bitkom

Version 1.0 - Stand 29.10.2024

Umsetzungsleitfaden zur KI-Verordnung

Compliance in der Praxis – Schritt für Schritt

Inhalt

Geleitwort

Der vorliegende Umsetzungsleitfaden wurde federführend von den Mitgliedern des Arbeitskreises Artificial Intelligence des Bitkom erstellt. Besonderer Dank gilt den folgenden Autorinnen und Autoren, die sich mit viel Engagement der Erstellung des Leitfadens gewidmet haben:

- Sandra Baum (Bundesdruckerei GmbH),
- Dr. Frank Beer INFODAS Gesellschaft für Systementwicklung und Informationsverarbeitung mbH),
- Eric Behrendt (TÜV Informationstechnik GmbH),
- Susan Bischoff (Morrison & Foerster LLP),
- Arnd Böken (GvW Graf von Westphalen Rechtsanwälte Steuerberater Partnerschaft mbB),
- Jan Breuer (Detecon International GmbH),
- Camilla Dalerci (Bundesdruckerei GmbH),
- Vasilios Danos (TÜV Informationstechnik GmbH),
- Prof. Dr. Heinz-Uwe Dettling Ernst & Young GmbH Wirtschaftsprüfungsgesellschaft),
- Marius Drabiniok (SKW Schwarz Rechtsanwälte),
- Markus Frowein (RWE AG),
- Dr. Axel Grätz (Oppenhoff & Partner Rechtsanwälte Steuerberater mbB),
- Valentino Halim (Oppenhoff & Partner Rechtsanwälte Steuerberater mbB),
- Janis Hecker (Bitkom),
- Dr. Rachel Hegemann (Deutsche Bahn AG),
- Benedict Huyeng (RWE AG),
- Sven Jacobs (Cisco Systems GmbH),
- Ali-Reza Khalaji (R+V Versicherung AG),
- Stephan Kress (Morrison & Foerster LLP),
- Dr. Christoph Krück (SKW Schwarz Rechtsanwälte),
- Malte Lange (Finanz Informatik GmbH & Co. KG),
- Dr. Kim Lauenroth (Fachhochschule Dortmund),
- Dr. Anastasia Linnik (Retresco GmbH),
- Stefan Mangold (Datev eG),

- Martin Meyer (Siemens Healthcare GmbH),
- Dilan Mienert (GÖRG Partnerschaft von Rechtsanwälten mbB),
- Lea Ludmilla Ossmann-Magiera (LL.M. Leiden) (Bitkom),
- Hung Pham (INFODAS Gesellschaft für Systementwicklung und Informationsverarbeitung mbH),
- Filipp Revinzon (Vay Technology GmbH),
- Lys Riemenschneider (Holisticon AG),
- Dr. Benedikt Rohrßen (Taylor Wessing Partnerschaftsgesellschaft mbB),
- Tim Sauerhammer (Reed Smith LLP),
- Jan-Dierk Schaal, LL.M (University of Melbourne) (SKW Schwarz Rechtsanwälte),
- Michael Schemel (UnternehmerTUM GmbH),
- Alexander Schmalenberger (Taylor Wessing Partnerschaftsgesellschaft mbB),
- Ferdinand Schwarz (SKW Schwarz Rechtsanwälte),
- Maria Stammwitz (Bundesdruckerei GmbH),
- Christiane Stützle (Morrison & Foerster LLP),
- Frank Wisselink (Deutsche Telekom AG).

1 Einleitung

Zur Zielsetzung und Anwendung dieses Leitfadens

Lea Ludmilla Ossmann-Magiera LL.M. (Leiden) (Bitkom)

Der vorliegende Umsetzungsleitfaden soll **Unternehmen**, die KI-Systeme in Verkehr bringen oder betreiben, dabei unterstützen, die rechtlichen Vorgaben der Verordnung (EU) 2022/1689 vom 13. Juni 2024 zur Festlegung harmonisierter Vorschriften für künstliche Intelligenz (nachfolgend KI-VO) erfolgreich und im Sinne des Gesetzgebers zu implementieren.

Um dieses Ziel zu erreichen, sind die einschlägigen Regelungen aus der KI-VO und die dazugehörigen Erwägungsgründe analysiert und in dem vorliegenden Dokument aufbereitet worden. Um den Leitfaden so praktikabel und funktional wie möglich zu gestalten, ist er in **Prüfungsschritte** gegliedert, die von den Leserinnen und Lesern nacheinander abgearbeitet werden können. Die Prüfungsschritte sind dabei als Fragen formuliert und es finden sich jeweils Hinweise, welcher Prüfungsschritt – je nach Antwort – als Nächstes folgt.

Konkret wird in den jeweiligen Prüfungsschritten in den rot umrandeten Kästen der **Gesetzestext der KI-VO** wiedergegeben. Danach folgen Ausführungen zu den gesetzlichen Voraussetzungen, die vor allem auf dem Gesetzestext selbst und den zugehörigen Erwägungsgründen der Verordnung basieren.

Die KI-VO enthält eine Reihe von unbestimmten Rechtsbegriffen, die vom Gesetzgeber) bewusst eingesetzt werden, um die gesetzlichen Regelungen flexibel zu gestalten und sie so anpassungsfähig an technologische Entwicklungen zu machen. Aus diesem Grund sind viele Regelungen der KI-VO in hohem Maße **auslegungsbedürftig.** Teilweise sind auch **Konkretisierungen** durch die EU-Kommission und andere Stakeholder – etwa **durch technische Norm- und Standardsetzung, delegierte Rechtsakte, Leitlinien oder Verhaltenskodizes** – ausdrücklich vorgesehen. In diesem Fällen nimmt der Umsetzungsleitfaden auf den neuesten Stand der entsprechenden Konkretisierungsmaßnahmen Bezug. Darüber hinaus werden Grenzfälle aufgeführt, bei denen sich aus dem Gesetzestext nicht eindeutig ergibt, ob ein KI-System, die Voraussetzungen erfüllt oder nicht. Diese praktischen Beispiele im Grenzbereich („Edge Cases“) sollen den Leserinnen und Lesern des Leitfadens dabei helfen, die rechtliche Einordnung ihres eigenen KI-Systems besser einschätzen zu können.

Da die konkretisierenden Maßnahmen im Zeitpunkt der Veröffentlichung des Umsetzungsleitfadens noch erarbeitet werden, ist dieser als **„lebendes Dokument“** zu betrachten, das im Zuge neuer konkretisierender Rechtsakte, neuer Standards und Verhaltenskodizes stetig weiterzuentwickeln und zu aktualisieren ist.

An den Anfang des Leitfadens sind einige grundlegende Informationen über die gesetzgeberischen Ziele, den Gesetzgebungsprozess und den Regulierungsansatz der KI-VO gestellt. Diese sollen dem allgemeinen Verständnis dienen und sind nicht Teil der konkreten

69%

der Unternehmen geben an, Unterstützung bei der Auseinandersetzung mit der KI-VO zu brauchen (Bitkom-Studie 2024)

Prüfungsschritte, die im Hinblick auf die „KI-Compliance“ im Unternehmen geprüft werden sollten.

Folgende **übergeordnete Fragen** werden im Umsetzungsleitfaden behandelt:

- Handelt es sich bei meinem System um ein **KI-System i.S.d. KI-VO**?
- Bin ich **Regulierungsadressat?**
- Ist der **räumliche Anwendungsbereich** eröffnet?
- In welche **Risikokategorie** fällt mein KI-System?
- Welche **Pflichten** muss ich als **Anbieter** bzw. als **Betreiber** von **Hochrisiko-KI-Systemen** erfüllen?
- Welche **Pflichten** muss ich in Bezug auf **KI-Systeme** mit **geringem Risiko** erfüllen?
- Welche Besonderheiten gelten für **KI-Systeme und -Modelle mit allgemeinem Verwendungszweck**?
- Wie weise ich **Konformität** mit den Anforderungen aus der KI-VO nach?
- Welche **Pflichten** muss ich **nach dem Inverkehrbringen** erfüllen?

2 Grundlagen KI-VO

Dr. Rachel Hegemann (Deutsche Bahn AG), Stefan Mangold (Datev eG), Martin Meyer (Siemens Healthcare GmbH), Lea Ludmilla Ossmann-Magiera (LL.M. Leiden) (Bitkom), Lys Riemenschneider (Holisticon AG)

Gesetzgeberische Ziele

Die **Verordnung (EU) 2024/1689** zur Festlegung harmonisierter Vorschriften für künstliche Intelligenz (nachfolgend KI-VO) zielt darauf ab, das Vertrauen in KI zu stärken, Sicherheit von KI-Systemen zu gewährleisten und Innovation zu fördern. Ihre Rechtsgrundlage findet sich in Art. 114 Abs. 1 AEUV, demgemäß das Europäische Parlament und der Rat Maßnahmen erlassen, welche die Errichtung und das Funktionieren des Binnenmarkts zum Gegenstand haben.

In **Art. 1 Abs. 1 KI-VO** wird der Zweck der Verordnung deshalb wie folgt beschrieben:

> „**Zweck dieser Verordnung** ist es, **das Funktionieren des Binnenmarkts** zu verbessern und die Einführung einer auf den Menschen ausgerichteten und **vertrauenswürdigen** künstlichen Intelligenz (KI) zu fördern und gleichzeitig ein **hohes Schutzniveau** in Bezug auf Gesundheit, Sicherheit und die in der Charta der Grundrechte verankerten Grundrechte, einschließlich Demokratie, Rechtsstaatlichkeit und Umweltschutz, vor schädlichen Auswirkungen von Systemen der künstlichen Intelligenz (KI-Systeme) in der Union zu gewährleisten und die **Innovation** zu unterstützen.

Diese Ziele sollen dadurch erreicht werden, dass einheitliche Regeln für die gesamte EU gelten (Maximalharmonisierung), gewisse Praktiken im Bereich KI verboten werden[1], besondere Anforderungen an sog. Hochrisiko-KI-Systeme gestellt werden[2], für bestimmte KI-Systeme einheitliche Transparenzvorschriften gelten[3], KI-Modelle mit allgemeinem Verwendungszweck besonderen Regeln unterworfen werden[4], eine effektive Durchsetzung

[1] Hierzu mehr in Abschnitt 4 und 5.
[2] Hierzu mehr in Abschnitt 6.
[3] Hierzu mehr in Abschnitt 7.
[4] Hierzu mehr in Abschnitt 7

der Verordnung gewährleistet wird und Maßnahmen zur Innovationsförderung mit besonderem Augenmerk auf KMU und Startups getroffen werden.[5]

Bei der Auslegung der einzelnen Regelungen der KI-VO ist immer der Zweck der KI-VO zugrunde zu legen und im Zweifel im Sinne der gesetzgeberischen Ziele zu entscheiden.

Der Gesetzgebungsprozess

Im Frühjahr 2021 legte die Europäische Kommission im Rahmen der Daten- und Cybersicherheitsstrategie als erster Gesetzgeber einen Entwurf des AI-Acts vor und stieß damit das Gesetzgebungsverfahren an.

Im Anschluss fand eine Überarbeitung durch den Rat der EU und das Europäische Parlament statt, aus der eine Veröffentlichung des Positionspapiers des Rates der EU bereits im Dezember 2022 hervorging, eine Abstimmung im Plenum des EU-Parlamentes zur Positionierung jedoch erst im Juni 2023 stattfinden konnte.

Neben den federführenden Ausschüssen des EU-Parlamentes für Binnenmarkt und Verbraucherschutz (IMCO) sowie für bürgerliche Freiheiten, Justiz und Inneres (LIBE) wurden auch Stellungnahmen von assoziierten Ausschüssen, EU-Institutionen wie EZB und beratender Gremien einbezogen.

Die Verzögerung resultierte einerseits aus mehreren Tausend Änderungsanträgen, aber auch aus dem Durchbruch der generativen AI ChatGPT und der daraus entsprungenen Debatte über den Umgang des AI Acts mit solchen Technologien, die in der ursprünglichen Version vom 21. April 2021 nicht berücksichtigt worden waren.

Nach Vorliegen des initialen Entwurfs der EU-Kommission, der Position des Rates der EU und des finalen Entwurfs des EU-Parlamentes starteten die Trilog-Verhandlungen der drei Organe. Ein 38-stündiger Marathon wurde am 08.12.2023 mit einer Einigung über die weltweit erste Regulierung von Künstlicher Intelligenz beendet.

Nach anschließendem formellem Einigungsverfahren von EU-Rat und EU-Parlament wurde der AI-Act am 21.05.2024 verabschiedet.

Am 12.07.2024 wurde er offiziell im EU-Amtsblatt veröffentlicht und trat 20 Tage später am 01.08.2024 in Kraft.

Eine Umsetzung wird schrittweise abhängig von der Risikoeinstufung mit folgenden Fristen erfolgen:

- Kurze Übergangsfrist von 6 Monaten (Februar 2025), in welcher allgemeine Vorschriften aus Kapitel I umgesetzt und verbotene Praktiken mit untragbarem Risiko aus Kapitel II eingestellt werden müssen (Art. 113 lit. a AI-Act).
- 12 Monate nach Inkrafttreten (August 2025) gelten Regelungen für AI-Systeme mit allgemeinem Verwendungszweck (GPAI), Art. 113 lit. b AI-Act.

[5] Hierzu mehr in Abschnitt 12.

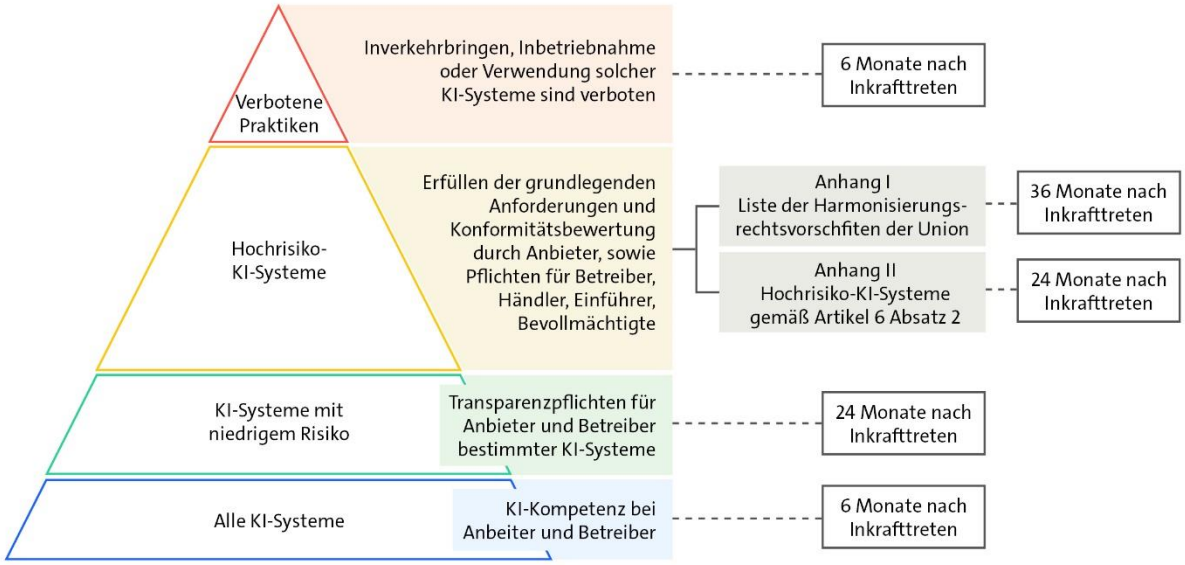
„Um ein verhältnismäßiges und wirksames verbindliches Regelwerk für KI-Systeme einzuführen, sollte **ein klar definierter risikobasierter Ansatz** verfolgt werden. Bei diesem Ansatz sollten **Art und Inhalt solcher Vorschriften auf die Intensität und den Umfang der Risiken zugeschnitte**n werden, die von KI-Systemen ausgehen können. Es ist daher notwendig, bestimmte **inakzeptable Praktiken** im Bereich der KI zu verbieten und **Anforderungen an Hochrisiko-KI-Systeme** und **Pflichten für die betreffenden Akteure** sowie **Transparenzpflichten für bestimmte KI-Systeme** festzulegen.“

- Es müssen zuständige Behörden der Mitgliedstaaten ernannt worden sein, Art. 70 AI-Act.
- 24 Monate nach Inkrafttreten (August 2026) beginnt die vollständige Geltung der Verordnung, Art. 111 AI-Act.
- Die Mitgliedstaaten müssen Vorschriften über Sanktionen, einschließlich Geldbußen, erlassen haben, Art. 57 AI-Act.
- 36 Monate nach Inkrafttreten (August 2027) greift die Umsetzungspflicht für die Harmonisierungsvorschriften aus Anhang I, wodurch weitere Hochrisiko-Anwendungsgruppen entstehen, Art. 113 lit. c AI-Act.
- Bis Ende 2030 treten Verpflichtungen für bestimmte KI-Systeme in Kraft, die Bestandteil der durch EU-Recht geschaffenen IT-Großsysteme in den Bereichen Freiheit, Sicherheit und Recht sind, wie z. B. das Schengener Informationssystem, Art. 111 AI-Act.

Der risikobasierte Ansatz

Die KI-VO folgt einem risikobasierten Regulierungsansatz, wonach KI-Systeme und -Modelle in unterschiedliche Risikoklassen mit divergierenden Rechtsfolgen kategorisiert werden. Aus diesem Ansatz heraus ergeben sich die wesentlichen materiellen Regelungskonzepte der KI-VO, die weitgehend nebeneinander bestehen.

Im **Erwägungsgrund 26 KI-VO** wird der risikobasierte Ansatz der KI-VO beschrieben:



Dieser risikobasierte Ansatz ist für KI-Systeme in der abgebildeten Risikopyramide dargestellt. Die angegebenen Fristen „nach Inkrafttreten“ besagen, in welchem Zeitraum nach Inkrafttreten der KI-VO die entsprechenden Vorschriften anwendbar werden.

Verbotene Praktiken im KI-Bereich

Gemäß Erwägungsgrund 28 KI-VO stehen insbesondere manipulative, ausbeuterische und soziale Kontrollpraktiken im Widerspruch zu den Werten der Union, weshalb solche besonders schädlichen und missbräuchlichen Verwendungsmöglichkeiten von KI als verboten gelten. Die KI-VO verbietet in diesem Zusammenhang in ihrem Art. 5 ausdrücklich bestimmte, mit einem unannehmbaren Risiko verbundene Praktiken im Zusammenhang mit KI-Systemen.

Hochrisiko-KI-Systeme

Die KI-Verordnung identifiziert zudem eine Reihe von Anwendungsfällen, in denen KI-Systeme als hochriskant eingestuft werden, da sie potenziell negative Auswirkungen auf die Gesundheit, die Sicherheit und die Grundrechte haben können. Diese Risikoklassifizierung basiert auf der Zweckbestimmung eines KI-Systems bzw. auf vernünftigerweise vorhersehbaren Fehlanwendungen. Die Funktion, die das KI-System ausführt, sowie der spezifische Zweck und die Modalitäten, für die das KI-System verwendet wird, sind entscheidend, um festzustellen, ob ein KI-System als hochriskant einzustufen ist oder nicht. Um die von ihnen ausgehenden Risiken zu verringern und einen vertrauenswürdigen KI-Einsatz zu ermöglichen, sollten derartige Hochrisiko-KI-Systeme gemäß Erwägungsgrund 46 KI-VO „nur dann auf dem Unionsmarkt in Verkehr gebracht, in Betrieb genommen oder verwendet werden, wenn sie bestimmte verbindliche Anforderungen erfüllen.“ Diese sogenannten grundlegenden Anforderungen an Hochrisiko-KI-Systeme sind vom Anbieter des KI-Systems zu erfüllen. Gemeint sind hiermit die Anforderungen in Kapitel III, Abschnitt 2, welche allerdings nur im Anhang VI KI-VO als „grundlegende Anforderungen“ bezeichnet werden.

Kapitel III legt auch die Anforderungen an andere Akteure neben dem Anbieter fest, nämlich an Betreiber, Händler, Einführer sowie Bevollmächtigte. Hier finden sich u. a. Regelungen,

unter welchen Bedingungen einer dieser Akteure selbst zum Anbieter eines Hochrisiko-KI-Systems wird.

Hervorzuheben ist bei den Hochrisiko-KI-Systemen, dass die KI-VO zwei Arten von Hochrisiko-KI-Systemen unterscheidet und für jede dieser zwei Arten zum Teil unterschiedliche Rechtsfolgen gelten – je nachdem ob ein KI-System im Zusammenhang mit Anhang I (Liste der Harmonisierungsrechtsvorschriften der Union) als Hochrisiko-KI-System eingestuft wird oder weil das KI-System einem der in Anhang III (Hochrisiko-KI-Systeme gemäß Artikel 6 Absatz 2) aufgeführten Fälle entspricht.

Solche KI-Systeme, die weder in den Bereich der verbotenen Praktiken fallen noch als Hochrisiko-KI-Systeme einzustufen sind, können als KI-Systeme mit niedrigem Risiko bezeichnet werden. Die KI-VO verwendet den Begriff „KI-Systeme mit niedrigem Risiko“ allerdings nicht, sondern es wird in Erwägungsgrund 165 von der „Entwicklung anderer KI-Systeme als Hochrisiko-KI-Systeme“ gesprochen.

Transparenzrisiken

Unabhängig von der Einstufung als Hochrisiko-KI-System können für bestimmte KI-Systeme darüber hinaus Transparenzpflichten gelten, die sich an Anbieter und aber teilweise auch an Betreiber richten.

Bestimmte KI-Systeme, die dazu bestimmt sind, mit natürlichen Personen zu interagieren oder Inhalte zu generieren, können spezifische Risiken der Nachahmung oder Täuschung darstellen, unabhängig davon, ob sie als Hochrisiko-KI-Systeme eingestuft werden oder nicht. Art. 50 KI-VO sieht für Anbieter und Betreiber bestimmter KI-Systeme Transparenzpflichten i.S.e. Kennzeichnung des „Ob“ des KI-Systems vor. Art. 50 KI-VO beinhaltet KI-Systeme zur direkten Interaktion mit Menschen (Abs. 1, z. B. Chatbots), zur Erzeugung synthetischer Inhalte (Abs. 2), biometrische Kategorisierungssysteme und Emotionserkennungssysteme (Abs. 3) und Deepfakes erzeugende KI-Systeme (Abs. 4).

Die vorstehende Unterteilung sinngemäß in „begrenzte“, „hohe“ und „inakzeptable“ Risiken darf indes nicht den Eindruck eines Rangverhältnisses von Risikoklassen mit alternativer Zuordnung suggerieren, da die Regelungskonzepte nebeneinander bestehen und voneinander unabhängig sind. Die Verbote und die Transparenzpflichten (vgl. Art. 50 Abs. 6 KI-VO) sind also unabhängig von der Einstufung als Hochrisiko-KI-System. Ein Chatbot kann beispielsweise sowohl den Transparenzpflichten als auch den Pflichten für Hochrisiko-KI-Systeme unterliegen; unabhängig davon kann eine bestimmte Verwendung des Chatbots nach Art. 5 KI-VO verboten sein.

Minimales Risiko

Alle weiteren KI-Systeme stellen ein minimales Risiko dar und unterliegen – mit Ausnahme allgemeiner Obligationen nach der KI-VO wie etwa der Förderung von KI-Kompetenz nach Art. 4 KI-VO – keinen besonderen Verpflichtungen über die derzeit geltenden Gesetze (z. B. DS-GVO) hinaus.

GPAI-Modelle als Basis der KI-Systeme

Über die KI-VO reguliert werden zudem GPAI-Modelle als Spitze der KI-Wertschöpfungskette. Sie sind wesentliche Komponenten von KI-Systemen, in die sie in der Regel integriert werden (vgl. ErwG 97). KI-Modelle mit allgemeinem Verwendungszweck sind in der Lage, ein breites Spektrum unterschiedlicher Aufgaben zu erfüllen (Art. 3 Nr. 63) und

können in eine Vielzahl nachgelagerter Systeme oder Anwendungen integriert werden. Auch von diesen Modellen gehen Gefahren aus, selbst wenn GPAI-Modelle in dieser frühen Entwicklungsstufe noch gar keinen festgelegten Bestimmungszweck aufweisen.

Grundsätzlich sind GPAI-Modelle von der erörterten Kategorisierung ausgenommen und unterliegen einer eigenen Risikoklassifizierung.

Hierbei werden KI-Modelle mit allgemeinem Verwendungszweck und mit systemischem Risiko unterschieden von den übrigen KI-Modellen mit allgemeinem Verwendungszweck. Die Differenzierung nach systemischem Risiko bestimmt die Pflichten für die jeweiligen Anbieter von GPAI-Modellen.

Der Produktregulierungsansatz

Das EU-KI-Gesetz gibt den Rahmen vor, in dem KI-Systeme als ein Produkt oder eine Dienstleistung angesehen werden, die gekauft oder verkauft werden können. Darin werden die verschiedenen Rollen entlang der Wertschöpfungskette und die entsprechenden Anforderungen an diese Rollen definiert. Ähnlich wie in der Maschinenverordnung (EU) 2023/1230 werden im EU-KI-Gesetz insbesondere die Systeme mit besonderem Risiko (verbotene und Hochrisikosysteme) klar umrissen, die zusätzliche Anforderungen und Konformitätsbewertungen erfüllen müssen, bevor sie auf den Markt gebracht werden können. Diese werden in Zukunft mit dem CE-Zeichen gekennzeichnet und in einer zentralen Datenbank registriert werden.

Der horizontale Ansatz

Gemäß dem in Kapitel 1, Artikel zwei definierten Anwendungsbereich ist das EU-KI-Gesetz eine horizontale Verordnung, d. h. es ist nicht auf einen bestimmten Sektor beschränkt. Obwohl es Details gibt, die die Beziehung zwischen dem EU-KI-Gesetz und bestimmten Sektoren spezifizieren, gilt der Anwendungsbereich für alle KI-Systeme und -Praktiken innerhalb der EU, unabhängig von der Branche.

3 Anwendbarkeit der KI-VO

Schritt 1.1: Handelt es sich um ein KI-System i.S.d. KI-VO?

Relevante(r) Artikel:

Art. 2 Abs. 1 lit. a, Art. 3 Nr. 1, 63, 66

Relevante(r) ErwG:

4, 12, 26, 46, 97, 99, 100

Konkretisierungsbedürftig:

Mittels Leitlinien der EU-Kommission

Jan Breuer (Detecon International GmbH), Valentino Halim (Oppenhoff & Partner Rechtsanwälte Steuerberater mbB), Dr. Rachel Hegemann (Deutsche Bahn), Ali-Reza Khalaji (R+V Versicherung AG), Dr. Anastasia Linnik (Retresco GmbH), Lea Ludmilla Ossmann-Magiera (LL.M. Leiden) (Bitkom), Alexander Schmalenberger (Taylor Wessing Partnerschaftsgesellschaft mbB)

Der (sachliche) Anwendungsbereich der KI-VO wird entscheidend vom Begriff des **KI-Systems** bestimmt. Die meisten Vorschriften der KI-VO finden nur auf KI-Systeme Anwendung, einige wenige gelten für **KI-Modelle mit allgemeinem Verwendungszweck** (General Purpose AI Models, „GPAI-Modelle"). Handelt es sich nicht um ein KI-System oder GPAI -Modell, unterliegt das fragliche (IT-)System bzw. (Software-)Anwendung nicht der KI-VO. Die Prüfung, ob das System bzw. die Anwendung den Anforderungen der KI-VO entspricht, kann dann an dieser Stelle abgebrochen werden.

Die gesetzliche Definition des KI-Systems

Die KI-VO führt erstmals eine Definition von KI ein. Dabei definiert das Gesetz nur den Begriff „KI-System" ausdrücklich. Nach **Art. 3 Nr. 1 KI-VO** ist ein KI-System:

> „ein **maschinengestütztes** System, das für einen in unterschiedlichem Grade **autonomen Betrieb** ausgelegt ist und das nach seiner Betriebsaufnahme **anpassungsfähig** sein **kann** und das aus den erhaltenen Eingaben für explizite oder implizite Ziele **ableitet,** wie Ausgaben wie etwa Vorhersagen, Inhalte, Empfehlungen oder Entscheidungen erstellt werden, die physische oder virtuelle **Umgebungen beeinflussen** können"

Die Merkmale der gesetzlichen Definition

Die **Erwägungsgründe (ErwG) 4 und 12** konkretisieren die gesetzliche Definition und einige ihrer Merkmale.

ErwG 4 kennzeichnet KI als eine **Reihe von Technologien**, die sich rasant entwickeln. Nach **ErwG** der Begriff „KI-System" klar definiert und eng mit internationalen Organisationen

abgestimmt werden. Dies zielt darauf ab, Rechtssicherheit, internationale Anschlussfähigkeit und hohe Akzeptanz sicherzustellen. Gleichzeitig soll die Definition des KI-Systems der schnellen technologischen Entwicklung in diesem Bereich Rechnung tragen und die wesentlichen Merkmale enthalten, um KI von einfachen, herkömmlichen Softwaresystemen und Programmierungsansätzen abgrenzen.

Weitere Erwägungsgründe adressieren die Risiken von KI-Systemen: ErwG 26 thematisieren die Einordnung von KI-Systemen in verschiedene Risikokategorien (risikobasierter Ansatz), ErwG 28 verbotene KI-Praktiken und ErwG 46 die Anforderungen an Hochrisiko-KI-Systeme.

Ein KI-System im Sinne der KI-VO ist nach der gesetzlichen Definition durch die folgenden Merkmale gekennzeichnet:

- KI-Systeme sind **maschinengestützt**.
- KI-Systeme sind auf einen in unterschiedlichem Grade **autonomen Betrieb** ausgelegt.
- Das KI-System kann nach Inbetriebnahme **Anpassungsfähigkeit** zeigen.
- Die Anwendung **leitet** aus den erhaltenen Eingaben – Informationen oder Daten– **ab**, wie Ausgaben erstellt werden.
- Die Ausgaben von KI-Systemen können physische oder virtuelle **Umgebungen beeinflussen**.

Zu den Merkmalen im Einzelnen:

Maschinengestützt sind alle Systeme, die „von Maschinen betrieben“ (ErwG 12 S. 6) werden, also auf (mindestens) einem Computer laufen. Das trifft auch auf jede herkömmliche Software oder sonstige IT-Anwendung zu. Damit handelt es sich nicht um ein unterscheidungskräftiges Merkmal, das zu einer klaren Kontur der Definition von KI-Systemen beiträgt.

Ähnliches gilt für die Anforderung, dass die Ausgaben von KI-Systemen physische oder virtuelle **Umgebungen beeinflussen** können. Jede Software beeinflusst zumindest virtuelle Umgebungen, ansonsten hätte sie keinen Sinn. Auch dieses Merkmal ist für die klare Unterscheidung von KI-Systemen und herkömmlicher Software nicht hilfreich.

Die Eigenschaft, nach Inbetriebnahme **anpassungsfähig** zu sein, bezieht sich auf die Lernfähigkeit eines KI-Systems während der Phase seiner Verwendung. Indes handelt es sich nicht um ein zwingendes Merkmal von KI-Systemen, wie sich aus dem klaren Wortlaut von Art. 3 Nr. 1 KI-VO ergibt („anpassungsfähig sein *kann*“). Auch andere Sprachfassungen der gesetzlichen Definition sind vergleichbar formuliert. Das bedeutet, dass auch solche Systeme als KI-Systeme gewertet werden können, die nach Inbetriebnahme nicht lernfähig sind. Ebenso umfasst die Definition KI-Systeme, deren zugrundeliegendes KI-Modell sich nach diesem Zeitpunkt laufend anpasst – etwa durch kontinuierliche Selbstoptimierung aufgrund der laufenden Auswertung der In- und Outputs.

KI-Systeme sind mit verschiedenen Graden an **Autonomie** ausgestattet. Hierfür muss das System bis zu einem gewissen Grad unabhängig von menschlichem Zutun agieren und ohne menschliche Eingriffe arbeiten können (ErwG 12 S. 10). Kein KI-System liegt dementsprechend vor, wenn es ausschließlich vorab von Menschen definierte Regeln automatisch ausführt (vgl. ErwG 12 S. 2). Unklar bleibt indes, was genau mit „einem gewissen Grad“ an Autonomie (ErwG 12 S. 10) gemeint ist. Nach dem Wortlaut der

Definition dürften auch Systeme mit minimalem – kaum messbaren – Autonomiegrad erfasst sein. Zudem lässt das *Explanatory Memorandum* zur 2023 überarbeiteten Definition der OECD eines KI-Systems, an die die Definition der KI-VO nahezu wörtlich angelehnt ist, darauf schließen, dass eine geringe Autonomiestufe genügen dürfte.[1] Auch das Autonomiemerkmal scheint kaum geeignet, den Begriff des KI-Systems entscheidend zu schärfen.

Als maßgebliches Merkmal von KI-Systemen i.S.d. KI-VO bleibt damit ihre Fähigkeit zur **Ableitung** (engl. *inference*) zu treffen. Diese bezieht sich auf den Prozess, wie aus den erhaltenen Eingaben die ausgegebenen Ergebnisse (z. B. Vorhersagen, Inhalte, Empfehlungen oder Entscheidungen) abzuleiten. Der EU-Gesetzgeber fasst darunter Techniken, die das Ableiten beim Aufbau eines KI-Systems ermöglichen, wie Konzepte für **maschinelles Lernen** sowie **logik- und wissensgestützte Konzepte**. Bei Ersteren wird aus Daten gelernt, wie bestimmte Ziele erreicht werden können, bei Letzteren aus kodierten Informationen oder symbolischen Darstellungen zu lösenden Aufgaben abgeleitet (ErwG 12 S. 5). Die Fähigkeit zur Ableitung geht über die einfache Datenverarbeitung hinaus und ermöglicht Lern-, Schlussfolgerungs- und Modellierungsprozesse (ErwG 12 S. 6). Auch insoweit deuten das *Explanatory Memorandum* zur OECD-Definition darauf hin, dass es genügt, wenn das System aus erhaltenen Eingaben (*input*) Ergebnisse (*output*) abzuleiten vermag.[2] Das dürfte auf die (weit) überwiegenden Zahl der als „KI“ vermarkteten Systeme und/oder Anwendungen zutreffen.

Trotz der Erläuterungen in den Erwägungsgründen bleiben einige Fragen bezüglich der Definition des KI-Systems offen. Nach der EU-Definition beschreiben KI-Systeme eher eine digitale Entscheidungsfindung oder die Automatisierung von Prozessen, die traditionell von Menschen durchgeführt werden. Eine klare Einordnung als KI-System auf der Grundlage „technologiebezogener“ Kriterien scheint derzeit noch nicht machbar.

Die EU-Kommission wird Leitlinien über die Definition bzw. Merkmale des KI-Systems veröffentlichen, um die einzelnen Merkmale zu konkreter zu bestimmen. Zum Zeitpunkt der Veröffentlichung dieses Leitfadens hat die EU-Kommission indes noch keinen Leitfaden veröffentlicht.

Beispielfälle

Die nachfolgende tabellarische Übersicht veranschaulicht anhand einiger Beispiele, welche Systeme, Anwendungen und/oder Technologien in der Praxis regelmäßig unter den Begriff des KI-Systems fallen und welche nicht. Zugleich zeigt die Übersicht Grenzfälle (*grey cases*) auf, in denen die Einordnung als KI-System zweifelhaft oder gänzlich unklar ist:

Beispiel	Kategorie	Bedingungen
Textverarbeitungsprogramme, Texteditoren, PDF-Viewer, OneNote, LaTeX, integrierte Entwicklungsumgebung (engl.	Keine KI	▪ Sammelt Text ▪ Macht keine Vorschläge ▪ Prüft die Rechtschreibung anhand eines festen Wörterbuchs
	Grenzfall	▪ Mögliche grammatikalische Fehler aufzeigen

Beispiel	Kategorie	Bedingungen
integrated development environment, IDE), DeepL Write	KI	▪ Gibt Text-/Codevorschläge auf der Grundlage dessen, was der Benutzer geschrieben hat, oder einer Datenbank.
Compiler/Programme, die Text in ausführbare Dateien/Anwendungen umwandeln	Keine KI	▪ Klassische Compiler ab 2024 – gcc, python, java. Keine bzw. deterministische Optimierung, deterministische Interpretation
	Grenzfall	▪ Optimierte Compiler auf der Grundlage vorheriger Daten oder angepasst an das Benutzersystem und/oder die Hardware.
Datenbanken und Storage (Excel/Hadoop/AWS buckets)	Keine KI	▪ Speichert und ruft Daten ab
	Grenzfall	▪ Filterung auf der Grundlage einfacher Funktionen wird an den Daten durchgeführt UND nicht nacheinander, nicht verschachtelt. Durchschnitt, Anzahl, Max (SQL-Basisfunktionalität)
	KI	▪ Komplexe Funktionen, die mehr als einen verschachtelten Durchschnitt, Addition, Maximum, Anzahl usw. enthalten. Funktionen auf die Daten mit einer Ausgabe (Hinweis: Dies sind im Wesentlichen die Bausteine eines neuronalen Netzes, d. h. eines gut etablierten Ansatzes für maschinelles Lernen)
Enterprise Resource Planning (ERP)-Systeme	Keine KI	▪ Einfache Verfolgung von Vermögenswerten, Ressourcen, Kosten usw.
	Grenzfall	▪ „Optimierungen“ werden von einer natürlichen Person konfiguriert.
	KI	▪ Vorschläge und Optimierung auf der Grundlage von Systemdaten/Benutzerdaten/ Eingabedaten/etc.
Funktionen	Grenzfall	▪ Numerische Näherungen/ Simulationen auf der Grundlage etablierter physikalischer Gleichungen mit (Problem, man kann immer noch einen sogenannten Face Swap bzw. Deep Fake auf reinen FEM/physikalisch basierten Simulationen erzeugen)

Beispiel	Kategorie	Bedingungen
Steuerungssysteme (mechanisch, elektrisch, hydraulisch usw.)	KI	▪ Komponenten, die explizit Methoden des maschinellen Lernens verwenden, wie neuronale Netze, Zufallswälder, Regression, Transformatoren, Parameterschätzung auf der Grundlage von Daten.
	Keine KI	▪ Maschinen, bei denen die auf einen Teil der Maschine wirkende Kraft physisch auf eine andere Form/Stelle der Maschine übertragen wird. Es gibt keine digitale Software.
	Grenzfall	▪ Maschine mit Software-integrierter Steuereinheit. Die Schwellenwerte für die Steuerung beruhen auf physikalischen Daten und Gleichungen. Die Annäherungen an die Gleichungen basieren auf klassischen numerischen Methoden (Runge-Kutta, Euler, Newton usw.). ▪ Physikalisch-basierte Regeltechnik/Steuerungssysteme – Zum Beispiel bestimmen Bremssysteme den Bremsdruck abgeleitet aus Grenzwerten physikalischer Gleichungen. ▪ Oder Wechselrichter-Umwandlung von Gleichstrom in Wechselstrom: Richtet sich nach den Schwellwerten. Die Schwellwerte werden nach physikalischen Formeln oder nach Erfahrungswerten ermittelt. Anhand der Messungen werden die Grenzwerte festgelegt. ▪ Viele Arten von Sensoren, wie Temperatur- oder Feuchtigkeitssensoren. Diese treffen Entscheidungen nach verschiedenen Zeitpunkten und Komponenten. Anzahl der Ereignisse. Einige auch wegen der Anzahl der Ereignisse. ▪ Note: Nach dem Stand der Technik handelt es sich dabei nicht um KI-Systeme, die jedoch aufgrund der Definition des EU-KI-Gesetzes technisch nicht ausgeschlossen werden können.
	KI	▪ Die Kontrollen beruhen auf maschinellem Lernen und Nutzungsdaten.
Suchmaschinen	KI	▪ Empfehlungen auf der Grundlage von Daten, z. B. Hyperlinks/Nutzung/ Relevanz/Sponsoren)
Dashboards/ Datenpräsentation	Grenzfall	▪ Die Filterung auf der Grundlage einfacher Funktionen wird an den Daten durchgeführt UND nicht nacheinander, nicht verschachtelt. Durchschnitt, Anzahl, Max (SQL-Basisfunktionalität)
	KI	▪ Verschachtelte Funktionen und/oder Sortierung von Daten auf der Grundlage von Funktionen des maschinellen Lernens.

Grenzfälle

Fortgeschrittene Algorithmen ohne Lernfähigkeit

- Komplexe Regel-basierte Systeme, die vordefinierte[6] Regeln verwenden, um Entscheidungen zu treffen. Diese können laut Richtlinie unter logik- und wissensgestützte Konzepte fallen, aber weisen keine Lernfähigkeit auf, welche als entscheidendes Merkmal von KI-Systemen hervorgehoben wird
- Hier gibt es viele Beispiele aus dem Bereich Manufacturing im Sinne von Prozessoptimierungsmodellen, die eigentlich ausschließlich nach menschlichen Regeln funktionieren, aber dennoch nicht ganz durchsichtig sind.
- Einfache Entscheidungsbaum- oder regelbasierte Automatisierungsskripte oder Software ohne Lernfähigkeit oder adaptives Verhalten.
- Beispiele:
  - **Regelbasierte Systeme** für Daten(vor)verarbeitung, Textgenerierung; (ältere) NLP-Systeme, z. B. für OCR, automatische Spracherkennung; regelbasierte Chatbots (sofern sie nicht über eine ML-basierte Komponente verfügen, die über ein ML-Modell mit einer Feedback-Schleife aus Interaktionen lernen kann) und Empfehlungssysteme.
  - No-Code-Anwendungen, z. B. RPA/Prozessautomatisierungstools, die auf einer vom Menschen definierten Logik basieren (z. B. Microsoft Power Automate, make.com).

Automatisierte Systeme mit begrenzter Adaptivität

- Roboter, die vordefinierte Bewegungsmuster ausführen und sich nur minimal an seine Umgebung anpassen, etwa durch einfache Sensoren, die Hindernisse erkennen und umgehen. Diese zeigen eine gewisse Autonomie und Reaktionsfähigkeit auf die Umgebung, jedoch fehlt auch hier die Fähigkeit zur Anpassung oder zum Lernen

Algorithmen zur Optimierung von Logistikrouten

- Basierend auf festen Parametern und historischen Daten führt der Code komplexe Berechnungen durch und kann Ergebnisse ableiten, aber er passt sich nicht an neue Daten an und lernt nicht kontinuierlich. Die Frage ist auch
- hier, ob diese Art der Optimierung die Anforderungen an die Lern- und Adaptionsfähigkeiten eines KI-Systems erfüllt.

Benutzerdefinierte Skripte und Makros

Makros in Tabellenkalkulationen wie Excel, die komplexe Strukturen aufweisen, aber ohne jegliche Form von maschinellem Lernen oder Adaptivität auskommen

Zwei damit zusammenhängende konzeptionelle Punkte/Fragen:

Determinismus: Man kann einige ML-Modelle technisch dazu zwingen, reproduzierbare Ergebnisse zu produzieren (über Seeds, Temperatur usw.). Sollten wir in diesem Fall zwischen von Menschen geschriebenen regelbasierten (deterministischen) und nicht-adaptiven ML-trainierten Algorithmen unterscheiden? Beide könnten technisch gesehen Verzerrungen enthalten und Entscheidungen fälschlicherweise beeinflussen.

Lernfähigkeit kann sich auf verschiedene Aspekte des Modelltrainings und -verhaltens beziehen. (1) Modelle werden auf von Menschen erzeugten Daten und/oder mit menschlichem Feedback trainiert. (2) Einmal trainiert, können sie von Menschen weiter trainiert/feinabgestimmt werden. (2) Modelle können ohne menschlichen Input weiter trainiert werden (z. B. im Fall von AGI). Die beiden letztgenannten Punkte sind ein Hinweis auf die „Anpassungsfähigkeit“ eines KI-Systems.

Ausnahmen

Zu beachten ist, dass Art. 2 bestimmte KI-Systeme aus dem Anwendungsbereich der KI-VO ausklammert. Zu diesen **Ausnahmen** gehören:

- Nationale Sicherheit, Verteidigung, Militär
- Forschung und Entwicklung
- Phase vor Inverkehrbringen
- Rein privater Gebrauch
- Open source (OS)*

*Für OS gelten eine Reihe von Rückausnahmen, etwa, wenn das OSKI-System in verbotenen Praktiken oder in Hochrisiko-Szenarien eingesetzt wird. Außerdem gelten für OS einige Transparenzpflichten.

Des Weiteren gilt die KI-VO nach Art. 111 KI-VO nicht für bestimmte KI-Systeme, die bereits auf dem Markt sind oder – in bestimmten Fällen – bis zum 2. August 2026 noch auf den Markt gebracht werden. Die Reichweite der Befreiung von der KI-Verordnung nach Art. 111 ist für Unternehmen, die (Hochrisiko-)KI-Systeme betreiben, von besonderer Bedeutung. Der Schwerpunkt liegt auf Absatz 2, der sich mit dem Bestandsschutz für solche Systeme befasst, die vor dem 2. August 2026 in Verkehr gebracht wurden. Diese Systeme unterliegen der Verordnung nur dann, wenn sie nach diesem Datum in ihrer Konzeption erheblich verändert werden (Art. 111 Abs. 2 KI-VO in Verbindung mit Art. 3 Nr. 23 KI-VO).

Eine wesentliche Änderung, die den Bestandsschutz beeinflusst, bezieht sich auf signifikante Anpassungen der Konzeption oder Zweckbestimmung des KI-Systems. Solche Änderungen könnten eine neue Konformitätsbewertung erforderlich machen, insbesondere wenn sie die Einhaltung der Verordnung beeinträchtigen könnten. Dazu zählen etwa Änderungen des Betriebssystems oder der Softwarearchitektur (Erwägungsgrund 128). Anpassungen, die durch das kontinuierliche Lernen des Systems erfolgen und bereits bei der ursprünglichen Konformitätsbewertung berücksichtigt wurden, stellen hingegen keine wesentliche Veränderung dar (Erwägungsgrund 128). Wie im allgemeinen Produktsicherheitsrecht könnte der Bestandsschutz auch dann enden, „… wenn i) [die] ursprüngliche Leistung, Verwendung oder Bauart geändert wurde, ohne dass dies bei der ursprünglichen Risikobewertung vorgesehen war, ii) sich die Art der Gefahr geändert oder das Risikoniveau im Vergleich zu den einschlägigen Harmonisierungsrechtsvorschriften der Union erhöht hat, iii) das Produkt zur Verfügung gestellt wird (oder in Betrieb genommen wird, wenn die Inbetriebnahme ebenfalls in den Anwendungsbereich der geltenden Rechtsvorschriften fällt).“ Damit besteht die Unsicherheit, ob der Bestandsschutz auch aus externen Gründen enden kann – also wegen einer geänderten Gefahrenlage.

Artikel 111 Absatz 2 impliziert, dass nicht nur Hochrisiko-KI-Systeme, sondern auch andere KI-Systeme, die keine signifikanten Risiken darstellen, von dieser Regelung profitieren könnten, sofern sie keine wesentlichen Änderungen erfahren. Dies folgt aus einem „erst

recht“ Schluss, da die Verordnung für Hochrisiko-Systeme strenger ist und ähnliche Prinzipien auf weniger riskante Systeme anwendbar sein sollten.

