufig nicht gelesen werden).

([39]) Verordnung (EU) 2024/1083 des Europäischen Parlaments und des Rates vom 11. April 2024 zur Schaffung eines gemeinsamen Rahmens für Mediendienste im Binnenmarkt und zur Änderung der Richtlinie 2010/13/EU (Europäisches Medienfreiheitsgesetz) (ABl. L, 2024/1083, 17.4.2024, ELI: http://data.europa.eu/eli/reg/2024/1083/oj).

([40]) Wenn Kinder zur Zielgruppe der Inhalte gehören, müssen diese Mitteilungen: i) kindgerecht, altersgerecht und für alle Kinder, einschließlich Kindern mit Behinderungen und/oder zusätzlichen Anforderungen an die Zugänglichkeit, leicht verständlich und leicht zugänglich sein; ii) übersichtlich, verständlich und möglichst einfach und prägnant dargeboten werden; iii) auf leicht überprüfbare Weise dargeboten werden, die einen sofortigen und intuitiven Zugang ermöglicht, und zwar dann, wenn sie Bedeutung erlangen; iv) in der/den Amtssprache(n) des Mitgliedstaats dargeboten werden, in dem der Dienst bereitgestellt wird; v) für Kinder ansprechend gestaltet sein. Dies kann auch die Verwendung von Grafiken, Videos und/oder Zeichen oder anderen Techniken umfassen; vi) den Kindern schrittweise und allmählich dargeboten werden, damit diese sie möglichst gut im Gedächtnis behalten.

7.2. **Erste Interaktion oder Aussetzung**

143. Nach Artikel 50 Absatz 5 der KI-Verordnung müssen die Informationen den betroffenen natürlichen Personen spätestens zum Zeitpunkt der ersten Interaktion oder Aussetzung zur Verfügung gestellt werden. Die erste Interaktion oder Aussetzung bezieht sich nicht nur auf die erste natürliche Person, die mit einem KI-System interagiert oder einem KI-System (bzw. dessen Ausgabe) ausgesetzt ist, sondern auch auf jede nachfolgende erste Interaktion mit oder Aussetzung gegenüber dem KI-System (bzw. dessen Ausgabe) durch eine andere natürliche Person. Was die erste Interaktion mit einem KI-System betrifft, so sollten die Informationen mindestens einmal zu Beginn einer interaktiven Sitzung bereitgestellt werden, wobei die unter den Randnummern 36 und 40 genannten spezifischen Kontexte zu berücksichtigen sind. In Bezug auf durch KI erzeugte oder manipulierte Inhalte, die in den Anwendungsbereich von Artikel 50 Absätze 2 und 4 der KI-Verordnung fallen, gilt die Informationspflicht für jede Ausgabe eines KI-Systems in Bezug auf jede natürliche Person, die den Inhalten ausgesetzt ist. Sie sollte im Sinne jedes Zeitpunkts verstanden werden, zu dem eine natürliche Person mit hinreichender Wahrscheinlichkeit der Ausgabe des KI-Systems ausgesetzt ist und die Offenlegung wahrnimmt, wobei die Besonderheiten der Offenlegung im Falle offensichtlich kreativer, künstlerischer und anderer Werke zu berücksichtigen sind (siehe Abschnitt 6.1.3). Darüber hinaus dürfen Betreiber natürliche Personen vor der ersten tatsächlichen Interaktion oder Aussetzung informieren (z. B. zu Beginn von Inhalten mit Deepfakes). Ist jedoch vernünftigerweise vorhersehbar, dass Personen Inhalte möglicherweise nicht von Anfang an wahrnehmen, so werden diese Personen durch die bloße Offenlegung zu Beginn der Inhalte nicht angemessen informiert und es sollte nach Möglichkeit zusätzlich eine Offenlegung zu späteren Zeitpunkten erfolgen. Dies kann gegebenenfalls umgesetzt werden, wobei dem Ziel der Transparenzpflichten Rechnung zu tragen ist, eine wirksame Bereitstellung der Informationen und eine sinnvolle Sensibilisierung der Nutzer in zunehmend immersiven und realistischen KI-Umgebungen zu gewährleisten, in denen die Nutzer zunehmend Schwierigkeiten haben, den Ursprung und die Wahrhaftigkeit der Inhalte und der Interaktion zu unterscheiden.

Beispiele für die Bereitstellung von Informationen bei der ersten Interaktion oder Aussetzung gemäß Artikel 50 Absatz 5 der KI-Verordnung:

— Erste Interaktion: beim Starten eines Gesprächs mit einem Chatbot oder beim Beginn der physischen Interaktion mit einem KI-System.