Es ist wichtig zu beachten, dass verbotene KI-Praktikanten gemäß Artikel 5 der KI-Verordnung ab dem 2. Februar 2025 nicht mehr betrieben werden dürfen, unabhängig davon, ob sie bereits in Verkehr gebracht wurden. Unternehmen müssen daher sicherstellen, dass ihre Systeme nicht unter die verbotenen Kategorien fallen, wie etwa solche, die für Massenüberwachung oder diskriminierende Praktiken genutzt werden könnten.

Zusammengefasst lassen sich folgende Betriebsausnahmen der KI-VO nennen:

3. Nationale Sicherheit und militärische Zwecke:
   Systeme, die ausschließlich für militärische und sicherheitsrelevante Zwecke eingesetzt werden, sind von der Verordnung ausgenommen. Dies betrifft also KI-Anwendungen, die von Streitkräften oder Nachrichtendiensten genutzt werden.

4. Rechtspflege und Strafverfolgung:
   Bestimmte KI-Systeme, die von Strafverfolgungsbehörden, der Justiz oder dem Zoll für die nationale Sicherheit oder die öffentliche Ordnung verwendet werden, können Ausnahmen genießen, sofern diese klar und eng begrenzt sind.

5. Forschung und Entwicklung:
   KI-Systeme, die ausschließlich für Forschungszwecke oder für den internen Testbetrieb verwendet werden, können von einigen regulatorischen Anforderungen befreit sein, solange diese nicht auf den Markt gebracht oder öffentlich zugänglich gemacht werden.

6. Ausnahmen für niedrigriskante Anwendungen:
   Systeme, die als „niedriges Risiko“ eingestuft werden (z.B. KI-Systeme zur Automatisierung von alltäglichen Prozessen wie Spamfilter oder Empfehlungssysteme), unterliegen weniger strengen Regelungen. Sie sind von umfassenden Vorschriften befreit, da sie weniger potenzielle Schäden verursachen.

7. Anwendung im Privatbereich:
   KI-Systeme, die von Privatpersonen ausschließlich für den persönlichen Gebrauch betrieben werden, sind ebenfalls von vielen Anforderungen der Verordnung ausgenommen. Diese Systeme dürfen allerdings keine Sicherheits- oder Datenschutzrisiken für andere Personen darstellen.

Diese Ausnahmen spiegeln den differenzierten Ansatz der KI-VO wider, der sicherstellt, dass nur KI-Systeme mit höherem Risiko strenger reguliert werden, während weniger risikobehaftete oder spezialisierte Anwendungen flexibler gehandhabt werden.

Zwischenergebnis

Wenn nach Prüfung der Voraussetzungen und ggf. unter Heranziehung der Liste mit Beispielen festgestellt werden kann, dass es sich um ein **KI-System i.S.d. KI-VO** handelt und keine Ausnahme greift, ist mit **Schritt 1.3** fortzufahren.

Sollten nicht klar, sein, ob es sich um ein **KI-System oder KI-Modell** handelt, ist mit **Schritt 1.2** fortzufahren.

Sollten die Voraussetzung von Art. 3 lit. 1 KI-VO nicht erfüllt oder eine Ausnahme aus Art. 2 einschlägig sein, ist die **Prüfung hier beendet.**

Schritt 1.2: Abgrenzungsfragen KI-System und General Purpose AI (GPAI) Modell

Marius Drabiniok (SKW Schwarz Rechtsanwälte), Ferdinand Schwarz (SKW Schwarz Rechtsanwälte)

Was versteht man unter General Purpose AI (GPAI) und welche weiteren Anforderungen müssen hierbei beachtet werden?

Relevante(r) Artikel:

3, 50, 51 ff.

Relevante(r) ErwG:

97, 99

Konkretisierungsbedürftig:

Systematische Abgrenzung zwischen KI-System und KI-Modell

Anforderungen an GPAI-Modelle

Definition des GPAI-Models in Art. 3 Nr. 63 der KI-V**O:**

„KI-Modell mit allgemeinem Verwendungszweck“ ein KI-Modell — einschließlich der Fälle, in denen ein solches KI-Modell mit einer großen Datenmenge unter umfassender Selbstüberwachung trainiert wird —, das eine erhebliche allgemeine Verwendbarkeit aufweist und in der Lage ist, unabhängig von der Art und Weise seines Inverkehrbringens ein breites Spektrum unterschiedlicher Aufgaben kompetent zu erfüllen, und das in eine Vielzahl nachgelagerter Systeme oder Anwendungen integriert werden kann, ausgenommen KI-Modelle, die vor ihrem Inverkehrbringen für Forschungs- und Entwicklungstätigkeiten oder die Konzipierung von Prototypen eingesetzt werden;

Definition des GPAI-Systems in Art. 3 Nr. 66 der KI-V**O:**

„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“.

Wann ein System (etwa eine Software) als ein KI-System anzusehen ist, wurde in Schritt 3.1 bereits umfassend dargestellt. Im vorliegenden Abschnitt soll demgegenüber die systematische Frage geklärt werden, was unter General Purpose AI (auch „GPAI“ genannt) zu verstehen ist und welche (weiteren) Pflichten in diesem Zusammenhang zu beachten sein können. Zu unterscheiden sind insoweit zunächst sog.

GPAI-Modelle und GPAI-Systeme. Eng hiermit verbunden ist die insoweit vorgelagerte Frage, was der Unterschied zwischen einem KI-System und einem KI-Modell ist.

Hintergrund der Fragestellung

Auch wenn GPAI nicht den Kern des vorliegenden Leitfadens darstellt, gibt es dennoch einige Konstellationen, in denen sich Abgrenzungsfragen zwischen KI-System und KI-Modell auftun. Da diese Fragen mitunter zu einigen Schwierigkeiten in der Praxis führen können, sollen nachstehend zumindest erste Anhaltspunkte zur Orientierung angeführt werden.

Die Begriffsbestimmung eines KI-Systems wurde bereits umfassend dargestellt. Allgemein kann man sich hierbei merken, dass die KI-Verordnung – von den Fällen eines lediglich minimalen Risikos abgesehen – sämtliche KI-Systeme (hierbei die Entwicklung sowie den anschließenden Betrieb) reguliert. Je nach Risikoklassifizierung sind in den jeweiligen Lebenszyklen des KI-Systems unterschiedliche Anforderungen zu beachten. Diese adressieren primär entweder den Anbieter oder den Betreiber des KI-Systems. (Ggf. hier Verweise auf weitere Kapitel einbauen)

Ein Unterfall eines KI-Systems stellt das sog. GPAI-System dar. Die Begriffe GPAI-System und GPAI-Modell werden hierbei häufig durcheinandergeworfen, obwohl beide Bezugsobjekte unterschiedlichen Regelungen folgen und sich auch aus technischer Sicht voneinander unterscheiden. Während es sich bei einem GPAI-System gerade um ein **KI-System** handelt, ist das GPAI-Modell als ein **KI-Modell** zu klassifizieren. An den Begriff „GPAI-System“ knüpft die KI-Verordnung allerdings keine unmittelbaren Rechtsfolgen, sodass die Klassifizierung als GPAI-System praktisch kaum eine Rolle spielen dürfte.

Auch wenn der Begriff des KI-Modells in der KI-Verordnung selbst nicht definiert wird, kann man sich die nachfolgenden, vereinfachten Grundsätze merken: Bei einem KI-System handelt es sich um die funktionsfähige und regelmäßig mit einer Benutzeroberfläche ausgestattete KI-Anwendung, während das KI-Modell das dahinterstehende (technische) Herzstück darstellt. Letzteres meint – am Beispiel eines Künstlichen Neuronalen Netzes (KNN) – bspw. die Architektur und Anzahl der Neuronen und Schichten sowie die dahinterstehenden Algorithmen und Gewichtungen. In Erwägungsgrund 97 der KI-Verordnung heißt es in Bezug auf GPAI-Modelle wörtlich:

„KI-Modelle mit allgemeinem Verwendungszweck können auf verschiedene Weise in Verkehr gebracht werden, unter anderem über Bibliotheken, Anwendungsprogrammierschnittstellen (API), durch direktes Herunterladen oder als physische Kopie. Diese Modelle können weiter geändert oder zu neuen Modellen verfeinert werden. Obwohl KI-Modelle wesentliche Komponenten von KI-Systemen sind, stellen sie für sich genommen keine KI-Systeme dar. Damit KI-Modelle zu KI-Systemen werden, ist die Hinzufügung weiterer Komponenten, zum Beispiel einer Nutzerschnittstelle, erforderlich. KI-Modelle sind in der Regel in KI-Systeme integriert und Teil davon.“

Ein System wird also gerade dann zum KI-System, wenn ein KI-Modell in eine Applikations- oder Systemumgebung integriert wird. Während das KI-System – bildlich gesprochen – die KI „zum Anfassen und Nutzen“ darstellt, ist das KI-Modell die dahinterstehende und für den Nutzer „unsichtbare“ Funktionsweise, welche ohne Integration in ein System nicht nutzbar ist. KI-Modelle sind damit zwar wesentliche Komponenten von KI-Systemen, denen aber ohne das Hinzufügen weiterer Komponenten die „Systemeigenschaft“ fehlt. Man kann sich an dieser Stelle bereits merken, dass die KI-Verordnung – neben KI-Systemen – solche (also nicht alle!) KI-Modelle reguliert, welche gerade als GPAI-Modelle anzusehen sind. Führt man sich diesen Umstand genauer vor Augen, treten einige Konstellationen auf, in denen gleichzeitig ein KI-System sowie das dahinterstehende KI-Modell in Verkehr gebracht und/oder genutzt werden, während in beiden Fällen unterschiedliche regulatorische Anforderungen zu beachten sein können.

Use Case zur Problemverdeutlichung

Zur Verdeutlichung der Abgrenzung kann man sich das nachfolgende – äußerst praxisrelevante – Szenario vorstellen:

Ein Unternehmen plant, einen Chatbot für Kunden zu entwickeln. Dieser Chatbot soll einfache Support-Anfragen beantworten können, das nötige Wissen zu den Produkten und Dienstleistungen des Unternehmens aufweisen und – bei schwierigen Fragen – den Kontakt zu einem Mitarbeiter des Kundensupports herstellen. Da das Unternehmen nicht über das nötige Knowhow zur Eigenentwicklung von KI-Komponenten verfügt, wird die Fremdlizenzierung verfügbarer Technologien in Erwägung gezogen.

Viele namhafte Anbieter bieten an, die von ihnen entwickelten KI-Modelle über Programmierschnittstellen (auch „*Application Programming Interface*“; kurz „*API*“) zu lizenzieren und in eigene Anwendungen einzubinden. Das Unternehmen entscheidet sich hierbei für ein generatives Sprachmodell mit breitem Anwendungsspektrum, welches die „Basis“ des Chatbots bilden soll. Damit der vorgesehene Chatbot am Ende des Tages auch über die relevanten Informationen zum Unternehmen verfügt, sollen weitere interne Datenbanken zum Einsatz kommen, die mit der KI „verknüpft“ werden. Hier spricht man von sog. „*Retrieval Augmented Generation*“ (kurz: „*RAG*“), oder „*Pre-Prompt Engineering*“.

Sofern das Vorhaben – wie geschildert – umgesetzt wird, stellt sich insbesondere die Frage, welche Rolle das Unternehmen (d. h. Anbieter und/oder Betreiber) in Bezug auf das jeweilige regulatorische Bezugsobjekt (KI-System und KI-Modell) einnimmt.

Unterscheidung KI-System und KI-Modell

In einem ersten Schritt muss zunächst die Frage beantwortet werden, wo hier die Unterschiede zwischen KI-System und KI-Modell liegen, und ob Letzteres gerade als ein von der KI-Verordnung reguliertes GPAI-Modell anzusehen ist.

KI-System

Unterstellt man in einem ersten Schritt, dass der Kunden-Chatbot sämtliche Kriterien aus Artikel 3 Nr. 1 der KI-Verordnung erfüllt, ist in der Chatbot-Applikation (ausgestattet mit einer funktionsfähigen Benutzeroberfläche) das hier relevante KI-System zu sehen. Läuft der Kunden-Chatbot nicht „Stand-alone“, sondern ist Teilkomponente eines größeren Kundensystems, können sich weitere Abgrenzungsfragen dahingehend stellen, ob die gesamte Software als KI-System zu qualifizieren ist oder abgrenzbare „Teilsysteme“ vorliegen. Daher ist es ebenso möglich, dass ein KI-Modell unterschiedliche (Teil-)Systeme „speist“ oder ein System auf voneinander abzugrenzenden KI-Modellen beruht. Für die ganz grundsätzliche Abgrenzung zwischen „System“ und „Modell“ spielen diese Fragen indes keine Rolle.

KI-Modell

Da ein KI-System nur mit einem integrierten KI-Modell arbeiten kann, ist die hinter dem Chatbot stehende Funktionsweise, welche hier über eine API von einem Drittanbieter bezogen wird, als das KI-Modell anzusehen. An dieser Bewertung würde sich auch dann nichts ändern, wenn das KI-Modell nicht über einen Drittanbieter in das System integriert wird, sondern fest in das System implementiert ist (etwa durch Rückgriff auf eine Open-Source-Lizenz).

GPAI-Modell

GPAI-Modelle weisen per Definition „eine erhebliche allgemeine Verwendbarkeit“ auf und können ein „breites Spektrum unterschiedlicher Aufgaben kompetent“ ausführen. Erforderlich ist auch, dass das Modell „in eine Vielzahl nachgelagerter Systeme oder Anwendungen integriert werden kann“. Wann diese Merkmale im Einzelfall erfüllt sind, wird in der KI-Verordnung nicht konkret beantwortet. In Erwägungsgrund 99 findet sich jedoch zumindest die nachfolgende Konkretisierung:

„Große generative KI-Modelle sind ein typisches Beispiel für ein KI-Modell mit allgemeinem Verwendungszweck, da sie eine flexible Erzeugung von Inhalten ermöglichen, etwa in Form von Text-, Audio-, Bild- oder Videoinhalten, die leicht ein breites Spektrum unterschiedlicher Aufgaben umfassen können.“

Dies bedeutet, dass jedenfalls die aktuell im Fokus stehenden großen Sprachmodelle (sog. Large Language Models – verkürzt auch LLMs) als GPAI-Modelle anzusehen sind. Daneben ist im jeweiligen Einzelfall zu prüfen, ob das KI-Modell, da es etwa dazu in der Lage ist, Texte, Bilder und/oder andere Inhalte zu erzeugen, eine erhebliche Allgemeinheit aufweist und insbesondere für verschiedene Zwecke eingesetzt werden kann.

Vorliegend ist davon auszugehen, dass es sich bei dem über eine API bezogenem KI-Modell auch gerade um ein GPAI-Modell handelt.

Rolle als Anbieter und/oder Betreiber

In einem nächsten Schritt muss sodann die Frage beantwortet werden, in welcher Rolle das betroffene Unternehmen in Bezug auf das KI-System sowie das dahinterstehende KI-Modell auftritt. Wer genau „Anbieter“ oder „Betreiber“ im Sinne der KI-VO ist, wird in Schritt 3.3 ausführlich geprüft; hier steht das jeweilige Bezugsobjekt im Mittelpunkt der Betrachtung, also ob die potenziell relevanten Handlungen des Unternehmens auf ein KI-Modell oder ein KI-System betreffen.

Hierfür soll in einem ersten Schritt zunächst die „einfachere“ Einordnung erfolgen: Bei dem Anbieter, welcher den Zugriff auf sein KI-Modell über eine API ermöglicht, handelt es sich um den Anbieter eines KI-Modells, und zwar in Bezug auf das konkret zur Verfügung gestellte KI-Modell. Warum dies gesondert zu betonen ist, soll gleich noch einmal aufgegriffen werden. Der Anbieter dieses KI-Modells ist jedoch nicht gleichzeitig auch Anbieter eines KI-Systems. Warum? Da dem KI-Modell – ohne weitere technische Integration – eine Benutzeroberfläche fehlt (vgl. hierzu Erwägungsgrund 97 der KI-Verordnung) und der Anbieter des KI-Modells zudem in keinerlei unmittelbarem Zusammenhang zum Kunden-Chatbot im beispielhaften Use Case steht. Stellt man also auf die maßgebliche Definition des Anbieters ab, wird zunächst nur das KI-Modell in Verkehr gebracht. Da es sich bei diesem KI-Modell (wie bereits dargestellt) auch gerade um ein GPAI-Modell handelt, muss der Anbieter alle hierfür geltenden regulatorischen Anforderungen entsprechend umsetzen (dazu näher unten bei Ziff. 5).

Nun die schwierigere Einordnung: In welcher Rolle tritt das Unternehmen auf, welches den Kunden-Chatbot bereitstellen möchte? Die Frage muss letztlich auf zwei verschiedenen Ebenen beantwortet werden:

KI-Modell

Da das Unternehmen ein „fremdes“ KI-Modell in eine eigene Anwendung integrieren möchte, stellt sich die Frage, ob es sich auch bei dem Unternehmen um den Anbieter eines KI-Modells handelt. Diese ganz grundsätzliche Frage ist äußerst umstritten und kann von Fall zu Fall unterschiedlich zu bewerten sein.

Wir haben in unserem beispielhaften Use Case bewusst ein einfacheres Beispiel gewählt, da man nach unserer Einschätzung zu dem Ergebnis kommen wird, dass das Unternehmen hier kein (eigenständiges) KI-Modell entwickelt und insoweit auch nicht als Anbieter anzusehen ist. Verfahren wie RAG oder Pre-Prompt Engineering leben gerade davon, dass externe Wissensquellen zur Entscheidungsfindung einer KI (nur) hinzugezogen werden, ohne dass an den eigentlichen Gewichtungen (also am Algorithmus der KI) selbst, Änderungen vorgenommen werden. Grade bei Integration von GPAI-Modell ist regelmäßig davon auszugehen, dass hier kein Wechsel der verantwortlichen Rollen stattfindet, da der Anbieter bzw. Betreiber des Systems schlicht auf ein „geschlossenes“ und nicht zweckgebundenes Modell zugreift.

Spricht man demgegenüber vom sog. „*Finetuning*“, also einem Verfahren, bei welchem das KI-Modell selbst mit unternehmenseigenen Trainingsdaten angelernt wird, kann man auch zu einer anderen rechtlichen Einschätzung gelangen.

Spätestens zu dem Zeitpunkt, an dem das zuvor in Verkehr gebrachte KI-Modell zu einem eigenständigen (neuen) KI-Modell (weiter-)entwickelt wird und dieses KI-Modell sodann in

Verkehr gebracht wird, wäre die Rolle des Anbieters eines KI-Modells denkbar. Als Richtschnur sollte man sich daran orientieren, ob das KI-Modell in seiner ursprünglichen Funktionsweise derart modifiziert wird, dass eine „wesentliche Veränderung“ im Sinne von Art. 3 Nr. 23 KI-Verordnung eintritt, also aus objektiven Gesichtspunkten eine Neubestimmung des Anbieters erforderlich ist, da die Konformität des KI-Systems mit den Anforderungen der KI-Verordnung neu bewertet werden muss.

Wann also die konkrete Schwelle zur (eigenständigen) Entwicklung – hier in Bezug auf das KI-Modell – überschritten wird, ist höchst umstritten und muss im jeweiligen Einzelfall bewertet werden.

KI-System

In Bezug auf das konkrete KI-System – hier also der Chatbot – wird man nach unserer Einschätzung demgegenüber zu dem Ergebnis kommen, dass das Unternehmen insoweit (!) als Anbieter anzusehen ist. Dieses Ergebnis lässt sich gut nachvollziehen, da nur das Unternehmen die Verantwortung für den konkreten Chatbot übernimmt. Betrachtet man bspw. Art. 50 Abs. 1 AI Act, heißt es dort wörtlich:

„Die Anbieter stellen sicher, dass KI-Systeme, die für die direkte Interaktion mit natürlichen Personen bestimmt sind, so konzipiert und entwickelt werden, dass die betreffenden natürlichen Personen informiert werden, dass sie mit einem KI-System interagieren, es sei denn, dies ist aus Sicht einer angemessen informierten, aufmerksamen und verständigen natürlichen Person aufgrund der Umstände und des Kontexts der Nutzung offensichtlich.“

Die vorgenannte Verpflichtung kann aus technischer Sicht auch nur durch den Anbieter umgesetzt werden, wobei hier ausdrücklich auf das KI-System Bezug genommen wird. Dies bedeutet, dass das Unternehmen als Anbieter eines KI-Systems anzusehen wäre, obwohl an dem eigentlichen KI-Modell selbst, keine wesentlichen Änderungen vorgenommen werden. Das genutzte KI-Modell wird jedoch in eine spezifische Anwendung integriert und somit auch einer konkret vorgesehenen Zweckbestimmung zugeführt. Diese Anwendung liegt ausschließlich im Verantwortungsbereich des Unternehmens, weshalb insoweit von der Rolle eines Anbieters auszugehen ist.

Die Frage, wer als Anbieter einer KI anzusehen ist, muss daher stets auf Basis des jeweiligen Bezugsobjekts (Modell oder System) bestimmt werden.

Besondere Pflichten bei GPAI-Systemen?

GPAI-Systeme sind – jedenfalls soweit es um die Risikoklassifizierung geht – als „gewöhnliche“ KI-Systeme anzusehen. In Erwägungsgrund 85 der KI-Verordnung heißt es hierzu wörtlich:

„KI-Systeme mit allgemeinem Verwendungszweck können als eigenständige Hochrisiko-KI-Systeme eingesetzt werden oder Komponenten anderer Hochrisiko-KI-Systeme sein.“

Während GPAI-Modelle einer gänzlich eigenständigen Regulierung unterliegen (hierzu sogleich), gelten für GPAI-Systeme grundsätzlich die gewöhnlichen „Spielregeln“ zur Risikoklassifizierung. Dies bedeutet, dass ein GPAI-System – in Abhängigkeit der jeweiligen Nutzung – etwa auch als Hochrisiko-KI-System eingestuft werden kann.

Daneben können spezifische Transparenzpflichten umzusetzen sein, welche wiederum für sämtliche einschlägigen KI-Systeme zu beachten sind. Diese werden in, Schritt 7.2.4 im Detail dargestellt. So heißt es in Art. 50 Abs. 2 der KI-Verordnung, welcher „*bestimmte KI-Systeme*“ adressiert, etwa wörtlich:

„Anbieter von KI-Systemen, einschließlich KI-Systemen mit allgemeinem Verwendungszweck, die synthetische Audio-, Bild-, Video- oder Textinhalte erzeugen, stellen sicher, dass die Ausgaben des KI-Systems in einem maschinenlesbaren Format gekennzeichnet und als künstlich erzeugt oder manipuliert erkennbar sind. […]“

Exkurs: Pflichten als Anbieter eines GPAI-Modells

Bei GPAI-Modellen existieren **ausschließlich** Pflichten für den Anbieter eines GPAI-Modells. Die Rolle des Betreibers existiert hier nicht, da ein KI-Modell (also auch ein GPAI-Modell) stets in ein System integriert werden muss, sodass es überhaupt nutzbar wird und „betrieben“ werden kann. Dies schließt natürlich nicht aus, dass der Anbieter eines KI-Modells gleichzeitig auch in „doppelter Rolle“ der Betreiber eines KI- bzw. GPAI-Systems sein kann (siehe oben unter 4.).

Welche Pflichten für Anbieter von GPAI-Modell gelten, muss danach differenziert werden, ob das GPAI-Modell ein systemisches Risiko aufweist, oder nicht. Es gelten also abweichende Risikostufen, als sie für KI-Systeme vorgesehen sind die Pflichten für Anbieter von GPAI-Modell werden in Schritt 7.2.4 im entsprechenden Exkurs dargestellt).

Der Hintergrund dieser Differenzierung im Gegensatz zu KI-Systemen lässt sich mit dem allgemeinen Verständnis zwischen KI-System und KI-Modell gut nachvollziehen. Da ein GPAI-Modell gerade kein KI-System darstellt, da ihm insbesondere eine Benutzeroberfläche fehlt, können die auf KI-Systeme gemünzten Merkmale bei der Risikoklassifizierung (welche häufig auf ganz spezielle Use Cases abstellen) nicht ohne weiteres herangezogen werden. Das GPAI-Modell kann gerade für eine Vielzahl von Zwecken verwendet werden und muss darüber hinaus zunächst in ein konkretes KI-System integriert werden.

In Art. 51 Abs. 1 AI Act wird sodann festgehalten, wann ein „systemisches Risiko“ anzunehmen ist. In der Norm heißt es wörtlich:

„Ein KI-Modell mit allgemeinem Verwendungszweck wird als KI-Modell mit allgemeinem Verwendungszweck mit systemischem Risiko eingestuft, wenn eine der folgenden Bedingungen erfüllt ist:

a) Es verfügt über Fähigkeiten mit hohem Wirkungsgrad, die mithilfe geeigneter technischer Instrumente und Methoden, einschließlich Indikatoren und Benchmarks, bewertet werden;

b) einem unter Berücksichtigung der in Anhang XIII festgelegten Kriterien von der Kommission von Amts wegen oder aufgrund einer qualifizierten Warnung des wissenschaftlichen Gremiums getroffenen Entscheidung zufolge verfügt es über Fähigkeiten oder eine Wirkung, die denen gemäß Buchstabe a entsprechen.“

Ein GPAI-Modell verfügt grundsätzlich dann über „Fähigkeiten mit hohem Wirkungsgrad“, sofern die kumulierte Menge der für sein Training verwendeten Rechnungen, gemessen in sog. Gleitkommaoperationen (auch „FLOPS“ genannt), mehr als 10^25 verfügt. Diese Frage

kann – wie bei der Nutzung von KI ganz generell – nur durch ein interdisziplinär besetztes Team mit der nötigen technischen Fachkunde geklärt werden.

Für Anbieter von GPAI-Modell **ohne systemisches Risiko** (eine Ausnahme gilt für „Open Source“ Modelle, für die Einschränkungen bestehen) gelten insbesondere die nachfolgenden Pflichten:

- Erstellen und Aktualisieren einer technischen Dokumentation des Modells, einschließlich seines Trainings- und Testverfahrens
- Erstellen, Aktualisieren und Bereitstellen von Informationen für Anbieter eines KI-Systems, die beabsichtigen, das KI-Modell mit allgemeinem Verwendungszweck in ihre KI-Systeme zu integrieren
- Entwicklung einer Strategie zur Einhaltung des EU-Urheberrechts insb. mit Blick auf die Beachtung von Nutzungsrechten beim Training von KI.

Für Anbieter von GPAI-Modell **mit systemischem Risiko** gelten zusätzlich zu den soeben genannten Vorgaben insbesondere die nachfolgenden (weiteren!) Pflichten:

- Durchführung einer Modellbewertung mit standardisierten Protokollen und Instrumenten, die dem Stand der Technik entsprechen
- Bewertung und Minderung von möglichen systemischen Risiken auf Unionsebene – einschließlich ihrer Ursachen –, die sich aus der Entwicklung, dem Inverkehrbringen oder der Verwendung ergeben
- Erfassung und Dokumentation von einschlägigen Informationen über schwerwiegende Vorfälle und mögliche Abhilfemaßnahmen
- Gewährleistung eines angemessenen Maßes an Cybersicherheit für die KI-Modelle mit allgemeinem Verwendungszweck mit systemischem Risiko und die Infrastruktur

Fazit

KI-Modelle (in der Form von GPAI-Modell) nehmen eine gewisse Sonderrolle in der KI-Verordnung ein und sind für viele Unternehmen nicht wirklich „greifbar“. Umso mehr müssen die jeweiligen Anforderungen sehr gründlich geprüft werden. Dies betrifft – neben der ganz grundsätzlichen Unterscheidung zwischen KI-System und KI-Modell – insbesondere die Frage, ob das Unternehmen im Einzelfall auch als Anbieter eines GPAI-Modell anzusehen sein kann.

Zwischenergebnis

Sollte die Prüfung ergeben, dass es sich um ein **KI-System** handelt, ist mit **Schritt 1.3** fortzufahren. Ergibt die Prüfung, dass es sich um ein **KI-Modell** handelt, ist mit **Schritt 5.2** fortzufahren.

Schritt 1.3: Bin ich Regulierungsadressat?

Susan Bischoff (Morrison & Foerster LLP), Christiane Stützle (Morrison & Foerster LLP)

Relevante(r) Artikel:

Art. 2 Abs. 1 lit. a, 16

Relevante(r) ErwG:

21, 22, 84

Konkretisierungsbedürftig:

Ja, siehe bei den einzelnen Adressaten

Als Nächstes ist zu prüfen, ob sich die KI-VO mit ihren rechtlichen Vorgaben an den Leser oder die Leserin richtet. Die Regelungen der KI-VO gelten nämlich nur für die in der Verordnung vorgesehenen Adressaten. Falls die Prüfung ergibt, dass Sie nicht in den Adressatenkreis fallen, kann die Prüfung schon an diesem Punkt abgebrochen werden.

Art. **2 Abs. 1 KI-VO** regelt, wer Adressat der Verordnung und ihren Regelungen somit unterworfen ist. Dies sind:

a) **Anbieter**, die in der Union KI-Systeme in Verkehr bringen oder in Betrieb nehmen oder KI-Modelle mit allgemeinem Verwendungszweck in Verkehr bringen, unabhängig davon, ob diese Anbieter in der Union oder in einem Drittland niedergelassen sind;

b) **Betreiber** von KI-Systemen, die ihren Sitz in der Union haben oder sich in der Union befinden;

c) **Anbieter** und Betreiber von KI-Systemen, die ihren Sitz in einem Drittland haben oder sich in einem Drittland befinden, wenn das vom System hervorgebrachte Ergebnis in der Union verwendet wird;

d) **Einführer** und **Händler** von KI-Systemen;

e) **Produkthersteller**, die KI-Systeme zusammen mit ihrem Produkt unter ihrem eigenen Namen oder ihrer Handelsmarke in Verkehr bringen oder in Betrieb nehmen;

f) **Bevollmächtigte von Anbietern**, die nicht in der Union niedergelassen sind;

g) **betroffene Personen**, die sich in der Union befinden.

Von diesen Akteuren sind **Anbieter** (unten 1.3.1), **Betreiber** (unten 1.3.2), **Einführer** (unten 1.3.3) und **Händler** (unten 1.3.4) eigenständige Regulierungsadressaten mit separaten Pflichten. **Produkthersteller** (unten 1.3.5) können unter bestimmten Voraussetzungen als Anbieter von Hochrisiko-KI-Systemen gelten und unterliegen damit ebenfalls den entsprechenden Anbieterpflichten, während **Bevollmächtigte von Anbietern** (unten 1.3.6) die ihnen konkret übertragenen Anbieteraufgaben wahrnehmen.

Trotz dieser klar abgegrenzten Adressatengruppen ist es ratsam, sich auch mit den Pflichten der Akteure zu befassen, deren Rolle man potenziell einmal einnehmen könnte. Denn unter der KI-VO können Akteure **mehr als eine Adressatenrolle** einnehmen und damit kumulativen Pflichtengruppen unterworfen sein oder auch **zu Anbietern „hochgestuft" werden**.

In **ErwG 83** wird das Verhältnis zwischen den verschiedenen Akteuren wie folgt erläutert:

> Angesichts des Wesens und der Komplexität der Wertschöpfungskette (…) ist es von wesentlicher Bedeutung, Rechtssicherheit zu gewährleisten und die Einhaltung dieser Verordnung zu erleichtern. Daher müssen die Rolle und die spezifischen Pflichten der relevanten Akteure entlang der Wertschöpfungskette, wie Einführer und Händler, die zur Entwicklung von KI-Systemen beitragen können, präzisiert werden. **In bestimmten Situationen könnten diese Akteure mehr als eine Rolle gleichzeitig wahrnehmen und sollten daher alle einschlägigen Pflichten, die mit diesen Rollen verbunden sind, kumulativ erfüllen**. So könnte ein Akteur beispielsweise gleichzeitig als Händler und als Einführer auftreten.

Von vornherein vom subjektiven Anwendungsbereich der KI-VO ausgenommen sind zum einen Behörden von Drittländern sowie internationale Organisationen, soweit sie KI-Systeme im Rahmen der Strafverfolgung und der justiziellen Zusammenarbeit mit der Union oder den Mitgliedstaaten verwenden und angemessene Garantien zum Schutz der Grundrechte und Grundfreiheiten bieten, Art. 2 Abs. 4 KI-VO. Zum anderen gilt die KI-VO auch nicht für Betreiber (unten 1.3.2), die natürliche Personen sind und KI-Systeme im Rahmen einer ausschließlich persönlichen und nicht beruflichen Tätigkeit verwenden, Art. 2 Abs. 10 KI-VO.

Schritt 1.3.1. Bin ich Anbieter?

Der **Anbieter** wird in **Art. 3 Nr. 3 KI-VO** definiert als eine natürliche oder juristische Person, Behörde, Einrichtung oder sonstige Stelle, die ein KI-System oder ein KI-Modell mit allgemeinem Verwendungszweck **entwickelt oder entwickeln lässt** und dieses **unter ihrem eigenen Namen** oder ihrer Handelsmarke in Verkehr bringt oder das KI-System unter ihrem eigenen Namen oder ihrer Handelsmarke in Betrieb nimmt, sei es entgeltlich oder unentgeltlich.

Zu beachten ist, dass es für die Qualifikation als „Anbieter" i.S.d. KI-VO nicht darauf ankommt, wo die Person, Behörde, Einrichtung oder Stelle ihren Sitz hat. Ein Anbieter muss nicht in der EU niedergelassen sein, so ausdrücklich Art. 2 Abs. 1 lit. a KI-VO. Vielmehr genügt es nach dem sog. **Marktortprinzip**, wenn das KI-System in der EU in den Verkehr gebracht oder in Betrieb genommen oder die vom KI-System hervorgebrachte Ausgabe (der Output) in der EU verwendet wird (Art. 2 Abs. 1 lit. a, lit. c KI-VO). Bei KI-Modellen mit allgemeinem Verwendungszweck gilt die KI-VO trotz Sitz des Anbieters in einem Drittland, wenn das Modell in der EU in den Verkehr gebracht wird (Art. 2 Abs. 1 lit. a KI-VO). Anbieter von Hochrisiko-KI-Systemen und KI-Modellen mit allgemeinem Verwendungszweck, die in Drittländern niedergelassen sind, müssen zusätzlich zu den sitzunabhängigen Anbieterpflichten einen in der Union niedergelassenen Bevollmächtigten benennen (Art. 22 Abs. 1, Art. 54 Abs. 1 KI-VO; zu diesen Bevollmächtigten unten 3.3.6.

Anbieter und damit Regulierungsadressat im Sinne der KI-VO ist daher, wer (*1*) eine taugliche Rolle bei der Entwicklung des KI-Systems oder des KI-Modells mit allgemeinem Verwendungszweck einnimmt und (2) eine tatbestandliche Nutzungshandlung in der Union in Bezug auf das System oder Modell vornimmt.

Relevante Artikel:

Art. 2 Abs. 1 lit. a, lit. c, Art. 3 Nr. 3, Art. 25

Relevante ErwG:

21, 22, 82, 84, 97

Konkretisierungsbedürftig:

- „Entwickeln" eines Systems/Modells: Auch durch (wesentliche) Modifizierung eines bestehenden KI-Systems/Modells mit der Folge, dass für den Modifizierenden Anbieterpflichten entstehen?

- „Entwickeln lassen" eines KI-Systems/Modells: Konkretisierung des erforderlichen Auftrags- und Spezifizierungs-charakters; Anwendbarkeit bei unternehmensinterner Bereitstellung?

(1) Beteiligung an der Entwicklung

Anbieter ist nur, wer das KI-System oder das KI-Modell mit allgemeinem Verwendungszweck entweder selbst entwickelt oder entwickeln lässt. Beide Varianten sind in der KI-VO nicht weiter spezifiziert und bedürfen der Konkretisierung.

Variante 1 – Entwickeln: Innerhalb der ersten Variante des Entwickelns ist gegenwärtig offen, ob die Modifizierung eines bestehenden Systems oder das Weitertrainieren eines bestehenden Modells ein tatbestandliches Entwickeln gewissermaßen eines „neuen" Systems oder Modells darstellen und eine Anbieterqualifikation auslösen kann. Hierdurch könnten insbesondere Betreiber zum Anbieter „aufsteigen" und unter den Pflichtenkatalog für Anbieter fallen. Sollte dies vom Gesetzgeber so gewollt sein, stellt sich ferner die Frage, *wann* eine solche Schwelle überschritten wäre, d. h. welche Intensität eine Modifizierung oder ein weiteres Trainieren haben müsste. Ausdrücklich geregelt ist ein solches Anbieter-Upgrade durch nachträgliche Änderungen nur für Hochrisiko-KI-Systeme (Art. 25 Abs. 1 lit. b KI-VO, s. unten). Da aber selbst im Hochrisikobereich eine Hochstufung zum Anbieter nur bei *wesentlichen* Veränderungen vorgesehen ist, dürfte für KI-Systeme und -Modelle unterhalb dieser Risikoschwelle erst recht nicht gelten, dass *unwesentliche* Änderungen und Modifizierungen zu einem solchen Wechsel der Akteurseigenschaft und einer Qualifizierung als Anbieter führen. Im Hinblick auf Feinabstimmung eines Modells (sog. Finetuning) durch die Zugabe weiterer Trainingsdaten ist aber ErwG 97 S. 5 KI-VO zu beachten. Dieser stellt fest, dass Modelle mit allgemeinem Verwendungszweck „zu neuen

Modellen verfeinert werden“ können (in der englischsprachigen Fassung „fine-tuned into new models“), wonach das Finetuning wohl grundsätzlich als tauglicher Akt zur Entwicklung eines neuen Modells anzusehen ist – indes ist aber stets eine Einzelfallbetrachtung des konkreten Umfangs erforderlich.

➔ *Praxisrelevante Grenzfälle*: wesentliche Modifizierungen, Finetuning, Weitertrainieren eines bestehenden KI-Systems/Modells

Variante 2 – Entwickeln lassen: Auch die zweite Variante des „Entwickeln lassen“ bedarf aus Gründen der Rechtssicherheit weiterer Konkretisierung. So ist gegenwärtig unklar, wie genau der wohl zugrunde zu legende Auftragscharakter zu bestimmen ist, etwa ob das KI-System oder KI-Modell mit allgemeinem Verwendungszweck speziell oder jedenfalls mit einem hohen Individualisierungsgrad gerade für den konkreten „Besteller“ und den konkret anvisierten Einsatzbereich entwickelt werden muss. Offen ist auch, ob die Bereitstellung eines solchen Systems oder Modells innerhalb eines Gesamtunternehmens oder Mutterkonzerns ebenfalls ein „Entwickeln lassen“ durch die empfangene Unternehmenseinheit darstellt. Es ist denkbar, dass der Gesetzgeber hier nur externe Entwicklungsaufträge im Sinn hatte und solche internen Bereitstellungen daher nicht unter das Tatbestandsmerkmal des „Entwickeln lassen“ fallen. Allerdings ist der bloße Wortlaut der KI-VO nicht hinreichend eindeutig und eine Klarstellung auch diesbezüglich erforderlich.

➔ *Praxisrelevante Grenzfälle*: Bezug eines bestehenden KI-Systems/Modells, das so bzw. nur mit geringen Unterschieden auch anderen Abnehmern bereitgestellt wird; unternehmens-/konzerninterner Bezug eines KI-Systems/Modells.

(2) Relevante Nutzungshandlung in der Union

Anbieter i.S.d. KI-VO ist darüber hinaus nur, wer das so entwickelte KI-System oder KI-Modell mit allgemeinem Verwendungszweck unter eigenem Namen oder eigener Handelsmarke **in Verkehr bringt** oder – dies nur im Fall eines KI-Systems – unter eigenem Namen oder eigener Handelsmarke **in Betrieb nimmt**. Da die relevanten Nutzungshandlungen, die eine Anbieter-Stellung auslösen können, nach KI-Systemen und KI-Modellen differenzieren, muss genau bestimmt werden, worauf sich eine konkrete Nutzung bezieht – auf ein KI-System oder auf ein KI-Modell (zur Unterscheidung zwischen KI-System und KI-Modell mit einem Praxisbeispiel für die Zuordnung von Anbieter- und Betreibereigenschaft s. oben Schritt 3.2). Auf die (Un-)Entgeltlichkeit einer solchen Handlung kommt es in jedem Fall nicht an.

Variante 1 – Inverkehrbringen (KI-System/KI-Modell mit allgemeinem Verwendungszweck): Art. 3 Nr. 9 KI-VO definiert das Inverkehrbringen als die erstmalige Bereitstellung eines KI-Systems oder eines KI-Modells mit allgemeinem Verwendungszweck auf dem Unionsmarkt. Die Bereitstellung auf dem Unionsmarkt wiederum wird konkretisiert als die entgeltliche oder unentgeltliche Abgabe eines KI-Systems oder eines KI-Modells mit allgemeinem Verwendungszweck zum Vertrieb oder zur Verwendung auf dem Unionsmarkt im Rahmen einer Geschäftstätigkeit (Art. 3 Nr. 10 KI-VO). Ein Inverkehrbringen, das eine Anbieterqualifikation auslösen kann, bedarf somit immer einer *externen* Bereitstellung des Systems oder Modells auf dem Unionsmarkt.

In Bezug auf KI-Modelle mit allgemeinem Verwendungszweck stellt ErwG 97 S. 7 KI-VO klar, dass die Pflichten für Anbieter gelten, sobald ein Modell in Verkehr gebracht wird; ein eigenes Modell gilt indes auch dann als in Verkehr gebracht, wenn es in ein eigenes KI-

System integriert und dieses System auf dem Unionsmarkt bereitgestellt oder auch nur in Betrieb genommen wird (ErwG 97 S. 8 KI-VO). Die Pflichten für Modelle greifen demnach nicht, wenn ein eigenes Modell für rein interne Verfahren verwendet wird, die für die Bereitstellung eines Produkts oder einer Dienstleistung an Dritte nicht wesentlich sind und die Rechte natürlicher Personen nicht beeinträchtigt werden (ErwG 97 S. 9 KI-VO), es sei denn, es handelt sich um Modelle mit systemischem Risiko (ErwG 97 S. 10 KI-VO). Auch Tätigkeiten, die dem Inverkehrbringen des Modells vorausgehen, lösen die Anbieterpflichten noch nicht aus; ErwG 97 S. 12 KI-VO nennt hier Forschungs- und Entwicklungstätigkeiten und die Konzipierung von Prototypen.

Variante 2 – Inbetriebnahme (KI-System): Unter einer Inbetriebnahme ist die Bereitstellung eines KI-Systems in der Union zum Erstgebrauch direkt an den Betreiber oder auch zum Eigengebrauch entsprechend seiner Zweckbestimmung zu verstehen (Art. 3 Nr. 11 KI-VO). Anders als beim Inverkehrbringen kann eine die Anbieterqualifikation auslösende Inbetriebnahme somit nicht nur in einer externen Bereitstellung des KI-Systems auf dem Unionsmarkt (an den Betreiber), sondern grundsätzlich auch in einer rein *internen* Verwendung eines KI-Systems (zum Eigengebrauch) liegen. Allerdings stellt nicht jede interne Nutzung eines KI-Systems, das eine eigene oder beauftragte Entwicklung ist, eine solche tatbestandliche Inbetriebnahme durch Eigengebrauch und damit eine die Anbietereigenschaft auslösende Nutzungshandlung dar. Ein solcher wesentlicher Eigengebrauch liegt nur dann vor, wenn der interne Gebrauch mit der Zweckbestimmung des KI-Systems, also seiner intendierten Einsatzform, identisch ist. Denn Art. 3 Nr. 12 KI-VO definiert eine solche Zweckbestimmung als die Verwendung, für die ein KI-System laut Anbieter bestimmt ist, was sich auch aus den besonderen Umständen und Bedingungen für die Verwendung gemäß den vom Anbieter bereitgestellten Informationen in den Betriebsanleitungen, im Werbe- oder Verkaufsmaterial und in diesbezüglichen Erklärungen sowie in der technischen Dokumentation ergeben soll. Nach dieser Definition stellen etwa Forschungs-, Entwicklungs- und Testtätigkeiten keinen wesentlichen Eigengebrauch dar, der als tatbestandliche Inbetriebnahme des Systems und damit als eine die Anbieterqualifikation auslösende Nutzung zu verstehen wäre.

➔ *Praxisrelevante Grenzfälle*: rein interner Eigengebrauch eines KI-Systems, das der Verwender selbst entwickelt hat oder hat entwickeln lassen – relevante Inbetriebnahme, wenn Gebrauch der (externen) Zweckbestimmung des Systems entspricht; Abgrenzung zu internen Entwicklungsgebrauch

Wichtig zu beachten: „Upgrade“ zum Anbieter

Andere Akteure, die nicht originär Anbieter sind, können durch bestimmte Handlungen zu Anbietern i.S.d. Art. 2 Abs. 1 lit. a KI-VO werden bzw. als solche gelten.

Für **Hochrisiko-KI-Systeme** regelt **Art. 25 KI-VO**, wann Händler, Einführer, Betreiber oder sonstige Dritte als Anbieter eines bereits in Verkehr gebrachten oder in Betrieb genommenen Hochrisiko-KI-Systems gelten und damit den Anbieterpflichten aus Art. 16 KI-VO unterliegen. In all diesen drei Fällen tritt der Akteur **an die Stelle des ursprünglichen Anbieters**, der fortan nicht mehr als Anbieter gilt, Art. 25 Abs. 2 KI-VO:

- Art. 25 Abs. 1 lit. a KI-VO: Dies ist zum einen der Fall, wenn ein Akteur das Hochrisiko-KI-System mit seinem Namen oder seiner Handelsmarke versieht (eine anderweitige

vertragliche Pflichtenaufteilung kann sich über diese gesetzliche Neubestimmung des Anbieters nicht hinwegsetzen).

- Art. 25 Abs. 1 lit. b KI-VO: Ein Akteur gilt auch dann als Anbieter, wenn er eine **wesentliche Veränderung** eines bereits in Verkehr gebrachten oder in Betrieb genommenen Hochrisiko-KI-Systems so vornimmt, dass es weiterhin ein Hochrisiko-KI-System bleibt. Eine wesentliche Veränderung ist nach Art. 3 Nr. 23 KI-VO eine solche Änderung des KI-Systems nach dessen Inverkehrbringen oder Inbetriebnahme, die in der vom Anbieter durchgeführten Konformitätsbewertung nicht vorgesehen oder geplant war und durch die Konformität des KI-Systems beeinträchtigt wird oder die zu einer Änderung der Zweckbestimmung führt, für die das KI-System bewertet wurde.
- Art. 25 Abs. 1 lit. c KI-VO: Schließlich gilt ein Akteur als Anbieter, wenn er die Zweckbestimmung eines bereits in Verkehr gebrachten oder in Betrieb genommenen KI-Systems so verändert, dass dieses zu einem Hochrisiko-KI-System wird.

Dass ein Akteur etwa durch wesentliche Veränderungen an einem bestehenden KI-System zu dessen Anbieter „aufrückt", regelt die KI-VO ausdrücklich nur in diesem Hochrisiko-KI-Bereich. **Unterhalb der Hochrisikoschwelle** ist dies nicht ausdrücklich normiert. Allerdings ist dort denkbar, dass Modifizierungen eines bestehenden KI-Systems oder -Modells etwa durch dessen Betreiber oberhalb einer (noch zu bestimmenden) Erheblichkeitsschwelle ein „Entwickeln" darstellen kann 8s. oben) und sich daraus – sofern auch ein tatbestandliches Inverkehrbringen oder Inbetriebnehmen vorliegt – gleichwohl eine Anbietereigenschaft nach der allgemeinen Anbieterdefinition ergibt.

Praxisbeispiel Grenzfall: Qualifikation als Anbieter durch Finetuning eines KI-Modells und Integration in KI-Systeme?

Ein Unternehmen, das Pressespiegel für Pharma-Unternehmen anbietet, möchte die Klassifizierung und Übersetzung von Fachtexten durch den Einsatz von KI zeiteffizienter gestalten. Da die entsprechende Expertise im Unternehmen vorhanden ist, wird die Entscheidung getroffen, ein vortrainiertes KI-Modell mit grundlegenden Fähigkeiten im Bereich der Texterkennung und Übersetzung zu lizenzieren und dieses für die spezifischen Belange des Unternehmens hinsichtlich der Besonderheiten der Fachsprache selbst zu verfeinern (was durch die Lizenz abgedeckt ist). Das vortrainierte KI-Modell, bei dem es sich um ein KI-Modell mit allgemeinem Verwendungszweck i.S.d. KI-VO handelt, bezieht das Unternehmen von einem Drittanbieter, der mithin Anbieter i.S.d. KI-VO im Hinblick auf dieses Modell ist.

Entwickeln/Entwickeln lassen eines KI-Modells durch das Unternehmen?

Bei dem lizenzierten KI-Modell handelt es sich um ein vom Drittanbieter standardmäßig angebotenes Modell, das nicht gesondert für das Unternehmen entwickelt oder auch nur spezifiziert wird. Es liegt daher kein „Entwickeln lassen" des KI-Modells durch das Unternehmen vor. Das KI-Modell wird sodann vom Unternehmen mit eigenen Trainingsdaten, d. h. fachspezifischen Texten, weiter trainiert (sog. *Finetuning,* s. dazu auch im Schritt 3.2). An dieser Stelle stellt sich für das Unternehmen die Frage, ob es sich bei dieser Weiterentwicklung des von einem Dritten zur Verfügung gestellten KI-Modells um ein „Entwickeln" eines (gewissermaßen neuen/anderen) KI-Modells durch das Unternehmen handelt, mit der Folge, dass das Unternehmen je nach späterer Verwendung selbst Anbieter dieses (weiterentwickelten) KI-Modells sein kann. Soweit die Fähigkeiten des

weitertrainierten Modells denen seines vortrainierten Standes entsprechen und das Unternehmen diese Textklassifizierungs- und Übersetzungsfähigkeiten durch das Finetuning lediglich an die fachsprachlichen Besonderheiten der eigenen Daten anpasst, könnte argumentiert werden, dass hierin kein Entwickeln eines (neuen) KI-Modells durch das Unternehmen liegt – denn das KI-Modell in seiner Grundfunktion und Struktur ist durch den Drittanbieter entwickelt worden. Mit Blick auf ErwG 97 S. 4 KI-VO könnte das Argument aber zu kurz gegriffen sein, jedenfalls in seiner Grundsätzlichkeit des Ausschlusses von Finetuning. Soweit es sich aber um nicht nur unwesentliche Modifizierungen des vortrainierten Modells handelt, könnte nach dem Wortlaut der KI-VO ein solches Entwickeln in Betracht kommen – ob dies vom Gesetzgeber am Ende tatsächlich gewollt ist und welche Erheblichkeitsschwellen hier greifen (bzw. ob sich eine Wesentlichkeit etwa bereits aus der Quantität neuer Trainingsdaten ergeben kann), ist derzeit noch offen.

Inverkehrbringen oder Inbetriebnahme durch das Unternehmen?

Geht man im Ergebnis der Risikobewertung davon aus, dass es sich bei dem Finetuning durch das Unternehmen um ein „Entwickeln“ eines KI-Modells mit allgemeinem Verwendungszweck handeln könnte, dann hängt die Einstufung als Anbieter entscheidend davon ab, ob ein tatbestandliches Inverkehrbringen durch das Unternehmen vorliegt. Das Unternehmen integriert das weiter trainierte KI-Modell in spezifische KI-Systeme für die konkrete unternehmenseigene Anwendung zur Klassifizierung und zur Übersetzung von Texten (zur Entstehung eines KI-Systems aus einem KI-Modell s. oben Schritt 1.2). Das KI-Modell selbst wird vom Unternehmen nicht extern am Markt bereitgestellt, sodass insoweit kein Inverkehrbringen des (weiter trainierten) Modells als solches durch das Unternehmen vorliegt. Allerdings könnte das KI-Modell dennoch als in Verkehr gebracht gelten, weil es vom Unternehmen in ein eigenes KI-System integriert und dieses System durch bestimmungsgemäßen Eigengebrauch (Textklassifikation; Übersetzung) in Betrieb genommen wird (vgl. ErwG 97 S. 8 KI-VO; genaue Bedeutung und Wirkungsweise dieser Feststellung im nicht-bindenden ErwG aber noch klarzustellen). Aufgrund einer solchen Inbetriebnahme des KI-Modells besteht das Risiko, dass das Unternehmen Anbieter des (weiter trainierten) KI-Modells ist. Darüber hinaus kann die Inbetriebnahme, der vom Unternehmen selbst entwickelten KI-Systeme durch bestimmungsgemäßen Eigengebrauch auch eine Anbietereigenschaft in Bezug auf diese KI-Systeme begründen.

Übersicht Pflichtenkatalog Anbieter

Welche Pflichten den Anbieter treffen, richtet sich nach der Art des konkreten KI-Systems bzw. KI-Modells. Die Pflichten von Anbietern von Hochrisiko-KI-Systemen sind in den Artikeln 16–22, 49 Abs. 1, Abs. 2, 72, 73 KI-VO normiert (hierzu im Einzelnen in den Schritten 3 und 4). Transparenzpflichten für KI-Systeme, die für die direkte Interaktion mit natürlichen Personen bestimmt sind, ergeben sich aus Art. 50 Abs. 1 KI-VO und Anbieter von KI-Systemen, die synthetische Audio-, Bild-, Video- oder Textinhalte erzeugen, müssen die Transparenzanforderungen aus Art. 50 Abs. 2 KI-VO umsetzen (hierzu im Einzelnen im Schritt 5.2). Für Anbieter von KI-Modellen mit allgemeinem Verwendungszweck gilt der Pflichtenkatalog aus Art. 53 KI-VO nebst der bereits erwähnten Pflicht zur Benennung eines in der Union niedergelassenen Bevollmächtigten aus Art. 54 KI-VO und für Anbieter von KI-Modellen mit allgemeinem Verwendungszweck und systemischem Risiko kommen außerdem die Anbieterpflichten aus Art. 55 KI-VO hinzu (hierzu im Einzelnen im Schritt 5.2).

Zwischenergebnis

Kann nach Prüfung der Voraussetzungen festgestellt werden, dass die Anbietereigenschaft i.S.d. KI-VO vorliegt, ist mit **Schritt 1.4** fortzufahren. Allerdings sollten Leserinnen und Leser, die die Anbietereigenschaft erfüllen, bedenken, dass sie zusätzlich auch Betreiber oder anderer Regulierungsadressat i.S.d. KI-VO sein und den entsprechenden Pflichten unterfallen können.

Sind die Voraussetzungen von Art. 2 Abs. 1 lit. a, lit. c KI-VO nicht erfüllt, ist mit **Prüfungsschritt 1.3.2** fortzufahren.

1.3.2. Bin ich Betreiber?

Der **Betreiber** wird in **Art. 3 Nr. 4 KI-VO** als natürliche oder juristische Person, Behörde, Einrichtung oder sonstige Stelle definiert, die ein KI-System **in eigener Verantwortung verwendet**.

Eine relevante Ausnahme gilt für Verwendungen von KI-Systemen im Rahmen einer persönlichen und nicht beruflichen Tätigkeit. Diese fallen nicht unter den Begriff des Betreibers und lösen daher keine Betreiberpflichten aus, Art. 3 Nr. 4, ErwG 13 KI-VO. Für derartige Verwendungen durch natürliche Personen gilt die KI-VO per se nicht, Art. 2 Abs. 10 KI-VO.

Die KI-VO gilt auch für Betreiber, die ihren Sitz nicht in der Union haben oder sich in der Union befinden, sofern die von dem KI-System hervorgebrachte Ausgabe (der Output) in der Union verwendet wird, Art. 2 Abs. 1 lit. c KI-VO.

Wichtig zu beachten: „Upgrade" zum Anbieter

Betreiber i.S.d. KI-VO können durch bestimmte Handlungen auch zum Anbieter eines KI-Systems „aufsteigen" oder als Anbieter eines KI-Modells mit allgemeinem Verwendungszweck anzusehen sein (s. im Detail oben im Schritt 3.3.1). Für Hochrisiko-KI-Systeme sieht Art. 25 Abs. 1 KI-VO drei konkrete Fallgruppen vor, in denen der ursprüngliche Anbieter durch den nunmehr als Anbieter geltenden Betreiber ersetzt wird. Für weniger riskante KI-Systeme sowie für KI-Modelle ist gegenwärtig offen, ob Modifizierungen durch einen Betreiber ein tatbestandliches „Entwickeln" darstellen und den Betreiber damit unter die allgemeine Anbieterdefinition fallen lassen können (hierzu oben im Schritt 3.3.1).

Übersicht Pflichtenkatalog Betreiber

Welche **Pflichten** den Betreiber treffen, richtet sich wiederum nach der Art des konkreten KI-Systems bzw. KI-Modells. Den Betreiber eines **Hochrisiko-KI-Systems** treffen im Wesentlichen folgende Pflichten:

- Art. 26 Abs. 1 KI-VO: Sicherstellen, dass das System insbesondere entsprechend der Betriebsanleitung verwendet wird
- Art. 26 Abs. 2 KI-VO: Aufsicht durch natürliche Personen mit hinreichender Kompetenz, Ausbildung und Befugnis
- Art. 26 Abs. 4 KI-VO: Eingabedaten unter der Betreiberkontrolle müssen der Zweckbestimmung des Systems entsprechen und ausreichend repräsentativ sein
- Art. 26 Abs. 5 KI-VO: Überwachung des Betriebs, Information des Anbieters, Verhalten bei Risiken und Vorfällen
- Art. 26 Abs. 6 KI-VO: Aufbewahrung von Protokollen
- Art. 26 Abs. 7 KI-VO: Information von Arbeitnehmer(vertretern)
- Art. 26 Abs. 8 KI-VO: Registrierungspflichten bestimmter Betreiber
- Art. 26 Abs. 11 KI-VO: Information natürlicher Personen, die der Verwendung des Systems unterliegen

Relevante(r) Artikel:

Art. 2 Abs. 1 lit. b, Art. 2 Abs. 1 lit. c, Art. 3 Nr. 4

Relevante(r) ErwG:

13, 93

Konkretisierungsbedürftig:

Ob Modifizierungen an bestehenden KI-Systemen/Modellen durch den Betreiber ein „Entwickeln" eines neuen Systems/Modells darstellen können mit der Folge, dass der Betreiber zum Anbieter wird

- Art. 26 Abs. 12 KI-VO: Zusammenarbeit mit Behörden
- Art. 27 KI-VO: Durchführung einer Grundrechte-Folgeabschätzung
- Art. 49 KI-VO: Registrierung in der EU-Datenbank

Darüber hinaus treffen den Betreiber **Informationspflichten** gegenüber natürlichen Personen, die von Emotionserkennungssystemen oder Systemen zur biometrischen Kategorisierung betroffen sind (Art. 50 Abs. 3 KI-VO) sowie **Transparenzpflichten** für KI-Systeme, die **Deepfakes** erzeugen (Art. 50 Abs. 4 KI-VO im Einzelnen im Schritt 7.2).

Zwischenergebnis

Kann nach Prüfung der Voraussetzungen festgestellt werden, dass die Betreibereigenschaft i.S.d. KI-VO vorliegt, ist mit **Schritt 1.4** fortzufahren. Allerdings sollten Leserinnen und Leser, bei denen die Betreibereigenschaft vorliegt, bedenken, dass sie zusätzlich auch anderer Regulierungsadressat i.S.d. KI-VO (z. B. Einführer) sein und unter die entsprechenden Pflichtenkataloge fallen können.

Sind die Voraussetzungen von Art. 2 Abs. 1 lit. b KI-VO nicht erfüllt, ist mit **Prüfungsschritt 1.3.3** fortzufahren.

1.3.3. Bin ich Einführer?

Relevante(r) Artikel:

Art. 2 Abs. 1 lit. d, Art. 3 Nr. 6, Art. 23

Relevante(r) ErwG:

83

Konkretisierungsbedürftig:

Ansässigkeit und Niederlassung in der Union

Der Einführer wird in Art. 3 Nr. 6 KI-VO als eine **in der Union ansässige oder niedergelassene natürliche oder juristische Person** definiert, die ein KI-System **in Verkehr bringt**, das den Namen oder die Handelsmarke einer in einem Drittland niedergelassenen natürlichen oder juristischen Person trägt. Das Inverkehrbringen ist wiederum gem. Art. 3 Abs. 9, Abs. 10 KI-VO als die erstmalige Bereitstellung des KI-Systems auf dem Unionsmarkt zu verstehen, also die entgeltliche oder unentgeltliche Abgabe des Systems zum Vertrieb oder zur Verwendung auf dem Unionsmarkt im Rahmen einer Geschäftstätigkeit.

Was unter einer solchen Ansässigkeit oder Niederlassung in der Union zu verstehen ist, wird in der KI-VO nicht konkretisiert – eine Verwendung dieser Begriffe ohne weitere Definition ist in europäischen Richtlinien und Verordnungen allerdings nicht unüblich. Ein beispielhafter Blick in die E-Commerce-Richtlinie (2000/31/EG vom 8. Juni 2000) verrät immerhin, dass der europäische Gesetzgeber dort eine **niedergelassene Einheit** (dort des Diensteanbieters) gemäß der Rechtsprechung des EuGH als die tatsächliche Ausübung einer Wirtschaftstätigkeit mittels einer festen Einrichtung auf unbestimmte Zeit definiert, Art. 2 lit. c, ErwG 19 E-Commerce-Richtlinie. Die **Ansässigkeit** in der EU wird vom EuGH (im Kontext der Dienstleistungsfreiheit des Art. 56 AEUV) allgemein als ein auf (unbestimmte) Dauer ausgerichteter Hauptaufenthalt verstanden, auch über eine Zweigniederlassung oder Nebenstelle. Im Sinne der Rechtssicherheit wäre es wünschenswert, wenn diese Begriffe auch in der KI-VO in ihrer Bedeutung konkretisiert würden.

Wichtig zu beachten: „Upgrade“ zum Anbieter und Kumulation von Akteurspflichten

Zu beachten ist, dass ein Einführer durch bestimmte Handlungen zum Anbieter eines Hochrisiko-KI-Systems „hochgestuft“ werden kann mit der Folge, dass er nunmehr auch Anbieterpflichten zu erfüllen hat, Art. 25 KI-VO (s. im Detail in Schritt 1.3.1).

Ferner ist zu berücksichtigen, dass ein Einführer gleichzeitig auch Händler (so ausdrücklich ErwG 83 S. 4 KI-VO) oder anderer Akteur i.S.d. KI-VO sein kann. In einem solchen Fall sind die Anforderungen mehrerer Akteure kumulativ zu erfüllen.

Übersicht Pflichtenkatalog Einführer

Den Einführer treffen Pflichten unter der KI-VO ausschließlich in Bezug auf **Hochrisiko-KI-Systemen**. Der Katalog des **Art. 23 KI-VO** umfasst folgende Pflichten:

- Art. 23 Abs. 1 (lit. a-d) KI-VO: Prüfpflicht, ob der Anbieter seine wesentlichen Pflichten erfüllt hat (Durchführung des Konformitätsbewertungsverfahrens, technische Dokumentation, CE-Kennzeichnung des Systems und Beifügung wesentlicher Unterlagen, Benennung eines Bevollmächtigten)
- Art. 23 Abs. 2 S. 1 KI-VO: Herstellen von Konformität des Systems durch den Einführer, wenn dieser Grund zur Annahme hat, dass diese nicht vorliegt
- Art. 23 Abs. 2 S. 2 KI-VO: Information des Anbieters, der Bevollmächtigten und der Marktüberwachungsbehörden über Risiken für die Gesundheit, Sicherheit oder Grundrechte von Personen
- Art. 23 Abs. 3 KI-VO: Angabe des Handelsnamens bzw. der Handelsmarke und der Kontaktanschrift auf der Verpackung bzw. in der dem Hochrisiko-KI-System beigefügten Dokumentation

- Art. 23 Abs. 4 KI-VO: Sicherstellung, dass Lagerungs- und Transportbedingungen die Konformität des Hochrisiko-KI-Systems nicht beeinträchtigen
- Art. 23 Abs. 5 KI-VO: Bereithaltung der Bescheinigung der notifizierenden Stelle sowie ggf. der Betriebsanleitungen und Konformitätserklärungen für einen Zeitraum von 10 Jahren nach dem Inverkehrbringen oder der Inbetriebnahme des Hochrisiko-KI-Systems
- Art. 23 Abs. 6 KI-VO: Übermittlung aller Informationen und Dokumentationen in leicht verständlicher Sprache zum Nachweis der Konformität des Hochrisiko-KI-Systems sowie der technischen Dokumentation an die nationalen Behörden auf deren begründete Anfrage
- Art. 23 Abs. 7 KI-VO: Zusammenarbeit mit nationalen Behörden bei Maßnahmen, die diese in Bezug auf ein vom Einführer in Verkehr gebrachtes Hochrisiko-KI-System ergreifen

Zwischenergebnis

Kann nach Prüfung der Voraussetzungen festgestellt werden, dass die Einführereigenschaft i.S.d. KI-VO vorliegt und keine Ausnahme greift, ist mit **Schritt 1.4** fortzufahren. Allerdings sollten Leserinnen und Leser, bei denen die Einführereigenschaft vorliegt, bedenken, dass sie zusätzlich auch anderer Regulierungsadressat i.S.d. KI-VO (z. B. Händler) sein und unter die entsprechenden Pflichtenkataloge fallen können.

Sind die Voraussetzungen von Art. 2 Abs. 1 lit. d KI-VO nicht erfüllt, ist mit **Prüfungsschritt 1.3.4** fortzufahren.

1.3.4 Bin ich Händler?

Relevante(r) Artikel:

Art. 2 Abs. 1 lit. d, Art. 3 Nr. 7, Art. 24

Relevante(r) ErwG:

83

Konkretisierungsbedürftig:

/

Der Händler wird in **Art. 3 Nr. 7 KI-VO** als eine natürliche oder juristische Person in der Lieferkette definiert, die **ein KI-System auf dem Unionsmarkt bereitstellt, mit Ausnahme des Anbieters oder des Einführers**. Unter einer solchen Bereitstellung auf dem Markt ist wiederum gem. Art. 3 Nr. 10 KI-VO die entgeltliche oder unentgeltliche Abgabe des KI-Systems zum Vertrieb oder zur Verwendung auf dem Unionsmarkt im Rahmen einer Geschäftstätigkeit zu verstehen.

Im Gegensatz zum Einführer ist Händler i.S.d. KI-VO nicht nur, wer in der Union ansässig oder niedergelassen ist. Vielmehr kommt es allein auf die Händler*tätigkeit* in der Union in Form der Bereitstellung des KI-Systems auf dem Unionsmarkt an.

Wichtig zu beachten: „Upgrade" zum Anbieter und Kumulation von Akteurspflichten

Zu beachten ist, dass ein Händler durch bestimmte Handlungen zum Anbieter eines Hochrisiko-KI-Systems **„hochgestuft"** werden kann mit der Folge, dass er nunmehr auch Anbieterpflichten zu beachten hat, Art. 25 KI-VO (s. im Detail oben im Schritt 3.3.1).

Ferner ist zu berücksichtigen, dass ein Händler **gleichzeitig auch ein** anderer Akteur i.S.d. KI-VO sein kann (ErwG 83 KI-VO). In einem solchen Fall sind die Anforderungen mehrerer Akteure kumulativ zu erfüllen.

Übersicht Pflichtenkatalog Händler

Den Händler treffen Pflichten unter der KI-VO ausschließlich in Bezug auf **Hochrisiko-KI-Systemen**. Der Katalog des **Art. 24 KI-VO** umfasst folgende Pflichten:

- Art. 24 Abs. 1 KI-VO: Prüfung der CE-Kennzeichnung des Systems und der erforderlich beizufügenden Unterlagen sowie der Einhaltung wesentlicher Pflichten durch Anbieter und ggf. Einführer des Systems vor dessen Bereitstellung auf dem Markt
- Art. 24 Abs. 2 S. 1 KI-VO: Herstellen von Konformität des Systems durch den Händler, wenn er Grund zur Annahme hat, dass diese nicht vorliegt
- Art. 24 Abs. 2 S. 2 KI-VO: Information des Anbieters bzw. des Einführers über Risiken für die Gesundheit, Sicherheit oder Grundrechte von Personen
- Art. 24 Abs. 3 KI-VO: Sicherstellung, dass Lagerungs- und Transportbedingungen die Konformität des Hochrisiko-KI-Systems nicht beeinträchtigen
- Art. 24 Abs. 4 KI-VO: Nimmt der Händler an, dass das System nicht konform ist, dann hat er selbst Korrekturmaßnahmen zu ergreifen, um die Konformität des Hochrisiko-KI-Systems herzustellen, es zurückzunehmen oder zurückzurufen oder hat sicherzustellen, dass der Anbieter, Einführer oder ggf. jeder relevante Akteur diese Maßnahmen ergreift; zusätzliche Informationspflichten bei Risiken für Gesundheit, Sicherheit oder Grundrechte von Personen
- Art. 24 Abs. 5 KI-VO: Übermittlung aller Informationen und Dokumentationen an nationale Behörden auf deren begründete Anfrage in Bezug auf vom Händler nach Art. 24 Abs. 1–4 KI-VO ergriffenen Maßnahmen zum Nachweis der Konformität des Systems

- Art. 24 Abs. 6 KI-VO: Zusammenarbeit mit Behörden in Bezug auf Maßnahmen, die diese im Hinblick auf ein vom Händler auf dem Markt bereitgestelltes Hochrisiko-KI-System ergreifen

Zwischenergebnis

Kann nach Prüfung der Voraussetzungen festgestellt werden, dass die Händlereigenschaft i.S.d. KI-VO vorliegt, ist mit **Schritt 1.4** fortzufahren. Allerdings sollten Leserinnen und Leser, bei denen die Händlereigenschaft vorliegt, bedenken, dass sie zusätzlich auch anderer Regulierungsadressat i.S.d. KI-VO (z. B. Produkthersteller) sein und unter die entsprechenden Pflichtenkataloge fallen können.

Sind die Voraussetzungen von Art. 2 Abs. 1 lit. d KI-VO nicht erfüllt, ist mit **Prüfungsschritt 1.3.5** fortzufahren**.**

1.3.5 Bin ich Produkthersteller?

Relevante(r) Artikel:

Art. 2 Abs. 1 lit. e, Art. 3 Nr. 8, Art. 25 Abs. 3,

Relevante(r) ErwG:

87

Konkretisierungsbedürftig:

/

Der Produkthersteller wird von der KI-VO neben Anbietern, Betreibern, Bevollmächtigten, Einführern und Händlern als möglicher „Akteur“ genannt, Art. 3 Nr. 8 KI-VO. Gemäß Art. 2 Abs. 1 lit. e KI-VO ist Produkthersteller, wer KI-Systeme zusammen mit seinem Produkt unter seinem eigenen Namen oder seiner Handelsmarke in Verkehr bringt oder in Betrieb nimmt.

Die KI-VO sieht keine originär eigenen Pflichten für den Produkthersteller vor. Allerdings kann der Produkthersteller **zum Anbieter eines Hochrisiko-KI-Systems werden** – ebenso wie Händler, Einführer, Betreiber oder sonstige Dritte (zu Anbieter „Upgrades“ bei diesen Akteuren im Detail im Schritt 1.3.1). In zwei Fällen **gilt der Produkthersteller gem. Art. 25 Abs. 3 KI-VO als Anbieter eines Hochrisiko-KI-Systems** und unterliegt damit den Anbieterpflichten nach Art. 16 KI-VO:

- Art. 25 Abs. 3 lit. a KI-VO: Der Produkthersteller gilt als Anbieter, wenn er das Hochrisiko-KI-System zusammen mit dem Produkt unter seinem Namen oder seiner Handelsmarke in Verkehr bringt.
- Art. 25 Abs. 3 lit. b KI-VO: Der Produkthersteller gilt als Anbieter, wenn er das Hochrisiko-KI-System unter seinem Namen oder seiner Handelsmarke in Betrieb nimmt, nachdem das Produkt in Verkehr gebracht wurde.

Zwischenergebnis

Kann nach Prüfung der Voraussetzungen festgestellt werden, dass die Produktherstellereigenschaft i.S.d. KI-VO vorliegt, ist mit **Schritt 1.4** fortzufahren.

Sind die Voraussetzungen von Art. 2 Abs. 1 lit. e KI-VO nicht erfüllt, ist mit **Prüfungsschritt 1.3.6** fortzufahren.

1.3.6. Bin ich Bevollmächtigter des Anbieters?

Relevante(r) Artikel:

Art. 2 Abs. 1 lit. f, Art. 3 Nr. 5, Art. 22, Art. 54

Relevante(r) ErwG:

82

Konkretisierungsbedürftig:

/

Die KI-VO gilt außerdem für Bevollmächtigte von Anbietern, die nicht in der Union niedergelassen sind, **Art. 2 Abs. 1 lit. f KI-VO**. Als Bevollmächtigter definiert **Art. 3 Nr. 5 KI-VO** eine in der Union ansässige oder niedergelassene natürliche oder juristische Person, die **vom Anbieter** eines KI-Systems oder eines KI-Modells mit allgemeinem Verwendungszweck **schriftlich dazu bevollmächtigt wurde und sich damit einverstanden erklärt hat**, in seinem Namen die in dieser Verordnung festgelegten Pflichten zu erfüllen bzw. Verfahren durchzuführen.

Dieser Bevollmächtigte spielt eine zentrale Rolle bei der Gewährleistung der Konformität der von den betreffenden Anbietern, die nicht in der Union niedergelassen sind, in der Union in Verkehr gebrachten oder in Betrieb genommenen Hochrisiko-KI-Systeme und indem er als ihr in der Union niedergelassener Ansprechpartner dient.

ErwG 82 S. 3 KI-VO macht deutlich, welchen Zweck diese Bevollmächtigten im Rahmen der KI-VO erfüllen:

Anders als bei den übrigen Akteuren entsteht die Eigenschaft des Bevollmächtigten also nicht durch die Zuweisung einer Rechtsstellung durch die KI-VO, sondern durch einen Bevollmächtigungsakt des Anbieters und die Annahme durch den Bevollmächtigten. Eine Pflicht zur Benennung eines Bevollmächtigten findet sich für Anbieter von Hochrisiko-KI-Systemen in Art. 22 KI-VO und für Anbieter von KI-Modellen mit allgemeinem Verwendungszweck in Art. 54 KI-VO.

Dementsprechend bestehen für den Bevollmächtigten auch keine eigenständigen Pflichten unter der KI-VO. Vielmehr **nimmt der Bevollmächtigte die Aufgaben wahr, die in seinem vom Anbieter erhaltenen Auftrag festgelegt sind** und dient damit zugleich neben oder anstelle des Anbieters als Ansprechpartner für die zuständigen Stellen und Behörden. Für Hochrisiko-KI-Systeme wird dies in Art. 22 Abs. 3 KI-VO und für KI-Modelle mit allgemeinem Verwendungszweck in Art. 54 Abs. 3 KI-VO ausdrücklich festgehalten und jeweils ein Mindestkatalog an Ermächtigungen durch einen solchen Auftrag vorgegeben. Der Bevollmächtigte hat den Marktüberwachungsbehörden bzw. dem Büro für Künstliche Intelligenz auf Anfrage eine Kopie des Auftrags bereitzustellen.

Zwischenergebnis

Kann nach Prüfung der Voraussetzungen festgestellt werden, dass die Bevollmächtigteneigenschaft i.S.d. KI-VO vorliegt, hat der Akteur die ihm vom Anbieter übertragenen Aufgaben wahrzunehmen und auf Anfrage mit den Behörden zu kooperieren. Da sich aus der KI-VO aber keine weiteren speziellen Pflichten für den Bevollmächtigten ergeben, ist die Prüfung unter diesem Leitfaden hier abzubrechen. Allerdings sollten Leserinnen und Leser, die Bevollmächtigte sind, sicherstellen, dass sie nicht auch zugleich anderer Akteur i.S.d. KI-VO sind, woraus sich eigene Pflichten ergeben können.

Schritt 1.4: Ist der räumliche Anwendungsbereich eröffnet?

Relevante(r) Artikel:

Art. 2 Abs. 1

Relevante(r) ErwG:

83

Konkretisierungsbedürftig:

Verhältnis zu Art. 53

Arnd Böken (GvW Graf von Westphalen Rechtsanwälte Steuerberater Partnerschaft mbB)
Art. 2 Abs. 1 KI-VO regelt neben dem persönlichen Anwendungsbereich auch den **räumlichen Anwendungsbereich**, also **wo** die Verordnung Anwendung findet:

a) Anbieter, die in der Union KI-Systeme in Verkehr bringen oder in Betrieb nehmen oder KI-Modelle mit allgemeinem Verwendungszweck in Verkehr bringen, unabhängig davon, ob diese Anbieter in der Union oder in einem Drittland niedergelassen sind;

b) Betreiber von KI-Systemen, die ihren Sitz in der Union haben oder sich in der Union befinden;

c) Anbieter und Betreiber von KI-Systemen, die ihren Sitz in einem Drittland haben oder sich in einem Drittland befinden, wenn das vom System hervorgebrachte Ergebnis in der Union verwendet wird;

d) Einführer und Händler von KI-Systemen;

e) Produkthersteller, die KI-Systeme zusammen mit ihrem Produkt unter ihrem eigenen Namen oder ihrer Handelsmarke in Verkehr bringen oder in Betrieb nehmen;

f) Bevollmächtigte von Anbietern, die nicht in der Union niedergelassen sind;

g) betroffene Personen, die sich in der Union befinden.

Die **ErwG 21 und 22** konkretisieren die Vorgaben in Art. 2 Abs. 1 KI-VO.

Der Gesetzgeber hat einen sehr weiten Anwendungsbereich bestimmt, um gleiche Wettbewerbsbedingungen für Unternehmen und einen wirksamen Schutz der Rechte und Freiheiten von Einzelpersonen zu gewährleisten. Im Ergebnis soll die KI-VO sämtliche Fälle erfassen, in denen das KI-System oder KI-Modell Bedeutung für Betroffene in der EU haben kann.

Gem. **ErwG 21** kommt es für die Anwendbarkeit der KI-VO nicht auf den Sitz des **Anbieters** an, sondern vor allem darauf, wo das durch das KI-System erbrachte Ergebnis **verwendet** wird. Statt des Territorialitätsprinzips, das an die Niederlassung des Regelungsadressaten anknüpft, gilt also für Anbieter das Marktortprinzip, bei dem es darauf ankommt, wo ein Produkt in Verkehr gebracht bzw. wo eine Dienstleistung angeboten wird.

Bei **Betreibern** kommt es dagegen auf den **Sitz** an, hier kommt das Territorialitätsprinzip zur Anwendung. Die Vorgaben er KI-VO gelten nur für Betreiber, die in der Union niedergelassen sind.

ErwG 22 a.E. geht darüber hinaus und erläutert, dass die Regelungen der KI-VO auch für Anbieter und Betreiber von KI-Systemen gelten sollen, die in einem Drittland niedergelassen sind, soweit beabsichtigt wird, das von diesem System erzeugte Ergebnis in der Union zu verwenden. Dies soll laut ErwG bezwecken, dass eine Umgehung der Verordnung verhindert wird und ein wirksamer Schutz in der Union ansässiger natürlicher Personen gewährleistet wird. Aus diesen Gründen ordnet Art. 2 Abs. 1 lit. c KI-VO an, dass bestimmte KI-Systeme unter die KI-VO fallen, selbst wenn sie **in der Union weder Verkehr gebracht noch in Betrieb genommen** oder verwendet werden. Dies soll laut ErwG 22 beispielsweise dann der Fall sein, wenn ein in der Union niedergelassener Akteur bestimmte Dienstleistungen an einen in einem Drittland niedergelassenen Akteur im Zusammenhang mit einer Tätigkeit vergibt, die von einem KI-System ausgeübt werden soll, das als hochriskant einzustufen wäre. Unter diesen Umständen könnte das von dem Akteur in einem Drittland betriebene KI-System Daten verarbeiten, die rechtmäßig in der Union erhoben und aus der Union übertragen wurden, und dem vertraglichen Akteur in der Union das aus dieser Verarbeitung resultierende Ergebnis dieses KI-Systems liefern, ohne dass dieses KI-System dabei in der Union in Verkehr gebracht, in Betrieb genommen oder verwendet würde. Auch diese Regelung dient dem Zweck, die Umgehung dieser Verordnung zu verhindern und einen wirksamen Schutz in der Union ansässiger natürlicher Personen zu gewährleisten. Nach den Erwägungsgründen bleibt offen, ob die Regelung nur gilt, wenn das vom KI-System erzeugte Ergebnis mit Wissen und Wollen von Anbieter und Betreiber in der EU verwendet wird oder auch eine Verwendung durch einen Dritten ausreicht, von der Anbieter und Betreiber nichts wissen.

Sonderproblem: „Versteckter“ erweiterter räumlicher Anwendungsbereich:

Eine wichtige Ausnahme bildet hierbei jedoch Art. 53 der KI-VO, der über den regulären räumlichen Anwendungsbereich hinausgeht. Während die Verordnung grundsätzlich das Marktortprinzip anwendet, greift Art. 53 auch dann, wenn das allgemeine KI-Modell (GPAI-Modell) außerhalb der EU trainiert wurde, aber später in der EU in Verkehr gebracht werden soll. Dies könnte auf eine versteckte Erweiterung des territorialen Anwendungsbereichs hinauslaufen, da Anbieter sicherstellen müssen, dass wirksam erklärte Vorbehalte zum Text- und Data-Mining auch außerhalb der EU beim Training berücksichtigt wurden. Diese Anforderung, die sich nicht direkt in den Artikeln der KI-VO wiederfindet, wird insbesondere durch Erwägungsgrund 106 untermauert. Für eine vertiefte Analyse dieser extraterritorialen Anwendung sei auf den „Exkurs: Pflichten für Anbieter von KI-Modellen mit allgemeinem Verwendungszweck“ verwiesen.

Beachte: Das Problem vertieft im „Exkurs: Pflichten für Anbieter von KI-Modellen mit allgemeinem Verwendungszweck“ behandelt.

Zwischenergebnis

Wenn nach Prüfung der Voraussetzungen festgestellt werden kann, dass der räumliche Anwendungsbereich eröffnet ist, ist mit **Schritt 2** fortzufahren.

Sollten die Voraussetzungen von **Art. 2 Abs. 1 KI-VO nicht erfüllt** sein, **ist die Prüfung abzubrechen.** Die KI-VO findet dann keine Anwendung.

4 Risikoklassifizierung

Lea Ludmilla Ossmann-Magiera LL.M. (Leiden) (Bitkom)

Nachdem festgestellt worden ist, dass es sich um ein KI-System i.S.d. KI-VO handelt und sowohl der persönliche als auch der räumliche Anwendungsbereich eröffnet sind, ist im nächsten Schritt zu prüfen, welcher Risikoklasse das KI-System zuzuordnen ist. Die Intensität der Prüfpflicht hängt dabei davon ab, welche Rolle das prüfende Unternehmen einnimmt. Als Anbieter von KI-Systemen ist man für die Einhaltung der anwendbaren Anforderungen der KI-VO voll verantwortlich. Bei Händlern ist die Prüfpflicht gemäß Art. 24 in gewisser Weise eingeschränkt und bezieht sich auf das Vorliegen aller Formalien (z. B. CE-Kennzeichnung) und eine ausreichende Prüfung der Informationen, die dem Händler zur Verfügung stehen.

Da die KI-VO einen risikobasierten Regulierungsansatz verfolgt (s. o.), richten sich die rechtlichen Anforderungen nach dem Risiko, das durch das konkrete KI-System hervorgerufen wird. Hierbei ist zwischen unannehmbarem Risiko (Verbote), hohem Risiko (strenge Anforderungen), geringem Risiko (weniger strenge Anforderungen) und minimalem Risiko (gar keine Anforderungen) zu unterscheiden.

Schritt 2: In welche Risikoklasse fällt mein KI-System?

2.1 Ist einer der Verbotstatbestände aus Art. 5 einschlägig?

Stephan Kress (Morrison & Foerster LLP), Dilan Mienert (GÖRG Partnerschaft von Rechtsanwälten mbB)

Verboten sind gem. **Art. 5 Abs. 1 KI-VO** eine Reihe von Praktiken. Im Folgenden werden diese einzeln dargestellt und behandelt.

Verbotstatbestand aus Art. 5 lit. a) KI-VO

a) das Inverkehrbringen, die Inbetriebnahme oder die Verwendung eines KI-Systems, das Techniken der unterschwelligen Beeinflussung außerhalb des Bewusstseins einer Person oder absichtlich manipulative oder täuschende Techniken mit dem Ziel oder der Wirkung einsetzt, das Verhalten einer Person oder einer Gruppe von Personen wesentlich zu beeinflussen, indem ihre Fähigkeit, eine fundierte Entscheidung zu treffen, spürbar beeinträchtigt wird, wodurch die Person veranlasst wird, eine Entscheidung zu treffen, die sie andernfalls nicht getroffen hätte, und zwar in einer Weise, die dieser Person, einer anderen Person oder einer Gruppe von Personen erheblichen Schaden zufügt oder zufügen kann;

Relevante(r) Artikel:

Art. 5 Abs. 1 lit. a)

Relevante(r) ErwG:

29

Konkretisierungsbedürftig:

Ja, durch juristische Auslegung

In Bezug auf den ersten Verbotstatbestand aus **Art. 5 Abs. 1 lit. a) KI-VO** ist gem. **ErwG 28 und 29** zu beachten, dass unterschwellige und manipulative Techniken nur dann verboten sind, wenn sie zum Ziel haben, Personen zu unerwünschten Verhaltensweisen zu bewegen oder sie zu täuschen, indem sie in einer Weise zu Entscheidungen angeregt werden, die ihre **Autonomie, Entscheidungsfindung und freie Auswahl untergräbt und beeinträchtigt.** Außerdem müssen sie das Ziel haben, menschliches Verhalten **maßgeblich nachteilig zu beeinflussen**, und **große Schäden**, insbesondere erhebliche nachteilige Auswirkungen auf die physische und psychische Gesundheit oder auf die finanziellen Interessen zu verursachen. Als Beispiele nennt ErwG 29 Computer-Schnittstellen oder virtuelle Realität.

Das Verbot nach dem ersten Verbotstatbestand bezieht sich auf den Einsatz, die Inbetriebnahme und das Inverkehrbringen der Systeme, nicht jedoch auf deren Entwicklung. Die Anwendung des Verbots setzt voraus, dass die Manipulation zielgerichtet erfolgt und eine klare Zurechenbarkeit vorliegt. Erforderlich ist eine Einflussnahme innerhalb der Kontrollmacht des Anbieters. Nicht erfasst vom ersten Verbotstatbestand sind wirtschaftliche oder materielle Beeinträchtigungen.

Verbotstatbestand aus Art. 5 Abs. 1 lit. b) KI-VO

Relevante(r) Artikel:

Art. 5 Abs. 1 lit. b)

Relevante(r) ErwG:

29

Konkretisierungsbedürftig:

Ja durch juristische Auslegung)

b) das Inverkehrbringen, die Inbetriebnahme oder die Verwendung eines KI-Systems, das eine Schwäche oder Schutzbedürftigkeit einer Person oder einer bestimmten Gruppe von Personen aufgrund ihres Alters, einer Behinderung oder einer bestimmten sozialen oder wirtschaftlichen Situation mit dem Ziel oder der Wirkung ausnutzt, das Verhalten einer Person oder einer dieser Gruppe angehörenden Person in einer Weise wesentlich zu beeinflussen, die dieser Person oder einer anderen Person erheblichen Schaden zufügt oder mit hinreichender Wahrscheinlichkeit zufügen wird

Der zweite Verbotstatbestand aus **Art. 5 Abs. 1 lit. b) KI-VO** ist ebenfalls in **ErwG 29** konkretisiert. Als Personengruppen, die besonders schwach und schutzbedürftig sind, wird auf Menschen mit Behinderungen i.S.d. RL (EU) 2019/882 verwiesen, auf alte Menschen und Personen, die in extremer Armut leben, und ethnische oder religiöse Minderheiten. Auch hier muss das Ziel des Einsatzes des KI-Systems darin bestehen, diesen Personen einen **erheblichen Schaden** zuzufügen. Diese Schäden müssen nicht eintreten, sondern nur mit hinreichender Wahrscheinlichkeit zu erwarten sein.

Der zweite Verbotstatbestand, der sich ebenfalls auf den Einsatz, die Inbetriebnahme und das Inverkehrbringen der Systeme bezieht, ist umfassender ausgestaltet als der erste Verbotstatbestand, da er nicht nur die subtile Beeinflussung verbietet, sondern auch das Ausnutzen von Schwächen bzw. Bedürfnissen, die durch Alter, Behinderung oder spezifische soziale und wirtschaftliche Umstände entstehen. Anders als das Verbot im Sinne von Art. 5 Abs. 1 lit. a) KI-VO, erfasst dieses Verbot auch wirtschaftliche Schäden.

Im Zusammenhang mit den ersten beiden Verbotstatbeständen ist die **RL 2005/29/EG** über unlautere Geschäftspraktiken im binnenmarktinternen Geschäftsverkehr zwischen Unternehmen und Verbrauchern zu sehen. Die KI-VO ergänzt die dortigen Vorschriften. **Ausnahmen** werden gemacht für rechtmäßige Praktiken im Zusammenhang mit medizinischen Behandlungen, etwa der psychologischen Behandlung einer psychischen Krankheit oder der physischen Rehabilitation (bei Zustimmung). Darüber hinaus sollten übliche und rechtmäßige Geschäftspraktiken, beispielsweise im Bereich der Werbung, die im Einklang mit den geltenden Rechtsvorschriften stehen, als solche nicht als schädliche manipulative KI-Praktiken gelten.

Verbotstatbestand aus Art. 5 Abs. 1 lit. c) KI-VO

c) das Inverkehrbringen, die Inbetriebnahme oder die Verwendung von KI-Systemen zur Bewertung oder Einstufung von natürlichen Personen oder Gruppen von Personen über einen bestimmten Zeitraum auf der Grundlage ihres sozialen Verhaltens oder bekannter, abgeleiteter oder vorhergesagter persönlicher Eigenschaften oder Persönlichkeitsmerkmale, wobei die soziale Bewertung zu einem oder beiden der folgenden Ergebnisse führt:

i) Schlechterstellung oder Benachteiligung bestimmter natürlicher Personen oder Gruppen von Personen in sozialen Zusammenhängen, die in keinem Zusammenhang zu den Umständen stehen, unter denen die Daten ursprünglich erzeugt oder erhoben wurden;

ii) Schlechterstellung oder Benachteiligung bestimmter natürlicher Personen oder Gruppen von Personen in einer Weise, die im Hinblick auf ihr soziales Verhalten oder dessen Tragweite ungerechtfertigt oder unverhältnismäßig ist

Relevante(r) Artikel:

Art. 5 Abs. 1 lit. c)

Relevante(r) ErwG:

31

Konkretisierungsbedürftig:

Ja durch juristische Auslegung)

Art. 5 Abs. 1 lit. c KI-VO wird durch ErwG 31 konkretisiert. Dieser Erwägungsgrund stellt klar, dass KI-Systeme, die eine soziale Bewertung natürlicher Personen durch öffentliche oder private Akteure ermöglichen, verboten sind, wenn sie zu diskriminierenden Ergebnissen, ungerechtfertigter Schlechterstellung oder Ausgrenzung führen. Solche Systeme bewerten Personen auf Basis von Datenpunkten zu sozialem Verhalten oder persönlichen Eigenschaften und können dadurch Menschenwürde, Nichtdiskriminierung, Gleichheit und Gerechtigkeit verletzen. Zulässig bleiben jedoch rechtmäßige Bewertungspraktiken, die im Einklang mit dem Unionsrecht und nationalem Recht stehen.

Auch der dritte Verbotstatbestand bezieht sich auf den Einsatz, die Inbetriebnahme und das Inverkehrbringen der Systeme, nicht jedoch auf deren Entwicklung. Der Schwerpunkt dieser Regelung liegt auf der Regulierung des sog. „Social Scoring“ durch KI-Systeme. Das Verbot findet Anwendung, wenn eine Verbindung zwischen dem Einsatz des KI-Systems und der daraus resultierenden Benachteiligung besteht. Vor diesem Hintergrund ist die Analyse der Verbindung essenziell für die Anwendung dieses Verbotstatbestandes.

Verbotstatbestand aus Art. 5 Abs. 1 lit. d) KI-VO

Relevante(r) Artikel:

Art. 5 Abs. 1 lit. d)

Relevante(r) ErwG:

31, 42

Konkretisierungsbedürftig:

Nein

d) das Inverkehrbringen, die Inbetriebnahme für diesen spezifischen Zweck oder die Verwendung eines KI-Systems zur Durchführung von Risikobewertungen in Bezug auf natürliche Personen, um das Risiko, dass eine natürliche Person eine Straftat begeht, ausschließlich auf der Grundlage des Profiling einer natürlichen Person oder der Bewertung ihrer persönlichen Merkmale und Eigenschaften zu bewerten oder vorherzusagen; dieses Verbot gilt nicht für KI-Systeme, die dazu verwendet werden, die durch Menschen durchgeführte Bewertung der Beteiligung einer Person an einer kriminellen Aktivität, die sich bereits auf objektive und überprüfbare Tatsachen stützt, die in unmittelbarem Zusammenhang mit einer kriminellen Aktivität stehen, zu unterstützen;

In Bezug auf den Verbotstatbestand aus **Art. 5 Abs. 1 lit. d) KI-VO** stellt **ErwG 31** klar, dass inakzeptable Bewertungspraktiken, die zu einer solchen Schlechterstellung oder Benachteiligung führen, verboten werden. Eine Ausnahme gilt für rechtmäßige Praktiken zur Bewertung natürlicher Personen, die im Einklang mit dem Unionsrecht und dem nationalen Recht zu einem bestimmten Zweck durchgeführt werden.

Die unter diese Regelung fallenden Praktiken sind unter dem Stichwort **Predictive Policing** bekannt. Art. 5 Abs. 1 lit. d) KI-VO verbietet den Einsatz von KI-Systemen, deren Zweck die Bewertung oder Vorhersage des Risikos der Begehung einer Straftat durch eine natürliche Person ausschließlich auf der Grundlage der Erstellung eines Profils oder der Bewertung von Persönlichkeitsmerkmalen und Eigenschaften ist. KI-Systeme, die eingesetzt werden, um zu ermitteln, an welchen Orten eine erhöhte Wahrscheinlichkeit künftiger Einbrüche besteht, um dem durch eine entsprechende Einsatzplanung öffentlicher Streifen entgegenzuwirken, fallen nicht unter das Verbot.