— Erste Aussetzung: zu Beginn eines Videos mit Deepfakes, zu Beginn einer durch KI erzeugten oder manipulierten Textveröffentlichung zu Angelegenheiten von öffentlichem Interesse oder beim Auffinden von durch KI manipulierten Deepfakes beim Scrollen in sozialen Medien.

7.3. **Einhaltung von geltenden Barrierefreiheitsanforderungen**

144. Anbieter und Betreiber von KI-Systemen, die unter die verschiedenen in Artikel 50 Absätze 1 bis 4 der KI-Verordnung aufgeführten Transparenzpflichten fallen, müssen sich der geltenden Barrierefreiheitsanforderungen (z. B. gemäß der Richtlinie 2016/2102 ([41]) und der Richtlinie 2019/882 ([42])) bewusst sein und bewerten, ob ihr Produkt, ihre Dienstleistung oder ihr Inhalt diesen Vorschriften entsprechen muss. In einem solchen Fall müssen sie die Zugänglichkeit der gemäß Artikel 50 der KI-Verordnung übermittelten Informationen im Einklang mit den geltenden Barrierefreiheitsanforderungen sicherstellen. Artikel 50 der KI-Verordnung enthält keine gesonderten oder zusätzlichen Barrierefreiheitsanforderungen.

8. **DURCHSETZUNG VON ARTIKEL 50 DER KI-VERORDNUNG**

145. In diesem Abschnitt soll Folgendes präzisiert werden: i) die Auswirkungen der Befolgung eines Praxisleitfadens in Bezug auf die Einhaltung der Transparenzpflichten für durch KI erzeugte Inhalte; ii) die Rolle der zuständigen Aufsichtsbehörden bei der Durchsetzung von Artikel 50 der KI-Verordnung; iii) die Sanktionen, die bei Verstößen gegen diese Pflichten verhängt werden; iv) der Zeitpunkt, zu dem die Pflichten nach Artikel 50 der KI-Verordnung in Kraft treten.

([41]) Richtlinie (EU) 2016/2102 des Europäischen Parlaments und des Rates vom 26. Oktober 2016 über den barrierefreien Zugang zu den Websites und mobilen Anwendungen öffentlicher Stellen (Text von Bedeutung für den EWR) (ABl. L 327 vom 2.12.2016, S. 1).

([42]) Richtlinie (EU) 2019/882 des Europäischen Parlaments und des Rates vom 17. April 2019 über die Barrierefreiheitsanforderungen für Produkte und Dienstleistungen (Text von Bedeutung für den EWR), PE/81/2018/REV/1 (ABl. L 151 vom 7.6.2019, S. 70).

8.1. **Auswirkungen der Befolgung eines als angemessen bewerteten Praxisleitfadens**

146. Anbieter und Betreiber von KI-Systemen, die in den Anwendungsbereich von Artikel 50 Absätze 2 und 4 der KI-Verordnung fallen, können die Einhaltung ihrer jeweiligen Transparenzpflichten für durch KI erzeugte Inhalte gemäß diesen Bestimmungen nachweisen, indem sie einen Praxisleitfaden befolgen, der gemäß Artikel 50 Absatz 7 der KI-Verordnung als angemessen bewertet wird ([43]). Ein solcher Leitfaden ersetzt weder die KI-Verordnung noch diese Leitlinien. Er ergänzt sie, indem er den einzigen unionsweit anerkannten praktischen Rahmen für Anbieter und Betreiber generativer KI-Systeme bietet, um die Einhaltung dieser Pflichten nachzuweisen, und zwar unabhängig von ihrem Sitz, ihrem Betrieb und den zuständigen Marktüberwachungsbehörden.

147. Während Anbieter und Betreiber die Einhaltung dieser Pflichten auch durch geeignete alternative Mittel nachweisen können, ist die Befolgung eines als angemessen erachteten Praxisleitfadens eine einfache, vorhersehbare und rechtlich sichere Art und Weise, die Einhaltung nachzuweisen. Bei Unterzeichnern eines Praxisleitfadens, der gemäß Artikel 50 Absatz 7 der KI-Verordnung als angemessen bewertet wird ([44]), werden die Kommission und die zuständigen Marktüberwachungsbehörden ihre Aufsichtstätigkeiten auf die Bewertung konzentrieren, ob diese Unterzeichner den Praxisleitfaden befolgt und die darin enthaltenen Maßnahmen umgesetzt haben. Da die befolgenden Anbieter und Betreiber in Bezug auf die Maßnahmen, die sie ergreifen, um Artikel 50 Absätze 2, 4 und 5 der KI-Verordnung nachzukommen, transparent sein werden, werden sie von einem größeren Vertrauen der Kommission, der anderen zuständigen Marktüberwachungsbehörden und anderer Interessenträger (einschließlich der breiten Öffentlichkeit) profitieren. Jede von den Unterzeichnern eines als angemessen erachteten Praxisleitfadens getroffene Ausnahme von den Abschnitten führt dazu, dass diese Anbieter und Betreiber den Vorteil verlieren, den Nachweis der Einhaltung in dieser Hinsicht zu erleichtern.