Verbotstatbestand aus Art. 5 Abs. 1 lit. e) KI-VO

e) das Inverkehrbringen, die Inbetriebnahme für diesen spezifischen Zweck oder die Verwendung von KI-Systemen, die Datenbanken zur Gesichtserkennung durch das ungezielte Auslesen von Gesichtsbildern aus dem Internet oder von Überwachungsaufnahmen erstellen oder erweitern;

Relevante(r) Artikel:

Art. 5 Abs. 1 lit. e)

Relevante(r) ErwG:

43

Konkretisierungsbedürftig:

Nein

ErwG 43 konkretisiert den Verbotstatbestand aus **Art. 5 Abs. 1 lit. e) KI-VO.** Dieser verbietet, Datenbanken zur Gesichtserkennung durch das ungezielte Auslesen von Gesichtsbildern aus dem Internet oder von Videoüberwachungsaufnahmen zu erstellen oder zu erweitern. Kern dieser Regelung ist das Verbot des Inverkehrbringens, der Inbetriebnahme oder der Nutzung solcher KI-Systeme.

Verbotstatbestand aus Art. 5 Abs. 1 lit. f) KI-VO

f) das Inverkehrbringen, die Inbetriebnahme für diesen spezifischen Zweck oder die Verwendung von KI-Systemen zur Ableitung von Emotionen einer natürlichen Person am Arbeitsplatz und in Bildungseinrichtungen, es sei denn, die Verwendung des KI-Systems soll aus medizinischen Gründen oder Sicherheitsgründen eingeführt oder auf den Markt gebracht werden;

Relevante(r) Artikel:

Art. 5 Abs. 1 lit. f)

Relevante(r) ErwG:

44

Konkretisierungsbedürftig:

Nein

ErwG 44 konkretisiert den Verbotstatbestand aus **Art. 5 Abs. 1 lit. f) KI-VO,** der es verbietet, mithilfe von KI-Systemen Emotion zu erkennen oder abzuleiten. Hintergrund der Regelung ist, dass KI-Systeme, die Emotionen oder Absichten natürlicher Personen auf der Grundlage ihrer biometrischen Daten erkennen oder ableiten können, ggf. diskriminierende Ergebnisse hervorbringen, da sich Gefühlsausdrücke je nach Kultur oder Situation erheblich unterscheiden. Dieses Verbot sollte nicht für KI-Systeme gelten, die ausschließlich aus medizinischen oder sicherheitstechnischen Gründen in Verkehr gebracht werden. Beispiele wären hier z. B. Systeme, die für therapeutische Zwecke bestimmt sind, oder Überwachungstechnik in Fahrzeugen, die Müdigkeit der Fahrzeugführer erkennen, um Sicherheitsmaßnahmen einzuleiten.

Verbotstatbestand aus Art. 5 Abs. 1 lit. g) KI-VO

g) das Inverkehrbringen, die Inbetriebnahme für diesen spezifischen Zweck oder die Verwendung von Systemen zur biometrischen Kategorisierung, mit denen natürliche Personen individuell auf der Grundlage ihrer biometrischen Daten kategorisiert werden, um ihre Rasse, ihre politischen Einstellungen, ihre Gewerkschaftszugehörigkeit, ihre religiösen oder weltanschaulichen Überzeugungen, ihr Sexualleben oder ihre sexuelle Ausrichtung zu erschließen oder abzuleiten; dieses Verbot gilt nicht für die Kennzeichnung oder Filterung rechtmäßig erworbener biometrischer Datensätze, wie z. B. Bilder auf der Grundlage biometrischer Daten oder die Kategorisierung biometrischer Daten im Bereich der Strafverfolgung;

Relevante(r) Artikel:

Art. 5 Abs. 1 lit. g)

Relevante(r) ErwG:

30, 40, 41

Konkretisierungsbedürftig:

Nein

In Bezug auf den Verbotstatbestand aus **Art. 5 Abs. 1 lit. g) KI-VO** sind gem. **ErwG 30 folgende Ausnahmen zum Verbot der biometrischen Kategorisierung** zu beachten: für die rechtmäßige Kennzeichnung, Filterung oder Kategorisierung biometrischer Datensätze, die im Einklang mit dem Unionsrecht oder dem nationalen Recht anhand biometrischer Daten erworben wurden, wie das Sortieren von Bildern nach Haar- oder Augenfarbe, was beispielsweise im Bereich der Strafverfolgung verwendet werden kann.

Verbotstatbestand aus Art. 5 Abs. 1 lit. h) KI-VO

Relevante(r) Artikel:

Art. 5 Abs. 1 lit. h)

Relevante(r) ErwG:

32, 33, 34, 35, 36, 37, 38, 39

Konkretisierungsbedürftig:

Nein

h) die Verwendung biometrischer Echtzeit-Fernidentifizierungssysteme in öffentlich zugänglichen Räumen zu Strafverfolgungszwecken, außer wenn und insoweit dies im Hinblick auf eines der folgenden Ziele unbedingt erforderlich ist:

i) gezielte Suche nach bestimmten Opfern von Entführung, Menschenhandel oder sexueller Ausbeutung sowie Suche nach vermissten Personen;

ii) Abwenden einer konkreten, erheblichen und unmittelbaren Gefahr für das Leben oder die körperliche Unversehrtheit natürlicher Personen oder einer tatsächlichen und bestehenden oder tatsächlichen und vorhersehbaren Gefahr eines Terroranschlags;

iii) Aufspüren oder Identifizieren einer Person, die der Begehung einer Straftat verdächtigt wird, zum Zwecke der Durchführung von strafrechtlichen Ermittlungen, der Verfolgung oder der Vollstreckung einer Strafe für die in Anhang II aufgeführten Straftaten, die in dem betreffenden Mitgliedstaat nach dessen Recht mit einer Freiheitsstrafe oder einer freiheitsentziehenden Maßregel der Sicherung im Höchstmaß von mindestens vier Jahren bedroht ist;

Die Absätze 2 bis 7 des Art. 5 sowie **ErwG 32**konkretisiert den Verbotstatbestand aus **Art. 5 Abs. 1 lit. h) KI-VO** zu biometrischer Echtzeit-Fernidentifizierung natürlicher Personen in öffentlich zugänglichen Räumen zu Strafverfolgungszwecken, die nur für die in jenem Buchstaben genannte Zwecke erfolgen darf. Die Praktik greift besonders in die Rechte und Freiheiten der betroffenen Personen ein, da sie die **Privatsphäre** eines großen Teils der Bevölkerung beeinträchtigt, ein Gefühl der ständigen Überwachung weckt und indirekt **von der Ausübung der Versammlungsfreiheit und anderer Grundrechte abhalten** kann. **Technische Ungenauigkeiten** von KI-Systemen, die für die biometrische Fernidentifizierung natürlicher Personen bestimmt sind, können zu verzerrten Ergebnissen führen und eine **diskriminierende Wirkung** haben. Außerdem wird auf die begrenzten Möglichkeiten weiterer Kontrollen oder Korrekturen im Zusammenhang mit der Verwendung solcher in Echtzeit betriebener Systeme hingewiesen. Die **Ausnahmen** werden in **ErwG 33** erläutert. Gem.

Folgende weitere Einschränkungen bzw. Vorgaben bestehen in Bezug auf die biometrische Echtzeit-Fernidentifizierung:

- Die Verwendung biometrischer Echtzeit-Fernidentifizierungssysteme darf nur zur Bestätigung der Identität der speziell betroffenen Person erfolgen.

- Eine vorherige Genehmigung der zuständigen Justizbehörde oder Verwaltungsbehörde ist einzuholen.
- Vor dem Einsatz ist durch die zuständige Strafverfolgungsbehörde eine **Grundrechte-Folgenabschätzung** durchzuführen und das KI-System muss gemäß Artikel 49 in der EU-Datenbank registriert werden.

Die Kommission veröffentlicht Jahresberichte über die Verwendung biometrischer Echtzeit-Fernidentifizierungssysteme in öffentlich zugänglichen Räumen zu Strafverfolgungszwecken.

Grenzfälle und Beispiele verbotener Praktiken

Bei folgenden Fällen könnte unter bestimmten Umständen eine nähere Prüfung des Vorliegens eines Verbotstatbestands angezeigt sein:

- Die Gestaltung virtueller Realitäten, die Nutzer zu einem potenziell schädlichen Verhalten animieren könnte, da VR ein hohes Maß an Kontrolle darüber ermöglicht, welche Reize den Nutzern angeboten werden.
- Der KI-gesteuerte Algorithmus einer Online-Plattform wertet gezielt das Nutzungsverhalten von Kindern aus, um sie unterschwellig zu einem Konsumverhalten anzuregen, das ggf. als schädlich eingestuft werden könnte.
- Ein KI-gestütztes Programm zur Bewertung der Bonität von natürlichen Personen schließt von bestimmten gruppenbezogenen Merkmalen auf eine schlechtere Bonität dieser Gruppe.
- Ein KI-System, das die (Glücks)Spielgewohnheiten der Nutzer analysiert, um ihnen gezielt Anreize für weitergehendes Spiel zu geben, könnte die Entscheidungsfreiheit der Nutzer untergraben und sie in eine Spielsucht treiben. Dies wäre nach Art. 5 Abs. 1 lit. a) verboten, wenn das System bewusst darauf abzielt, süchtiges Verhalten zu fördern und erhebliche finanzielle oder psychische Schäden verursacht.
- Ein KI-System, das Bürgern basierend auf ihrem sozialen Verhalten (z. B. Online-Aktivitäten, Zahlungsverhalten) Punkte vergibt, die Auswirkungen auf ihre Zugangschancen zu öffentlichen Dienstleistungen haben, wäre nach Art. 5 Abs. 1 lit. c) KI-VO verboten, wenn es zu diskriminierenden Ergebnissen und ungerechtfertigter Benachteiligung führt.
- Ein KI-System, das Bürgern basierend auf ihrem sozialen Verhalten (z. B. Online-Aktivitäten, Zahlungsverhalten) Punkte vergibt, die Auswirkungen auf ihre Zugangschancen zu öffentlichen Dienstleistungen haben, wäre nach Art. 5 Abs. 1 lit. e) KI-VO verboten.
- Ein KI-System, das in Vorstellungsgesprächen die Emotionen von Bewerbern analysiert, um deren Eignung für eine Stelle zu bewerten, könnte gegen Art. 5 Abs. 1 lit. f) KI-VO verstoßen.

Zwischenergebnis

Sollte einer der Verbotstatbestände erfüllt sein, geht die Prüfung bei **Schritt 3** weiter.

Wenn keiner der Verbotstatbestände auf das in Rede stehende KI-System zutrifft, ist die Prüfung im **Prüfungsschritt 2.2** fortzusetzen.

2.2 Ist das System hochriskant nach Art. 6 Abs. 1 i.V.m. Anhang I?

Dr. Benedikt Rohrßen (Taylor Wessing Partnerschaftsgesellschaft mbB)

Relevante(r) Artikel:

Art. 6 Abs.1 i.V.m. Anhang I

Art. 6 Abs. 5 und 6

Relevante(r) ErwG:

46, 47, 48, 49

Konkretisierungsbedürftig:

Ja, (Leitlinien der Kommission zur praktischen Umsetzung des Art. 6 gem. Art. 96 und umfassende Liste praktischer Beispiele für Anwendungsfälle für KI-Systeme, die hochriskant oder nicht hochriskant sind (Art. 6 Abs. 5))

Art. 6 KI-VO regelt die **Risikoklassifizierung** von KI-Systemen. Diese Klassifizierung hat **durch den Anbieter** zu erfolgen. Gem. **Art. 6 Abs. 1 KI-VO** gilt ein KI-System als hochriskant, wenn folgende Voraussetzungen **kumulativ** erfüllt sind:

a) a) das KI-System soll als Sicherheitsbauteil eines unter die in Anhang I aufgeführten Harmonisierungsrechtsvorschriften der Union fallenden Produkts verwendet werden oder das KI-System ist selbst ein solches Produkt;

b) b) das Produkt, dessen Sicherheitsbauteil gemäß Buchstabe a das KI-System ist, oder das KI-System selbst als Produkt muss einer Konformitätsbewertung durch Dritte im Hinblick auf das Inverkehrbringen oder die Inbetriebnahme dieses Produkts gemäß den in Anhang I aufgeführten Harmonisierungsrechtsvorschriften der Union unterzogen werden.

Aus ErwG 46 ergibt sich, dass als hochriskant nur solche KI-Systeme eingestuft werden sollten, die **erhebliche schädliche Auswirkungen auf die Gesundheit, die Sicherheit und die Grundrechte von Personen in der Union haben**, wodurch eine mögliche Beschränkung des internationalen Handels so gering wie möglich bleibt. Gem. **ErwG 48** ist das Ausmaß der nachteiligen Auswirkungen des KI-Systems auf die durch die **Charta geschützten Grundrechte** ist bei der Einstufung eines KI-Systems als hochriskant von besonderer Bedeutung.

Als **Beispiele** nennt ErwG 48 die Würde des Menschen, die Achtung des Privat- und Familienlebens, den Schutz personenbezogener Daten, die Freiheit der Meinungsäußerung und die Informationsfreiheit, die Versammlungs- und Vereinigungsfreiheit, das Recht auf die Nichtdiskriminierung, das Recht auf Bildung, der Verbraucherschutz, die Arbeitnehmerrechte, die Rechte von Menschen mit Behinderungen, die Gleichstellung der Geschlechter, Rechte des geistigen Eigentums, das Recht auf einen wirksamen Rechtsbehelf und ein faires Gerichtsverfahren, das Verteidigungsrecht, die Unschuldsvermutung sowie das Recht auf eine gute Verwaltung. Kinder verfügen darüber hinaus über Sonderrechte.

Voraussetzung für die Einstufung eines KI-Systems als Hochrisiko-KI-System ist **zunächst**, dass das KI-System entweder als Sicherheitsbauteil für ein Produkt verwendet wird, das Regelungsgegenstand der in Anhang I aufgeführten Harmonisierungsrechtsakte ist, oder selbst ein solches Produkt ist.

Anhang I Abschnitt A listet folgende Rechtsakte auf, die zum New Legislative Framework (NLF) angehören:

1. **Richtlinie 2006/42/EG** des Europäischen Parlaments und des Rates vom 17. Mai 2006 über **Maschinen** und zur Änderung der Richtlinie 95/16/EG [aufgehoben durch die Maschinenverordnung]

2. **Richtlinie 2009/48/EG** des Europäischen Parlaments und des Rates vom 18. Juni 2009 über die Sicherheit von Spielzeug

3. **Richtlinie 2013/53/EU** des Europäischen Parlaments und des Rates vom 20. November 2013 über **Sportboote und Wassermotorräder** und zur Aufhebung der Richtlinie 94/25/EG

4. **Richtlinie 2014/33/EU** des Europäischen Parlaments und des Rates vom 26. Februar 2014 zur Angleichung der Rechtsvorschriften der Mitgliedstaaten über Aufzüge und Sicherheitsbauteile für **Aufzüge**

5. Richtlinie 2014/34/EU des Europäischen Parlaments und des Rates vom 26. Februar 2014 zur Harmonisierung der Rechtsvorschriften der Mitgliedstaaten für Geräte und Schutzsysteme zur bestimmungsgemäßen Verwendung in **explosionsgefährdeten Bereichen**

6. **Richtlinie 2014/53/EU** des Europäischen Parlaments und des Rates vom 16. April 2014 über die Harmonisierung der Rechtsvorschriften der Mitgliedstaaten über die Bereitstellung von **Funkanlage**n auf dem Markt und zur Aufhebung der Richtlinie 1999/5/EG

7. **Richtlinie 2014/68/EU** des Europäischen Parlaments und des Rates vom 15. Mai 2014 zur Harmonisierung der Rechtsvorschriften der Mitgliedstaaten über die Bereitstellung von **Druckgeräten** auf dem Markt

8. **Verordnung (EU) 2016/424** des Europäischen Parlaments und des Rates vom 9. März 2016 über **Seilbahnen** und zur Aufhebung der Richtlinie 2000/9/EG

9. **Verordnung (EU) 2016/425** des Europäischen Parlaments und des Rates vom 9. März 2016 über **persönliche Schutzausrüstungen** und zur Aufhebung der Richtlinie 89/686/EWG des Rates

10. **Verordnung (EU) 2016/426** des Europäischen Parlaments und des Rates vom 9. März 2016 über **Geräte zur Verbrennung gasförmiger Brennstoffe** und zur Aufhebung der Richtlinie 2009/142/EG

11. Verordnung (EU) 2017/745 des Europäischen Parlaments und des Rates vom 5. April 2017 über Medizinprodukte, zur Änderung der Richtlinie 2001/83/EG, der Verordnung (EG) Nr. 178/2002 und der Verordnung (EG) Nr. 1223/2009 und zur Aufhebung der Richtlinien 90/385/EWG und 93/42/EWG des Rates

12. Verordnung (EU) 2017/746 des Europäischen Parlaments und des Rates vom 5. April 2017 über In-vitro-Diagnostika und zur Aufhebung der Richtlinie 98/79/EG und des Beschlusses 2010/227/EU der Kommission

Anhang I Abschnitt B listet folgende Rechtsvorschriften **auf:**

13. **Verordnung (EG) Nr. 300/2008** des Europäischen Parlaments und des Rates vom 11. März 2008 über gemeinsame Vorschriften für die Sicherheit in der **Zivilluftfahrt** und zur Aufhebung der Verordnung (EG) Nr. 2320/2002

14. Verordnung (EU) Nr. 168/2013 des Europäischen Parlaments und des Rates vom 15. Januar 2013 über **die Genehmigung und Marktüberwachung von zwei- oder dreirädrigen und vierrädrigen Fahrzeugen**

15. Verordnung (EU) Nr. 167/2013 des Europäischen Parlaments und des Rates vom 5. Februar 2013 über die **Genehmigung und Marktüberwachung von land- und forstwirtschaftlichen Fahrzeugen**

16. **Richtlinie 2014/90/EU** des Europäischen Parlaments und des Rates vom 23. Juli 2014 über **Schiffsausrüstung** und zur Aufhebung der Richtlinie 96/98/EG des Rates

17. Richtlinie (EU) 2016/797 des Europ. Parlaments und des Rates vom 11. Mai 2016 über **die Interoperabilität des Eisenbahnsystems in der EU**

18. **Verordnung (EU) 2018/858** des Europäischen Parlaments und des Rates vom 30. Mai 2018 über die **Genehmigung und die Marktüberwachung von Kraftfahrzeugen und Kraftfahrzeuganhängern** sowie von Systemen, Bauteilen und selbstständigen technischen Einheiten für diese Fahrzeuge, zur Änderung der Verordnungen (EG) Nr. 715/2007 und (EG) Nr. 595/2009 und zur Aufhebung der Richtlinie 2007/46/EG

19. **Verordnung (EU) 2019/2144** des Europäischen Parlaments und des Rates vom 27. November 2019 über die **Typgenehmigung von Kraftfahrzeugen und Kraftfahrzeuganhängern** sowie von Systemen, Bauteilen und selbstständigen technischen Einheiten für diese Fahrzeuge im Hinblick auf ihre allgemeine Sicherheit und den Schutz der Fahrzeuginsassen und von ungeschützten Verkehrsteilnehmern, zur Änderung der Verordnung (EU) 2018/858 des Europäischen Parlaments und des Rates und zur Aufhebung der Verordnungen (EG) Nr. 78/2009, (EG) Nr. 79/2009 und (EG) Nr. 661/2009 des Europäischen Parlaments und des Rates sowie der Verordnungen (EG) Nr. 631/2009, (EU) Nr. 406/2010, (EU) Nr. 672/2010, (EU) Nr. 1003/2010, (EU) Nr. 1005/2010, (EU) Nr. 1008/2010, (EU) Nr. 1009/2010, (EU) Nr. 19/2011, (EU) Nr. 109/2011, (EU) Nr. 458/2011, (EU) Nr. 65/2012, (EU) Nr. 130/2012, (EU) Nr. 347/2012, (EU) Nr. 351/2012, (EU) Nr. 1230/2012 und (EU) 2015/166 der Kommission

20. **Verordnung (EU) 2018/1139** des Europäischen Parlaments und des Rates vom 4. Juli 2018 zur Festlegung gemeinsamer Vorschriften für die **Zivilluftfahrt und zur Errichtung einer Agentur der Europäischen Union für Flugsicherhei**t sowie zur Änderung der Verordnungen (EG) Nr. 2111/2005, (EG) Nr. 1008/2008, (EU) Nr. 996/2010, (EU) Nr. 376/2014 und der Richtlinien 2014/30/EU und 2014/53/EU des Europäischen Parlaments und des Rates, und zur Aufhebung der Verordnungen (EG) Nr. 552/2004 und (EG) Nr. 216/2008 des Europäischen Parlaments und des Rates und der Verordnung (EWG) Nr. 3922/91 des Rates, insoweit die Konstruktion, Herstellung und Vermarktung von Luftfahrzeugen gemäß Artikel 2 Absatz 1 Buchstaben a und b in Bezug auf unbemannte Luftfahrzeuge sowie deren Motoren, Propeller, Teile und Ausrüstung zur Fernsteuerung betroffen sind.

Aus **ErwG 49** ergibt sich **Änderungsbedarf** hinsichtlich folgender Rechtsakte, um sicherzustellen, dass die Kommission beim Erlass weiterer Rechtsakte basierend auf diesen Rechtsakten die spezifischen Anforderungen an Hochrisiko-KI-Systeme aus der KI-VO berücksichtigt, ohne bestehende Mechanismen zu beeinträchtigen:

- Verordnung (EG) Nr. 300/2008
- Verordnung (EU) Nr. 167/2013

- Verordnung (EU) Nr. 168/2013
- Richtlinie 2014/90/EU
- Richtlinie (EU) 2016/797
- Verordnung (EU) 2018/858
- Verordnung (EU) 2018/1139
- Verordnung (EU) 2019/2144

Um als hochriskant eingestuft zu werden, muss **zudem** entweder das KI-System selbst als Produkt oder das Produkt, dessen Sicherheitsbauteil das KI-System ist, einer Konformitätsbewertung durch Dritte gemäß dieser Harmonisierungsrechtsvorschriften unterliegen.

Gem. ErwG 50 impliziert die für diese Produkte notwendige Vorabkonformitätsbewertung durch Dritte die Einstufung des jeweiligen Produkts als „hochriskant" gemäß der KI-VO. Darunter fallen namentlich Maschinen, Spielzeuge und Medizinprodukte sowie beispielhaft Aufzüge, Funkanlagen, Druckgeräte und Seilbahnen. Aus der Einstufung eines KI-Systems als hochriskant gemäß der KI-VO ergibt sich allerdings nicht, dass das gesamte Produkt, das die KI-Komponente enthält, oder das KI-System gemäß den einschlägigen Harmonisierungsrechtsvorschriften als hochriskant betrachtet wird, ErwG 51.

Andere Hochrisiko-KI-Systeme hingegen, die nicht Sicherheitsbauteile sind oder selbst Produkte sind, d. h. „**eigenständige KI-Systeme**", sind als **hochriskant einzustufen, wenn sie** aufgrund ihrer vorgesehenen Verwendung ein hohes Risiko darstellen können, die Gesundheit, Sicherheit oder Grundrechte von Menschen zu gefährden und sofern sie in Bereichen eingesetzt werden, die in der KI-VO festgelegt sind. Bei dieser Einstufung wird sowohl die Schwere des möglichen Schadens als auch die Wahrscheinlichkeit seines Eintretens berücksichtigt.

Die Methode und die Kriterien zur Bestimmung dieser hochriskanten Systeme sind die gleichen, die auch für zukünftige Änderungen der Liste dieser Systeme verwendet werden. Insofern hat die Europäische Kommission die Befugnis, diese Liste durch delegierte Rechtsakte zu aktualisieren, um mit der schnellen technologischen Entwicklung und den Veränderungen in der Nutzung von KI-Systemen Schritt zu halten, vgl. ErwG 52.

Beispiele Hochrisiko-KI-Systeme (Anhang I)

Gem. Art. 6 Abs. 5 KI-VO stellt die Kommission nach Konsultation des Europäischen Ausschusses für künstliche Intelligenz spätestens 18 Monate nach Inkrafttreten der Verordnung **Leitlinien zur praktischen Umsetzung des Artikel 6** gemäß **Art. 96 KI-VO** und eine **umfassende Liste praktischer Beispiele** für Anwendungsfälle für KI-Systeme, die hochriskant oder nicht hochriskant sind, bereit.

Zwischenergebnis

Ist das KI-System nach Art. 6 Abs. 1 i.V.m. Anhang I als hochriskant einzustufen, gelten die Anforderungen an Hochrisiko-KI-Systeme gem. der Art. 8 – 15 KI-VO. Wirtschaftsakteure haben je nach ihrer Rolle ihren spezifischen Pflichten aus den Art. 16 – 27 KI-VO gerecht zu werden. Die Prüfung geht mit **Schritt 4.1 für Anbieter und Schritt 4.2 für Betreiber** weiter.

Ist das KI-System dagegen nicht nach Art. 6 Abs. 1 i.V.m. Anhang I als hochriskant einzustufen, geht die Prüfung mit **Schritt 2.3** weiter.

2.3 Ist das System hochriskant nach Art. 6 Abs. 2 i.V.m. Anhang III?

Dr. Benedikt Rohrßen (Taylor Wessing Partnerschaftsgesellschaft mbB)

Art. 6 Abs. 2 KI-VO regelt die zweite Alternative, nach der KI-Systeme als hochriskant einzustufen sind:

> (2) Zusätzlich zu den in Absatz 1 genannten Hochrisiko-KI-Systemen gelten die in Anhang III genannten KI-Systeme als hochriskant.

1. **Biometrik**, soweit ihr Einsatz nach einschlägigem Unionsrecht oder nationalem Recht zugelassen ist:
   a) **biometrische Fernidentifizierungssysteme**. Dazu gehören nicht KI-Systeme, die bestimmungsgemäß für die biometrische Verifizierung, deren einziger Zweck darin besteht, zu bestätigen, dass eine bestimmte natürliche Person die Person ist, für die sie sich ausgibt, verwendet werden sollen;
   b) KI-Systeme, die bestimmungsgemäß für die **biometrische Kategorisierung** nach sensitiven oder geschützten Attributen oder Merkmalen oder auf der Grundlage von Rückschlüssen auf diese Attribute oder Merkmale verwendet werden sollen;
   c) KI-Systeme, die bestimmungsgemäß zur **Emotionserkennung** verwendet werden sollen.
2. **Kritische Infrastruktur**
   a) KI-Systeme, die bestimmungsgemäß als Sicherheitsbauteile im Rahmen der Verwaltung und des Betriebs **kritischer digitaler Infrastruktur, des Straßenverkehrs sowie der Wasser-, Gas-, Wärme- und Stromversorgung** verwendet werden sollen
3. **Allgemeine und berufliche Bildung**
   a) KI-Systeme, die bestimmungsgemäß zur Feststellung **des Zugangs oder der Zulassung oder zur Zuweisung natürlicher Personen zu Einrichtungen** aller Ebenen der allgemeinen und **beruflichen Bildung** verwendet werden sollen;
   b) KI-Systeme, die bestimmungsgemäß für **die Bewertung von Lernergebnisse**n verwendet werden sollen, einschließlich des Falles, dass diese Ergebnisse dazu dienen, den Lernprozess natürlicher Personen in Einrichtungen oder Programmen aller Ebenen der allgemeinen und beruflichen Bildung zu steuern;
   c) KI-Systeme, die bestimmungsgemäß zum Zweck **der Bewertung des angemessenen Bildungsniveaus**, das eine Person im Rahmen von oder innerhalb von Einrichtungen der allgemeinen und beruflichen Bildung erhalten wird oder zu dem sie Zugang erhalten wird, verwendet werden sollen;

Relevante(r) Artikel:

Art. 6 Abs.2 i.V.m. Anhang III 6 Abs.2 i.V.m. Anhang III

Art. 6 Abs. 3, 4, 5, 6

Relevante(r) ErwG:

46, 47, 48

Konkretisierungsbedürftig:

Ja, Leitlinien der Kommission zur praktischen Umsetzung des Art. 6 gem. Art. 96 und umfassende Liste praktischer Beispiele für Anwendungsfälle für KI-Systeme, die hochriskant oder nicht hochriskant sind (Art. 6 Abs. 5); delegierte Rechtsakte der Kommission gem. Art. 97 KI-VO zur Änderung des Art. 6 Abs. 3 Unterabs. 2 (Art. 6 Abs. 6 und 7 KI-VO)6 und 7 KI-VO)

d) KI-Systeme, die bestimmungsgemäß zur **Überwachung und Erkennung von verbotenen Verhalten von Schülern bei Prüfungen** im Rahmen von oder innerhalb von Einrichtungen der allgemeinen und beruflichen Bildung verwendet werden sollen.

4. **Beschäftigung, Personalmanagement und Zugang zur Selbstständigkeit**

   a) KI-Systeme, die bestimmungsgemäß für die Einstellung oder Auswahl natürlicher Personen verwendet werden sollen, insbesondere um gezielte Stellenanzeigen zu schalten, Bewerbungen zu sichten oder zu filtern und Bewerber zu bewerten;

   b) KI, die bestimmungsgemäß für Entscheidungen, die die Bedingungen von Arbeitsverhältnissen, Beförderungen und Kündigungen von Arbeitsvertragsverhältnissen beeinflussen, für die Zuweisung von Aufgaben aufgrund des individuellen Verhaltens oder persönlicher Merkmale oder Eigenschaften oder für die Beobachtung und Bewertung der Leistung und des Verhaltens von Personen in solchen Beschäftigungsverhältnissen verwendet werden soll.

5. **Zugänglichkeit und Inanspruchnahme grundlegender privater und grundlegender öffentlicher Dienste und Leistungen**:

   a) KI-Systeme, die bestimmungsgemäß von Behörden oder im Namen von Behörden verwendet werden sollen, um zu beurteilen, ob natürliche Personen Anspruch auf **grundlegende öffentliche Unterstützungsleistungen** und -dienste, einschließlich Gesundheitsdiensten, haben und ob solche Leistungen und Dienste zu gewähren, einzuschränken, zu widerrufen oder zurückzufordern sind;

   b) KI-Systeme, die bestimmungsgemäß für die **Kreditwürdigkeitsprüfung** und Kreditpunktebewertung natürlicher Personen verwendet werden sollen, mit Ausnahme von KI-Systemen, die zur Aufdeckung von Finanzbetrug verwendet werden;

   c) KI-Systeme, die bestimmungsgemäß für **die Risikobewertung und Preisbildung** in Bezug auf natürliche Personen im Fall von **Kranken- und Lebensversicherungen** verwendet werden sollen;

   d) KI-Systeme, die bestimmungsgemäß zur **Bewertung und Klassifizierung von Notrufen** von natürlichen Personen oder für die Entsendung oder Priorisierung des Einsatzes von Not- und Rettungsdiensten, einschließlich Polizei, Feuerwehr und medizinischer Nothilfe, sowie für Systeme für die Triage von Patienten bei der Notfallversorgung verwendet werden sollen.

6. **Strafverfolgun**g, soweit ihr Einsatz nach einschlägigem Unionsrecht oder nationalem Recht zugelassen ist:

   a) a) KI-Systeme, die bestimmungsgemäß von Strafverfolgungsbehörden oder in deren Namen oder von Organen, Einrichtungen und sonstigen Stellen der Union zur Unterstützung von Strafverfolgungsbehörden oder in deren Namen zur Bewertung des Risikos einer natürlichen Person, zum Opfer von Straftaten zu werden verwendet werden sollen;

   b) b) KI-Systeme, die bestimmungsgemäß von Strafverfolgungsbehörden oder in deren Namen oder von Organen, Einrichtungen und sonstigen Stellen der Union zur Unterstützung von Strafverfolgungsbehörden als Lügendetektoren oder ähnliche Instrumente verwendet werden sollen;

c) c) KI-Systeme, die bestimmungsgemäß von Strafverfolgungsbehörden oder in deren Namen oder von Organen, Einrichtungen und sonstigen Stellen der Union zur Unterstützung von Strafverfolgungsbehörden zur Bewertung der Verlässlichkeit von Beweismitteln im Zuge der Ermittlung oder Verfolgung von Straftaten verwendet werden sollen;

d) d) KI-Systeme, die bestimmungsgemäß von Strafverfolgungsbehörden oder in deren Namen oder von Organen, Einrichtungen und sonstigen Stellen der Union zur Unterstützung von Strafverfolgungsbehörden zur Bewertung der Wahrscheinlichkeit, dass eine natürliche Person eine Straftat begeht oder erneut begeht, nicht nur auf der Grundlage der Erstellung von Profilen natürlicher Personen gemäß Artikel 3 Absatz 4 der Richtlinie (EU) 2016/680 oder zur Bewertung persönlicher Merkmale und Eigenschaften oder vergangenen kriminellen Verhaltens von natürlichen Personen oder Gruppen verwendet werden sollen;

e) e) KI-Systeme, die bestimmungsgemäß von Strafverfolgungsbehörden oder in deren Namen oder von Organen, Einrichtungen und sonstigen Stellen der Union zur Unterstützung von Strafverfolgungsbehörden zur Erstellung von Profilen natürlicher Personen gemäß Artikel 3 Absatz 4 der Richtlinie (EU) 2016/680 im Zuge der Aufdeckung, Ermittlung oder Verfolgung von Straftaten verwendet werden sollen.

7. **Migration, Asyl und Grenzkontrolle**, soweit ihr Einsatz nach einschlägigem Unionsrecht oder nationalem Recht zugelassen ist:

a) KI-Systeme, die bestimmungsgemäß von zuständigen Behörden als **Lügendetektoren** und ähnliche Instrumente verwendet werden sollen;

Die Erwägungsgründe 54 ff. der KI-VO erläutern die in Anhang III der KI-VO genannten Hochrisiko-KI-Systeme näher:

Zunächst gelten **biometrische Systeme** wie **Fernidentifizierungs-, Kategorisierungs- und Emotionserkennungssysteme** als hochriskant. Denn solche KI-Systeme, die biometrische Daten wie Gesichts- und Emotionserkennung nutzen, sind besonders missbrauchsanfällig und können durch technische Ungenauigkeiten herbeigeführte Verzerrungen schwerwiegende, auch diskriminierende Auswirkungen auf Datenschutz und Privatsphäre entfalten.

KI-Systeme zur biometrischen Identifizierung und Authentifizierung hingegen, die ausschließlich dazu bestimmt sind, zu bestätigen, dass eine Person diejenige ist, für die sie sich ausgibt, sowie zur **Identitätsbestätigung, um Zugang zu einem Dienst zu erhalten, ein Gerät zu entriegeln oder sicheren Zugang zu Räumlichkeiten** zu gewähren, sind nicht als hochriskant einzustufen. Auch biometrische Systeme, deren alleiniger Zweck darin liegt, die Durchführung von Maßnahmen zur Cybersicherheit und zum Schutz personenbezogener Daten zu ermöglichen, gelten nicht als Hochrisikosysteme.

Daneben gelten KI-Systeme, die als **Sicherheitsbauteile** in sicherheitskritischen Bereichen wie kritischer digitaler Infrastruktur, dem Verkehrssektor oder der Energiewirtschaft (Wasser-, Gas-, Wärme- und Stromversorgung) eingesetzt werden, als hochriskant, da Ausfälle oder Störungen solcher Bauteile zu erheblichen Gesundheits- und Lebensgefahren für Personen, Risiken für das Eigentum sowie Störungen bei der Durchführung sozialer und wirtschaftlicher Tätigkeiten führen können. Dazu gehören etwa Systeme für die

Überwachung des Wasserdrucks oder Feuermelder-Kontrollsysteme. Bauteile hingegen, die ausschließlich im Rahmen der Cybersicherheit eingesetzt werden, stellen keine Sicherheitskomponenten i.S.d. Anhang III Nr. 2 lit. a) und sind daher nicht als hochriskant einzustufen.

Auch wird der Einsatz von KI-Systemen zur **Beurteilung von Leistungen**, zur **Klassifizierung von Personen** oder dem Zugang zu und der Nutzung bestimmter **grundlegender privater und öffentlicher Dienstleistungen** deshalb als hochriskant eingestuft, da eine ungenaue oder voreingenommene Bewertung signifikante Auswirkungen auf die Berufsaussichten, gesellschaftliche Teilhabe, Lebensqualität und Existenzgrundlagen der betroffenen Personen haben kann. Solche bewertenden Hochrisiko-KI-Systeme analysieren beispielsweise Daten wie Einkommen, Beschäftigungshistorien und Bonität, um die Kreditwürdigkeit von Unternehmen oder Personen vorherzusagen. Auch prüfen sie Lebensläufe und berufliche Qualifikationen von Bewerbern für bestimmte Stellen und treffen möglicherweise eine Vorauswahl qualifizierter Kandidaten oder bewerten Arbeitskräfte für Beförderungen oder Kündigungen. Der Einsatz solcher KI-Systemen wird **insbesondere in den folgenden Bereichen als hochriskant** betrachtet:

- **Bildung**: Auch wenn der Einsatz von KI-Systemen zu Bildungszwecken für eine dem digitalen Fortschritt gerecht werdende Förderung von wesentlicher Bedeutung ist, kann der Einsatz von KI-Systemen im Bildungsbereich bei fehlerhafter Konzeption unsachgemäße Entscheidungen über den Zugang zur Bildung von Personen gefolgt von erheblichen Auswirkungen auf deren Lebensführung begründen. So kann beispielsweise die Sicherung des notwendigen Lebensbedarfs, aber auch das Recht auf allgemeine und berufliche Bildung sowie Nichtdiskriminierung tangiert werden.

- **Beschäftigung, Personalmanagement und Zugang zur Selbstständigkeit:** KI-Systeme, die in den Bereichen Beschäftigung, Personalmanagement und Zugang zur Selbstständigkeit eingesetzt werden, können die Karriereaussichten, Lebensgrundlagen und Arbeitnehmerrechte der betroffenen Personen erheblich beeinflussen und historische Diskriminierungsmuster beispielsweise gegenüber Frauen, bestimmten Altersgruppen oder Personen mit einer bestimmten Herkunft fortsetzen. Der Einsatz solcher KI-Systeme zur Überwachung oder Bewertung der Leistung von Personen in Arbeitsvertragsverhältnissen können zudem die Grundrechte auf Datenschutz und Privatsphäre beeinträchtigen.

- **Staatliche Unterstützungsleistungen und -dienste**: Wenn KI-Systeme eingesetzt werden, um zu entscheiden, ob der Zugang zu und die Nutzung von grundlegenden privaten und öffentlichen Unterstützungsleistungen und -diensten wie Gesundheitsdienste, Sozialversicherungsleistungen oder soziale Dienste gewährt, verweigert, gekürzt, widerrufen oder zurückgefordert werden, können diese Systeme erhebliche Auswirkungen auf die Lebensgrundlagen der Betroffenen haben, die zur uneingeschränkten Teilhabe an der Gesellschaft auf diese Dienste angewiesen sind. Zudem können Grundrechte wie das Recht auf sozialen Schutz, Nichtdiskriminierung, Menschenwürde und wirksamen Rechtsbehelf verletzt werden.

- **Kreditpunktebewertung** oder Bewertung der Kreditwürdigkeit natürlicher Personen: Die Bewertung der Bonität und Kreditwürdigkeit von Personen ist ausschlaggebend für den Zugang zu Finanzmitteln oder wesentlichen Dienstleistungen wie Wohnraum, Strom und Telekommunikation. KI-Systeme, die für diese Zwecke eingesetzt werden, sind insoweit

als hochriskant einzustufen, als sie den Zugang zu solchen Mitteln und Dienstleistungen diskriminierend beeinflussen können.

KI-Systeme hingegen, die zur Betrugserkennung bei Finanzdienstleistungen oder zur Berechnung von Eigenkapitalanforderungen bei Banken und Versicherungen gemäß Unionsrecht eingesetzt werden, sind nicht als Hochrisiko-Systeme einzustufen.

- Bewertung von natürlichen Personen für den Abschluss von **Kranken- und Lebensversicherunge**n: Auch können KI-Systeme, die zur Risikobewertung und Preisgestaltung bei Kranken- und Lebensversicherungen für Einzelpersonen verwendet werden, erhebliche Auswirkungen auf deren Existenzgrundlage haben. Bei unsachgemäßer Gestaltung, Entwicklung und Anwendung können diese Systeme schwerwiegende Folgen für Leben und Gesundheit der Betroffenen sowie finanzielle Ausgrenzung und Diskriminierung nach sich ziehen.

- Bewertung und Klassifizierung von **Notrufen** sowie Systeme für die **Triage bei Notfallversorgung**: KI-Systeme, die bei der Bewertung und Einstufung von Notrufen sowie bei der Bereitstellung und Priorisierung von Not- und Rettungsdiensten sowie bei der Triage von Patienten in Notfallsituationen verwendet werden, treffen ebenfalls situationskritische Entscheidungen, die erhebliche Auswirkungen auf das Leben, die Gesundheit und das Eigentum von Menschen haben können.

- **Strafverfolgung**: KI-Systeme, die im Auftrag von oder durch Strafverfolgungsbehörden oder EU-Behörden zur Strafverfolgung eingesetzt werden, z. B. als Lügendetektoren oder zur Bewertung der Zuverlässigkeit von Beweisen im Strafverfahren, gelten als hochriskant, da solche Systeme die Gefahr der Diskriminierung oder anderer ungerechter Praktiken bergen, wenn sie vor Inverkehrbringung oder Inbetriebnahme nicht ordnungsgemäß konzipiert und getestet, korrekt trainiert oder die Anforderungen an Leistung, Genauigkeit und Robustheit nicht eingehalten werden. Dies kann die Ausübung grundlegender rechtlicher Garantien beeinträchtigen, weshalb zur Wahrung des öffentlichen Vertrauens und um wirksamen Rechtsschutz sowie Rechenschaftspflicht zu sichern, solche KI-Systeme als hochriskant einzustufen sind. In diesem Zusammenhang ist es besonders wichtig, die Auswirkungen auf die Verteidigungsrechte von Verdächtigen zu berücksichtigen, einschließlich der Herausforderungen, verständliche Informationen über die Funktionsweise solcher Systeme zu erhalten und deren Ergebnisse vor Gericht anzufechten, insbesondere für die betroffenen Personen selbst. Insoweit darf der Einsatz von KI-Systemen durch Strafverfolgungsbehörden nicht zu Ungleichheit oder Ausgrenzung führen.

Nicht als Hochrisiko-KI-Systeme gelten jedoch Systeme, die speziell für Verwaltungsverfahren in Steuer- und Zollbehörden sowie in Zentralstellen für Geldwäsche-Verdachtsanzeigen, die Verwaltungsaufgaben zur Analyse von Informationen gemäß dem Unionsrecht zur Bekämpfung der Geldwäsche durchführen, bestimmt sind.

- **Migration, Asyl und Grenzkontrolle**: KI-Systeme, die in den Bereichen Migration, Asyl und Grenzkontrolle eingesetzt werden, beeinflussen Menschen in besonders prekären Situationen stark, da ihre Rechte und Freiheiten von den Entscheidungen der Behörden abhängen können. Die Genauigkeit, Nichtdiskriminierung und Transparenz solcher KI-Systeme sind daher von entscheidender Bedeutung, um die Grundrechte der Betroffenen zu wahren, einschließlich ihrer Rechte auf Freizügigkeit, Nichtdiskriminierung, Datenschutz und internationalen Schutz. Deshalb sind auch KI-Systeme, die von zuständigen Behörden in diesen Bereichen eingesetzt werden, als hochriskant

einzustufen, sofern deren Einsatz gesetzlich zulässig ist. Dies gilt insbesondere für den Einsatz als Lügendetektoren, zur Risikobewertung von Einreisenden oder Asylbewerbern, zur Unterstützung bei der Prüfung von Asyl- und Visumanträgen sowie zur Identifizierung von Personen im Zusammenhang mit Migration und Grenzkontrollen, jedoch nicht für die Überprüfung von Reisedokumenten. KI-Systeme im Bereich Migration, Asyl und Grenzkontrolle, die unter diese Verordnung fallen, müssen den geltenden Verfahrensvorschriften der Europäischen Union entsprechen, wie sie in Verordnung (EG) Nr. 810/2009 des Europäischen Parlaments und des Rates sowie in Richtlinie 2013/32/EU und anderen relevanten Gesetzen festgelegt sind. Ihr Einsatz darf keinesfalls dazu dienen, internationale Verpflichtungen zu umgehen oder das Prinzip der Nichtzurückweisung zu verletzen. Sie sollen auch nicht legale und sichere Wege in die EU blockieren, einschließlich des Rechts auf internationalen Schutz.

- **Rechtspflege und demokratische Prozesse**: Auch sind bestimmte KI-Systeme, die für die Rechtspflege und demokratische Prozesse bestimmt sind, aufgrund ihrer potenziell bedeutenden Auswirkungen auf Demokratie, Rechtsstaatlichkeit, individuelle Freiheiten und das Recht auf einen fairen Rechtsprozess als hochriskant einzustufen.

Besonders KI-Systeme, die von Justizbehörden genutzt werden, um Sachverhalte zu ermitteln, Rechtsvorschriften auszulegen und auf konkrete Sachverhalte anzuwenden, sind als hochriskant zu betrachten, um möglichen Risiken wie Verzerrungen und Fehler zu begegnen. KI-Systeme, die in diesem Sinne zur Durchführung alternativer Streitbeilegungen genutzt werden und dabei Rechtswirkungen für die Parteien entfalten, sind ebenfalls als hochriskant einzustufen.

Zwar können solche KI-Systeme die richterliche Entscheidung unterstützen, aber nicht ersetzen; die letzte Entscheidung muss immer in menschlicher Hand bleiben. Jedoch sind KI-Systeme, die rein administrative Aufgaben wie die Anonymisierung von Gerichtsunterlagen oder die Verwaltung von Kommunikation übernehmen und die tatsächliche Rechtspflege nicht direkt beeinflussen, nicht als hochriskant einzustufen.

Zur Sicherung des Wahlrechts gemäß der Charta der Grundrechte und zur Vermeidung nachteiliger Auswirkungen auf Demokratie und Rechtsstaatlichkeit gelten auch KI-Systeme, die das Ergebnis von Wahlen oder Referenden beeinflussen können, als hochriskant, außer wenn sie nicht unmittelbar auf natürliche Personen einwirken, wie bei der Organisation und Optimierung politischer Kampagnen.

Erwägungsgrund 63 stellt fest, dass die **Einstufung eines KI-Systems als Hochrisiko-KI-System** gemäß der KI-VO **nicht automatisch** bedeutet, dass seine Verwendung gemäß anderen EU-Rechtsvorschriften oder nationalen Gesetzen, die mit dem EU-Recht vereinbar sind, rechtmäßig ist. Dies betrifft insbesondere den **Schutz personenbezogener Daten** oder die **Verwendung von Lügendetektoren** und ähnlichen Instrumenten zur Ermittlung des emotionalen Zustands von Personen. Die Verwendung eines solchen KI-Systems muss weiterhin den spezifischen Anforderungen der Charta der Grundrechte der Europäischen Union, des anwendbaren EU-Sekundärrechts und nationalen Gesetzen entsprechen.