148. Von Anbietern und Betreibern, die keinen Praxisleitfaden unterzeichnet haben, der gemäß Artikel 50 Absatz 7 der KI-Verordnung als angemessen erachtet wird, wird erwartet, dass sie durch andere geeignete Mittel nachweisen, wie sie ihren Pflichten gemäß Artikel 50 Absätze 2, 4 und 5 der KI-Verordnung nachgekommen sind. Darüber hinaus wird von diesen Anbietern und Betreibern erwartet, dass sie erläutern, wie mit den von ihnen umgesetzten Maßnahmen die Einhaltung ihrer Pflichten aus der KI-Verordnung sichergestellt wird. So sollten sie beispielsweise eine Lückenanalyse durchführen, bei der die von ihnen umgesetzten Maßnahmen mit den Maßnahmen verglichen werden, die in einem als angemessen bewerteten Praxisleitfaden festgelegt sind. Da die zuständigen Behörden (einschließlich des Büros für Künstliche Intelligenz) weniger Verständnis darüber haben werden, wie Anbieter und Betreiber, die keinen Praxisleitfaden unterzeichnet haben, die Einhaltung ihrer Pflichten gemäß Artikel 50 Absätze 2, 4 und 5 der KI-Verordnung sicherstellen, werden sie bei der Überwachung der Einhaltung wahrscheinlich detailliertere Informationen benötigen. Beispielsweise können Anbieter, die keinen Praxisleitfaden befolgen, auch einer größeren Zahl von Auskunftsersuchen und Anträgen auf Zugang unterliegen, um die Wirksamkeit, Interoperabilität, Belastbarkeit und Zuverlässigkeit der technischen Lösungen zu bewerten, die zur Einhaltung von Artikel 50 Absatz 2 der KI-Verordnung umgesetzt werden. Betreiber können solchen Ersuchen auch in Bezug auf ihre Kennzeichnungspraktiken gemäß Artikel 50 Absatz 4 der KI-Verordnung unterliegen.

149. Die zuständigen Behörden können Verpflichtungen, die im Einklang mit einem als angemessen erachteten Praxisleitfaden umgesetzt werden, je nach den besonderen Umständen als mildernden Umstand bei der Festsetzung der Höhe von Geldbußen berücksichtigen ([45]).

150. Wird ein Praxisleitfaden gemäß Artikel 50 Absatz 7 der KI-Verordnung als nicht angemessen erachtet, kann die Kommission einen Durchführungsrechtsakt erlassen, in dem gemeinsame Vorschriften für die Umsetzung der Pflichten aus Artikel 50 Absätze 2, 4 und 5 der KI-Verordnung festgelegt werden, die für alle einschlägigen Anbieter und Betreiber von KI-Systemen gelten würden, die in den Anwendungsbereich dieser Bestimmungen fallen ([46]).

([43]) Siehe einschlägige Informationen über den Verhaltenskodex für die Transparenz von KI-generierten Inhalten | Gestaltung der digitalen Zukunft Europas, veröffentlicht am 10. Juni 2026.

([44]) Siehe Stellungnahme der Kommission zur Bewertung des Verhaltenskodex zur Transparenz von KI-generierten Inhalten | Gestaltung der digitalen Zukunft Europas.

([45]) Artikel 99 Absatz 7 Buchstabe e der KI-Verordnung.

([46]) Artikel 50 Absatz 7 der KI-Verordnung.