(**Mit anderen Worten:** Die KI-Verordnung selbst stellt **keine eigenständige Rechtsgrundlage** für die Verarbeitung personenbezogener Daten dar, es sei denn, dies wird ausdrücklich in der KI-Verordnung angegeben.)

Beispiele Hochrisiko-KI-Systeme (Anhang III)

Gem. Art. 6 Abs. 5 KI-VO stellt die Kommission nach Konsultation des Europäischen Ausschusses für künstliche Intelligenz spätestens 18 Monate nach Inkrafttreten der Verordnung **Leitlinien zur praktischen Umsetzung des Artikel 6** gemäß **Art. 96 KI-VO** und eine **umfassende Liste praktischer Beispiele** für Anwendungsfälle für KI-Systeme, die hochriskant oder nicht hochriskant sind, bereit.

Beispiele für die in Anhang III genannten Anwendungsfälle mit **hohem Risiko**:

- Bestimmte **kritische Infrastrukturen**, z. B. in Bereichen wie Straßenverkehr und Wasser-, Gas-, Wärme- und Stromversorgung;
- allgemeine und berufliche **Bildung**, z. B. Bewertung von Lernergebnissen, Steuerung des Lernprozesses und Überwachung von Prüfungen;
- **Beschäftigung**, Personalmanagement und Zugang zu selbstständiger Erwerbstätigkeit, z. B. Veröffentlichung gezielter Stellenanzeigen, Analyse und Filterung von Bewerbungen sowie Auswahl und Bewertung von Bewerbern;
- Zugang zu wichtigen privaten und öffentlichen **Dienstleistungen** und zu **Sozialleistungen** (z. B. Gesundheitsversorgung), Bewertung der Kreditwürdigkeit natürlicher Personen sowie Risikobewertung und Preisfestsetzung im Zusammenhang mit Lebens- und Krankenversicherungen;
- bestimmte Systeme, die in den Bereichen **Strafverfolgung**, Grenzkontrolle, Justizverwaltung und demokratische Prozesse eingesetzt werden;
- Bewertung und Klassifizierung von **Notrufen**;
- Systeme zur **biometrischen Identifizierung**, Kategorisierung und Emotionserkennung (außerhalb der verbotenen Kategorien).

Hingegen **nicht** enthalten sind die **Empfehlungssysteme** sehr großer Online-Plattformen, weil sie bereits von anderen Rechtsvorschriften (Gesetz über digitale Dienste, Gesetz über digitale Märkte) erfasst werden.

Zwischenergebnis

Ist das KI-System nach Art. 6 Abs. 2 i.V.m. Anhang III als hochriskant einzustufen, geht die Prüfung mit **Schritt 2.4** weiter.

Ist das KI-System dagegen nicht nach Art. 6 Abs. 2 i.V.m. Anhang III als hochriskant einzustufen, geht die Prüfung mit **Schritt 2.5** weiter.

2.4 Ist eine Ausnahme i.S.v. Art. 6 Abs. 3 KI-VO gegeben?

Dr. Benedict Rohrßen (Taylor Wessing Partnerschaftsgesellschaft mbB)

Art. 6 Abs. 3 KI-VO regelt, unter welchen Voraussetzungen ein KI-System, das in einem der Bereiche in Anhang III eingesetzt wird, **ausnahmsweise nicht als hochriskant** einzustufen ist:

> (3) Abweichend von Absatz 2 gilt ein KI-System nicht als hochriskant, wenn es kein erhebliches Risiko der Beeinträchtigung in Bezug auf die Gesundheit, Sicherheit oder Grundrechte natürlicher Personen birgt, indem es unter anderem nicht das Ergebnis der Entscheidungsfindung wesentlich beeinflusst. Dies ist der Fall, wenn eine oder mehrere der folgenden Bedingungen erfüllt sind:
>
> a) das KI-System ist dazu bestimmt, eine eng gefasste Verfahrensaufgabe durchzuführen;
>
> b) das KI-System ist dazu bestimmt, das Ergebnis einer zuvor abgeschlossenen menschlichen Tätigkeit zu verbessern;
>
> c) das KI-System ist dazu bestimmt, Entscheidungsmuster oder Abweichungen von früheren Entscheidungsmustern zu erkennen, und ist nicht dazu gedacht, die zuvor abgeschlossene menschliche Bewertung, ohne eine angemessene menschliche Überprüfung zu ersetzen oder zu beeinflussen; oder
>
> d) das KI-System ist dazu bestimmt, eine vorbereitende Aufgabe für eine Bewertung durchzuführen, die für die Zwecke der in Anhang III aufgeführten Anwendungsfälle relevant ist.
>
> Ungeachtet des Unterabsatzes 1 gilt ein in Anhang III aufgeführtes KI-System immer dann als hochriskant, wenn es ein Profiling natürlicher Personen vornimmt.

Relevante(r) Artikel:

Art. 6 Abs. 3, 4, 5, 6

Relevante(r) ErwG:

46, 47, 48

Konkretisierungsbedürftig:

Ja, (Leitlinien der Kommission zur praktischen Umsetzung des Art. 6 gem. Art. 96 und umfassende Liste praktischer Beispiele für Anwendungsfälle für KI-Systeme, die hochriskant oder nicht hochriskant sind (Art. 6 Abs. 5); delegierte Rechtsakte der Kommission gem. Art. 97 KI-VO zur Änderung des Art. 6 Abs. 3 Unterabs. 2 (Art. 6 Abs. 6 und 7 KI-VO) 6 und 7 KI-VO))

Gem. **Art. 6 Abs. 6 KI-VO** erlässt die Kommission gemäß Artikel 97 delegierte Rechtsakte zur Änderung der in Absatz 3 Unterabsatz 1 des vorliegenden Artikels festgelegten Bedingungen. Die Kommission kann gemäß Artikel 97 delegierte Rechtsakte, mit denen neue Bedingungen zu den in Absatz 3 Unterabsatz 1 genannten Bedingungen hinzugefügt oder diese geändert werden, nur dann erlassen, wenn konkrete und zuverlässige Beweise für das Vorhandensein von KI-Systemen vorliegen, die in den Anwendungsbereich von Anhang III fallen, jedoch kein erhebliches Risiko der Beeinträchtigung in Bezug auf die Gesundheit, Sicherheit oder Grundrechte natürlicher Personen bergen.

Gem. Art. 6 Abs. 7 KI-VO erlässt die Kommission gemäß Art. 97 **delegierte Rechtsakte**, um eine der in Art. 6 Abs. 3 festgelegten **Bedingungen zu streichen**, wenn konkrete und zuverlässige Beweise dafür vorliegen, dass dies für die Aufrechterhaltung des Schutzniveaus in Bezug auf Gesundheit, Sicherheit und Grundrechte in der Union erforderlich ist. Eine Änderung der in Absatz 3 Unterabsatz 1 festgelegten Bedingungen darf das allgemeine Schutzniveau in Bezug auf Gesundheit, Sicherheit und Grundrechte in der Union nicht senken. Beim Erlass der delegierten Rechtsakte stellt die Kommission die Kohärenz mit den gemäß Artikel 7 Absatz 1 erlassenen delegierten Rechtsakten sicher und trägt den Marktentwicklungen und den technologischen Entwicklungen Rechnung.

Ein Anbieter eines unter **Anhang III KI-VO** aufgeführten Hochrisiko-KI-Systems, das eine oder mehrere der in **Art. 6 Abs. 3 KI-VO genannten Bedingungen** erfüllt und damit **ausnahmsweise kein hohes Risiko** darstellt, hat vor Inverkehrbringung und Inbetriebnahme des KI-Systems eine **Dokumentation der Bewertung** zu erstellen und diese den zuständigen nationalen Behörden auf Anfrage zur Verfügung zu stellen. Zudem hat er das KI-System in der EU-Datenbank zu **registrieren**.

Für die **praktische Umsetzung** der Bedingungen des Art. 6 Abs. 3 KI-VO, gemäß denen die im Anhang III aufgeführten Hochrisiko-KI-Systeme kein hohes Risiko darstellen, stellt die **Kommission Leitlinien** zur praktischen Umsetzung und eine Liste praktischer Fallbeispiele für die Anwendung von KI-Systemen bereit.

Zwischenergebnis

Greift die Ausnahme nach Art. 6 Abs. 3, gelten die Anforderungen an Hochrisiko-KI-Systeme gem. der Art. 8 – 15 KI-VO. Wirtschaftsakteure haben je nach ihrer Rolle ihren spezifischen Pflichten aus den Art. 16 – 27 KI-VO gerecht zu werden und die Prüfung geht mit **Schritt 2.5** weiter.

Ist das KI-System dagegen nach Art. 6 Abs. 2 i.V.m. Anhang III als hochriskant einzustufen, geht die Prüfung mit **Schritt 4** weiter.

2.5. Geht vom KI-System ein geringes Risiko aus?

Dr. Benedict Rohrßen (Taylor Wessing Partnerschaftsgesellschaft mbB)

Wenn das KI-System weder unter Art. 6 Abs. 1 i.V.m. Anhang I noch unter Art. 6 Abs. 2 i.V.m. Anhang III subsumiert werden kann, und damit nicht den strengen Anforderungen an Hochrisiko-KI-Systeme unterliegt, ist als Nächstes zu prüfen, ob vom KI-System ein geringes/begrenztes oder nur ein minimales/kein Risiko ausgeht. KI-Systeme mit geringem Risiko unterliegen etwa den Transparenzpflichten des Art. 50 KI-VO, während KI-Systeme, von denen bloß ein minimales Risiko ausgeht, nicht von der Verordnung erfasst sind.

KI-Systeme mit geringem/begrenztem Risiko sind solche, die mit natürlichen Personen interagieren und eine Manipulationsgefahr bergen, wie beispielsweise Chatbots oder Empfehlungssysteme. Diese Systeme unterliegen besonderen Transparenzpflichten gem. Art. 50 KI-VO. Nutzer müssen darüber informiert werden, dass sie mit einem KI-System interagieren, es sei denn, dies ist aus dem Nutzungskontext heraus offensichtlich. Diese Transparenzpflicht gilt auch für Systeme zur Emotionserkennung, biometrischen Kategorisierung und für sogenannte Deepfakes, bei denen klar erkennbar sein muss, dass es sich um künstlich erzeugte oder manipulierte Inhalte handelt. Eine Ausnahme gilt für Strafverfolgungszwecke. Erfüllt ein solches System mit geringem/begrenztem Risiko die Kriterien für ein Hochrisiko-KI-System, muss dieses zusätzlich zu den Transparenzverpflichtungen auch die Anforderungen an Hochrisiko-KI-Systeme erfüllen, Art. 50 Abs. 6 KI-VO.

Im Gegensatz dazu stehen **KI-Systeme mit minimalem/keinem Risiko**, die vom Anwendungsbereich der Verordnung ausgenommen sind. Zu dieser Kategorie zählen weitverbreitete administrative oder interne Anwendungen wie Spamfilter, KI-gestützte Videospiele oder Bestandsverwaltungssysteme. Diese Systeme erfordern keine speziellen regulatorischen Maßnahmen, da sie als wenig risikobehaftet gelten. Insoweit können diese unter Einhaltung des allgemein geltenden Rechts entwickelt und verwendet werden, ohne dass es zusätzlicher spezifischer Verpflichtungen bedarf. Die Anbieter solcher Systeme können sich jedoch freiwillig dazu verpflichten, Verhaltenskodizes einzuhalten, um die Anforderungen an vertrauenswürdige KI-Systeme zu erfüllen, auch wenn dies nicht gesetzlich vorgeschrieben ist, vgl. Art. 95 KI-VO.

Durch den risikobasierten Ansatz stellt die KI-Verordnung sicher, dass Systeme mit einem höheren Risiko strenger reguliert werden, während Systeme mit geringem oder minimalem Risiko entweder speziellen Transparenzpflichten unterliegen oder gänzlich von der Verordnung ausgenommen sind. Dies soll ein Gleichgewicht zwischen der Förderung von Innovation und dem Schutz der Nutzer schaffen

Art. 50 KI-VO legt Transparenzpflichten für Anbieter und Betreiber von KI-Systemen mit geringem Risiko fest. Nutzer müssen darüber informiert werden, dass sie mit einer KI interagieren, um sich schützen zu können.

Dementsprechend müssen KI-Systeme, die mit Menschen interagieren, gem. Art. 50 Abs. 1 so gestaltet sein, dass eine durchschnittlich informierte und aufmerksame Person erkennt, dass sie mit einer KI zu tun hat. Dies betrifft z. B. Chatbots, die in Kundenservices eingesetzt werden. Auch künstlich erzeugte Deepfakes, Audios, Videos, Bilder oder Texte müssen als von KI generiert gekennzeichnet sein.

Relevante(r) Artikel:

Art. 50, Art. 95

Relevante(r) ErwG:

132, 134, 165

Konkretisierungsbedürftig:

Ja (Praxisleitfäden durch das Büro für Künstliche Intelligenz zur Erleichterung der wirksamen Umsetzung der Transparenzpflichten gem. Art. 50 Abs. 7 KI-VO; Leitlinien der Kommission für die praktische Umsetzung der Transparenzpflichten gem. Art. 96 Abs. 1 d) KI-VO, ErwG 135; Verhaltenskodizes für die freiwillige Anwendung der in Kapitel III Abschnitt 2 genannten Anforderungen (Anforderungen an Hochrisiko-KI-Systeme), Art 95 KI-VO)1 d) KI-VO, ErwG 135; Verhaltenskodizes für die freiwillige Anwendung der in Kapitel III Abschnitt 2 genannten Anforderungen (Anforderungen an Hochrisiko-KI-Systeme), Art 95 KI-VO)

Diese Transparenzpflichten gelten laut Art. 50 Abs. 4 KI-VO auch für Betreiber, mit Ausnahmen für künstlerische und satirische Zwecke, wo jedoch ebenfalls kenntlich gemacht werden muss, dass eine Manipulation vorliegt. Betreiber müssen zudem informieren, wenn Emotionserkennungssysteme oder Systeme zur biometrischen Kategorisierung eingesetzt werden, Art. 50 Abs. 3 KI-VO.

Zwischenergebnis

Ist das KI-System als gering riskant einzustufen, sind die Transparenzpflichten gem. Art. 50 KI-VO einzuhalten und die Prüfung mit **Schritt 5** weiter.

Ist das KI-System dagegen nur minimal riskant einzustufen, ist die Prüfung hier beendet. Dann findet die KI-VO keine Anwendung.

5 Umgang mit verbotenen KI-Systemen

Stephan Kress (Morrison & Foerster LLP), Dilan Mienert (GÖRG Partnerschaft von Rechtsanwälten mbB)

Schritt 3: Einstellen verbotener Praktiken

Die Verbote gelten **ab 6 Monaten nach Inkrafttreten der KI-VO**. Sollte ein Anbieter feststellen, dass er mit Inverkehrbringen, Inbetriebnahme oder Einsatz seines KI-Systems einen Verbotstatbestand erfüllt, hat er die Praktik somit bis spätestens zum **2. Februar 2025** einzustellen.

Das Einstellen kann dabei auf verschiedene Weise erfolgen:

- Das betroffene KI-System wird nicht weiterverwendet, sondern wird gelöscht bzw. vernichtet. Hier kann es ggf. zu vertraglichen Komplikationen kommen, wenn das KI-System an Dritte lizenziert ist, jedoch sollten sich derartige Konstellationen über die Prinzipien der Unmöglichkeit bzw. der (Teil-)Nichtigkeit von gegen Gesetz verstoßenden Rechtsgeschäften lösen lassen.
- Das betroffene KI-System wird dergestalt verändert, dass der Tatbestand des Verbots nicht länger erfüllt ist.

6 Compliance-Anforderungen für Hochrisiko-KI-Systeme

Schritt 4.1: Welche Pflichten muss ich als Anbieter eines Hochrisiko-KI-Systems erfüllen?

Nachdem festgestellt worden ist, dass es sich um ein Hochrisiko-KI-System i.S.d. KI-VO handelt, ist im nächsten Schritt zu prüfen, welche Anforderungen vom Anbieter zu erfüllen sind.

Die KI-VO sieht im Abschnitt 2, also in den **Art. 8 bis 15 KI-VO**, eine Reihe von Pflichten vor, die vom Anbieter zu erfüllen sind. Nach Art. 8 Abs. 1 müssen Hochrisiko-KI-Systeme alle in den Art. 9–15 KI-VO genannten Anforderungen erfüllen.

Anbieter haben neben den **spezifischen Anforderungen an Hochrisiko-KI-Systeme** nach den Art. 8 bis 15 KI-VO (Abschnitt 2 *Anforderungen an Hochrisiko-KI-Systeme* des Kapitels III Hochrisiko-Systeme der KI-VO) auch die **einschlägigen vertikal-sektorspezifischen Vorschriften der EU** sicherzustellen, die sog. Harmonisierungsrechtsvorschriften. Die Notwendigkeit, die Anforderungen der KI-VO mit den Harmonisierungsrechtsvorschriften zu verschränken, folgt aus der horizontalen Ausrichtung der KI-VO: Sie gilt für KI grundsätzlich, unabhängig davon, ob und in welche Produkte sie eingebettet ist. **Zugleich gelten für in der EU vertriebene Produkte je nach Produktkategorie** sog. Harmonisierungsrechtsvorschriften. Harmonisierungsrechtsvorschriften stellen grundlegende Sicherheits- und Gesundheitsanforderungen auf. Gesetzlich definiert sind **Harmonisierungsrechtsvorschriften** als „*Rechtsvorschriften der Gemeinschaft zur Harmonisierung der Bedingungen für die Vermarktung von Produkten*“ (Art. 2 Nr. 21 Verordnung (EU) Nr. 765/2008). Konkret listet die Marktüberwachungsverordnung Nr. 2019/1020 in Anhang I die Harmonisierungsrechtsvorschriften. Harmonisierungsrechtsvorschriften gelten **z. B.** für Maschinen und Medizinprodukte.

Was gilt es zu tun?

Anbieter von Produkten, die ein Hochrisiko-KI-System enthalten, müssen sicherstellen, dass sie

1. die **Anforderungen,** die an **Hochrisiko-KI-Systeme** gestellt werden, gemäß der KI-VO (Art. 9 ff. KI-VO) und
2. die jeweils **produktkategoriespezifischen**, in Anhang I Abschnitt A aufgelisteten **Harmonisierungsrechtsvorschriften** einhalten.

Schritt 4.1.1: Ist Art. 8 Abs. 2 einschlägig?

Prof. Dr. Heinz-Uwe Dettling (Ernst & Young GmbH Wirtschaftsprüfungsgesellschaft), Dr. Benedikt Rohrßen (Taylor Wessing Partnerschaftsgesellschaft mbB)

2) Enthält ein Produkt ein KI-System, für das die Anforderungen dieser Verordnung und die Anforderungen der in Anhang I Abschnitt A aufgeführten Harmonisierungsrechtsvorschriften der Union gelten, so sind die Anbieter dafür verantwortlich, sicherzustellen, dass ihr Produkt alle geltenden Anforderungen der geltenden Harmonisierungsrechtsvorschriften der Union vollständig erfüllt. Bei der Gewährleistung der Erfüllung der in diesem Abschnitt festgelegten Anforderungen durch die in Absatz 1 genannten Hochrisiko-KI-Systeme und **im Hinblick auf die Gewährleistung der Kohärenz**, der Vermeidung von Doppelarbeit und der Minimierung zusätzlicher Belastungen haben die Anbieter die Wahl, die erforderlichen Test- und Berichterstattungsverfahren, Informationen und Dokumentationen, die sie im Zusammenhang mit ihrem Produkt bereitstellen, gegebenenfalls in Dokumentationen und Verfahren zu integrieren, die bereits bestehen und gemäß den in Anhang I Abschnitt A aufgeführten Harmonisierungsrechtsvorschriften der Union vorgeschrieben sind.

Relevante(r) Artikel:

Art. 8 Abs. 2, Art. 11 Abs. 2, Art. 16, Art. 25 Abs. 3, Art 43 Abs. 3 und 4, Art. 46 Abs. 7, Art. 72 Abs. 4, Art 74 Abs. 3 und 4, Anhang I Abschnitt A.

Relevante(r) ErwG:

64

Konkretisierungsbedürftig:

Leitlinien der Kommission für die praktische Umsetzung der KI-VO, insbesondere detaillierte Informationen über das Verhältnis der KI-VO zu den in Anhang I aufgeführten Harmonisierungsrechtsvorschriften der Union sowie zu anderen einschlägigen Rechtsvorschriften der Union, insbesondere auch in Bezug auf deren kohärente Durchsetzung, Art. 96 Abs. 1 e) KI-VO1 e) KI-VO

Anbieter haben neben den **spezifischen Anforderungen an Hochrisiko-KI-Systeme** nach den Art. 8 bis 15 KI-VO (Abschnitt 2 *Anforderungen an Hochrisiko-KI-Systeme* des Kapitels III Hochrisiko-Systeme der KI-VO) auch die **einschlägigen vertikal-sektorspezifischen Vorschriften der EU** sicherzustellen, die sog. Harmonisierungsrechtsvorschriften. Die Notwendigkeit, die Anforderungen der KI-VO mit den Harmonisierungsrechtsvorschriften zu verschränken, folgt aus der horizontalen Ausrichtung der KI-Verordnung: Sie gilt für KI grundsätzlich, unabhängig davon, ob und in welche Produkte sie eingebettet ist. **Zugleich gelten für in der EU vertriebene Produkte je nach Produktkategorie** sog. Harmonisierungsrechtsvorschriften. Harmonisierungsrechtsvorschriften stellen grundlegende Sicherheits- und Gesundheitsanforderungen auf. Gesetzlich definiert sind **Harmonisierungsrechtsvorschriften** als *„Rechtsvorschriften der Gemeinschaft zur Harmonisierung der Bedingungen für die Vermarktung von Produkten“* (Art. 2 Nr. 21 Verordnung (EU) Nr. 765/2008). Konkret listet die Marktüberwachungsverordnung Nr. 2019/1020 in Anhang I die Harmonisierungsrechtsvorschriften. Harmonisierungsrechtsvorschriften gelten **z. B.** für Maschinen und Medizinprodukte.

Art. 8 Abs. 2 der KI-VO gilt allerdings **nur für einen Teil der Harmonisierungsrechtsvorschriften**, nämlich nur für diejenigen Produkte, die aufgrund der 12 in Anhang I Abschnitt A der KI-VO aufgelisteten Rechtsakte reguliert sind. Hierzu gehören ebenfalls beispielsweise die Richtlinie 2006/42/EG über Maschinen, die mit Wirkung ab dem 20. Januar 2027 durch die Verordnung (EU) 2023/1230 über Maschinen ersetzt wird, und die Verordnung (EU) 2017/745 über Medizinprodukte.

Art. 8 Abs. 2 der KI-VO steht darüber hinaus insbesondere in **Zusammenhang** mit Art. 11 Abs. 2 der KI-VO (technische Dokumentation), Art. 16 und 25 Abs. 3 der KI-VO (Verantwortlichkeiten entlang der KI-Wertschöpfungskette), Art. 43 Abs. 3 und 4 sowie Art. 46 Abs. 7 (Konformitätsbewertung und Ausnahmen von der Konformitätsbewertung), Art. 72 Abs. 4 (Produktbeobachtung) und Art. 74 Abs. 3 und 4 (Marktüberwachung).

Was gilt es zu tun?

Anbieter von Produkten, die ein Hochrisiko-KI-System enthalten, müssen sicherstellen, dass sie

1. die Anforderungen, die an Hochrisiko-KI-Systeme gestellt werden, gemäß der KI-Verordnung (Art. 9 ff. KI-VO) und
2. die jeweils produktkategoriespezifischen, in Anhang I Abschnitt A aufgelisteten Harmonisierungsrechtsvorschriften einhalten.

Wie sind diese Anforderungen einzuhalten?

Art. 8 Abs. 2 KI-VO sieht vor, die **Anforderungen** der KI-Verordnung in die bestehenden sektorspezifischen Sicherheitsvorschriften zu **integrieren** und den mit der Produktsicherheit verbundenen Aufwand gering zu halten, Doppelarbeit zu vermeiden. Insoweit können Anbieter nach Art. 8 Abs. 2 S. 2 KI-VO bei Erfüllung dieser „Doppelanforderungen“ die nach der KI-VO notwendigen Test- und Berichtserstattungsverfahren, Informations- und Dokumentationsanforderungen in bestehende Konformitätsbewertungsverfahren integrieren. Mit anderen Worten: sie können die Dokumentationen und Verfahren, die sie bereits gemäß den EU-Harmonisierungsrechtsvorschriften verwenden, auch zur Erfüllung der Anforderungen der KI-Verordnung heranziehen. Anbieter sollten daher prüfen und festlegen, wie sie die **„zwei Fliegen“ der produktrechtlichen und KI-rechtlichen Regularien jeweils „mit einer Klappe“ schlagen** können. Das ist durch die KI-VO ausdrücklich erwünscht.

Erwägungsgrund 64 erläutert hierzu, dass mehr als ein Rechtsakt der Harmonisierungsrechtsvorschriften der Union auf ein Produkt anwendbar sein können, da die Bereitstellung oder Inbetriebnahme nur erfolgen kann, wenn das Produkt allen geltenden Harmonisierungsrechtsvorschriften der Union entspricht. Dies gelte als allgemeine Regel auf der Grundlage des neuen Rechtsrahmens, wie in der Bekanntmachung der Kommission „Leitfaden für die Umsetzung der Produktvorschriften der EU 2022 (Blue Guide)“ dargelegt. Die Gefahren von KI-Systemen, die unter die Anforderungen der KI-VO fallen, decken andere Aspekte ab als die bestehenden Harmonisierungsrechtsvorschriften der Union, weshalb die Anforderungen der KI-VO das bestehende Regelwerk der Harmonisierungsrechtsvorschriften der Union ergänzen. So bergen etwa Maschinen oder Medizinprodukte mit einer KI-Komponente möglicherweise Risiken, die von den grundlegenden Gesundheits- und Sicherheitsanforderungen der einschlägigen harmonisierten Rechtsvorschriften der Union nicht erfasst werden, da diese sektoralen Rechtsvorschriften keine spezifischen KI-Risiken behandeln. Dies erfordert die gleichzeitige

und ergänzende Anwendung mehrerer Rechtsakte. Um Kohärenz zu gewährleisten und unnötigen Verwaltungsaufwand sowie unnötige Kosten zu vermeiden, sollten die Anbieter eines Produkts, das ein oder mehrere Hochrisiko-KI-Systeme enthält, für die die Anforderungen dieser Verordnung und der in einem Anhang dieser Verordnung aufgeführten und auf dem neuen Rechtsrahmen beruhenden Harmonisierungsvorschriften der Union gelten, in Bezug auf betriebliche Entscheidungen darüber flexibel sein, wie die Konformität eines Produkts, das ein oder mehrere Hochrisiko-KI-Systeme enthält, bestmöglich mit allen geltenden Anforderungen dieser harmonisierten Rechtsvorschriften der Union sichergestellt werden kann. Diese Flexibilität könnte beispielsweise bedeuten, dass der Anbieter beschließt, einen Teil der gemäß der KI-VO erforderlichen Test- und Berichterstattungsverfahren, Informationen und Unterlagen in bereits bestehende Dokumentationen und Verfahren zu integrieren, die nach den auf dem neuen Rechtsrahmen beruhenden und in einem Anhang dieser Verordnung aufgeführten geltenden Harmonisierungsrechtsvorschriften der Union erforderlich sind. Dies sollte in keiner Weise die Verpflichtung des Anbieters untergraben, alle geltenden Anforderungen zu erfüllen.

Dazu bestimmt etwa Art. 11 Abs. 2 der KI-VO, dass **eine einzige technische Dokumentation** erstellt wird, die alle in Art. 11 Abs. 1 der KI-VO genannten Informationen sowie die nach diesen Rechtsakten erforderlichen Informationen enthält, **wenn** ein Hochrisiko-KI-System, das mit einem Produkt verbunden ist, das unter die in Anhang I Abschnitt A auf geführten Harmonisierungsrechtsvorschriften der Union fällt, in Verkehr gebracht oder in Betrieb genommen wird.

Art. 25 Abs. 3 KI-VO bestimmt, dass der Produkthersteller im Falle von Hochrisiko-KI-Systemen, bei denen es sich um **Sicherheitsbauteile von Produkten** handelt, die unter die in Anhang I Abschnitt A aufgeführten Harmonisierungsrechtsvorschriften der Union fallen, **als Anbieter des Hochrisiko-KI-Systems** gilt und in den beiden nachfolgenden Fällen den **Pflichten nach Art. 16 KI-VO** unterliegt:

a) Das Hochrisiko-KI-System wird zusammen mit dem Produkt unter dem Namen oder der Handelsmarke des Produktherstellers in Verkehr gebracht;

b) das Hochrisiko-KI-System wird unter dem Namen oder der Handelsmarke des Produktherstellers in Betrieb genommen, nachdem das Produkt in Verkehr gebracht wurde.

Art. 43 Abs. 3 KI-VO bestimmt zum **Konformitätsbewertungsverfahren**, dass der Anbieter bei den Hochrisiko-KI-Systemen, die unter die in Anhang I Abschnitt A aufgeführten Harmonisierungsrechtsakte der Union fallen, die einschlägigen Konformitätsbewertungsverfahren, die nach diesen Rechtsakten erforderlich sind, befolgt. Die in Abschnitt 2 dieses Kapitels III der KI-VO festgelegten Anforderungen gelten für diese Hochrisiko-KI-Systeme und werden in diese Bewertung einbezogen. Anhang VII Nummern 4.3, 4.4 und 4.5 sowie Nummer 4.6 Absatz 5 der KI-VO finden ebenfalls Anwendung. Für die Zwecke dieser Bewertung sind die notifizierten Stellen, die gemäß diesen Rechtsakten notifiziert wurden, berechtigt, die Konformität der Hochrisiko-KI-Systeme mit den in Abschnitt 2 des Kapitels III der KI-VO festgelegten Anforderungen zu kontrollieren, sofern im Rahmen des gemäß diesen Rechtsakten durchgeführten Notifizierungsverfahrens geprüft wurde, dass diese notifizierten Stellen die in Artikel 31 Absätze 4, 5, 10 und 11 der KI-VO festgelegten Anforderungen erfüllen. Wenn ein in Anhang I Abschnitt A aufgeführter Rechtsakt es dem Hersteller des Produkts ermöglicht, auf eine Konformitätsbewertung durch Dritte zu verzichten, sofern dieser Hersteller alle

harmonisierten Normen, die alle einschlägigen Anforderungen abdecken, angewandt hat, so darf dieser Hersteller nur dann von dieser Möglichkeit Gebrauch machen, wenn er auch harmonisierte Normen oder gegebenenfalls gemeinsame Spezifikationen gemäß Art. 41 der KI-VO, die alle in Abschnitt 2 des Kapitels III der KI-VO festgelegten Anforderungen abdecken, angewandt hat.

Art. 43 Abs. 4 der KI-VO bestimmt generell, dass Hochrisiko-KI-Systeme, die bereits Gegenstand eines Konformitätsbewertungsverfahren gewesen sind, im Falle einer **wesentlichen Änderung** einem **neuen Konformitätsbewertungsverfahren** unterzogen werden, unabhängig davon, ob das geänderte System noch weiter in Verkehr gebracht oder vom derzeitigen Betreiber weitergenutzt werden soll. Bei Hochrisiko-KI-Systemen, die nach dem Inverkehrbringen oder der Inbetriebnahme weiterhin dazulernen, gelten Änderungen des Hochrisiko-KI-Systems und seiner Leistung, die vom Anbieter zum Zeitpunkt der ursprünglichen Konformitätsbewertung vorab festgelegt wurden und in den Informationen der technischen Dokumentation gemäß Anhang IV Nummer 2 Buchstabe f der KI-VO enthalten sind, nicht als wesentliche Veränderung;

Art. 46 Abs. 7 der KI-VO regelt zu Ausnahmen vom Konformitätsbewertungsverfahren, dass für Hochrisiko-KI-Systeme im Zusammenhang mit Produkten, die unter die in Anhang I Abschnitt A aufgeführten Harmonisierungsrechtsvorschriften der Union fallen, nur die in diesen Harmonisierungsrechtsvorschriften der Union festgelegten Ausnahmen von den Konformitätsbewertungsverfahren gelten.

Art. 72 Abs. 4 der KI-VO bestimmt zur **Produktbeobachtung**, dass die Anbieter bei Hochrisiko-KI-Systemen, die unter die in Anhang I Abschnitt A der KI-VO aufgeführten Harmonisierungsrechtsvorschriften der Union fallen, und für die auf der Grundlage dieser Rechtsvorschriften bereits ein System zur Beobachtung nach dem Inverkehrbringen sowie ein entsprechender Plan festgelegt wurden, zur Gewährleistung der Kohärenz, zur Vermeidung von Doppelarbeit und zur Minimierung zusätzlicher Belastungen die Möglichkeit haben, unter Verwendung des Musters nach Art. 72 Abs. 3, gegebenenfalls die in den Art. 72 Abs. 1, 2 und 3 genannten erforderlichen Elemente in die im Rahmen dieser Vorschriften schon vorhandenen Systeme und Pläne **zu integrieren**, sofern ein gleichwertiges Schutzniveau erreicht wird

Art. 74 Abs. 3 der KI-VO bestimmt, dass als **Marktüberwachungsbehörde** für die Zwecke der KI-VO bei Hochrisiko-KI-Systemen und damit in Zusammenhang stehenden Produkten, auf die die in Anhang I Abschnitt A aufgeführten Harmonisierungsrechtsvorschriften der Union Anwendung finden, die in jenen Rechtsakten für die Marktüberwachung benannte Behörde gilt. Jedoch können die Mitgliedstaaten hiervon abweichend und unter geeigneten Umständen eine andere einschlägige Behörde benennen, die die Funktion der Marktüberwachungsbehörde übernimmt, sofern sie die Koordinierung mit den einschlägigen sektorspezifischen Marktüberwachungsbehörden, die für die Durchsetzung der in Anhang I aufgeführten Harmonisierungsrechtsvorschriften der Union zuständig sind, sicherstellen.

Gemäß Art. 74 Abs. 4 der KI-VO gelten die behördlichen Überwachungs-Verfahren gemäß den Artikeln 79 bis 83 der KI-VO **nicht** für KI-Systeme, die im Zusammenhang mit Produkten stehen, auf die die in Anhang I Abschnitt A aufgeführten Harmonisierungsrechtsvorschriften der Union Anwendung finden, wenn in diesen Rechtsakten bereits Verfahren, die ein gleichwertiges Schutzniveau sicherstellen und dasselbe Ziel haben, vorgesehen sind. In diesen Fällen kommen stattdessen die einschlägigen sektorspezifischen Verfahren zur Anwendung.

Beispielhaft bedeutet dies für Maschinen, die mit einem integrierten (eingebetteten) Hochrisiko-KI-System ausgestattet sind: Anbieter müssen sicherstellen, dass die Maschinen allen relevanten Sicherheitsanforderungen entsprechen, einschließlich der spezifischen Risiken, die durch die KI-Komponente entstehen, d. h. es gelten sowohl die Anforderungen der KI-Verordnung als auch die der Maschinenrichtlinie (künftig: Maschinenverordnung) als sektorspezifischer Harmonisierungsrechtsvorschriften. Zur Vermeidung von Doppelarbeit und Minimierung des Konformitätsaufwands sowie zwecks effizienter Einhaltung beider Rechtsrahmen können die Anbieter die Konformitätsbewertungsverfahren, die nach der Maschinenverordnung vorgeschrieben sind, auch zur Erfüllung dieser Anforderung nach der KI-Verordnung nutzen. Bei wesentlichen Änderungen an den KI-Systemen, die über die ursprüngliche technische Dokumentation und die Ex-ante-Konformitätsbewertung hinausgehen, ist jedoch eine erneute Konformitätsbewertung erforderlich, vgl. Art. 43 Abs. 4 KI-VO.

Art. 8 Abs. 2 der KI-VO steht darüber hinaus insbesondere in **Zusammenhang** mit Art. 11 Abs. 2 (technische Dokumentation), Art. 16 und 25 Abs. 3 (Verantwortlichkeiten entlang der KI-Wertschöpfungskette), Art. 43 Abs. 3 und 4 sowie Art. 46 Abs. 7 (Konformitätsbewertung und Ausnahmen von der Konformitätsbewertung), Art. 72 Abs. 4 (Produktbeobachtung) und Art. 74 Abs. 3 und 4 (Marktüberwachung) der KI-VO.

Zwischenergebnis

Sollte Art. 8 Abs. 2 KI-VO einschlägig sein, d. h. für das Produkt, in das ein Hochrisiko-KI-System integriert ist, zugleich Regelungen auf der Grundlage der im Anhang I Abschnitt A aufgelisteten Rechtsakte gelten, so sind die Anbieter dafür verantwortlich, sicherzustellen, dass ihr Produkt alle geltenden Anforderungen der geltenden Harmonisierungsrechtsvorschriften der Union vollständig entspricht.

Sollte Art. 8 Abs. 2 KI-VO nicht einschlägig sein, ist mit **Schritt 4.1.2** fortzufahren.

Schritt 4.1.2: Wie ist das Risikomanagementsystem auszugestalten?

Dr. Benedikt Rohrßen (Taylor Wessing Partnerschaftsgesellschaft mbB, Annegrit Seyerlein-Klug (neurocat GmbH)

Art. 9 KI-VO regelt, wie das Risikomanagementsystem auszugestalten ist.

(1) Für Hochrisiko-KI-Systeme wird ein Risikomanagementsystem eingerichtet, angewandt, dokumentiert und aufrechterhalten.

(2) Das Risikomanagementsystem versteht sich als ein kontinuierlicher iterativer Prozess, der während des gesamten Lebenszyklus eines Hochrisiko-KI-Systems geplant und durchgeführt wird und eine regelmäßige systematische Überprüfung und Aktualisierung erfordert. Es umfasst folgende Schritte:

a) die Ermittlung und Analyse der bekannten und vernünftigerweise vorhersehbaren Risiken, die vom Hochrisiko-KI-System für die Gesundheit, Sicherheit oder Grundrechte ausgehen können, wenn es entsprechend seiner Zweckbestimmung verwendet wird;

b) die Abschätzung und Bewertung der Risiken, die entstehen können, wenn das Hochrisiko-KI-System entsprechend seiner Zweckbestimmung oder im Rahmen einer vernünftigerweise vorhersehbaren Fehlanwendung verwendet wird;

c) die Bewertung anderer möglicherweise auftretender Risiken auf der Grundlage der Auswertung der Daten aus dem in Artikel 72 genannten System zur Beobachtung nach dem Inverkehrbringen;

d) die Ergreifung geeigneter und gezielter Risikomanagementmaßnahmen zur Bewältigung der gemäß Buchstabe a ermittelten Risiken.

(3) Die in diesem Artikel genannten Risiken betreffen nur solche Risiken, die durch die Entwicklung oder Konzeption des Hochrisiko-KI-Systems oder durch die Bereitstellung ausreichender technischer Informationen angemessen gemindert oder behoben werden können.

Relevante(r) Artikel:

Art. 9

Relevante(r) ErwG:

65

Konkretisierungsbedürftig:

Ja (Durchführungsrechtsakt der Kommission zur Erstellung eines Musterplans für Beobachtung nach Inverkehrbringen sowie Liste der aufzunehmenden Elemente (Art. 9 Abs. 2 lit. c) i.V.m. Art. 72 Abs. 3 KI-VO)

Durchführungsrechtsakt der Kommission zur Festlegung der Elemente des Plans für Test unter Realbedingungen (Art. 9 Abs. 7 i.V.m. Art. 60 KI-VO)

(1) Für Hochrisiko-KI-Systeme wird ein Risikomanagementsystem eingerichtet, angewandt, dokumentiert und aufrechterhalten.

(4) Bei den in Absatz 2 Buchstabe d genannten Risikomanagementmaßnahmen werden die Auswirkungen und möglichen Wechselwirkungen, die sich aus der kombinierten Anwendung der Anforderungen dieses Abschnitts ergeben, gebührend berücksichtigt, um die Risiken wirksamer zu minimieren und gleichzeitig ein angemessenes Gleichgewicht bei der Durchführung der Maßnahmen zur Erfüllung dieser Anforderungen sicherzustellen.

(5) Die in Absatz 2 Buchstabe d genannten Risikomanagementmaßnahmen werden so gestaltet, dass jedes mit einer bestimmten Gefahr verbundene relevante Restrisiko sowie das Gesamtrestrisiko der Hochrisiko-KI-Systeme als vertretbar beurteilt wird. Bei der Festlegung der am besten geeigneten Risikomanagementmaßnahmen ist Folgendes sicherzustellen:

a) soweit technisch möglich, Beseitigung oder Verringerung der gemäß Absatz 2 ermittelten und bewerteten Risiken durch eine geeignete Konzeption und Entwicklung des Hochrisiko-KI-Systems;

b) gegebenenfalls Anwendung angemessener Minderungs- und Kontrollmaßnahmen zur Bewältigung nicht auszuschließender Risiken;

c) Bereitstellung der gemäß Artikel 13 erforderlichen Informationen und gegebenenfalls entsprechende Schulung der Betreiber.

d) Zur Beseitigung oder Verringerung der Risiken im Zusammenhang mit der Verwendung des Hochrisiko-KI-Systems werden die technischen Kenntnisse, die Erfahrungen und der Bildungsstand, die vom Betreiber erwartet werden können, sowie der voraussichtliche Kontext, in dem das System eingesetzt werden soll, gebührend berücksichtigt.

(6) Hochrisiko-KI-Systeme müssen getestet werden, um die am besten geeigneten gezielten Risikomanagementmaßnahmen zu ermitteln. Durch das Testen wird sichergestellt, dass

Hochrisiko-KI-Systeme stets bestimmungsgemäß funktionieren und die Anforderungen dieses Abschnitts erfüllen.

(7) Die Testverfahren können einen Test unter realen Bedingungen gemäß Artikel 60 umfassen.

(8) Das Testen von Hochrisiko-KI-Systemen erfolgt zu jedem geeigneten Zeitpunkt während des gesamten Entwicklungsprozesses und in jedem Fall vor ihrem Inverkehrbringen oder ihrer Inbetriebnahme. Das Testen erfolgt anhand vorab festgelegter Parameter und probabilistischer Schwellenwerte, die für die Zweckbestimmung des Hochrisiko-KI-Systems geeignet sind.

(9) Bei der Umsetzung des in den Absätzen 1 bis 7 vorgesehenen Risikomanagementsystems berücksichtigen die Anbieter, ob angesichts seiner Zweckbestimmung das Hochrisiko-KI-System wahrscheinlich nachteilige Auswirkungen auf Personen unter 18 Jahren oder gegebenenfalls andere Gruppen schutzbedürftiger Personen haben wird.

(10) Bei Anbietern von Hochrisiko-KI-Systemen, die den Anforderungen an interne Risikomanagementprozesse gemäß anderen einschlägigen Bestimmungen der Rechtsvorschriften der Union unterliegen, können die in den Absätzen 1 bis 9 enthaltenen Aspekte Bestandteil der nach diesen Rechtsvorschriften festgelegten Risikomanagementverfahren sein oder mit diesen Verfahren kombiniert werden.

Hochrisiko-KI-Systeme bieten **zahlreiche Vorteile**, bergen jedoch auch **potenzielle Risiken** für Gesundheit, Sicherheit und Grundrechte. Deshalb ist es unerlässlich, diese Risiken korrekt zu identifizieren und zu mindern. Ein effektives Risikomanagement ist nicht nur eine rechtliche Verpflichtung, sondern auch eine organisatorische Notwendigkeit.

Risikomanagement umfasst den systematischen Prozess der Ermittlung, Bewertung und Steuerung von Risiken, die mit dem Einsatz von Künstlicher Intelligenz über ihren gesamten Lebenszyklus hinweg verbunden sind. Die Integration eines solchen Systems soll rechtliche und finanzielle Risiken minimieren und gleichzeitig Innovation und verantwortungsvollen Einsatz von KI fördern.

Die KI-Verordnung misst der Etablierung eines robusten Risikomanagementsystems **hohe Bedeutung** bei: Artikel 9 der KI-Verordnung legt dem Anbieter die Pflicht auf, ein Risikomanagementsystem zu errichten, das den **gesamten Lebenszyklus** von Hochrisiko-KI-Systemen umfasst. Ziel ist es, gesundheits-, sicherheits- und grundrechtsrelevante Gefahren solcher Systeme zu berücksichtigen und effektiv zu minimieren. Das Risikomanagementsystem soll dem Anbieter dabei helfen, Risiken zu erkennen, die sich sowohl bei bestimmungsgemäßer Nutzung als auch bei vorhersehbarer Fehlanwendung des

KI-Systems ergeben können, und geeignete Minderungsmaßnahmen zu ergreifen. Hierbei ist stets der aktuelle Stand der Technik zu berücksichtigen.

Die **Implementierung eines effektiven Risikomanagementsystems** umfasst folgende **Schritte**:

1. Risikoidentifizierung: Um geeignete Risikominderungsmaßnahmen zu ergreifen, ist es zunächst wesentlich, potenzielle Risiken des Hochrisiko-KI-Systems zu identifizieren. Dies betrifft nicht nur bekannte Risiken, sondern auch die Berücksichtigung möglicher Fehlanwendungen von Hochrisiko-KI-Systemen, die aufgrund vorhersehbaren menschlichen Verhaltens auftreten können.

2. Risikoanalyse- und -bewertung: Nach der Identifikation potenzieller Risiken sind diese zu bewerten. Dies ist zentraler Bestandteil für die Ermittlung geeigneter Risikominderungsmaßnahmen. Dabei sind sowohl die potenziellen Auswirkungen als auch die Wahrscheinlichkeit eines Risikos abzuschätzen.

3. Risikominderung: Basierend auf der Bewertung müssen Maßnahmen zur Risikominderung ergriffen werden. Anbieter haben die am besten geeigneten Risikomanagementmaßnahmen zu ermitteln, wofür gegebenenfalls Sachverständige und externe Interessengruppen hinzugezogen werden sollen. Zudem sind Testverfahren zur Ermittlung der am besten geeigneten Maßnahmen durchzuführen. Bei der Festlegung der am besten geeigneten Risikomanagementmaßnahmen müssen Anbieter außerdem ihre Entscheidungen dokumentieren und erläutern. Empfehlenswert sind spezifische Schulungsmaßnahmen, um vorhersehbaren Fehlanwendungen wirksam entgegenzuwirken.

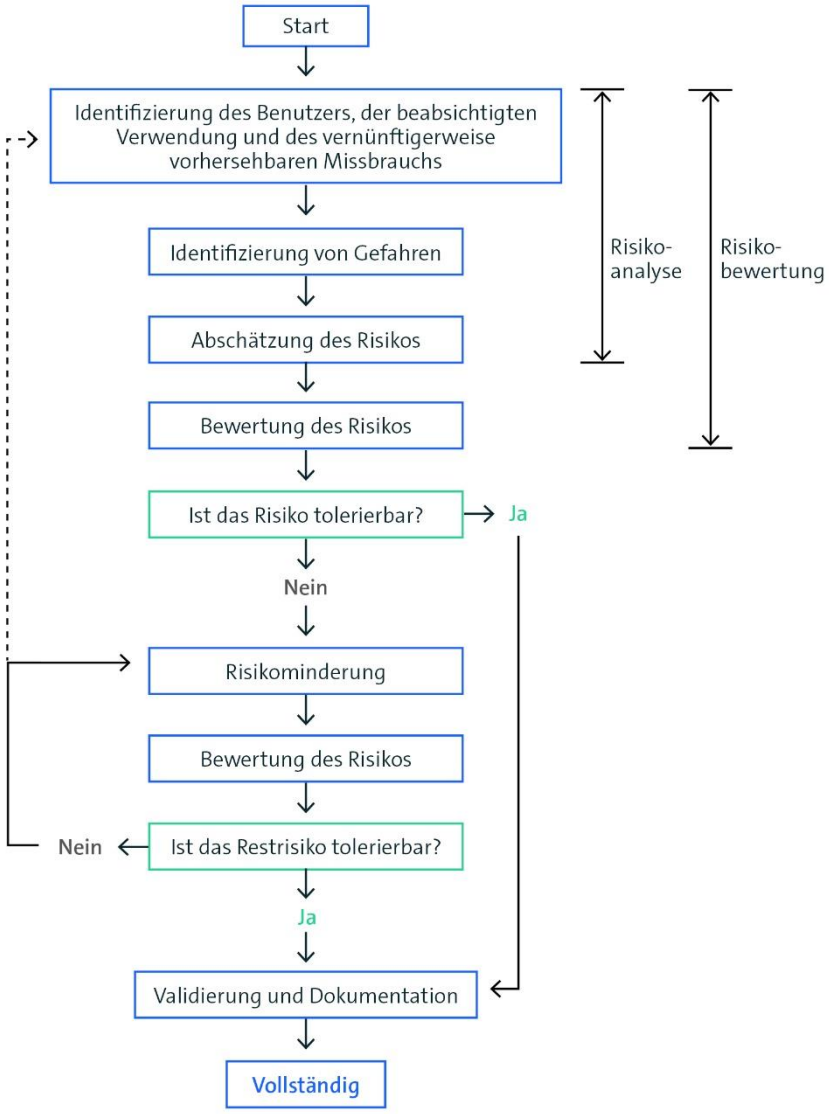
4. Alle bekannten oder vorhersehbaren Umstände bezüglich der Verwendung des Hochrisiko-KI-Systems, die zu Risiken für Gesundheit, Sicherheit oder Grundrechte führen könnten, sind in der Betriebsanleitung aufzuführen. Dies stellt sicher, dass der Betreiber sich dieser Umstände bewusst ist und sie bei der Nutzung des Systems berücksichtigen kann.

5. Kontinuierliche Aktualisierung: Risikomanagement ist ein fortlaufender, iterativer Prozess. Neue Risiken können jederzeit auftreten, daher müssen Daten und Erkenntnisse kontinuierlich gesammelt und bewertet werden, um per regelmäßigem Überprüfungs- und Aktualisierungsprozess sicherzustellen, dass das Risikomanagementsystem stets auf dem neuesten Stand bleibt.

Ein **umfassendes Risikomanagementsystem** gemäß der KI-VO ist unerlässlich, um die Sicherheit, Gesundheit und Grundrechte der Nutzer zu schützen. Anbieter müssen nicht nur ihre gesetzlichen Verpflichtungen erfüllen, sondern auch praktische Maßnahmen zur Implementierung dieser Anforderungen ergreifen. Ein solides Risikomanagementsystem hilft, die Risiken im Zusammenhang mit Hochrisiko-KI-Systemen umfassend zu analysieren und effektiv zu bewältigen.

Um Doppelarbeit und unnötigen Verwaltungsaufwand zu vermeiden, kann das gemäß der KI-VO erforderliche Risikomanagementsystem in ein bereits bestehendes Risikomanagementverfahren nach einschlägigem sektoralen Unionsrecht integriert werden, sofern es die Anforderungen der KI-Verordnung erfüllt.

Die Europäische Kommission bevorzugt ein Risikomanagement für AI-Produkte ähnlich ISO/IEC Guide 51 Safety aspects — Guidelines for their inclusion in standards:



Zwischenergebnis

Ist das Risikomanagementsystem eingerichtet, ist mit **Schritt 4.1.3** fortzufahren.

Schritt 4.1.3: Wie ist die Daten-Governance zu gestalten?

Relevante(r) Artikel:

Art. 10

Relevante(r) ErwG:

27, 67, 68, 70

Konkretisierungsbedürftig:

Ja

(1) Hochrisiko-KI-Systeme, in denen Techniken eingesetzt werden, bei denen KI-Modelle mit Daten trainiert werden, müssen mit Trainings-, Validierungs- und Testdatensätzen entwickelt werden, die den in den Absätzen 2 bis 5 genannten Qualitätskriterien entsprechen, wenn solche Datensätze verwendet werden.

(2) Für Trainings-, Validierungs- und Testdatensätze gelten Daten-Governance- und Datenverwaltungsverfahren, die für die Zweckbestimmung des Hochrisiko-KI-Systems geeignet sind. Diese Verfahren betreffen insbesondere

a) die einschlägigen konzeptionellen Entscheidungen,

b) die Datenerhebungsverfahren und die Herkunft der Daten und im Falle personenbezogener Daten den ursprünglichen Zweck der Datenerhebung,

c) crelevante Datenaufbereitungsvorgänge wie Annotation, Kennzeichnung, Bereinigung, Aktualisierung, Anreicherung und Aggregierung,

d) die Aufstellung von Annahmen, insbesondere in Bezug auf die Informationen, die mit den Daten erfasst und dargestellt werden sollen,

e) eine Bewertung der Verfügbarkeit, Menge und Eignung der benötigten Datensätze,

f) eine Untersuchung im Hinblick auf mögliche Verzerrungen (Bias), die die Gesundheit und Sicherheit von Personen beeinträchtigen, sich negativ auf die Grundrechte auswirken oder zu einer nach den Rechtsvorschriften der Union verbotenen Diskriminierung führen könnten, insbesondere wenn die Datenausgaben die Eingaben für künftige Operationen beeinflussen,

g) geeignete Maßnahmen zur Erkennung, Verhinderung und Abschwächung möglicher gemäß Buchstabe f ermittelter Verzerrungen,

h) die Ermittlung relevanter Datenlücken oder Mängel, die der Einhaltung dieser Verordnung entgegenstehen, und wie diese Lücken und Mängel behoben werden können.

(3) Die Trainings-, Validierungs- und Testdatensätze müssen im Hinblick auf die Zweckbestimmung relevant, hinreichend repräsentativ und so weit wie möglich fehlerfrei und vollständig sein. Sie müssen die geeigneten statistischen Merkmale, gegebenenfalls auch bezüglich der Personen oder Personengruppen, für die das Hochrisiko-KI-System bestimmungsgemäß verwendet werden soll, haben. Diese Merkmale der Datensätze können auf der Ebene einzelner Datensätze oder auf der Ebene einer Kombination davon erfüllt werden.

(4) Die Datensätze müssen, soweit dies für die Zweckbestimmung erforderlich ist, die entsprechenden Merkmale oder Elemente berücksichtigen, die für die besonderen geografischen, kontextuellen, verhaltensbezogenen oder funktionalen Rahmenbedingungen, unter denen das Hochrisiko-KI-System bestimmungsgemäß verwendet werden soll, typisch sind.

(5) Soweit dies für die Erkennung und Korrektur von Verzerrungen im Zusammenhang mit Hochrisiko-KI-Systemen im Einklang mit Absatz 2 Buchstaben f und g dieses Artikels unbedingt erforderlich ist, dürfen die Anbieter solcher Systeme ausnahmsweise besondere Kategorien personenbezogener Daten verarbeiten, wobei sie angemessene Vorkehrungen für den Schutz der Grundrechte und Grundfreiheiten natürlicher Personen treffen müssen. Zusätzlich zu den Bestimmungen der Verordnungen (EU) 2016/679 und (EU) 2018/1725 und der Richtlinie (EU) 2016/680 müssen alle folgenden Bedingungen erfüllt sein, damit eine solche Verarbeitung stattfinden kann:

a) Die Erkennung und Korrektur von Verzerrungen kann durch die Verarbeitung anderer Daten, einschließlich synthetischer oder anonymisierter Daten, nicht effektiv durchgeführt werden;

b) die besonderen Kategorien personenbezogener Daten unterliegen technischen Beschränkungen einer Weiterverwendung der personenbezogenen Daten und modernsten Sicherheits- und Datenschutzmaßnahmen, einschließlich Pseudonymisierung;

c) die besonderen Kategorien personenbezogener Daten unterliegen Maßnahmen, mit denen sichergestellt wird, dass die verarbeiteten personenbezogenen Daten gesichert, geschützt und Gegenstand angemessener Sicherheitsvorkehrungen sind, wozu auch strenge Kontrollen des Zugriffs und seine Dokumentation gehören, um Missbrauch zu verhindern und sicherzustellen, dass nur befugte Personen Zugang zu diesen personenbezogenen Daten mit angemessenen Vertraulichkeitspflichten haben;

d) die besonderen Kategorien personenbezogener Daten werden nicht an Dritte übermittelt oder übertragen, noch haben diese Dritten anderweitigen Zugang zu diesen Daten;

e) die besonderen Kategorien personenbezogener Daten werden gelöscht, sobald die Verzerrung korrigiert wurde oder das Ende der Speicherfrist für die personenbezogenen Daten erreicht ist, je nachdem, was zuerst eintritt;

f) die Aufzeichnungen über Verarbeitungstätigkeiten gemäß den Verordnungen (EU) 2016/679 und (EU) 2018/1725 und der Richtlinie (EU) 2016/680 enthalten die Gründe, warum die Verarbeitung besonderer Kategorien personenbezogener Daten für die Erkennung und Korrektur von Verzerrungen unbedingt erforderlich war und warum dieses Ziel mit der Verarbeitung anderer Daten nicht erreicht werden konnte.

(6) Bei der Entwicklung von Hochrisiko-KI-Systemen, in denen keine Techniken eingesetzt werden, bei denen KI-Modelle trainiert werden, gelten die Absätze 2 bis 5 nur für Testdatensätze.

Sandra Baum (Bundesdruckerei GmbH), Camilla Dalerci (Bundesdruckerei GmbH), Sven Jacobs (Cisco Systems GmbH), Dr. Christoph Krück (SKW Schwarz Rechtsanwälte), Malte Lange (Finanz Informatik GmbH & Co. KG), Maria Stammwitz (Bundesdruckerei GmbH)

Art. 10 KI-VO regelt, wie die Daten-Governance auszugestalten ist.

Eine Daten-Governance umfasst strukturierte Richtlinien und Prinzipien, um sicherzustellen, dass Daten verantwortungsbewusst, ethisch und in Übereinstimmung mit regulatorischen Standards eingesetzt werden. Eine dementsprechende Governance erfüllt mehrere Zwecke, unter anderem ist sie Teil des Risikomanagements, von Transparenz und Accountability, der Einhaltung gesetzlicher Bestimmungen und regulatorischer Vorgaben sowie einer Vertrauensbasis mit Geschäftspartnern und Aufsichtsbehörden.

Im Rahmen der KI-Verordnung sind Regelungen zu Daten und der Daten-Governance in Art. 10 niedergelegt: Sofern Hochrisiko-KI-Systeme mit Daten trainiert werden, müssen Trainings-, Validierungs- und Testdatensätze verwendet werden, die den dort genannten Qualitätskriterien entsprechen. Insbesondere gelten für Trainings-, Validierungs- und Testdatensätze Daten-Governance- und Datenverwaltungsverfahren, die für die Zweckbestimmung des jeweiligen Hochrisiko-KI-Systems geeignet sind (Absatz 2), die verwendeten Datensätze müssen im Hinblick auf die Zweckbestimmung relevant, hinreichend repräsentativ und so weit wie möglich fehlerfrei und vollständig sein (Absatz 3) und, soweit dies für die Zweckbestimmung erforderlich ist, die entsprechenden Merkmale oder Elemente berücksichtigen, die für die besonderen geografischen, kontextuellen,

verhaltensbezogenen oder funktionalen Rahmenbedingungen, unter denen das Hochrisiko-KI-System bestimmungsgemäß verwendet werden soll, typisch sind (Absatz 4). Bei der Entwicklung von Hochrisiko-KI-Systemen, in denen keine Techniken eingesetzt werden, bei denen KI-Modelle trainiert werden, gelten diese Anforderungen nur für Testdatensätze (Absatz 6).

Diese Anforderungen werden in den Erwägungsgründen ergänzt teilweise Einzelheiten zu den Elementen einer Daten-Governance, wie beispielsweise, dass der Schutz der Privatsphäre und des Datenschutzes gewährleistet werden muss (ErwG 27), dass Daten-Governance- und Datenverwaltungsverfahren bei personenbezogenen Daten Transparenz in Bezug auf den ursprünglichen Zweck der Datenerhebung umfassen sollten oder dass Datensätze auch die geeigneten statistischen Merkmale haben sollten, auch bezüglich der Personen oder Personengruppen, auf die das Hochrisiko-KI-System bestimmungsgemäß angewandt werden soll, unter besonderer Berücksichtigung der Minderung möglicher Verzerrungen in den Datensätzen, die die Gesundheit und Sicherheit von Personen beeinträchtigen, sich negativ auf die Grundrechte auswirken oder zu einer nach dem Unionsrecht verbotenen Diskriminierung führen könnten, insbesondere wenn die Datenausgaben die Eingaben für künftige Operationen beeinflussen (Rückkopplungsschleifen) (ErwG 67).

Neben den Anforderungen zur Daten-Governance aus der KI-Verordnung, unter anderem, dass Datensätze relevant, repräsentativ, fehlerfrei und vollständig sein müssen, sind bei personenbezogenen Daten auch die Vorgaben der DSGVO (insbesondere das Vorliegen einer Rechtsgrundlage) und bei nicht-personenbezogenen Daten die Vorgaben der Datenverordnung zu beachten.

Kernkomponenten eines solchen Governance-Systems sind die Identifizierung von Zielen einer Daten-Governance, einschließlich entsprechender Prozess, die Etablierung einer klaren Führungsstruktur mit definierten Rollen und Verantwortlichkeiten, die Identifizierung und Minderung potenzieller Risiken, die Implementierung robuster operativer Kontrollen, einschließlich eines Lebenszyklusmanagements, die Bereitstellung kontinuierlicher Schulungen für Personal und die regelmäßige Überprüfung und Aktualisierung der Daten-Governance.

Im Rahmen der Implementierung ist ein mehrstufiges Verfahren oft ratsam, welches aus der Planung und Vorbereitung, Entwicklung und Integration, Ausführung und Überwachung und Überprüfung und Verbesserung bestehen sollte.

Die Anforderungen an die Daten-Governance können durch die Inanspruchnahme Dritter erfüllt werden, die zertifizierte Compliance-Dienste anbieten, einschließlich der Überprüfung der Daten-Governance, der Datensatzintegrität und der Datenschulungs-, Validierungs- und Testverfahren, sofern die Einhaltung der Datenanforderungen dieser Verordnung gewährleistet ist.

Praxisbeispiel

Ein Praxisbeispiel im öffentlichen Raum kann der Einsatz von künstlicher Intelligenz in der Verkehrsplanung einer Stadt sein, etwa im Rahmen von „Smart Cities“. KI-gestützte Videoanalysen helfen einerseits bei der Früherkennung von sicherheitskritischen Ereignissen und damit z. B. bei der Minimierung von Unfällen, andererseits bei einer umweltfreundlichen und effizienten Stadtplanung, z. B. durch die Verkürzung von Fahrzeiten und bei der Planung von Parkplätzen.

In der Regel verwenden diese KI-Systeme für prädiktive Analysen das Ergebnis mehrerer Datenquellen, bspw. Sensor- und Kameradaten, historische Aufzeichnungen, Wetterdaten und GPS-Daten. Die Herausforderung besteht im Beispiel darin, dass die Daten geografisch repräsentativ sein sollten und das Fahrverhalten fehlerfrei widerspiegeln, dabei jedoch datenschutzrechtliche Rahmenbedingungen berücksichtigen müssen.

Im Praxisbeispiel sollte besonders auf die **Datenquellen**, auf die Vermeidung von **Bias** (können bestimmte Stadtteile benachteiligt werden?) und **Datenschutz** (werden persönliche Daten der Bürgerinnen und Bürger, wie Kfz-Kennzeichen oder Bewegungsdaten rechtzeitig aus den Datensätzen entfernt?) geachtet werden, um Data Governance im Sinne der KI-Verordnung umzusetzen. Dazu gehören Verfahren zur **Pseudonymisierung oder Anonymisierung von Daten**, sowie eine klare **Kommunikation** an die Bürgerinnen, Bürger und Stakeholder über die Datenquellen, den Zweck und den Zeitpunkt der Datenerhebung und -verarbeitung.

Zwischenergebnis

Ist die Daten-Governance gewährleistet, ist mit **Schritt 4.1.4** fortzufahren.

Schritt 4.1.4: Wie ist die technische Dokumentation auszugestalten?

Dr. Kim Lauenroth (Fachhochschule Dortmund), Dr. Benedikt Rohrßen (Taylor Wessing Partnerschaftsgesellschaft mbB)

Ziel der KI-VO ist es, einen sicheren Binnenmarkt zu schaffen, der die Einführung menschenzentrierter und vertrauenswürdiger KI-Systeme fördert, während gleichzeitig Gesundheit, Sicherheit und Grundrechte gewahrt bleiben. Dazu ist eine umfassende **technische Dokumentation** unerlässlich, **bevor** das System in Verkehr gebracht oder in Betrieb genommen wird. Die technische Dokumentation dient als zentrales Instrument zur **Beobachtung und Beurteilung der Konformität** des Hochrisiko-KI-Systems mit den geltenden Anforderungen der KI-VO sowie zu ihrer effektiven Überwachung durch die zuständigen nationalen Behörden.

Art. 11 KI-VO regelt, wie die technische Dokumentation auszugestalten ist. Die allgemeinen Anforderungen werden in Absatz 1 definiert. Die Konkretisierung dieser Anforderungen erfolgt in Anhang IV.

1. Die technische Dokumentation eines Hochrisiko-KI-Systems wird erstellt, bevor dieses System in Verkehr gebracht oder in Betrieb genommen wird, und ist auf dem neuesten Stand zu halten. Die technische Dokumentation wird so erstellt, dass aus ihr der Nachweis hervorgeht, wie das Hochrisiko-KI-System die Anforderungen dieses Abschnitts erfüllt, und dass den zuständigen nationalen Behörden und den notifizierten Stellen die Informationen in klarer und verständlicher Form zur Verfügung stehen, die erforderlich sind, um zu beurteilen, ob das KI-System diese Anforderungen erfüllt. Sie enthält zumindest die in Anhang IV genannten Angaben. KMU, einschließlich Start-up-Unternehmen, können die in Anhang IV aufgeführten Elemente der technischen Dokumentation in vereinfachter Weise bereitstellen. Zu diesem Zweck erstellt die Kommission ein vereinfachtes Formular für die technische Dokumentation, das auf die Bedürfnisse von kleinen Unternehmen und Kleinstunternehmen zugeschnitten ist. Entscheidet sich ein KMU, einschließlich Start-up-Unternehmen, für eine vereinfachte Bereitstellung der in Anhang IV vorgeschriebenen Angaben, so verwendet es das in diesem Absatz genannte Formular. Die notifizierten Stellen akzeptieren das Formular für die Zwecke der Konformitätsbewertung.

2. Wird ein Hochrisiko-KI-System, das mit einem Produkt verbunden ist, das unter die in Anhang I Abschnitt A aufgeführten Harmonisierungsrechtsvorschriften der Union fällt, in Verkehr gebracht oder in Betrieb genommen, so wird eine einzige technische Dokumentation erstellt, die alle in Absatz 1 genannten Informationen sowie die nach diesen Rechtsakten erforderlichen Informationen enthält.

3. Die Kommission ist befugt, wenn dies nötig ist, gemäß Artikel 97 delegierte Rechtsakte zur Änderung des Anhangs IV zu erlassen, damit die technische Dokumentation in Anbetracht des technischen Fortschritts stets alle Informationen enthält, die erforderlich sind, um zu beurteilen, ob das System die Anforderungen dieses Abschnitts erfüllt.

Relevante(r) Artikel:

Art. 11, 18

Relevante(r) ErwG:

66, 71

Konkretisierungsbedürftig:

Ja, Bereitstellung von einem Formular für die technische Dokumentation für KMU durch Kommission, Art. 11 Abs. 1 KI-VO

Befugnis der Kommission, delegierte Rechtsakte zur Änderung des Anhangs IV zu erlassen, um die Anforderungen an die technische Dokumentation an den Stand der Technik anzupassen.

Die technische Dokumentation in der KI-VO ist auf den ersten Blick mehrdeutig formuliert und bezieht sich sowohl auf das Hochrisiko-KI-System an sich, sowie den Herstellungs- und Betriebsprozess. Ganz allgemein kann festgehalten werden, dass die in Abschnitt 2 definierten Anforderungen auf erprobten und guten Praktiken des professionellen Designs und Engineerings digitaler Lösungen bestehen. **Die wesentliche Aufgabe** der technischen Dokumentation besteht darin, den Nachweis zu erbringen, dass das KI-System konform zu den Anforderungen aus Abschnitt 2 der KI-VO entworfen, hergestellt und betrieben werden kann. Weiterhin werden in Artikel 11 bzw. Anhang IV implizit auch konkrete Anforderungen an ein Hochrisiko-KI-System formuliert.

Für KMU und Start-up-Unternehmen wird ein vereinfachtes Formular als Option benannt. Dieses Formular ist, Stand heute, aber noch nicht verfügbar.

Im Folgenden werden die zuvor genannten Aspekte kompakt dargestellt. Da sich Artikel 11 auf das gesamte Kapitel III 2 der KI-VO (Hochrisiko-KI-Systeme) bezieht, werden im Folgenden Verweise auf das gesamte Kapitel III 2 formuliert.

Technische Dokumentation des Hochrisiko-KI-Systems an sich

Die technische Dokumentation muss das Hochrisiko-KI-System an sich umfassen darstellen. Hierzu werden die folgenden **Bestandteile zur Beschreibung** genannt:

1. Allgemeine Beschreibung des KI-Systems (Anhang IV, Abs. 1)
2. Entwurfsspezifikation (Anhang IV, Abs. 2b)
3. Systemarchitektur (Anhang IV, Abs. 2c)
4. Datenanforderungen mit allgemeiner Beschreibung der Daten (Anhang IV, Abs. 2d)
5. Anforderungen an Genauigkeit, Robustheit und Cybersicherheit (Artikel 15)
6. Betriebsanleitung (Artikel 13 Abs. 3)
7. ggf. Beschreibung der vorab bestimmten Änderungen an dem KI-System und seiner Leistung (Anhang IV, Abs. 2f)
8. ergriffene Cybersicherheitsmaßnahmen (Anhang IV, Abs. 2h)
9. Detaillierte Informationen über die Überwachung, Funktionsweise und Kontrolle des KI-Systems (Anhang IV, Abs. 3)
10. Aufstellung der verwendeten Normen (Anhang IV, Abs. 7)
11. Kopie der EU-Konformitätserklärung gemäß Artikel 47 (Anhang IV, Abs. 8)

Technische Dokumentation mit Blick auf den Entwicklungs- bzw. Herstellungsprozess

Die technische Dokumentation muss den Entwicklungs- bzw. Herstellungsprozess des Hochrisiko-KI-Systems darstellen. Hierdurch werden **implizit Vorgaben** gemacht, wie der **Entwicklungsprozess** zu gestalten ist. **Folgende Aspekte** hierbei genannt:

1. Beschreibung des Entwicklungsprozesses (Anhang IV, Abs. 2a)
2. Risikomanagementsystem als Teil des Herstellungsprozesses (Artikel 9)

3. Systematisches Testen im Sinne des Softwaretests (Artikel 9, Abs. 6 - 9) inklusive Dokumentation der Testverfahren und der Testprotokolle (Anhang IV, Abs. 2g)
4. Daten-Governance (Artikel 10)
5. Darstellung der Trainingsmethoden und -techniken (Anhang IV Abs. 2d)
6. Beschreibung einschlägiger Änderungen über den gesamten Lebenszyklus (Anhang IV, Abs. 6)

Technische Dokumentation mit Blick auf den Betriebsprozess

Die technische Dokumentation muss den Betriebsprozess des Hochrisiko-KI-Systems darstellen. Hierdurch werden **implizit Vorgaben** gemacht, wie der **Betrieb** zu gestalten ist. **Folgende Aspekte** hierbei genannt:

1. Transparenz und Bereitstellung von Informationen für die Betreiber (Artikel 13)
2. Detaillierte Beschreibung des Systems zur Beobachtung und Bewertung der Leistung des KI-Systems in der Phase nach dem Inverkehrbringen (Anhang IV, Abs. 9)

Anforderungen an das Hochrisiko-KI-System

Im Kontext der technischen Dokumentation und dem damit verbundenen Kapitel III 2 werden eine Reihe von sehr konkreten Anforderungen formuliert, die ein Hochrisiko-KI-System umsetzen muss. **Folgende Funktionen** werden hierbei genannt:

1. Protokollierungsfunktionen zur Erfüllung der Aufzeichnungspflichten (Artikel 12 bzw. 13 Abs. 3f).
2. Wirksame menschliche Aufsicht (Artikel 14)
3. „Stop-Taste“ als Funktionalität im Notfall (Artikel 14 Abs. 4e).

Für Hochrisiko-KI-Systeme im Zusammenhang mit einem Produkt, das unter die in Abschnitt A von Anhang I der KI-VO aufgeführten zwölf EU-Harmonisierungsrechtsvorschriften fällt, wie z. B. Maschinen oder Medizinprodukte, genügt die Erstellung einer technischen Dokumentation, die sowohl den Anforderungen der KI-VO als auch der einschlägigen Harmonisierungsrechtsvorschriften gerecht wird. Dies spiegelt den Ansatz der Verordnung wider, Kohärenz zu gewährleisten, Doppelarbeit zu vermeiden und zusätzlichen Aufwand für Anbieter zu minimieren, die ihre Produkte einer Konformitätsbewertung unterziehen und bereits bestehende Produktanforderungen einhalten müssen.

Zwischenergebnis

Ist die technische Dokumentation eingerichtet, ist mit **Schritt 4.1.5** fortzufahren.

Schritt 4.1.5: Wie sind die Aufzeichnungspflichten zu erfüllen?

Benedict Huyeng (RWE AG)

Art. 12 KI-VO regelt, wie die Aufzeichnungspflichten zu erfüllen sind.

(1) Die Technik der Hochrisiko-KI-Systeme muss die automatische Aufzeichnung von Ereignissen (im Folgenden „Protokollierung“) während des Lebenszyklus des Systems ermöglichen.

(2) Zur Gewährleistung, dass das Funktionieren des Hochrisiko-KI-Systems in einem der Zweckbestimmung des Systems angemessenen Maße rückverfolgbar ist, ermöglichen die Protokollierungsfunktionen die Aufzeichnung von Ereignissen, die für Folgendes relevant sind:

a) die Ermittlung von Situationen, die dazu führen können, dass das Hochrisiko-KI-System ein Risiko im Sinne des Artikels 79 Absatz 1 birgt oder dass es zu einer wesentlichen Änderung kommt,

b) die Erleichterung der Beobachtung nach dem Inverkehrbringen gemäß Artikel 72 und

c) die Überwachung des Betriebs der Hochrisiko-KI-Systeme gemäß Artikel 26 Absatz 5.

(3) Die Protokollierungsfunktionen der in Anhang III Nummer 1 Buchstabe a genannten Hochrisiko-KI-Systeme müssen zumindest Folgendes umfassen:

a) Aufzeichnung jedes Zeitraums der Verwendung des Systems (Datum und Uhrzeit des Beginns und des Endes jeder Verwendung);

b) die Referenzdatenbank, mit der das System die Eingabedaten abgleicht;

c) die Eingabedaten, mit denen die Abfrage zu einer Übereinstimmung geführt hat;

d) die Identität der gemäß Artikel 14 Absatz 5 an der Überprüfung der Ergebnisse beteiligten natürlichen Personen

Relevante(r) Artikel:

Art. 12

Relevante(r) ErwG:

71

Konkretisierungsbedürftig:

Bereitstellung von einem Formular für die Technische Dokumentation für KMU durch Kommission, Art. 11 Abs. 1 KI-VO

Befugnis der Kommission, delegierte Rechtsakte zur Änderung des Anhangs IV zu erlassen, um die Anforderungen an die technische Dokumentation an den Stand der Technik anzupassen.

Erwägungsgrund 71

Erwägungsgrund 71 betont die Bedeutung umfassender Informationen und Transparenz im Zusammenhang mit Hochrisiko-KI-Systemen, um ihre Nachvollziehbarkeit und die Einhaltung

der geltenden Vorschriften zu gewährleisten. Der Erwägungsgrund erkennt an, dass dies die Führung von Aufzeichnungen und die Erstellung einer umfassenden technischen Dokumentation erfordert. Anbieter von Hochrisiko-KI-Systemen sollten eine detaillierte technische Dokumentation erstellen, die alle relevanten Informationen über das System enthält. Dazu gehören die allgemeinen Merkmale, Fähigkeiten und Grenzen, die verwendeten Algorithmen, die Trainings-, Test- und Validierungsverfahren sowie die Dokumentation des Risikomanagementsystems. Die technische Dokumentation sollte klar und umfassend sein und während der gesamten Lebensdauer des Systems auf dem neuesten Stand gehalten werden. Dies umfasst Updates und Änderungen, um sicherzustellen, dass die Dokumentation immer aktuell und genau bleibt.

Bezüglich der Aufzeichnung von Ereignissen wird betont, dass Hochrisiko-KI-Systeme technisch in der Lage sein sollten, Ereignisse während ihrer Lebensdauer automatisch aufzuzeichnen (Protokollierung). Dies ermöglicht die Rückverfolgung der Funktionsweise des Systems, die Identifizierung potenzieller Probleme und die Bewertung seiner Leistung im Laufe der Zeit. Die protokollierten Informationen können verschiedene Formen annehmen, wie Zeitstempel, Benutzeraktivitäten, Eingabedaten, Systemausgaben und anomale Ereignisse oder Fehler. Die umfassende Dokumentation und Protokollierung sollen es den zuständigen Behörden ermöglichen, die Konformität von Hochrisiko-KI-Systemen mit den geltenden Vorschriften zu bewerten und potenzielle Risiken zu identifizieren.

Darüber hinaus erleichtern die verfügbaren Informationen und Aufzeichnungen auch die Beobachtung und Überwachung des Systems nach seiner Markteinführung. Dies ermöglicht eine fortlaufende Bewertung seiner Leistung und der Auswirkungen auf die Endbenutzer. Erwägungsgrund 71 unterstreicht die Verantwortung der Anbieter von Hochrisiko-KI-Systemen, sicherzustellen, dass die erforderlichen Informationen bereitgestellt und auf dem neuesten Stand gehalten werden. Dies umfasst die Zusammenarbeit mit den Behörden und die Bereitstellung von Informationen für Überprüfungen oder Audits. Insgesamt zielt der Erwägungsgrund darauf ab, Transparenz und Rechenschaftspflicht im Zusammenhang mit Hochrisiko-KI-Systemen zu fördern, um so den Schutz der Endbenutzer zu gewährleisten.

Artikel 12

Artikel 12 der KI-Verordnung befasst sich mit der Protokollierung und Rückverfolgbarkeit von Hochrisiko-KI-Systemen. Dieser Artikel zielt darauf ab, sicherzustellen, dass Ereignisse während des Lebenszyklus eines Hochrisiko-KI-Systems automatisch aufgezeichnet werden, um eine Rückverfolgung seiner Funktionsweise zu ermöglichen und potenzielle Risiken zu identifizieren.

Hochrisiko-KI-Systeme müssen über Protokollierungsfunktionen verfügen, die es ermöglichen, relevante Ereignisse aufzuzeichnen. Dies umfasst Situationen, die zu Risiken im Sinne des Artikels 79 Absatz 1 führen können, sowie wesentliche Änderungen am System. Die Protokollierung soll auch die Beobachtung nach der Markteinführung gemäß Artikel 72 erleichtern und die Überwachung des Betriebs gemäß Artikel 26 Absatz 5 unterstützen.

Für Hochrisiko-KI-Systeme gemäß Anhang III Nummer 1 Buchstabe a der Verordnung müssen spezifische Ereignisse aufgezeichnet werden, darunter:

- Beginn und Ende der Nutzung des Systems (Datum und Uhrzeit).
- Referenzdatenbank, mit der das System die Eingabedaten abgleicht.
- Eingabedaten, die zu einer Übereinstimmung geführt haben.

- Identität der natürlichen Personen, die an der Überprüfung der Ergebnisse beteiligt sind.

Diese Aufzeichnungen sollen ein klares Bild der Funktionsweise des Systems bieten und es ermöglichen, bei Bedarf Verantwortlichkeiten zu klären und Probleme zu beheben.

Praxisanleitung

Um die Anforderungen des Artikels 12 in der Praxis umzusetzen, können Unternehmen, die Hochrisiko-KI-Systeme entwickeln oder einsetzen, folgende Schritte unternehmen:

1. Implementierung von Protokollierungsfunktionen: Integrieren Sie in das KI-System Mechanismen zur automatischen Aufzeichnung der in Artikel 12 aufgeführten Ereignisse. Dies kann durch die Entwicklung oder Anpassung von Softwarelösungen erfolgen, die Daten während des Betriebs des Systems erfassen und speichern.
2. Sicherstellung der Rückverfolgbarkeit: Stellen Sie sicher, dass die protokollierten Daten es ermöglichen, den Entscheidungspfad und die Funktionsweise des Systems rückwirkend nachzuvollziehen. Dies umfasst die Aufzeichnung von Eingabedaten, Verarbeitungsabläufen und Ergebnissen, einschließlich aller Zwischenwerte und Schwellenwerte, die das System verwendet.
3. Spezifische Anforderungen berücksichtigen: Beachten Sie die spezifischen Anforderungen je nach Art des Hochrisiko-KI-Systems, wie in Anhang III Nummer 1 Buchstabe a aufgeführt. Stellen Sie sicher, dass die Protokollierung mindestens den Zeitraum der Verwendung, die verwendeten Referenzdatenbanken, die relevanten Eingabedaten und die Identität der überprüfenden Personen umfasst.
4. Sichere Datenspeicherung: Implementieren Sie Maßnahmen zur sicheren und geschützten Speicherung der protokollierten Daten. Dies kann die Verschlüsselung von Daten, Zugangskontrollen und die Aufbewahrung in sicheren Datenbanken umfassen, um die Vertraulichkeit und Integrität der Informationen zu wahren.
5. Zugriff und Überprüfung: Ermöglichen Sie autorisierten Personen, wie z. B. Entwicklern, Auditoren oder Regulierungsbehörden, den Zugriff auf die protokollierten Daten, um Überprüfungen durchzuführen. Entwickeln Sie Verfahren für die regelmäßige Überprüfung der Protokolle, um potenzielle Probleme oder Abweichungen vom erwarteten Verhalten des Systems zu identifizieren.
6. Integration in den Lebenszyklus: Betrachten Sie die Protokollierung als integralen Bestandteil des gesamten Lebenszyklus des KI-Systems. Stellen Sie sicher, dass die Protokollierungsfunktionen bereits in der Design- und Entwicklungsphase berücksichtigt werden und während des Betriebs, der Wartung und bei eventuellen Aktualisierungen oder Änderungen am System fortgeführt werden.

Beispiel

Angenommen, Sie entwickeln ein Hochrisiko-KI-System für die medizinische Diagnose, das unter Anhang III Nummer 1 Buchstabe a fällt. In diesem Fall sollten die Protokollierungsfunktionen wie folgt implementiert werden:

- Aufzeichnung der Nutzungszeiten: Das System protokolliert automatisch das Datum und die Uhrzeit des Beginns und Endes jeder Verwendungssitzung.
- Referenzdatenbank: Es wird aufgezeichnet, welche Referenzdatenbank (z. B. ein bestimmter Datensatz medizinischer Bilder) für die jeweilige Abfrage verwendet wurde.
- Eingabedaten: Alle Eingabedaten, die zu einer Diagnose führen, werden protokolliert, z. B. Symptome, Laborergebnisse oder Bilddaten.
- Überprüfungsergebnisse: Die Identität der Personen, die gemäß Artikel 14 Absatz 5 an der Überprüfung der Ergebnisse beteiligt sind (z. B. Ärzte), wird zusammen mit ihren Überprüfungsergebnissen aufgezeichnet.

Beispielhafter Umsetzungsprozess

1. Integration von Protokollierungsfunktionen: Das medizinische Diagnosesystem wird mit einem Protokollierungsmodul ausgestattet, das während der Systemausführung im Hintergrund läuft.
2. Aufzeichnung der Nutzungszeiten: Das Modul erfasst automatisch Zeitstempel, wenn das System aktiviert und deaktiviert wird, und speichert diese in einer sicheren Datenbank.
3. Protokollierung der Referenzdatenbank und Eingabedaten: Für jede Diagnoseabfrage wird die verwendete Referenzdatenbank sowie alle spezifischen Eingabedaten, wie Symptome oder Bildaufnahmen, aufgezeichnet.
4. Überprüfungsprozess: Wenn das System eine Diagnose vorschlägt, wird eine Benachrichtigung an einen überprüfenden Arzt gesendet. Die Identität des Arztes und seine Korrektur oder Bestätigung des Ergebnisses werden zusammen mit den entsprechenden Eingabedaten protokolliert.
5. Sichere Datenspeicherung: Alle protokollierten Daten werden verschlüsselt und in einer sicheren Cloud-Umgebung gespeichert, die nur autorisierten Benutzern Zugang gewährt.
6. Regelmäßige Überprüfungen: Ein Datenanalyst überprüft regelmäßig die Protokolle, um sicherzustellen, dass das System wie vorgesehen funktioniert, und um potenzielle Probleme oder unerwartete Ereignisse zu identifizieren.
7. Integration in den Lebenszyklus: Die Protokollierungsfunktionen werden während der gesamten Entwicklung und beim Einsatz des Systems beibehalten, einschließlich Updates und Verbesserungen, um eine lückenlose Rückverfolgbarkeit zu gewährleisten.

Durch die Umsetzung dieser Schritte können Unternehmen sicherstellen, dass ihre Hochrisiko-KI-Systeme den Anforderungen von Artikel 12 der KI-Verordnung entsprechen und eine umfassende Protokollierung und Rückverfolgbarkeit gewährleisten.

Zwischenergebnis

Sind die Aufzeichnungspflichten erfüllt, ist mit **Schritt 4.1.6** fortzufahren.

Schritt 4.1.6: Wie sind die Transparenzpflichten zu erfüllen?

Dilan Mienert (GÖRG Partnerschaft von Rechtsanwälten mbB), Benedict Huyeng (RWE AG); Markus Frowein (RWE AG), Dr. Axel Grätz (Oppenhoff & Partner Rechtsanwälte Steuerberater mbB)

Art. 13 KI-VO regelt, wie die Transparenzpflichten zu erfüllen sind.

Relevante(r) Artikel:

Art. 13

Relevante(r) ErwG:

66, 72

Konkretisierungsbedürftig:

Maßnahmen zur Gewährleistung der Transparenz (durch juristische Auslegung)

(1) Hochrisiko-KI-Systeme werden so konzipiert und entwickelt, dass ihr Betrieb hinreichend transparent ist, damit die Betreiber die Ausgaben eines Systems angemessen interpretieren und verwenden können. Die Transparenz wird auf eine geeignete Art und in einem angemessenen Maß gewährleistet, damit die Anbieter und Betreiber ihre in Abschnitt 3 festgelegten einschlägigen Pflichten erfüllen können.

(2) Hochrisiko-KI-Systeme werden mit Betriebsanleitungen in einem geeigneten digitalen Format bereitgestellt oder auf andere Weise mit Betriebsanleitungen versehen, die präzise, vollständige, korrekte und eindeutige Informationen in einer für die Betreiber relevanten, barrierefrei zugänglichen und verständlichen Form enthalten.

(3) Die Betriebsanleitungen enthalten mindestens folgende Informationen:

a) den Namen und die Kontaktangaben des Anbieters sowie gegebenenfalls seines Bevollmächtigten;

b) die Merkmale, Fähigkeiten und Leistungsgrenzen des Hochrisiko-KI-Systems, einschließlich

i) seiner Zweckbestimmung

ii) des Maßes an Genauigkeit — einschließlich diesbezüglicher Metriken —, Robustheit und Cybersicherheit gemäß Artikel 15, für das das Hochrisiko-KI-System getestet und validiert wurde und das zu erwarten ist, sowie aller bekannten und vorhersehbaren Umstände, die sich auf das erwartete Maß an Genauigkeit, Robustheit und Cybersicherheit auswirken können;

iii) aller bekannten oder vorhersehbaren Umstände bezüglich der Verwendung des Hochrisiko-KI-Systems im Einklang mit seiner Zweckbestimmung oder einer vernünftigerweise vorhersehbaren Fehlanwendung, die zu den in Artikel 9 Absatz 2 genannten Risiken für die Gesundheit und Sicherheit oder die Grundrechte führen können,

iv) gegebenenfalls der technischen Fähigkeiten und Merkmale des Hochrisiko-KI-Systems, um Informationen bereitzustellen, die zur Erläuterung seiner Ausgaben relevant sind;

v) gegebenenfalls seiner Leistung in Bezug auf bestimmte Personen oder Personengruppen, auf die das System bestimmungsgemäß angewandt werden soll;

vi) gegebenenfalls der Spezifikationen für die Eingabedaten oder sonstiger relevanter Informationen über die verwendeten Trainings-, Validierungs- und Testdatensätze, unter Berücksichtigung der Zweckbestimmung des Hochrisiko-KI-Systems;

vii) gegebenenfalls Informationen, die es den Betreibern ermöglichen, die Ausgabe des Hochrisiko-KI-Systems zu interpretieren und es angemessen zu nutzen;

c) Änderungen des Hochrisiko-Systems und seiner Leistung, die der Anbieter zum Zeitpunkt der ersten Konformitätsbewertung vorab bestimmt hat;

d) die in Artikel 14 genannten Maßnahmen zur Gewährleistung der menschlichen Aufsicht, einschließlich der technischen Maßnahmen, die getroffen wurden, um den Betreibern die Interpretation der Ausgaben von Hochrisiko-KI-Systemen zu erleichtern;

e) die erforderlichen Rechen- und Hardware-Ressourcen, die erwartete Lebensdauer des Hochrisiko-KI-Systems und alle erforderlichen Wartungs- und Pflegemaßnahmen einschließlich deren Häufigkeit zur Gewährleistung des ordnungsgemäßen Funktionierens dieses KI-Systems, auch in Bezug auf Software-Updates;

f) gegebenenfalls eine Beschreibung der in das Hochrisiko-KI-System integrierten Mechanismen, die es den Betreibern ermöglicht, die Protokolle im Einklang mit Artikel 12 ordnungsgemäß zu erfassen, zu speichern und auszuwerten.

Die Erwägungsgründe 66 und 72

Die Erwägungsgründe 66 und 72 der KI-Verordnung heben hervor, wie wichtig strenge Anforderungen an Hochrisiko-KI-Systeme sind, um deren sichere und transparente Nutzung sicherzustellen. Diese Anforderungen betreffen unter anderem das Risikomanagement, die Qualität und Relevanz der verwendeten Datensätze, die technische Dokumentation, die Aufzeichnungspflichten sowie die menschliche Überwachung. Solche Maßnahmen sind

notwendig, um die potenziellen Risiken für Gesundheit, Sicherheit und Grundrechte zu minimieren, ohne dabei den Handel unangemessen einzuschränken.

Ein zentrales Element ist die Transparenzpflicht für Hochrisiko-KI-Systeme, die gewährleistet, dass Betreiber diese Systeme verstehen und sicher anwenden können. Bevor Hochrisiko-KI-Systeme in Verkehr gebracht oder genutzt werden dürfen, müssen sie so gestaltet sein, dass ihre Funktionsweise klar nachvollziehbar ist. Betreiber müssen in der Lage sein, die Leistungsfähigkeit des Systems zu bewerten und dessen Stärken und Schwächen zu erkennen.

Hierzu gehört auch, dass umfassende und leicht verständliche Betriebsanleitungen bereitgestellt werden, die detaillierte Informationen über die Merkmale, Fähigkeiten und Leistungsgrenzen des Systems enthalten. Hierbei muss über die Zweckbestimmung des KI-Systems, das Maß an Genauigkeit, einschließlich diesbezüglicher Kennzahlen, Robustheit und Cybersicherheit und aller bekannten oder vorhersehbaren Umstände, die auf das erwartete Maß Einfluss haben können, informiert werden.

Die Betriebsanleitungen sollten die Umstände der Nutzung des Systems beschreiben und Änderungen erläutern, die vorab auf Konformität geprüft wurden. Ebenso wichtig sind die Maßnahmen der menschlichen Überwachung und solche, die den Betreibern helfen, die Ausgaben des KI-Systems zu interpretieren. Dabei müssen sowohl die bekannten als auch die vorhersehbaren Umstände berücksichtigt werden, soweit die vorhersehbare oder vernünftigerweise vorhersehbare Fehlanwendung zu Risiken für die Gesundheit, Sicherheit oder Grundrechte führen kann.

Die Transparenz, einschließlich klarer und anschaulicher Beispiele in den Betriebsanleitungen, soll den Betreibern fundierte Entscheidungen ermöglichen und sie dabei unterstützen, das geeignete System auszuwählen und korrekt zu nutzen. Anbieter sind verpflichtet, sicherzustellen, dass die Dokumentation umfassend, zugänglich und verständlich ist, wobei die Bedürfnisse und das Wissen der Zielbetreiber berücksichtigt werden müssen. Die Betriebsanleitungen sollten in einer Sprache verfasst sein, die von den Betreibern leicht verstanden werden kann.

Inhalt Artikel 13

Artikel 13 der KI-Verordnung setzt diese Transparenzanforderungen detailliert um. Dadurch wird dazu beigetragen, dass Bedenken hinsichtlich der Undurchsichtigkeit und Komplexität von Hochrisiko-KI-Systemen adressiert und Betreiber bei der Erfüllung ihrer Pflichten unterstützt werden.

Die Intention des Artikels 13 ist die Schaffung von Transparenz bezüglich der Konzeption und Entwicklung von Hochrisiko-KI-Systemen. Zudem soll gewährleistet werden, dass die Betreiber die mit diesen Systemen verbundenen Ausgaben und Ergebnisse adäquat interpretieren und nutzen können. Dies ist von entscheidender Bedeutung, um potenzielle Risiken zu minimieren und die Verpflichtungen von Anbietern und Betreibern zu erfüllen.

Im Rahmen der Konzeption von Hochrisiko-KI-Systemen ist ein angemessenes Maß an Transparenz zu gewährleisten. Dies impliziert, dass die Betriebsprozesse dieser Systeme für die Betreiber nachvollziehbar sein sollten, einschließlich der Funktionalität und der Stärken und Grenzen des Systems. Auf diese Weise ist es Betreibern möglich, die Ergebnisse des Systems zu evaluieren, potenzielle Fehler zu identifizieren und sicherzustellen, dass es im Einklang mit den geltenden rechtlichen und ethischen Rahmenbedingungen operiert.