8.2. **Marktüberwachungsbehörden**

151. Die von den Mitgliedstaaten benannten Marktüberwachungsbehörden, das Büro für Künstliche Intelligenz ([47]) und der Europäische Datenschutzbeauftragte ([48]) sind für die Überwachung und Durchsetzung der Vorschriften für KI-Systeme zuständig, die in ihre Zuständigkeit fallen, einschließlich der Transparenzpflichten nach Artikel 50 der KI-Verordnung. Diese Durchsetzung erfolgt im Rahmen des Systems der Marktüberwachung und der Konformität von Produkten, das mit der Verordnung (EU) 2019/1020 und der KI-Verordnung eingerichtet wurde. Die Durchsetzungsbefugnisse der Marktüberwachungsbehörden in Bezug auf KI-Systeme sind in diesen Verordnungen festgelegt. Diese Behörden können von sich aus oder aufgrund einer Beschwerde, zu deren Einreichung jede betroffene Person oder jede andere natürliche oder juristische Person, die Grund zu der Annahme hat, dass solche Verstöße vorliegen, berechtigt ist, Durchsetzungsmaßnahmen in Bezug auf die in Artikel 50 der KI-Verordnung genannten Pflichten ergreifen ([49]). Die Zuständigkeiten der Marktüberwachungsbehörden nach der KI-Verordnung berühren nicht die Befugnisse und Aufgaben anderer Aufsichtsbehörden oder Grundrechtebehörden nach ihrem eigenen Rechtsrahmen in Situationen, in denen sowohl Artikel 50 der KI-Verordnung als auch dieser Rahmen Anwendung findet (z. B. Datenschutz, Verbraucherschutz).

8.3. **Sanktionen**

152. Anbieter und Betreiber, die die geltenden Transparenzpflichten nach Artikel 50 der KI-Verordnung nicht erfüllen, können mit einer Geldbuße von bis zu 15 000 000 EUR oder, wenn es sich bei dem Täter um ein Unternehmen handelt, von bis zu 3 % seines weltweiten Jahresumsatzes des vorangegangenen Geschäftsjahres belegt werden, je nachdem, welcher Betrag höher ist. Gegen Organe, Einrichtungen und sonstige Stellen der EU, die gegen die Transparenzpflichten verstoßen, können Geldbußen von bis zu 750 000 EUR verhängt werden. Bei kleinen und mittleren Unternehmen (KMU), einschließlich Start-up-Unternehmen, beträgt jede Geldbuße bis zu den oben genannten Prozentsätzen oder Beträgen, je nachdem, welcher Betrag niedriger ist. Die Geldbußen müssen wirksam, verhältnismäßig und abschreckend sein und den Interessen von KMU und kleinen Midcap-Unternehmen, einschließlich Start-up-Unternehmen, sowie ihrer wirtschaftlichen Tragfähigkeit Rechnung tragen. Bei der Entscheidung über die Verhängung einer Geldbuße und bei der Entscheidung über die Höhe der Geldbuße in jedem Einzelfall sollten die zuständigen Behörden alle relevanten Umstände der jeweiligen Situation und gegebenenfalls eine Reihe von Faktoren bewerten, die die Schwere des Verstoßes beeinflussen (z. B. Art, Schwere und Dauer des Verstoßes und seiner Folgen, Fahrlässigkeit oder Vorsätzlichkeit, andere erschwerende oder mildernde Umstände, Umfang der Zusammenarbeit mit den Marktüberwachungsbehörden usw.).

8.4. **Anwendungsbeginn**

153. Gemäß Artikel 113 der KI-Verordnung gilt Artikel 50 der KI-Verordnung ab dem 2. August 2026. Dies setzt voraus, dass alle in den Anwendungsbereich fallenden KI-Systeme, die in der Union in Verkehr gebracht oder in Betrieb genommen werden, zu diesem Zeitpunkt dieser Bestimmung entsprechen, unabhängig von ihrem Datum des Inverkehrbringens oder der Inbetriebnahme ([50]). Die Verordnung zur Änderung der KI-Verordnung („KI-Omnibus“), die kürzlich vom Unionsgesetzgeber angenommen wurde, sieht eine gezielte Bestandsschutzregelung nur in Bezug auf die Kennzeichnungs- und Feststellungspflichten gemäß Artikel 50 Absatz 2 der KI-Verordnung für generative KI-Systeme vor, die vor dem 2. August 2026 in Verkehr gebracht oder in Betrieb genommen wurden. Den Anbietern dieser bestehenden Systeme wird eine Übergangsfrist eingeräumt, um ihre Systeme bis zum 2. Dezember 2026 mit den Anforderungen in Einklang zu bringen. Systeme, die teilweise interaktiv und teilweise generativ sind, können diesen Übergangszeitraum nur in Bezug auf die Kennzeichnungspflicht nach Artikel 50 Absatz 2 der KI-Verordnung in Anspruch nehmen, während die Einhaltung der Offenlegungspflicht für KI-Systeme, die direkt mit natürlichen Personen interagieren, ab dem 2. August 2026 sichergestellt werden muss.