Um die zuvor genannten Ziele zu erreichen, ist es erforderlich, dass Anbieter von Hochrisiko-KI-Systemen während des gesamten Entwicklungsprozesses eine transparente und detaillierte Dokumentation erstellen. Zu den zu dokumentierenden Elementen zählen Protokolle über die verwendeten Algorithmen, Trainingsdaten sowie sämtliche Änderungen und Aktualisierungen, die am System vorgenommen werden. Die Dokumentation sollte für die Betreiberinnen und Betreiber leicht zugänglich sein und ihnen die Möglichkeit bieten, den Entscheidungsprozess des Systems nachzuvollziehen und zu verstehen.

Des Weiteren sieht Artikel 13 die Bereitstellung von Betriebsanleitungen für Hochrisiko-KI-Systeme vor. Die Anleitungen sollten demnach präzise, vollständig, korrekt und unmissverständlich verfasst sein und in einem für die Betreiber leicht zugänglichen digitalen Format vorliegen. Die Inhalte der Betriebsanleitungen müssen die Bedürfnisse und vorhersehbaren Kenntnisse der Zielbetreiber berücksichtigen und in einer barrierefreien, leicht verständlichen Form präsentiert werden.

Die praktischen Auswirkungen dieses Artikels sind von weitreichender Natur. Die Gewährleistung von Transparenz sowie die Bereitstellung detaillierter Betriebsanleitungen ermöglichen es Betreibern von Hochrisiko-KI-Systemen, fundiertere Entscheidungen zu treffen. Dies ermöglicht eine Evaluierung der Zuverlässigkeit und Genauigkeit der Systemausgaben sowie die Erkennung potenzieller Probleme und unerwünschter Vorurteile. Dies fördert das Vertrauen in die Verwendung von KI-Systemen und hilft dabei, rechtliche und ethische Risiken zu minimieren. Zudem erleichtert die Transparenz im Entwurf die Einhaltung anderer gesetzlicher Anforderungen, wie beispielsweise Datenschutz- und Haftungsbestimmungen. Durch die verbesserte Verständlichkeit der Systemausgaben können Betreiber besser sicherstellen, dass die Rechte und Privatsphäre der Betroffenen geschützt werden.

Beispielhafte Betriebsanleitung gem. Art. 13 Abs. 3 KI-VO

Use Case: „Entwicklung und Einsatz eines Hochrisiko-KI-Systems für die medizinische Diagnoseunterstützung“

Ziel:

Entwicklung und Einsatz eines KI-basierten Systems zur Unterstützung von Ärzten bei der Diagnose und Behandlung von Patienten, unter Berücksichtigung der Anforderungen an Hochrisiko-KI-Systeme gemäß der KI-Verordnung.

Beteiligte:

- Entwicklerteam: Verantwortlich für die Konzeption, Entwicklung und Wartung des Hochrisiko-KI-Systems.
- Medizinisches Fachpersonal: Ärzte und andere medizinische Fachkräfte, die das Hochrisiko-KI-System im klinischen Alltag nutzen.
- Patienten: Individuen, deren medizinische Daten vom Hochrisiko-KI-System verarbeitet werden, um Diagnoseunterstützung zu bieten.

Schritte:

1. **Bedarfsanalyse:** Das Entwicklerteam führt eine gründliche Analyse der Anforderungen und Ziele des medizinischen Fachpersonals durch. Dazu gehören die Identifizierung häufiger oder komplexer Diagnoseprobleme, die Bestimmung relevanter Eingabedaten (z. B. Symptome, Laborwerte, Bildgebung) und die Definition der gewünschten Ausgaben des Systems (z. B. Rangliste möglicher Diagnosen, Empfehlungen für weitere Tests).
2. **Datensammlung und -aufbereitung**: Eine umfangreiche Menge an medizinischen Daten wird gesammelt und aufbereitet, einschließlich elektronischer Gesundheitsakten, Forschungsergebnisse und Fachliteratur. Diese Daten werden strukturiert, gekennzeichnet und so aufbereitet, dass sie für das Training und die Validierung des Hochrisiko-KI-Systems geeignet sind.
3. **Systemdesign und -entwicklung:** Das Hochrisiko-KI-System wird gemäß ethischen und rechtlichen Rahmenbedingungen entworfen und entwickelt. Dazu gehören die Auswahl geeigneter Algorithmen, die Festlegung der Systemarchitektur und die Implementierung von Sicherheitsmaßnahmen. Das System soll Ärzten bei der Diagnose und der Entwicklung von Behandlungsplänen assistieren.
4. **Training und Validierung:** Das Hochrisiko-KI-System wird mithilfe der aufbereiteten Daten trainiert und validiert, um seine Genauigkeit und Robustheit sicherzustellen. Verschiedene Techniken des maschinellen Lernens werden angewendet, und das System wird umfassend getestet, um seine Leistung zu bewerten und mögliche Schwachstellen zu identifizieren.
5. **Integration und Einsatz:** Das trainierte Hochrisiko-KI-System wird in die klinische Arbeitsumgebung integriert, beispielsweise in Form einer Software-Schnittstelle oder eines webbasierten Tools. Ärzte können nun das System bei der Diagnose und Behandlung von Patienten nutzen, wobei das System als unterstützendes Werkzeug dient und die menschliche Entscheidungsfindung ergänzt.
6. **Überwachung, Wartung und Aktualisierung:** Das Entwicklerteam überwacht kontinuierlich die Leistung des Systems im Echtbetrieb. Dazu gehört die Überprüfung der Genauigkeit der Diagnosevorschläge, die Bewertung der Robustheit gegenüber verschiedenen klinischen Szenarien und die Sicherstellung der Cybersicherheit. Regelmäßige Wartung und Aktualisierungen werden durchgeführt, um die Leistung zu verbessern und sich an neue medizinische Erkenntnisse oder Änderungen in den klinischen Leitlinien anzupassen.

Beispielhafte Betriebsanleitung für das Hochrisiko-KI-System zur medizinischen Diagnoseunterstützung:

Titel: Betriebsanleitung für das Hochrisiko-KI-System zur medizinischen Diagnoseunterstützung

Inhaltsverzeichnis:

1. Einführung

1.1 Beteiligte Parteien und ihre Rollen, inkl. Namen und Kontaktangaben des Anbieters

1.2 Zweck und Ziel des Hochrisiko-KI-Systems

1.3 Anwendungsbereich und rechtlicher Rahmen, einschließlich KI-VO

2. Technische Übersicht und Systemarchitektur

2.1 Beschreibung der Systemkomponenten und ihrer Interaktionen

2.2 Datenquellen und -aufbereitung

2.3 Genutzte Algorithmen, Modelle und Trainingsmethoden

3. Leistungsmerkmale und Grenzen

3.1 Zweckbestimmung und Anwendungsfälle

3.2 Genauigkeit, Robustheit und Cybersicherheitsmaßnahmen

3.3 Grenzen der Leistungsfähigkeit und mögliche Auswirkungen auf die Patientenversorgung

3.4 Umgang mit Unsicherheiten und Fehlern

4. Einsatz und Nutzung

4.1 Integration in die klinische Arbeitsumgebung

4.2 Arten von Eingabedaten und Beispielausgaben

4.3 Interpretation der Systemvorschläge und empfohlene Maßnahmen

4.4 Schulungs- und Supportmaterialien für medizinisches Fachpersonal

5. Überwachung, Wartung und Aktualisierungen

5.1 Überwachung der Systemleistung und -genauigkeit

5.2 Protokollierung und Aufbewahrung von Daten

5.3 Häufigkeit und Art von Wartungs- und Aktualisierungsmaßnahmen

5.4 Verfahren bei identifizierten Risiken oder Fehlfunktionen

6. Ethik und Datenschutz

6.1 Maßnahmen zur Gewährleistung von Fairness, Transparenz und Nichtdiskriminierung

6.2 Datenschutzbestimmungen, einschließlich Datenspeicherung, -löschung und -schutz

6.3 Umgang mit potenziellen Risiken für die Gesundheit und Grundrechte von Patienten

7. Menschliche Aufsicht und Verantwortung

7.1 Maßnahmen zur Gewährleistung angemessener menschlicher Aufsicht

7.2 Technische Hilfestellungen für die Interpretation der Systemausgaben

Detaillierte Beschreibung der Abschnitte:

Im Abschnitt „Leistungsmerkmale und Grenzen“ werden die Zweckbestimmung und Anwendungsfälle des Systems erläutert, beispielsweise die Unterstützung bei der Diagnose seltener Krankheiten oder die Erkennung von Interaktionen zwischen Medikamenten. Metriken der Genauigkeit, wie die Sensitivität, Spezifität und Genauigkeit der Diagnosevorschläge, werden bereitgestellt. Mögliche Umstände, die die Leistung beeinflussen könnten, wie seltene Krankheitsverläufe oder unvollständige Patientendaten, werden ebenfalls offengelegt.

Der Abschnitt „Einsatz und Nutzung“ bietet praktische Anleitungen für das medizinische Fachpersonal. Er beschreibt die Arten von Eingabedaten, die benötigt werden (z. B. Symptome, Laborwerte), und stellt Beispielausgaben des Systems bereit. Es werden Richtlinien für die Interpretation der Systemvorschläge gegeben, einschließlich der Berücksichtigung von Unsicherheiten und der Gewichtung verschiedener Faktoren. Schulungs- und Supportmaterialien, wie Benutzerhandbücher oder Tutorials, werden ebenfalls in diesem Abschnitt bereitgestellt.

Der Abschnitt „Ethik und Datenschutz“ behandelt ausführlich die Maßnahmen, die ergriffen wurden, um ethische Prinzipien zu wahren. Dazu gehören Fairness bei der Datenverarbeitung, Transparenz hinsichtlich der Systemfunktionen und der Entscheidungsfindung sowie der Schutz vor Diskriminierung. Datenschutzbestimmungen werden detailliert beschrieben, einschließlich der Art und Dauer der Datenspeicherung, der Rechte der Patienten im Zusammenhang mit ihren Daten und der Einhaltung relevanter Datenschutzbestimmungen.

Die Betriebsanleitung bietet eine umfassende und detaillierte Anleitung für den Einsatz und die Wartung des Hochrisiko-KI-Systems zur medizinischen Diagnoseunterstützung. Sie gewährleistet Transparenz, ethisches Verhalten und die Einhaltung rechtlicher Rahmenbedingungen, während sie gleichzeitig praktische Informationen für die effektive Nutzung des Systems im klinischen Alltag bietet.

Zwischenergebnis

Sind die Transparenzpflichten erfüllt, ist mit **Schritt 4.1.7** fortzufahren.

Schritt 4.1.7: Wie ist die menschliche Aufsicht zu gestalten?

Prof. Dr. Heinz-Uwe Dettling (Ernst & Young GmbH Wirtschaftsprüfungsgesellschaft), Jan-Dierk Schaal, LL.M (University of Melbourne) (SKW Schwarz Rechtsanwälte), Alexander Schmalenberger (Taylor Wessing Partnerschaftsgesellschaft mbB)

Art. 14 KI-VO regelt, wie die menschliche Aufsicht zu gestalten ist.

Relevante(r) Artikel:

Art. 14, 26

Relevante(r) ErwG:

66, 73

Konkretisierungsbedürftig:

Betreibereigenschaft

Qualifikationserfordernis

(1) Hochrisiko-KI-Systeme werden so konzipiert und entwickelt, dass sie während der Dauer ihrer Verwendung — auch mit geeigneten Instrumenten einer Mensch-Maschine-Schnittstelle — von natürlichen Personen wirksam beaufsichtigt werden können.

(2) Die menschliche Aufsicht dient der Verhinderung oder Minimierung der Risiken für Gesundheit, Sicherheit oder Grundrechte, die entstehen können, wenn ein Hochrisiko-KI-System im Einklang mit seiner Zweckbestimmung oder im Rahmen einer vernünftigerweise vorhersehbaren Fehlanwendung verwendet wird, insbesondere wenn solche Risiken trotz der Einhaltung anderer Anforderungen dieses Abschnitts fortbestehen.

(3) Die Aufsichtsmaßnahmen müssen den Risiken, dem Grad der Autonomie und dem Kontext der Nutzung des Hochrisiko-KI-Systems angemessen sein und werden durch eine oder beide der folgenden Arten von Vorkehrungen gewährleistet:

a) Vorkehrungen, die vor dem Inverkehrbringen oder der Inbetriebnahme vom Anbieter bestimmt und, sofern technisch machbar, in das Hochrisiko-KI-System eingebaut werden;

i) Vorkehrungen, die vor dem Inverkehrbringen oder der Inbetriebnahme des Hochrisiko-KI-Systems vom Anbieter bestimmt werden und dazu geeignet sind, vom Betreiber umgesetzt zu werden.

1. Sinn und Zweck des Erfordernisses der menschlichen Aufsicht

Das Erfordernis der menschlichen Aufsicht durch Anbieter und Betreiber von KI-Systemen ist eines der zentralen Instrumente der KI-Verordnung, um das Ziel einer vertrauenswürdigen KI als menschenzentrierte Technologie zu erreichen, die den Menschen dienen und letztlich das menschliche Wohlergehen verbessern soll. Den Hintergrund bildet die Natur von KI-Systemen als mit verschiedenen Graden der Autonomie ausgestatteten Systemen, die bis zu einem gewissen Grad unabhängig von menschlichem Zutun agieren und in der Lage sind, ohne menschliches Eingreifen zu arbeiten, und deren hohe Leistungsfähigkeit zur Erfüllung vielfältiger Funktionen, darunter Vorhersagen, Inhalte, Empfehlungen oder Entscheidungen. Gerade wegen der hohen Leistungsfähigkeit von KI besteht das Risiko eines automatischen oder übermäßigen Vertrauens in die von einem Hochrisiko-KI-System hervorgebrachte Ausgabe (sog. „Automatisierungsbias“) und eine damit einhergehende Verselbstständigung der KI, wenn diese über einen längeren Zeitraum ohne menschliche Kontrolle oder Beeinflussung sich autark weiterentwickelt.

2. Regelungsüberblick und Anwendungsbereich

Die Pflichten der Anbieter in Zusammenhang mit der menschlichen Aufsicht sind insbesondere in Artikel 14 der KI-Verordnung, die Pflichten der Betreiber in Artikel 26 der KI-Verordnung geregelt. Nähere Details zur menschlichen Aufsicht können in harmonisierten Normen gemäß Artikel 40 der KI-Verordnung, in Durchführungsrechtsakten der Kommission zur Festlegung gemeinsamer Spezifikationen gemäß Artikel 41 der KI-Verordnung und in Leitlinien der Kommission gemäß Art. 96 der KI-Verordnung festgelegt werden.

Das Konzept der menschlichen Aufsicht in der KI-Verordnung umfasst die folgenden Kern-Elemente:

- die Betrauung von natürlichen Personen, die über die erforderliche Kompetenz, Ausbildung und Befugnis verfügen, mit der Funktion der menschlichen Aufsicht durch Betreiber von Hochrisiko-KI-Systemen entsprechend der den Systemen beigefügten Betriebsanleitungen (Art. 26 Abs. 1 und 2 der KI-Verordnung)
- Angemessenheit der Aufsichtsmaßnahmen in Hinblick auf die Risiken, den Grad der Autonomie und den Kontext der Nutzung des Hochrisiko-KI-Systems (Art. 14 Abs. 3 KI-Verordnung)
- Zurverfügungstellung des Hochrisiko-KI-Systems durch den Anbieter in einer Weise, dass den natürlichen Personen des Betreibers, denen die menschliche Aufsicht übertragen wurde, angemessen und verhältnismäßig ermöglicht wird (Art. 14 Abs. 4 der KI-Verordnung):
    a) Verständnis und Überwachung des Hochrisiko-KI-Systems
    b) Bewusstsein des Automatisierungsbias;
    c) richtige Interpretation der Ausgabe des Hochrisiko-KI-Systems;
    d) Nichtverwendung des Hochrisiko-KI-Systems oder einzelner seiner Ausgaben in bestimmten Situationen;
- e) Eingriff in den Betrieb oder Stopp des Betriebs des Hochrisiko-KI-Systems („Stopptaste“).

Die Anforderungen der menschlichen Aufsicht bei Hochrisiko-KI-Systemen für Anbieter gemäß Artikel 14 und für Betreiber gemäß Artikel 26 der KI-Verordnung dienen dazu, Risiken für Gesundheit, Sicherheit und Grundrechte wirksam zu mitigieren.

Diese Anforderungen gelten insbesondere für das Risikomanagement, die Qualität und Relevanz der verwendeten Datensätze, die technische Dokumentation, die Aufzeichnungspflichten, die Transparenz und die Bereitstellung von Informationen für die Betreiber sowie die Robustheit, Genauigkeit und Sicherheit der Systeme.

Artikel 14 der KI-Verordnung bezieht sich auf alle Hochrisiko-KI-Systeme.

3. Maßnahmen menschlicher Aufsicht

a) Maßnahmen, die bei Erstellung des KI-Systems zu berücksichtigen sind

Diese Systeme müssen so konzipiert und entwickelt werden, dass sie während ihrer gesamten Einsatzdauer von natürlichen Personen wirksam beaufsichtigt werden können. Bei der Erstellung von Hochrisiko-KI-Systemen müssen Anbieter dafür sicherstellen, dass geeignete technische Maßnahmen (Mensch-Maschine-Schnittstellen) integriert sind, die den menschlichen Aufsichtsprozess ermöglichen und unterstützen. Insbesondere sollten solche Maßnahmen gewährleisten, dass das System integrierten Betriebseinschränkungen unterliegt, über die sich das System selbst nicht hinwegsetzen kann. Es ist außerdem unerlässlich, Mechanismen zu integrieren, um die Aufsichtspersonen zu beraten und zu informieren, damit sie fundierte Entscheidungen darüber treffen können, ob, wann und wie einzugreifen ist, um negative Folgen oder Risiken zu vermeiden, oder das System anzuhalten, wenn es nicht wie beabsichtigt funktioniert. Dabei wird für den Anbieter die Herausforderung bestehen, innerhalb der Betriebsanleitung sämtliche hierfür erforderlichen Informationen in einer Art zusammenzufassen, die der menschlichen Aufsicht das notwendige Verständnis des Systems vermittelt (*explainable AI*). Zudem ist das KI-System zwingend derart auszugestalten, dass in seinen Betrieb korrektiv eingegriffen und es jederzeit ausgeschaltet werden kann (Erwägungsgrund 73). Gerade bei Hochrisiko-KI-Systemen, die ein Sicherheitsbauteil oder Bestandteil eines Produktes sind, ist diese Ausschaltfunktion derart zu gestalten, dass die Sicherheit und Funktionsweise des Produktes auch ohne die weitere Zuarbeit des KI-Systems nicht gefährdet ist.

b) Maßnahmen, die bei Betrieb des KI-Systems zu berücksichtigen sind

> (2) Die Betreiber übertragen natürlichen Personen, die über die erforderliche Kompetenz, Ausbildung und Befugnis verfügen, die menschliche Aufsicht und lassen ihnen die erforderliche Unterstützung zukommen.
>
> (3) Die Pflichten nach den Absätzen 1 und 2 lassen sonstige Pflichten der Betreiber nach Unionsrecht oder nationalem Recht sowie die Freiheit der Betreiber bei der Organisation ihrer eigenen Ressourcen und Tätigkeiten zur Wahrnehmung der vom Anbieter angegebenen Maßnahmen der menschlichen Aufsicht unberührt.

Es muss sichergestellt werden, dass das KI-System auf den menschlichen Bediener reagiert und dass die natürlichen Personen, denen die menschliche Aufsicht übertragen wurde, über die erforderliche Kompetenz, Ausbildung und Befugnis verfügen, um diese Aufgabe wahrzunehmen.

Insoweit stellt das Erfordernis der menschlichen Aufsicht eine wesentliche Grundlage der Qualitätssicherung und der weiteren Pflichten im Rahmen eines Incident Management Systems dar. Art. 26 Abs. 5 KI-Verordnung verpflichtet die Betreiber im Rahmen ihrer eigenen Qualitätssicherung dazu, den Anbieter des Hochrisiko-KI-Systems über Vorfälle, die ein Risiko im Sinne des Art. 79 KI-Verordnung begründen, unverzüglich zu informieren, damit diese ihrer Meldepflicht nach Art. 73 KI-Verordnung nachkommen können. Daneben sind aber auch Betreiber verpflichtet, Einführer oder Händler sowie die zuständigen Marktüberwachungsbehörden über schwerwiegende Vorfälle zu informieren. Die Kenntnis über eben solche Vorfälle wird der Betreiber intern jedoch gerade durch eine effektive menschliche Aufsicht erlangen.

In dem Zusammenhang wird sich häufig die Frage stellen, wer Betreiber des KI-Systems ist, mithin die menschliche Aufsicht sicherstellen muss. So stellt sich beispielsweise bei einem mit KI betriebenen Kraftfahrzeug die Frage, wie die jeweiligen Rollen zuzuordnen sind. Anbieter der im Fahrzeug verbauten KI wird wohl in der Regel der Hersteller sein. Größere Schwierigkeiten bereiten die Zuordnung der Betreiberrolle. Diese wird häufig dem Betrieb zuzuordnen sein, in dessen geschäftlichen Umfeld das Fahrzeug eingesetzt wird, mithin der Halter oder auch Leasingnehmer. Wird die menschliche Aufsicht dann durch den Fahrer ausgeübt, mit allen sich daraus ergebenden Folgen, oder ist dieser lediglich als Bediener zu qualifizieren und der Betreiber kann die menschliche Aufsicht an eine weitere Person delegieren, die dann allerdings – wenn sie selbst nicht zugleich der Fahrer ist – auf die Erfahrungen des Fahrers beim Betrieb des KI-Systems zugreifen können müsste? Diese Rollenverteilung gilt es seitens des Betreibers im Rahmen seines Quality-Management-

Systems zu konzeptionieren und festzulegen. Von besonderer Relevanz ist dies für die sich aus der Rollenzuordnung ergebenen Anforderungen an die Kompetenz und Ausbildung der aufsichtsverantwortlichen natürlichen Person. Da die Betreiber typischerweise KI-Laien sind, andererseits aber so gut wie jedes Unternehmen als Betreiber in Betracht kommt, können hier keine hohen Anforderungen gestellt und insbesondere keine Hochschul-IT-Ausbildung verlangt werden. Voraussichtlich dürften entsprechende Schulungen genügen.

Ungeklärt ist in dem Zusammenhang zudem, ob die aufsichtsverantwortliche natürliche Person ein weisungsgebundener Mitarbeiter des Betreibers sein muss oder auch ein freier Mitarbeiter oder ein Mitarbeiter beispielsweise eines auf menschliche Aufsicht spezialisierten Dienstleisters sein kann. Näheres zu den Anforderungen an die aufsichtsverantwortliche natürliche Person sollte in Dokumenten der Kommission festgelegt werden.

i) Maßnahmenkatalog des Art. 14 Abs. 4 KI-VO

(4) Für die Zwecke der Durchführung der Absätze 1, 2 und 3 wird das Hochrisiko-KI-System dem Betreiber so zur Verfügung gestellt, dass die natürlichen Personen, denen die menschliche Aufsicht übertragen wurde, angemessen und verhältnismäßig in der Lage sind,

a) die einschlägigen Fähigkeiten und Grenzen des Hochrisiko-KI-Systems angemessen zu verstehen und seinen Betrieb ordnungsgemäß zu überwachen, einschließlich in Bezug auf das Erkennen und Beheben von Anomalien, Fehlfunktionen und unerwarteter Leistung;

b) sich einer möglichen Neigung zu einem automatischen oder übermäßigen Vertrauen in die von einem Hochrisiko-KI-System hervorgebrachte Ausgabe („Automatisierungsbias“) bewusst zu bleiben, insbesondere wenn Hochrisiko-KI-Systeme Informationen oder Empfehlungen ausgeben, auf deren Grundlage natürliche Personen Entscheidungen treffen;

c) die Ausgabe des Hochrisiko-KI-Systems richtig zu interpretieren, wobei beispielsweise die vorhandenen Interpretationsinstrumente und -methoden zu berücksichtigen sind;

d) in einer bestimmten Situation zu beschließen, das Hochrisiko-KI-System nicht zu verwenden oder die Ausgabe des Hochrisiko-KI-Systems außer Acht zu lassen, außer Kraft zu setzen oder rückgängig zu machen;

e) in den Betrieb des Hochrisiko-KI-Systems einzugreifen oder den Systembetrieb mit einer „Stopptaste“ oder einem ähnlichen Verfahren zu unterbrechen, was dem System ermöglicht, in einem sicheren Zustand zum Stillstand zu kommen

Artikel 14 Absatz 4 der KI-Verordnung fordert, dass die menschliche Aufsicht die folgenden Aspekte umfasst:

- Verstehen der Fähigkeiten und Grenzen des Hochrisiko-KI-Systems sowie der daraus resultierenden Risiken.
- Überwachung des Betriebs des Systems, um Anomalien, Fehlfunktionen und unerwartete Leistungen zu erkennen und zu beheben.
- Vermeidung eines automatischen oder übermäßigen Vertrauens in die Ausgaben des KI-Systems.
- Eignung zum richtigen Verständnis und Einschätzung der Interpretationsspielräume hinsichtlich der Ausgabe des KI-Systems.

- Geeignete Ausbildung und Befugnisse der Aufsichtspersonen, um die Überwachung effektiv durchzuführen und die Entscheidung zu treffen, in den Betrieb des KI-Systems einzugreifen oder dieses abzuschalten.

Insoweit ist Betreibern von Hochrisiko-KI-Systemen zu empfehlen, bereits Vertretungsnotwendigkeiten oder eine freiwillige Einführung eines 4-Augen-Prinzips im Rahmen der Konzeption der menschlichen Aufsicht zu berücksichtigen.

Die enge Verzahnung von Anbieter- und Betreiberpflichten in Zusammenhang mit der menschlichen Aufsicht kommt auch in den Pflichten zum Inhalt der Betriebsanleitung zum Ausdruck. Nach Art. 13 Abs. 3 Buchstabe d KI-VO muss die Betriebsanleitung auch jeweils produktspezifische Informationen zu den in Art. 14 KI-VO genannten Maßnahmen zur Gewährleistung der menschlichen Aufsicht enthalten. Dies umfasst auch Informationen über die technischen Maßnahmen, die getroffen wurden, um den Betreibern die Interpretation der Ausgaben von Hochrisiko-KI-Systemen zu erleichtern. Die Informationen zur menschlichen Aufsicht in den Betriebsanleitungen sind das für die Aufsichtsperson des Betreibers maßgebliche Dokument.

ii) Sondervorschriften für Systeme zur biometrischen Fernidentifizierung (Art. 14 Abs. 5 KI-VO)

Angesichts der bedeutenden Konsequenzen für Personen im Falle eines falschen Treffers durch bestimmte biometrische Fernidentifizierungssysteme ist es angezeigt, für diese Systeme eine verstärkte Anforderung im Hinblick auf die menschliche Aufsicht vorzusehen. Der Betreiber darf keine Maßnahmen oder Entscheidungen aufgrund des vom System hervorgebrachten Identifizierungsergebnisses treffen, solange dies nicht von mindestens zwei natürlichen Personen getrennt überprüft und bestätigt wurde. Diese Personen müssen hinreichend im Umgang mit dem KI-System geschult sein und über die Kompetenz und Autorität verfügen, die Ausgabe des KI-Systems im Rahmen der Entscheidung nicht zu berücksichtigen. Diese Personen könnten von einer oder mehreren Einrichtungen stammen und die Person umfassen, die das System bedient oder verwendet. Diese Anforderung sollte keine unnötigen Belastungen oder Verzögerungen mit sich bringen, und es könnte ausreichen, dass die getrennten Überprüfungen durch die verschiedenen Personen automatisch in die vom System erzeugten Protokolle aufgenommen werden (Erwägungsgrund 73). Angesichts der Besonderheiten der Bereiche Strafverfolgung, Migration, Grenzkontrolle und Asyl gilt dieses Erfordernis in diesen Bereichen nicht, wenn die Geltung dieser Anforderung nach Unionsrecht oder nationalem Recht unverhältnismäßig ist (vgl. Erwägungsgrund 73).

4. Verantwortlichkeit der Beaufsichtigenden

Eine wichtige Frage, die sich im Kontext der menschlichen Aufsicht stellt, ist die rechtliche Verantwortlichkeit der Aufsichtspersonen für Fehlentscheidungen der KI. Es muss geklärt werden, in welchem Umfang die natürlichen Personen, die die Aufsicht über das KI-System ausüben, für die Handlungen und Entscheidungen des Systems rechtlich zur Verantwortung gezogen werden können. Dies umfasst sowohl zivilrechtliche als auch strafrechtliche Aspekte und erfordert klare Regelungen, um die Verantwortlichkeiten und Haftungen eindeutig zu definieren und Missbrauch oder unangemessene Belastungen der Aufsichtspersonen zu vermeiden. Derzeit richtet sich die Verantwortlichkeit nach den allgemeinen Vorschriften,

mithin im Angestelltenverhältnis nach dem Grundsatz des innerbetrieblichen Schadensausgleichs und im Beamtenverhältnis nach den beamtenrechtlichen Regelungen.

Die Haftung für Schäden Dritter richtet sich ebenfalls nach den allgemeinen Regelungen. Insoweit ist vorausgehend die Frage zu klären, ob der Betreiber gegenüber dem Dritten haftet. Für den Betreiber wird insoweit eine Haftung nach dem Produkthaftungsrecht in der Regel nicht in Betracht kommen, da ein Versagen der menschlichen Aufsicht kein Produktfehler ist. Allerdings kommt eine Haftung des Anbieters nach Produkthaftungsrecht in Betracht, wenn er nicht oder unzureichend die erforderlichen Maßnahmen zu Ermöglichung der menschlichen Aufsicht integriert hat.

Zusammenfassung

Die menschliche Aufsicht gemäß Artikel 14 und 26 der KI-Verordnung ist ein wesentlicher Bestandteil der Sicherheits- und Risikomanagementstrategien für Hochrisiko-KI-Systeme. Durch die Implementierung technischer und organisatorischer Maßnahmen sowohl bei der Erstellung als auch beim Betrieb dieser Systeme kann gewährleistet werden, dass sie sicher und im Einklang mit den geltenden Vorschriften verwendet werden. Besondere Aufmerksamkeit gilt dabei den Systemen zur biometrischen Fernidentifizierung, für die strengere Überprüfungsmechanismen erforderlich sind, um die Rechte der betroffenen Personen zu schützen. Die Verantwortlichkeit der Aufsichtspersonen sollte vom Betreiber im Innenverhältnis klar geregelt werden, um Rechtssicherheit und effektiven Schutz zu gewährleisten.

Zwischenergebnis

Ist die menschliche Aufsicht gewährleistet, geht es weiter mit **Schritt 4.1.8.**

Schritt 4.1.8: Wie sind Genauigkeit, Robustheit und Cybersicherheit auszugestalten?

Eric Behrendt (TÜV Informationstechnik GmbH), Annegrit Seyerlein-Klug (neurocat GmbH), Ferdinand Schwarz (SKW Schwarz Rechtsanwälte)

Relevante(r) Artikel:

Art 15
Art. 42 Abs. 2
Art. 9

Relevante(r) ErwG:

74, 75, 76, 77, 114, 115

Konkretisierungsbedürftig:

Entwicklung von Standards zur Erfüllung der Anforderungen an die Sicherheit von KI-Systemen durch CEN-CENELEC.

(1) Hochrisiko-KI-Systeme werden so konzipiert und entwickelt, dass sie ein angemessenes Maß an Genauigkeit, Robustheit und Cybersicherheit erreichen und in dieser Hinsicht während ihres gesamten Lebenszyklus beständig funktionieren.

(2) Um die technischen Aspekte der Art und Weise der Messung des angemessenen Maßes an Genauigkeit und Robustheit gemäß Absatz 1 und anderer einschlägiger Leistungsmetriken anzugehen, fördert die Kommission in Zusammenarbeit mit einschlägigen Interessenträgern und Organisationen wie Metrologie- und Benchmarking-Behörden gegebenenfalls die Entwicklung von Benchmarks und Messmethoden.

(3) Die Maße an Genauigkeit und die relevanten Genauigkeitsmetriken von Hochrisiko-KI-Systemen werden in den ihnen beigefügten Betriebsanleitungen angegeben.

(4) Hochrisiko-KI-Systeme müssen so widerstandsfähig wie möglich gegenüber Fehlern, Störungen oder Unstimmigkeiten sein, die innerhalb des Systems oder der Umgebung, in der das System betrieben wird, insbesondere wegen seiner Interaktion mit natürlichen Personen oder anderen Systemen, auftreten können. In diesem Zusammenhang sind technische und organisatorische Maßnahmen zu ergreifen. Die Robustheit von Hochrisiko-KI-Systemen kann durch technische Redundanz erreicht werden, was auch Sicherungs- oder Störungssicherheitspläne umfassen kann.

Art. 15 KI-VO normiert zwingende Anforderungen an die Cybersicherheit, Robustheit und Genauigkeit bei **Hochrisiko-KI-Systemen**, die von den betroffenen Regelungsadressaten im Rahmen der KI-Verordnung zu erfüllen sind. Die Vorschrift enthält somit den zentralen Pflichtenkatalog in Bezug auf die Abwehr von KI-spezifischen Sicherheitsrisiken.

> Hochrisiko-KI-Systeme, die nach dem Inverkehrbringen oder der Inbetriebnahme weiterhin dazulernen, sind so zu entwickeln, dass das Risiko möglicherweise verzerrter Ausgaben, die künftige Vorgänge beeinflussen („Rückkopplungsschleifen“), beseitigt oder so gering wie möglich gehalten wird und sichergestellt wird, dass auf solche Rückkopplungsschleifen angemessen mit geeigneten Risikominderungsmaßnahmen eingegangen wird.
>
> (5) Hochrisiko-KI-Systeme müssen widerstandsfähig gegen Versuche unbefugter Dritter sein, ihre Verwendung, Ausgaben oder Leistung durch Ausnutzung von Systemschwachstellen zu verändern. Die technischen Lösungen zur Gewährleistung der Cybersicherheit von Hochrisiko-KI-Systemen müssen den jeweiligen Umständen und Risiken angemessen sein. Die technischen Lösungen für den Umgang mit KI-spezifischen Schwachstellen umfassen gegebenenfalls Maßnahmen, um Angriffe, mit denen versucht wird, eine Manipulation des Trainingsdatensatzes („data poisoning“) oder vortrainierter Komponenten, die beim Training verwendet werden („model poisoning“), vorzunehmen, Eingabedaten, die das KI-Modell zu Fehlern verleiten sollen („adversarial examples“ oder „model evasions“), Angriffe auf vertrauliche Daten oder Modellmängel zu verhüten, zu erkennen, darauf zu reagieren, sie zu beseitigen und zu kontrollieren.

Die Regelung darf in diesem Kontext allerdings **nicht isoliert** betrachtet werden. Vielmehr ist die technisch-organisatorische **Sicherheit** des KI-Systems eine **elementare Voraussetzung**, zur Gewährleistung von **Vertrauenswürdigkeit** sowie der **Sicherheit** (eng. „Safety“) und **Unversehrtheit** der fundamentalen (Grund-)Rechte der (End-)Nutzer (vgl. nur Art. 1 Abs. 1 KI-VO). „Sicherheit“ hat in Terminologie und Systematik der KI-VO somit eine doppelte Bedeutung, da sie zugleich Ziel („Safety der KI-Nutzer“) und Mittel („Sicherheit des KI-Systems“) der KI-VO betrifft. Deshalb ist stets eine ganzheitliche, systemübergreifende und kontinuierliche Betrachtung von „KI-Sicherheit“ nach Maßgabe des risikobasierten Ansatzes der KI-VO geboten, die sich maßgeblich an den genannten Schutzzielen der KI-VO orientiert.

Für **KI-Modelle**, insbesondere die in Art. 51 ff. KI-VO adressierten GPAI-Modell, gelten teilweise abweichende Regelungen im Hinblick auf ihre (Cyber-)Sicherheit. Diese sind *per se* nicht Gegenstand dieses Leitfadens. Dennoch darf das zugrundeliegende KI-Modell bei der Umsetzung der KI-Systemsicherheit keinesfalls ausgeklammert werden. Es ist Komponente des KI-Systems und damit ebenso notwendiger Bestandteil jeder Risikobetrachtung, genauso wie die ihm zugrundeliegende ITK-Infrastruktur.

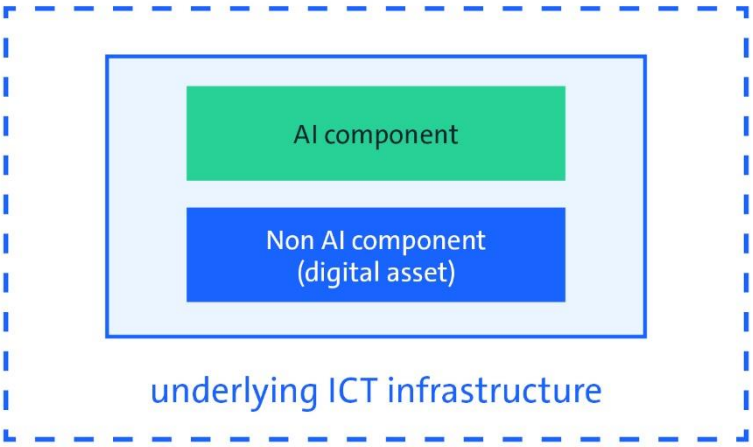


***Abbildung**: Relevante Komponenten bei der Umsetzung von KI-Sicherheit*

Schritt 1: Welche Begriffe sind relevant und wie sind diese im (Gesamt-)Kontext einzuordnen?

Bevor die konkreten Anforderungen näher betrachtet werden, sind im **ersten Schritt** die in Art. 15 der KI-VO herangezogenen **Begriffe** näher einzuordnen. Im Bereich der (Cyber-)Sicherheit liefert die KI-VO (fast) keine Definitionen, sodass sich das Begriffsverständnis nur durch Auslegung und Rückgriff aus allgemeinen Grundsätzen ermitteln lässt:

- **Cybersicherheit** wird vom Verordnungsgeber ganz offenbar eng verstanden und bezieht sich maßgeblich auf den Schutz vor externen Angriffen und Cyberattacken (vgl. etwa Art. 15 Abs. 5 KI-VO und Erwägungsgrund 76). Dementgegen wird „Cybersicherheit" im herkömmlichen Sinn grds. weiter verstanden (wie etwa im Rahmen des Cyber Security Act, vgl. Art. 2 Nr. 1 CSA) und umfasst daher den generellen Schutz eines Systems in einem Netzwerk (einschließlich der darin verarbeiteten Daten) vor Bedrohungen aller Art. Damit steht jedenfalls fest, dass „Cybersicherheit" im Sinne der KI-VO nicht als „Oberbegriff" für (informations-)technische Sicherheit eingesetzt wird, sondern ein eigenständiges Schutzziel in Bezug auf KI-Systeme darstellt. Da in der Praxis (auch in Bezug auf KI-Sicherheit) das „weitere" Verständnis gebräuchlich ist, kann dies für Verwirrungen sorgen und sollte daher im Blick behalten werden.

- **Robustheit** betrifft die ordnungsgemäße Funktionsfähigkeit des KI-Systems innerhalb der system- bzw. modelseitig festgelegten Grenzen. Das Kriterium weist eine Schnittmenge zum Schutzziel der Verfügbarkeit im Rahmen der „klassischen" IT-Sicherheit auf und bezweckt insbesondere (aber auch nicht nur) die Gewährleistung eines störungsfreien Betriebs des KI-Systems in seiner Einsatzumgebung.

- **Genauigkeit** wiederum ist sehr KI-spezifisch. Sie ist ein messbares Qualitätsmerkmal. Sie betrifft die ordnungsgemäße Funktionsfähigkeit des KI-Systems im Hinblick auf die Genauigkeit, Reproduzierbarkeit und Verlässlichkeit der Ausgaben und Entscheidungen eines KI-Systems bzw. der KI-Komponente. Sie ist teilweise verknüpft mit der Integrität. Die Qualität der (Trainings-)Daten wird durch die Gewährleistung von Integrität bzw. Richtigkeit der Datenquellen gestärkt. Mangelnde Datenqualität ist ein Sicherheitsrisiko.

Diese Anforderungen an die Sicherheit des KI-Systems können im Einzelnen nicht immer trennscharf voneinander abgegrenzt werden. Vielmehr greifen sie häufig ineinander über und bedingen sich teilweise gegenseitig. Im Ergebnis sind sie kumulative Voraussetzung für die Gewährleistung der **Vertraulichkeit** von KI-Systemen. Aus Sicht der Verantwortlichen ist mithin stets darauf zu achten, dass den Begriffen je nach Einzelfall und Kontext eine unterschiedliche Bedeutung und Schutzrichtung zukommen kann.

Schritt 2: Welche Vorgaben der KI-VO sind in Bezug auf die Sicherheit von KI-Systeme zu beachten?

In einem **zweiten Schritt** sind aus Sicht des Pflichtadressaten nun alle sicherheitsbezogenen Vorschriften zu identifizieren. Im Rahmen der KI-VO sind aufgrund des integrativen und ganzheitlichen Bewertungsprozesses insbesondere die folgenden Vorschriften von Relevanz:

- **Art. 15 KI-VO** regelt – wie oben bereits dargestellt – die wesentlichen und KI-spezifischen Anforderungen zur informationstechnischen Sicherheit eines KI-Systems: Cybersicherheit, Robustheit und Genauigkeit (zu deren **Ausgestaltung** weiter unten bei **Schritt 5**).
- Die Sicherheit des KI-Systems ist integraler Bestandteil eines **Risikomanagementsystems** nach **Art. 9 KI-VO** (→ siehe dazu Kapitel 7). Art. 9 Abs. 2 lit. a) KI-VO schreibt insoweit ausdrücklich vor, dass die Ermittlung, Analyse und anschließende Bewertung der für das jeweilige KI-System bestehenden Risiken in Bezug auf **die Sicherheit** zu erfolgen haben. Die Identifikation geeigneter technischer Maßnahmen zur Vermeidung dieser Risiken knüpft an diese Risikobewertung an (siehe bspw. Art. 5 Abs. 2 lit. a) KI-VO).
- Die Vorgaben zur **Data-Governance** gem. **Art. 10 KI-VO** (→ siehe Kap. 19) stehen in unmittelbarer Beziehung zur **Genauigkeit** und Robustheit i.S.v. Art. 15 KI-VO. Eine vernünftige Data-Governance wird als zentraler Faktor betrachtet, um sicherzustellen, dass das KI-System bestimmungsgemäße Ergebnisse hervorbringt (vgl. Erwägungsgrund 67). Sie sind bei Umsetzung der Anforderungen insb. nach Art. 15 KI-VO zu berücksichtigen.
- Das **Transparenzprinzip** ist im Kontext von Sicherheit ebenfalls relevant. Transparenz „vermittelt" die (Cyber-)Sicherheit des KI-Systems gegenüber dem Endnutzer. Indem der Verantwortliche entlang der Wertschöpfungskette über die getroffenen Maßnahmen zur Sicherheit des KI-Systems umfassend **informiert,** wird zudem die **Kontrolle** und die **Aktualität** der Maßnahmen (mittelbar) sichergestellt. Regelungsadressaten haben hierzu Angaben in der Bedienungsanleitung bereitzustellen, vgl. **Art. 13 Abs. 3 lit. b) ii) KI-VO**.
- **Protokollierungspflichten** in **Art. 12** KI-VO (dazu ausf. Kapitel 10) sind im Kontext von Sicherheit bzw. Sicherheitsmanagement von wesentlicher Bedeutung. Sie gewährleisten, dass das Funktionieren des KI-Systems zurückverfolgbar ist. Viele der hier anerkannten Sicherheitsmanagementkontrollstandards (z. B. ISO 27000) adressieren die Bedeutung von Event-Logs für die Zwecke der Informationssicherheit und der Cyberabwehr.

Für KI-Systeme **ohne hohes Risiko** ist eine freiwillige Einhaltung der genannten Vorgaben im Wege der Ko- bzw. Selbstregulierung eröffnet (→ siehe Kapitel 18). Hierzu werden **Verhaltenskodizes** auch zu den soeben genannten sicherheitsbezogenen Anforderungen erstellt, an denen sich Unternehmen orientieren können. Dies kann gerade dann sinnvoll

sein, wenn sich bereits andeutet, dass das KI-System künftig zu einem Hochrisiko-KI-System „hochklassifiziert“ werden könnte.

Schritt 3: Welche Standards bzw. Normen können im Hinblick auf die Sicherheit von KI-Systemen herangezogen werden?

Die Umsetzung von (Cyber-)Sicherheit ist niemals nur ein KI-spezifisches Thema. KI-Systeme werden in der Regel in eine IT-Umgebung bzw. –Anwendung oder digitale Produkte eingebettet, die bei der Evaluierung und Umsetzung erforderlicher Sicherheitsmaßnahmen keinesfalls außen vor bleiben dürfen. In Bezug auf das gesamte IT-System können daher zahlreiche ergänzende regulatorische (Cyber-)Sicherheitsvorgaben zu beachten sein.

Es ist daher durchaus empfehlenswert, vorab ein „**Mapping**“ der geltenden Regularien im Sicherheitsbereich für das KI-System durchzuführen und dieses stetig zu aktualisieren (wie es etwa im Rahmen der ISO 27001 Zertifizierung zur Compliance vorgeschrieben ist). Hierdurch wird nicht nur die eine umfassende Compliance mit den geltenden Vorschriften gewährleistet. Vielmehr wird gleichzeitig eine möglichst frühzeitige **Harmonisierung** mit „überlappenden“ gesetzlichen Anforderungen sichergestellt. Ein Großteil der hier relevanten Regularien bietet hierfür Harmonisierungsinstrumente, so auch die KI-VO (vgl. nur Erwägungsgrund 124).

Für das KI-System relevante (cyber-)sicherheitsrechtliche Vorgaben können – wobei es sich gerade *nicht* um eine abschließende Aufzählung handelt – insbesondere enthalten:

- Der **CRA** enthält einen umfassenden Pflichtenkatalog im Hinblick auf die Cybersicherheit von Produkten mit digitalen Elementen. Erfasst werden hierbei letztlich alle Software- und Hardwareprodukte (ausgenommen sind grundsätzlich nur SaaS- und Cloud-Services). Da KI-Systeme demgemäß häufig dem Geltungsbereich des CRA unterfallen dürften, sind zusätzlich zur KI-VO die dort geregelten Anforderungen an die Cybersicherheit des Produkts zu berücksichtigen. Der Gesetzgeber eröffnet hier allerdings bestimmte Harmonisierungsinstrumente: Für Hochrisiko-KI-Systeme, die die Anforderungen des CRA erfüllen, gilt eine Vermutung, dass diese mit den Cybersicherheitsanforderungen in Art. 15 KI-VO in Einklang stehen. Die Vermutung gilt jedoch nur für das Kriterium der Cybersicherheit, nicht für die weiteren relevanten Anforderungen der KI-VO, wie insbesondere der Genauigkeit und Robustheit gemäß Art. 15 KI-VO. Zudem gelten weitere Einschränkungen bei bestimmten kritischen digitalen Produkten.