([47]) Als Marktüberwachungsbehörde für KI-Systeme, die auf von demselben Anbieter bereitgestellten KI-Modellen mit allgemeinem Verwendungszweck beruhen (Artikel 75 Absatz 1 der KI-Verordnung). Dies umfasst die Pflicht für Anbieter gemäß Artikel 50 Absätze 1 und 2 der KI-Verordnung. In Bezug auf Artikel 50 Absätze 3 und 4 der KI-Verordnung ist das Büro für Künstliche Intelligenz nur dann zuständig, wenn der jeweilige Anbieter von KI-Modellen und -Systemen mit allgemeinem Verwendungszweck auch Betreiber des Systems ist.

([48]) Als Marktüberwachungsbehörde für die Organe, Einrichtungen und sonstigen Stellen der EU.

([49]) Artikel 85 der KI-Verordnung.

([50]) Die besondere Bestandsschutzregelung in Artikel 111 Absatz 2 der KI-Verordnung gilt für Hochrisiko-KI-Systeme, die vor dem Geltungsbeginn der Hochrisikovorschriften in Verkehr gebracht wurden, nur in Bezug auf die Konformität dieser Systeme mit den Anforderungen und Pflichten für Hochrisiko-KI-Systeme. Unterliegt das Hochrisiko-KI-System auch einer oder mehreren Transparenzpflichten nach Artikel 50 der KI-Verordnung, so gilt diese besondere Bestandsschutzregelung nicht, da die verschiedenen Pflichten für Hochrisiko-KI-Systeme und die Transparenz nach Artikel 50 kumulativ gelten und der Bestandsschutz nur im Hinblick auf die Einhaltung der Hochrisikobestimmungen gerechtfertigt und begrenzt ist.

154. Durch KI erzeugte oder manipulierte Ausgaben, die in den Anwendungsbereich von Artikel 50 Absatz 2 der KI-Verordnung fallen, und Deepfakes, die in den Anwendungsbereich von Artikel 50 Absatz 4 Unterabsatz 1 der KI-Verordnung fallen und vor dem 2. August 2026 erzeugt oder manipuliert wurden, müssen nicht rückwirkend gekennzeichnet werden. Gleiches gilt für Texte zu Angelegenheiten von öffentlichem Interesse, die in den Anwendungsbereich von Artikel 50 Absatz 4 Unterabsatz 2 der KI-Verordnung fallen und vor dem 2. August 2026 durch KI erzeugt oder manipuliert und veröffentlicht wurden. Umgekehrt müssen Texte, die vor dem 2. August 2026 erzeugt oder manipuliert wurden und an oder nach diesem Datum veröffentlicht werden, gekennzeichnet werden. Angesichts des mit der KI-Verordnung verfolgten Ziels der Transparenz und der Stärkung des Vertrauens in das Informationsökosystem und dessen Integrität werden Betreiber von KI-Systemen und andere Akteure, die bereits bestehende nicht gekennzeichnete Deepfakes besitzen oder verbreiten, dennoch ermutigt, dies zu tun, ohne dass von ihnen zu diesem Zweck unverhältnismäßige Anstrengungen erwartet werden (z. B. Prüfung bereits bestehender Inhaltsdatenbanken oder Änderung bereits gedruckter Produktverpackungen).

9. **ÜBERPRÜFUNG UND AKTUALISIERUNG DER LEITLINIEN**

155. Diese Leitlinien stellen eine erste Auslegung mit praktischen Beispielen für die in Artikel 50 der KI-Verordnung festgelegten Transparenzpflichten dar. Die Kommission wird diese Leitlinien überprüfen, sobald dies angesichts der praktischen Erfahrungen bei der Umsetzung der Transparenzpflichten und des Tempos der technologischen, gesellschaftlichen und regulatorischen Entwicklungen in diesem Bereich erforderlich ist. Dazu gehören auch einschlägige Erfahrungen aus Durchsetzungsmaßnahmen im Rahmen der Marktüberwachung und Auslegungen des Artikels 50 der KI-Verordnung durch den EuGH. Im Verlauf einer solchen Überprüfung kann die Kommission beschließen, diese Leitlinien zurückzuziehen oder zu ändern. Die Kommission ermutigt Anbieter und Betreiber von KI-Systemen, nationale Marktüberwachungsbehörden über das KI-Gremium, das KI-Beratungsforum, die Forschungsgemeinschaft und Organisationen der Zivilgesellschaft, zu diesem Prozess beizutragen, indem sie auf künftige Aufforderungen zur öffentlichen Konsultation reagieren.