- Der **CSA** schafft einen EU-weiten einheitlichen Zertifizierungsrahmen für die Cybersicherheit bei ITK-Dienstleistungen und -Produkten. Da der Großteil von KI-Systemen von diesem Rahmen erfasst wird, können Unternehmen eine entsprechende Zertifizierung für das KI-Vorhaben in Betracht ziehen. Auch dies kann wiederum eine Konformitätsvermutung begründen: Gemäß Art. 42 Abs. 2 KI-VO besteht durch eine Zertifizierung eines KI-Systems unter dem CSA die Vermutung der Konformität mit den *Cybersicherheits*anforderungen der KI-VO. Auch hier gilt wiederum: Diese bezieht sich bloß auf die Teilanforderung der Cybersicherheit; die anderen Pflichten nach Art. 15 KI-VO bleiben unberührt (vgl. Erwägungsgrund 77).

- Die **NIS-2-RL** enthält bestimmte (Cyber-)Sicherheitspflichten für Anbieter wesentlicher Dienste. Wird das KI-System als Komponente eines solchen wesentlichen Dienstes eingesetzt, können die Vorschriften der NIS-2-RL bzw. der nationalen Umsetzungsgesetze

ergänzende Vorgaben enthalten. Da der Anwendungsbereich der NIS-RL nicht unterschätzt werden darf, sollten KI-Akteure auch diese Vorgaben im Blick behalten.

- Wird das KI-System im Rahmen einer **kritischen Infrastruktur** (KRITIS) eingesetzt (bspw. im Gesundheitswesen oder in der Energieversorgung), können zusätzlich zu den angeführten Rechtsakten ggf. weitere nationale Pflichten nach dem IT-Sicherheitsgesetz gelten, insbesondere als Teil der besonderen Pflichten nach dem **BSIG**.
- Die **DSGVO** enthält insbesondere in **Art. 32** konkrete Anforderungen an die Sicherheit der Datenverarbeitung, was ebenfalls zu zusätzlichen Pflichten für das verantwortliche Unternehmen führen kann, soweit das KI-System **personenbezogene Daten** verarbeitet.
- Handelt es sich um sektorale KI-Systeme, können weitere **sektorspezifische** (Cyber-)Sicherheitsvorgaben Anwendung finden. So können sich z. B. für eine Hochrisiko-KI-Anwendung in der medizinischen Diagnostik besondere Vorgaben aus dem **MPR** ergeben. Für KI-Anwendungen im Finanzbereich, etwa ein Tool zur Analyse der Kreditwürdigkeit, können sich aus dem **DORA** ergeben.

Schritt 4: Welche Vorgaben außerhalb der KI-VO sind in Bezug auf die Sicherheit von KI-Systemen relevant?

Im vierten Schritt sollte aus Unternehmenssicht geprüft werden, ob eine Compliance mit den (cyber-)sicherheitsbezogenen Anforderungen der KI-VO (und ggf. auch anderer flankierenden Vorgaben – siehe soeben) durch die Einhaltung von Standards möglich ist. Die KI-VO sieht hierfür insbesondere das Instrument der harmonisierten Normen vor (→ siehe Kap. 19). Die Erfüllung und Einhaltung dieser Standards führt zu einer sog. Konformitätsvermutung.

Für die oben skizzierten (cyber-)sicherheitsrelevanten Vorgaben und Verfahrensschritte in Bezug auf KI-Systeme (etwa Schaffung eines Risikomanagements nach Art. 9 KI-VO, die Anforderungen des Art. 15 KI-VO) werden derzeit im Zuge eines Normungsauftrags durch CEN-CENELEC harmonisierte Standards erarbeitet. Allerdings sind hier noch viele Fragen ungeklärt und es bestehen zahlreiche Abgrenzungsschwierigkeiten zu bestehenden Standards.

Es bietet es sich daher an, zunächst auf bereits vorhandenen Standards aufzusetzen. Die Methodik der Risikobewertung und des Managements für Cybersecurity ist in Standards beschrieben und entspricht grundsätzlich den Anforderungen der KI-VO. Zu achten ist darauf, wie die KI-Systeme dort zu berücksichtigen sind. Auch die Umsetzung der klassischen Cybersecurity Maßnahmen wie beispielsweise in ISO/IEC 27 Serien oder IEC 62443 oder sind eine passende Vorbereitung. Die Anpassungen an die harmonisierten Standards zur Sicherheit von KI-Systemen sollten in der Folge dann mit wenig Aufwand möglich sein. Daher können von den Regelungsadressaten beispielsweise schon jetzt herangezogen werden:

- ISO/IEC-Standards zu KI (abrufbar hier) v.a. im Hinblick auf das Risikomanagement
- Übersicht der ENISA zur Standardisierungsabdeckung im Bereich KI (abrufbar hier)
- Guidelines bzw. Studien zur Sicherheit von KI(-Modellen) des BSI
- Guidelines bzw. Studien des ENISA

Schritt 5: Rechtskonforme Umsetzung der KI-spezifischen Sicherheitsanforderungen, insbesondere aus Art. 15 KI-VO

Die Anforderungen an die Sicherheit eines KI-Systems müssen der systemspezifischen Anwendung und den Betriebsumständen sowie sich daraus ergebenden Risiken angemessen sein (vgl. Erwägungsgrund 74). Dieser Implementierungsprozess ist stark einzelfallgeprägt und erfordert umfassende konkret KI-systembezogene Vorarbeiten durch das jeweilige mit deren Umsetzung betraute Team. Erforderlich ist eine kontinuierliche Bewertung, Umsetzung, und Aktualisierung der Sicherheitsmaßnahmen für den gesamten Lebenszyklus des Systems.

Art. 15 Abs. 1 KI-VO:

„Hochrisiko-KI-Systeme werden so konzipiert und entwickelt, dass sie ein angemessenes Maß an Genauigkeit, Robustheit und Cybersicherheit erreichen und in dieser Hinsicht während ihres gesamten Lebenszyklus beständig funktionieren."

Risikoanalyse als Ausgangspunkt

Die **initiale Risikobewertung** ist Ausgangspunkt für sämtliche (cyber-)sicherheitsbezogene Maßnahmen im Rahmen der KI-Verordnung (vgl. auch Erwägungsgrund 76 und 77).

Im Rahmen der sicherheitsbezogenen Risikoanalyse ist zu berücksichtigen, dass bestimmte KI-Systeme im Bereich des maschinellen Lernens einzigartige KI-spezifische Merkmale im **Unterschied zu klassischen IT-Systemen** aufweisen. Diese ermöglichen neuartige Angriffe, die im Design („security by design"; „privacy by design"), der Entwicklung, im operativen Betrieb, mit und ohne direkten Zugriff über den Life-Cycle, zu berücksichtigen sind. KI-Systeme sind aber auch IT-Systeme, sodass klassische Sicherheitsmaßnahmen ebenso geboten sind (siehe sogleich).

Darüber hinaus ist zu berücksichtigen, dass KI-Systeme in der Praxis nicht isoliert, sondern in häufig in ein umfassenderes IT-System eingebettet sind, was aus verschiedenen Komponenten besteht und zusätzliche einzuführende Verteidigungsschichten („defence in debt") notwendig macht. Die Risikoanalyse sollte daher u. a. eine Definition von KI-spezifischen Bedrohungen beinhalten, die technische und organisatorische Sicherheitsrichtlinien adressiert, wie u. a.:

- Ableitung von KI-spezifischen Sicherheitszielen
- Definition von KI-spezifischen funktionalen Sicherheitsanforderungen
- Definition von KI-spezifischen Sicherheitsanforderungen (SARs)

Da diese junge Technologie sich ständig und rasant entwickelt, ist es erforderlich, bei der Suche nach geeigneten Maßnahmen, flexibel zu agieren, den **Stand der Technik** fortwährend zu berücksichtigen, und ggf. mehrere Techniken aufeinander abstimmen und anzupassen.

Umsetzung „klassischer" Sicherheitsmaßnahmen bei KI-Systemen

Im Rahmen der klassischen IT-Sicherheitsmaßnahmen haben sich u. a. die internationale Standardserie ISO/IEC 27, insbesondere mit ISO/IEC 27001/27002, und der BSI-Grundschutz etabliert und als hilfreiches Kompendium erwiesen. Diese sind auch im Zusammenhang mit KI-Systemen prinzipiell eine verlässliche Grundlage.

In Übereinstimmung mit Dokumentationspflichten ist es wichtig alle relevanten Fakten und Entscheidungen während der Entwicklung des Systems und der Art und Weise, wie das System funktioniert festzuhalten, und im Betrieb bspw. mit Protokolldateien und Logdateien zur Überwachung des Systembetriebs zu dokumentieren und regelmäßig auf Anomalien zu überprüfen. Es ist empfehlenswert, die zu ergreifenden Maßnahmen und Verantwortlichkeiten zwischen Entwicklung und Betrieb zu definieren und abzugrenzen.

Technische Schutzmaßnahmen sollten/müssen auf verschiedenen Ebenen angewendet um die verschiedenen Angriffsszenarien und Fehlerquellen zu mitigieren. Dazu gehören die klassische Netzwerksicherheit, z. B. durch den Einsatz von Firewalls, Schutz von Ein- und Ausgabe des KI-Systems vor Manipulationen auf der Hardware-, Betriebssystem- und Softwareebene, bis zu Maßnahmen, die Zugriffe und Änderungen auf das KI-System mit Rechten und Zugängen über eine Authentifizierung auf einem angemessenen Sicherheitsniveau steuern. Schließlich muss sichergestellt werden, dass die schnellstmögliche Installation von Sicherheits-Patches, die der jeweiligen Bedrohungslage angemessen sind, möglich ist.

Neben den allgemeinen klassischen Sicherheitsmaßnahmen können auch andere allgemeine Maßnahmen helfen, KI-spezifische Bedrohungen zu bekämpfen, wie in den bekannten und ggf. flankierend einschlägigen Empfehlungen und Richtlinien (EU-NIS-2, EU-CRE, DORA, RED) vorgeschlagen: im Entwicklungsprozess von KI-Systemen etwa die Durchführung von Hintergrundüberprüfungen der (Kern-)Entwickler, die Dokumentation und der kryptografische Schutz wichtiger Informationen während des gesamten KI-Lebenszyklus wie bspw. die verwendeten Datensätze, der Schutz im Rahmen von Vorverarbeitungshandlungen wie die vor-trainierten Modelle oder das Trainingsverfahren selbst, und Redundanz des Systems bzw. von Systemkomponenten.

Ausgestaltung KI-spezifischer Sicherheitsmaßnahmen

Wie skizziert, handelt es sich bei Cybersicherheit, Robustheit und Genauigkeit nach **Art. 15 Abs. 1 KI-VO** um eigenständige und formal **abzugrenzende Kriterien**, in denen das Gesetz die spezifische KI-Systemsicherheit adressiert. Im herkömmlichen Sinne betreffen all diese Anforderungen letztlich die *Cyber*sicherheit (siehe oben). Da Compliance im Bereich der KI-Sicherheit aber gerade nicht nur aus Art. 15 KI-VO folgt, sondern aus dem **Zusammenspiel** aller Regelungen der KI-Verordnung (siehe oben Schritt 2) und den regulatorischen Vorgaben für Systemsicherheit im Allgemeinen ergibt (siehe oben Schritt 3), ist diese Trennung in der Praxis letztlich nicht trennscharf möglich. Sie kann sich aber dort auswirken, wo es um den Inhalt und die Reichweite von Zertifizierungs- und/oder Konformitätsbewertungen geht.

Ein erster **Überblick** zum wesentlichen Inhalt und Zusammenspiel KI-spezifischer (Cyber-)Sicherheitsrisiken- und Kontrollen kann der nachstehenden Abbildung entnommen werden:

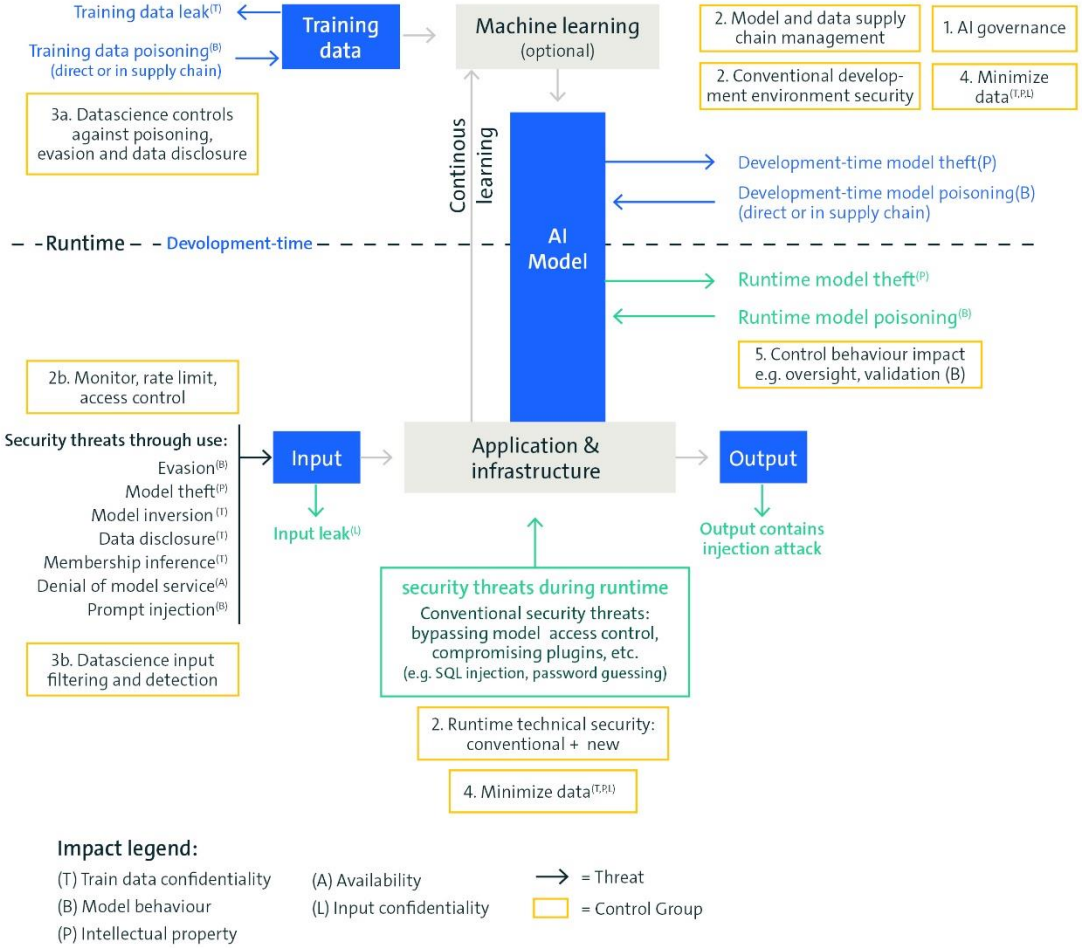


***Abbildung**: Übersicht der KI-spezifischen Sicherheitsrisiken- und Kontrollen*

Für eine detaillierte Orientierung enthält das öffentliche verfügbare „AI Risk Repository“ (siehe hier) des MIT eine umfangreiche Datenbank zu KI-spezifischen Risiken.

Im Hinblick auf Art. 15 KI-VO gelten Hochrisiko-KI-Systeme folgende Grundsätze:

Die Anforderung der **Cybersicherheit** bei KI-Systemen in der KI-VO ist maßgeblich auf den Schutz vor **externen Angriffen** fokussiert. Durch sie soll sichergestellt werden, dass KI-Systeme **widerstandsfähig** gegen die Versuche unbefugter Dritter sind, ihre Verwendung, Ausgaben oder Leistung zu verändern oder ihre Sicherheitsmerkmale zu beeinträchtigen. Welche **technischen** Maßnahmen hierfür von verpflichteten Unternehmen umzusetzen sind bzw. sein können, ist in Art. 15 Abs. 5 UAbs. 3 KI-VO angedeutet (vgl. Erwägungsgrund 76):

Art. 15 Abs. 5 UAbs. 2 KI-VO:

„Die technischen Lösungen für den Umgang mit KI-spezifischen Schwachstellen umfassen gegebenenfalls Maßnahmen, um Angriffe, mit denen versucht wird, eine Manipulation des Trainingsdatensatzes („data poisoning“) oder vortrainierter Komponenten, die beim Training

verwendet werden („model poisoning"), vorzunehmen, Eingabedaten, die das KI-Modell zu Fehlern verleiten sollen („adversarial examples" oder „model evasions"), Angriffe auf vertrauliche Daten oder Modellmängel zu verhüten, zu erkennen, darauf zu reagieren, sie zu beseitigen und zu kontrollieren.

Diese Lösungen sind **nicht zwingend zu erfüllen** („gegebenenfalls"), sondern hängen wie auch sonst von Einsatzzweck, Umgebung und den Ergebnissen der Risikobewertung ab. Nur so kann identifiziert werden, welche Angriffsvektoren spezielle Maßnahmen erfordern, die über den „klassischen" Schutz vor Cyberattacken hinausgehen. Das MIT stellt eine hilfreiche Datenbank zur Verfügung, in der KI-spezifische Risiken erfasst werden (vgl. hier). Obgleich Art. 15 Abs. 5 im Vergleich zu Art. 15 Abs. 4 UAbs. 1 KI-VO **organisatorische** Maßnahmen nicht explizit nennt, sind diese keinesfalls exkludiert, sondern ebenfalls geboten.

Die **Robustheit** nach Art. 15 Abs. 4 KI-VO verlangt technisch-organisatorische Maßnahmen, um Widerstandsfähigkeit gegenüber Störungen innerhalb des Systems herzustellen:

Art. 15 Abs. 4 UAbs. 1 KI-VO:

„Hochrisiko-KI-Systeme müssen so widerstandsfähig wie möglich gegenüber Fehlern, Störungen oder Unstimmigkeiten sein, die innerhalb des Systems oder der Umgebung, in der das System betrieben wird, insbesondere wegen seiner Interaktion mit natürlichen Personen oder anderen Systemen, auftreten können. In diesem Zusammenhang sind technische und organisatorische Maßnahmen zu ergreifen."

Der Fokus dieses Schutzziels liegt – im Gegensatz zu gezielten externen Angriffen – auf der generellen Funktionsfähigkeit des Systems innerhalb seiner Umgebung und Grenzen. Es beschreibt die Fähigkeit eines KI-Systems, das **Leistungsniveau** und die **Funktion** auch im Störfall aufrechtzuerhalten (vgl. ISO/IEC 22989:2022 – KI-Definitionen). Als technische Maßnahmen werden in Art. 15 Abs. 4 UAbs. 2 KI-VO (nur) Sicherungs- oder Störungspläne angeführt. Damit werden insbesondere Lösungen adressiert, die es dem System ermöglichen, seinen Betrieb bei Anomalien in einer (für seine Nutzungszwecke) sicheren Weise sicher zu unterbrechen (vgl. Erwägungsgrund 75). Art. 15 Abs. 4 UAbs. 3 KI-VO stellt ferner klar, dass Robustheit gerade auch die Modellebene und die **Entwicklung** betrifft und hier konkrete Maßnahmen vorzusehen sind, welche das Risiko verzerrter Ausgaben und somit Eintritt von sog. „**Rückkopplungsschleifen**" mitigieren.

Die **Genauigkeit** beschreibt den **Grad** der tatsächlich korrekten Ausgaben eines KI-Systems im Verhältnis zu den erwarteten Ausgaben bzw. Ergebnissen. Es handelt sich um eine Größe zur Bestimmung der **Qualität** des KI-Systems. Die Genauigkeit wird maßgeblich von der Qualität der Daten, den zugrundeliegenden Bewertungsmethoden und vielen weiteren system- bzw. modellimmanenten Faktoren beeinflusst. Wird bspw. durch geringe Qualität des Models eine zu große Streuung erzeugt oder handelt es sich um „schlechte" Daten, ist dies eine potenzielle Schwachstelle für die (Cyber-)Sicherheit und birgt folglich hohe Risiken für die Rechte der betroffenen Personen. Die KI-Verordnung verlangt ein „angemessenes" Maß an Genauigkeit. Eine konkrete Benchmark oder ein Prozess für dessen Ermittlung sind nicht vorgegeben. Die KI-VO sieht nur vor, die Entwicklung von Benchmarks und Messmethoden für KI-Systeme durch damit befasste Organisationen zu fördern (vgl. Art. 15 Abs. 2 KI-VO). In der Umsetzung kommt es hier daher auch wieder auf eine systematische Bewertung auf Basis der Risikoanalyse an. Ein angemessener Grad an „Genauigkeit" bestimmt sich durch zahlreiche einzelfallbezogene Faktoren wie beispielsweise die Definition des Einsatzwecks, die Gestaltung des Modells und seiner Parameter, die Qualität, Kontrolle

und Auswahl der (Trainings-)Daten, eine im Einzelfall geeignete Messmethode, usw. Auch die Genauigkeit hat damit zahlreiche Wechselwirkungen zu anderen Vorgaben der KI-VO.

Zwischenergebnis

Sind Genauigkeit, Robustheit und Cybersicherheit gewährleistet, ist mit **Schritt 4.1.9** fortzufahren.

Schritt 4.1.9: Wie ist das Qualitätsmanagement zu gestalten?

Dr. Benedikt Rohrßen (Taylor Wessing Partnerschaftsgesellschaft mbB)

Relevante(r) Artikel:

Art 17

Relevante(r) ErwG:

-

Konkretisierungsbedürftig:

-

Art. 17 KI-VO regelt, wie das Qualitätsmanagementsystem auszugestalten ist.

(1) Anbieter von Hochrisiko-KI-Systemen richten ein Qualitätsmanagementsystem ein, das die Einhaltung dieser Verordnung gewährleistet. Dieses System wird systematisch und ordnungsgemäß in Form schriftlicher Regeln, Verfahren und Anweisungen dokumentiert und umfasst mindestens folgende Aspekte:

a) ein Konzept zur Einhaltung der Regulierungsvorschriften, was die Einhaltung der Konformitätsbewertungsverfahren und der Verfahren für das Management von Änderungen an dem Hochrisiko-KI-System miteinschließt;

b) Techniken, Verfahren und systematische Maßnahmen für den Entwurf, die Entwurfskontrolle und die Entwurfsprüfung des Hochrisiko-KI-Systems;

c) Techniken, Verfahren und systematische Maßnahmen für die Entwicklung, Qualitätskontrolle und Qualitätssicherung des Hochrisiko-KI-Systems;

d) Untersuchungs-, Test- und Validierungsverfahren, die vor, während und nach der Entwicklung des Hochrisiko-KI-Systems durchzuführen sind, und die Häufigkeit der Durchführung;

e) die technischen Spezifikationen und Normen, die anzuwenden sind und, falls die einschlägigen harmonisierten Normen nicht vollständig angewandt werden oder sie nicht alle relevanten Anforderungen gemäß Abschnitt 2 abdecken, die Mittel, mit denen gewährleistet werden soll, dass das Hochrisiko-KI-System diese Anforderungen erfüllt;

f) Systeme und Verfahren für das Datenmanagement, einschließlich Datengewinnung, Datenerhebung, Datenanalyse, Datenkennzeichnung, Datenspeicherung, Datenfilterung, Datenauswertung, Datenaggregation, Vorratsdatenspeicherung und sonstiger Vorgänge in Bezug auf die Daten, die im Vorfeld und für die Zwecke des Inverkehrbringens oder der Inbetriebnahme von Hochrisiko-KI-Systemen durchgeführt werden

g) das in Artikel 9 genannte Risikomanagementsystem;

h) die Einrichtung, Anwendung und Aufrechterhaltung eines Systems zur Beobachtung nach dem Inverkehrbringen gemäß Artikel 72;

i) Verfahren zur Meldung eines schwerwiegenden Vorfalls gemäß Artikel 73;

j) die Handhabung der Kommunikation mit zuständigen nationalen Behörden, anderen einschlägigen Behörden, auch Behörden, die den Zugang zu Daten gewähren oder erleichtern, notifizierten Stellen, anderen Akteuren, Kunden oder sonstigen interessierten Kreisen;

k) Systeme und Verfahren für die Aufzeichnung sämtlicher einschlägigen Dokumentation und Informationen;

l) Ressourcenmanagement, einschließlich Maßnahmen im Hinblick auf die Versorgungssicherheit;

m) einen Rechenschaftsrahmen, der die Verantwortlichkeiten der Leitung und des sonstigen Personals in Bezug auf alle in diesem Absatz aufgeführten Aspekte regelt.

2) Die Umsetzung der in Absatz 1 genannten Aspekte erfolgt in einem angemessenen Verhältnis zur Größe der Organisation des Anbieters. Die Anbieter müssen in jedem Fall den Grad der Strenge und das Schutzniveau einhalten, die erforderlich sind, um die Übereinstimmung ihrer Hochrisiko-KI-Systeme mit dieser Verordnung sicherzustellen.

(3) Anbieter von Hochrisiko-KI-Systemen, die Pflichten in Bezug auf Qualitätsmanagementsysteme oder eine gleichwertige Funktion gemäß den sektorspezifischen Rechtsvorschriften der Union unterliegen, können die in Absatz 1 aufgeführten Aspekte als Bestandteil der nach den genannten Rechtsvorschriften festgelegten Qualitätsmanagementsysteme einbeziehen.

(4) Bei Anbietern, die Finanzinstitute sind und gemäß den Rechtsvorschriften der Union über Finanzdienstleistungen Anforderungen in Bezug auf ihre Regelungen oder Verfahren der internen Unternehmensführung unterliegen, gilt die Pflicht zur Einrichtung eines Qualitätsmanagementsystems — mit Ausnahme des Absatzes 1 Buchstaben g, h und i des vorliegenden Artikels — als erfüllt, wenn die Vorschriften über Regelungen oder Verfahren der internen Unternehmensführung gemäß dem einschlägigen Unionsrecht über Finanzdienstleistungen eingehalten werden. Zu diesem Zweck werden die in Artikel 40 genannten harmonisierten Normen berücksichtigt.

Art. 17 KI-VO konkretisiert einen Teil der Pflichten der Anbieter von Hochrisiko-KI-Systemen (vgl. Art. 16 lit c KI-VO). Gem. Art. 17 KI-VO müssen Anbieter ein **Qualitätsmanagementsystem** einrichten. Es soll sicherstellen, dass ihre Hochrisiko-KI-Systeme die **Anforderungen der KI-VO** erfüllen. **Folgende Aspekte** muss das Qualitätsmanagementsystem umfassen (Art. 17 Abs. 1 KI-VO):

- **Regulierungskonzept**: zur Einhaltung der Regulierungsvorschriften. Es soll auch die Einhaltung der Konformitätsbewertungsverfahren und des Verfahrens für das Management von Änderungen an dem Hochrisiko-KI-System erfassen.
- **Entwurfsvalidierung**: Techniken, Verfahren und systematische Maßnahmen für Entwurf, Entwurfskontrolle und -prüfung des Hochrisiko-KI-Systems**.**
- **Entwicklungs- und Qualitätskontrollen**: Techniken, Verfahren und systematische Maßnahmen für die Entwicklung sowie für Qualitätskontrolle und -sicherung des Hochrisiko-KI-Systems.
- **Prüf- und Validierungsverfahren**: Untersuchungs-, Test- und Validierungsprozeduren, die vor, während und nach der Entwicklung des Hochrisiko-KI-Systems durchzuführen sind (inklusive der Häufigkeit der Durchführung)
- **Technische Spezifikationen und Normen**: Anzuwendende technische Spezifikationen und Normen. Falls die harmonisierten Normen nicht vollständig angewandt werden oder nicht alle Anforderungen an Hochrisiko-KI-Systeme gem. Art. 8 – 15 KI-VO abdecken, müssen die Mittel zur Gewährleistung der Erfüllung der Anforderungen nach Art. 8 – 15 KI-VO erfasst werden.
- **Datenmanagementsysteme und -verfahren**: Für Datengewinnung, -erhebung, -analyse, -kennzeichnung, -speicherung, -filterung, -auswertung, -aggregation und Vorratsdatenspeicherung sowie sonstige datenbezogene Vorgänge **im Vorfeld und für die Zwecke des Inverkehrbringens oder der Inbetriebnahme** von Hochrisiko-KI-Systemen
- **Risikomanagementsystem** (Art. 9 KI-VO)
- System zur **Beobachtung nach dem Inverkehrbringen** (Art. 72 KI-VO)
- Verfahren zur **Meldung eines schwerwiegenden Vorfalls** (Art. 73 KI-VO)
- **Handhabung der Kommunikation**: mit zuständigen nationalen sowie solchen Behörden, die den Zugang zu Daten gewähren oder erleichtern, notifizierten Stellen, anderen Akteuren, Kunden oder sonstigen interessierten Kreisen.
- **Dokumentationssysteme**: Systeme und Verfahren für die Aufzeichnung sämtlicher einschlägigen Dokumentation und Informationen
- **Ressourcenmanagement**: einschließlich Maßnahmen im Hinblick auf die Versorgungssicherheit
- **Rechenschaftsrahmen**: welcher die Verantwortlichkeiten der Leitung und des sonstigen Personals für alle genannten Aspekte regelt

Die Umsetzung der oben genannten Aspekte erfolgt dabei **proportional zur Größe** der Organisation des Anbieters. Dies entbindet Anbieter jedoch nicht den notwendigen Grad an Strenge und Schutzniveau einzuhalten, der für die Übereinstimmung ihrer Hochrisiko-KI-Systeme mit dieser Verordnung erforderlich ist (Art. 17 Abs. 2 KI-VO).

So müssen gem. Erwägungsgrund 146 z. B. Kleinstunternehmen nur eine vereinfachte Version des Qualitätsmanagements einführen. Damit sollen Verwaltungsaufwand und Kosten für das Unternehmen geringer und gleichzeitig das Schutzniveau aufrechterhalten werden. Um Klarheit für Betroffene zu schaffen, wird die Kommission Leitlinien ausarbeiten, um die Anforderungen an diese vereinfachte Version des Qualitätsmanagements festzulegen.

Anbieter von Hochrisiko-KI-Systemen, die gemäß den sektorspezifischen EU-Rechtsvorschriften (z. B. die Maschinenrichtlinie (künftig: Maschinenverordnung)) bereits hinsichtlich Qualitätsmanagementsysteme oder ähnlichem verpflichtet sind, können freiwillig die oben aufgeführten Aspekte in ihre Systeme integrieren (Art. 17 Abs. 3 KI-VO).

Sofern **Finanzinstitute** Anbieter von Hochrisiko-KI-Systeme sind und sie gemäß den EU-Rechtsvorschriften über Finanzdienstleistungen der internen Unternehmensführung unterliegen (so z. B. gemäß der Verordnung über Aufsichtsanforderungen an Kreditinstitute und Wertpapierfirmen), gilt etwas anderes: Ihre Pflicht zur Einrichtung eines Qualitätsmanagementsystems **gilt bereits dann als erfüllt, wenn** die **Vorschriften der internen Unternehmensführung** gemäß den einschlägigen EU-Rechtsvorschriften über Finanzdienstleistung erfüllt sind. Hierbei werden die in Art. 40 KI-VO genannten harmonisierten Normen berücksichtigt. Art. 17 Abs. 4 KI-VO). Zudem gibt die BaFin in ihrem Rundschreiben vom Mai 2023 Mindestanforderungen an das Risikomanagement von Finanzinstituten vor, welche auch auf KI-Modelle anwendbar sind. So müssen KI-Modelle entsprechend nachvollziehbar sein und regelmäßig validiert werden, um die Datenqualität und die Eignung der Modelle sicherzustellen. Zudem ist auf eine hinreichende Erklärbarkeit zu achten.

Von diesem Prinzip gibt es jedoch eine **Ausnahme**. Denn selbst wenn Finanzinstitute die Vorschriften der internen Unternehmensführung erfüllen, müssen sie trotzdem **folgende Pflichten** einhalten:

- Einrichtung, Anwendung, Dokumentierung und Aufrechthaltung eines Risikomanagementsystems (Art. 9 KI-VO),
- Einrichtung, Anwendung und Aufrechterhaltung eines Systems zur Beobachtung nach dem Inverkehrbringen der Hochrisiko-KI-Systeme (Art. 72 KI-VO),
- Verfahren zur Meldung eines schwerwiegenden Vorfalls (Art. 73 KI-VO).
- Gem. Erwägungsgrund 81 sollten Anbieter von Hochrisiko-KI-Systemen zusammenfassend:
- Ein solides Qualitätsmanagementsystem einrichten,
- sicherstellen, dass das vorgeschriebene Konformitätsbewertungsverfahren durchgeführt wird,
- die erforderliche Dokumentation erstellen, sowie
- ein robustes System zur Überwachung nach dem Inverkehrbringen einrichten.

Anbieter, die bereits Qualitätsmanagementpflichten gemäß anderen EU-Vorschriften haben, sollen die Möglichkeit erhalten, die Elemente des Qualitätsmanagementsystems der KI-VO in ihr bestehendes System integrieren zu können.

Behörden, die Hochrisiko-KI-Systeme für den Eigengebrauch verwenden, können die Vorschriften für das Qualitätsmanagementsystem als Teil ihres nationalen oder regionalen Systems umsetzen, wobei die Besonderheiten ihres Bereichs zu berücksichtigen sind.

Zwischenergebnis

Ist das Qualitätsmanagementsystem etabliert, ist mit **Schritt 4.1.10** fortzufahren.

Schritt 4.1.10: Wie ist die Dokumentation aufzubewahren?

Dr. Benedikt Rohrßen (Taylor Wessing Partnerschaftsgesellschaft mbB)

Art. 18 KI-VO konkretisiert die Pflicht des Anbieters zur Aufbewahrung der Dokumentation (vgl. Art. 16 lit. d KI-VO). Die Pflicht zur Dokumentierung besteht **zehn Jahre lang** und beginnt, nachdem das Hochrisiko-KI-System in Verkehr gebracht oder in Betrieb genommen wurde.

(1) Der Anbieter hält für einen Zeitraum von zehn Jahren ab dem Inverkehrbringen oder der Inbetriebnahme des Hochrisiko-KI-Systems folgende Unterlagen für die zuständigen nationalen Behörden bereit:

a) die in Artikel 11 genannte technische Dokumentation;

b) die Dokumentation zu dem in Artikel 17 genannten Qualitätsmanagementsystem;

c) die Dokumentation über etwaige von notifizierten Stellen genehmigte Änderungen;

d) gegebenenfalls die von den notifizierten Stellen ausgestellten Entscheidungen und sonstigen Dokumente;

e) die in Artikel 47 genannte EU-Konformitätserklärung

(2) Jeder Mitgliedstaat legt die Bedingungen fest, unter denen die in Absatz 1 genannte Dokumentation für die zuständigen nationalen Behörden für den in dem genannten Absatz angegebenen Zeitraum bereitgehalten wird, für den Fall, dass ein Anbieter oder sein in demselben Hoheitsgebiet niedergelassener Bevollmächtigter vor Ende dieses Zeitraums in Konkurs geht oder seine Tätigkeit aufgibt.

(3) Anbieter, die Finanzinstitute sind und gemäß dem Unionsrecht über Finanzdienstleistungen Anforderungen in Bezug auf ihre Regelungen oder Verfahren der internen Unternehmensführung unterliegen, pflegen die technische Dokumentation als Teil der gemäß dem Unionsrecht über Finanzdienstleistungen aufzubewahrenden Dokumentation.

Relevante(r) Artikel:

Art. 18

Relevante(r) ErwG:

-

Konkretisierungsbedürftig:

Ja

Mitgliedstaaten legen Bedingungen für Bereithalten der Dokumentation fest (bei Konkurs oder Geschäftsaufgabe des Anbieters oder Bevollmächtigen vor Ende des Zeitraums), Art. 18 Abs. 2 KI-VO

Art. 18 KI-VO **konkretisiert die Pflicht** des Anbieters zur Aufbewahrung der Dokumentation (vgl. Art. 16 lit. d KI-VO). Die Pflicht zur Dokumentierung besteht **zehn Jahre lang** und beginnt, nachdem das Hochrisiko-KI-System in Verkehr gebracht oder in Betrieb genommen wurde.

Folgende Dokumente sind aufzubewahren:

- Technische Dokumentation zum KI-System (Art. 11 KI-VO)
- Dokumentation zum Qualitätssicherungssystem (Art. 17 KI-VO)
- Die EU-Konformitätserklärung (Art. 47 KI-VO)

Folgende Dokumente sind aufzubewahren, **falls vorhanden**:

- Dokumentation über genehmigte Änderungen von notifizierten Stellen
- Entscheidungen und sonstige Dokumente von notifizierten Stellen

Die hier genannte notifizierte Stelle eine Konformitätsbewertungsstelle, die gemäß der KI-VO und den anderen einschlägigen Harmonisierungsrechtsvorschriften der Union notifiziert wurde (Art. 3 Nr. 21 KI-VO).

Die **EU-Mitgliedstaaten legen selbst die Regeln fest**, wie diese Dokumente aufzubewahren sind, falls ein Anbieter oder sein Bevollmächtigter (der im selben Staat niedergelassen ist) innerhalb der zehn-Jahres-Frist in Konkurs geht oder die Geschäftstätigkeit aufgibt (Art. 18 Abs. 2 KI-VO).

Sofern **Finanzinstitute Anbieter** der Systeme und diese gemäß den EU-Rechtsvorschriften über Finanzdienstleistungen der internen Unternehmensführung unterliegen (so z. B. gemäß der Verordnung über Aufsichtsanforderungen an Kreditinstitute und Wertpapierfirmen), gilt etwas anderes. Diese müssen die **technische Dokumentation im Rahmen ihrer regulären Dokumentation gemäß den EU-Rechtsvorschriften aufbewahren** (Art. 18 Abs. 3 KI-VO).

Zwischenergebnis

Wenn die technische Dokumentation aufbewahrt ist, geht es weiter mit **Schritt 4.1.11.**

Schritt 4.1.11: Wie sind die automatisch erzeugten Protokolle aufzubewahren?

Relevante(r) Artikel:

Art. 19

Relevante(r) ErwG:

-

Konkretisierungsbedürftig:

Nein

Art. 19 KI-VO konkretisiert die Pflicht des Anbieters, **automatisch erzeugte Protokolle aufzubewahren** (vgl. Art. 16 lit. e KI-VO).

> (1) Anbieter von Hochrisiko-KI-Systemen bewahren die von ihren Hochrisiko-KI-Systemen automatisch erzeugten Protokolle gemäß Artikel 12 Absatz 1 auf, soweit diese Protokolle ihrer Kontrolle unterliegen. Unbeschadet des geltenden Unionsrechts oder nationalen Rechts werden die Protokolle für einen der Zweckbestimmung des Hochrisiko-KI-Systems angemessenen Zeitraum von mindestens sechs Monaten aufbewahrt, sofern in den geltenden Rechtsvorschriften der Union, insbesondere im Unionsrecht zum Schutz personenbezogener Daten, oder im geltenden nationalen Recht nichts anderes vorgesehen ist.
>
> (2) Anbieter, die Finanzinstitute sind und gemäß den Rechtsvorschriften der Union über Finanzdienstleistungen Anforderungen in Bezug auf ihre Regelungen oder Verfahren der internen Unternehmensführung, unterliegen, bewahren die von ihren Hochrisiko-KI-Systemen automatisch erzeugten Protokolle als Teil der gemäß dem einschlägigen Unionsrecht über Finanzdienstleistungen aufzubewahrenden Dokumentation auf

Anbieter von Hochrisiko-KI-Systemen müssen die automatisch erzeugten Protokolle (Art. 12 Abs. 1 KI-VO) aufbewahren. Dies ist aber **nur dann** der Fall, **wenn** die **Protokolle ihrer Kontrolle** unterliegen. Ob es hierfür ausreicht, auf diese zu speichern und verwalten zu können, lässt sich der Verordnung nicht entnehmen. Da die Protokolle aber automatisch erzeugt werden müssen, wird dies wohl ausreichend sein.

Die **Dauer der Aufbewahrung** richtet sich grundsätzlich nach den einschlägigen **EU-Rechtsvorschriften** (insbesondere zum Schutz personenbezogener Daten) bzw. den **nationalen Gesetzen**. Ansonsten besteht die Pflicht zur Aufbewahrung für **mindestens sechs Monate**. Dieser Zeitraum kann im Einzelfall länger sein, sofern dies angesichts der Zweckbestimmung des Hochrisiko-KI-Systems angemessen ist.

Sofern **Finanzinstitute Anbieter** der Systeme sind und diese gemäß den EU-Rechtsvorschriften über Finanzdienstleistungen der internen Unternehmensführung unterliegen (so z. B. gemäß der Verordnung über Aufsichtsanforderungen an Kreditinstitute und Wertpapierfirmen), gilt etwas anderes. Diese müssen die automatisch erzeugten Protokolle **im Rahmen ihrer regulären Dokumentation gemäß den EU-Rechtsvorschriften aufbewahren** (Art. 18 Abs. 3 KI-VO).

Zwischenergebnis:

Sind die Logs aufbewahrt, geht es weiter mit **Schritt 4.1.12.**

Schritt 4.1.12: Welche Korrekturmaßnahmen und Informationspflichten obliegen dem Anbieter?

Relevante(r) Artikel:

Art. 20

Relevante(r) ErwG:

155

Konkretisierungsbedürftig:

Nein

Art. 20 KI-VO konkretisiert die Pflicht des Anbieters, die erforderlichen **Korrekturmaßnahmen** zu ergreifen und die **betroffenen Wirtschaftsakteure zu informieren** (vgl. Art. 20 lit. j KI-VO).

(1) Anbieter von Hochrisiko-KI-Systemen, die der Auffassung sind oder Grund zu der Annahme haben, dass ein von ihnen in Verkehr gebrachtes oder in Betrieb genommenes Hochrisiko-KI-System nicht dieser Verordnung entspricht, ergreifen unverzüglich die erforderlichen Korrekturmaßnahmen, um die Konformität dieses Systems herzustellen oder es gegebenenfalls zurückzunehmen, zu deaktivieren oder zurückzurufen. Sie informieren die Händler des betreffenden Hochrisiko-KI-Systems und gegebenenfalls die Betreiber, den Bevollmächtigten und die Einführer darüber.

(2) Birgt das Hochrisiko-KI-System ein Risiko im Sinne des Artikels 79 Absatz 1 und wird sich der Anbieter des Systems dieses Risikos bewusst, so führt er unverzüglich gegebenenfalls gemeinsam mit dem meldenden Betreiber eine Untersuchung der Ursachen durch und informiert er die Marktüberwachungsbehörden, in deren Zuständigkeit das betroffene Hochrisiko-KI-System fällt, und gegebenenfalls die notifizierte Stelle, die eine Bescheinigung für dieses Hochrisiko-KI-System gemäß Artikel 44 ausgestellt hat, insbesondere über die Art der Nichtkonformität und über bereits ergriffene relevante Korrekturmaßnahmen.

Wenn Anbieter von Hochrisiko-KI-Systemen **denken oder den Verdacht haben**, dass ihr System **nicht den Vorschriften der KI-VO** entspricht, müssen sie **unverzüglich Maßnahmen** ergreifen (Art. 20 Abs. 1 KI-VO). Das sind die Maßnahmen, die erforderlich sind, um die Konformität wieder herzustellen. Falls unmöglich, müssen sie KI-Systeme zurücknehmen, deaktivieren oder zurückrufen. Außerdem müssen sie alle Händler der Systeme informieren (sowie ggf. Betreiber, Vertreter und Importeure).

Wenn der **Anbieter des Hochrisiko-KI-Systems** merkt, dass sein Hochrisiko-KI-System ein Risiko für die Gesundheit oder Sicherheit oder Grundrechte von Personen darstellt (Art. 79 Abs. 1 KI-VO) muss der Anbieter **unverzüglich**:

- eine **Untersuchung der Ursachen** durchführen. Diese Untersuchung ist ggf. gemeinsam mit dem meldenden Betreiber durchzuführen.
- Die **Marktüberwachungsbehörden**, in deren Zuständigkeit das betroffene Hochrisiko-KI-System fällt, **informieren**.
- Ggf. auch die **notifizierte Stelle**, die eine Bescheinigung für dieses Hochrisiko-KI-System gemäß Art. 44 KI-VO ausgestellt hat, **informieren**.

Bei der Information der Marktüberwachungsbehörden (und ggf. der notifizierten Stelle) hat der Anbieter insbesondere über die Art der Nichtkonformität und über bereits ergriffene relevante Korrekturmaßnahmen zu berichten.

Damit diese Korrekturmaßnahmen **rechtzeitig** ergriffen werden können, sollten Anbieter ein **System zur Beobachtung nach der Markteinführung** haben (Erwägungsgrund 155). Falls nötig, sollte dieses System auch die Interaktion mit anderen KI-Systemen und Software analysieren. Sofern die Betreiber Strafverfolgungsbehörden sind, soll dieses System jedoch keine sensiblen Daten erfassen.

Anbieter müssen außerdem ein **System** einrichten, um **schwerwiegende Vorfälle an die zuständigen Behörden zu melden**. Zu solchen Vorfällen zählen Ereignisse, die zum Tod oder schweren Gesundheitsschäden führen, schwere und irreversible Störungen der Verwaltung und kritischer Infrastrukturen verursachen, gegen EU-Recht verstoßen (welche dem Schutz von Grundrechten dient) oder schwere Sach- oder Umweltschäden anrichten.

Zwischenergebnis

Sind Korrekturmaßnahmen ergriffen, geht es weiter mit **Schritt 4.1.13.**

Schritt 4.1.13: Wie gestaltet sich die Zusammenarbeit mit Behörden?

Relevante(r) Artikel:

Art. 21

Relevante(r) ErwG:

-

Konkretisierungsbedürftig:

Nein

Art. 21 KI-VO regelt näheres zur **Pflicht des Anbieters**, mit den zuständigen **Behörden zusammenzuarbeiten** (vgl. Art. 16 lit. k KI-VO).

> (1) Anbieter von Hochrisiko-KI-Systemen übermitteln einer zuständigen Behörde auf deren begründete Anfrage sämtliche Informationen und Dokumentation, die erforderlich sind, um die Konformität des Hochrisiko-KI-Systems mit den in Abschnitt 2 festgelegten Anforderungen nachzuweisen, und zwar in einer Sprache, die für die Behörde leicht verständlich ist und bei der es sich um eine der von dem betreffenden Mitgliedstaat angegebenen Amtssprachen der Institutionen der Union handelt.
>
> (2) Auf begründete Anfrage einer zuständigen Behörde gewähren die Anbieter der anfragenden zuständigen Behörde gegebenenfalls auch Zugang zu den automatisch erzeugten Protokollen des Hochrisiko-KI-Systems gemäß Artikel 12 Absatz 1, soweit diese Protokolle ihrer Kontrolle unterliegen.
>
> (3) Alle Informationen, die eine zuständige Behörde aufgrund dieses Artikels erhält, werden im Einklang mit den in Artikel 78 festgelegten Vertraulichkeitspflichten behandelt.

Art. 21 KI-VO regelt näheres zur **Pflicht des Anbieters**, mit den zuständigen **Behörden zusammenzuarbeiten** (vgl. Art. 16 lit. k KI-VO).

Hiernach müssen Anbieter von Hochrisiko-KI-Systemen den Behörden **auf begründete Anfrage**:

- **alle Informationen und Dokumente** , die zeigen, dass ihr System den Vorschriften (Art. 8 – 15 KI-VO) entspricht. Diese Informationen müssen in einer Sprache bereitgestellt werden, die für die Behörde leicht verständlich ist und als Amtssprache von der EU für das betroffene Mitgliedstaat anerkannt ist (Art. 21 Abs. 1 KI-VO).
- **Zugang zu automatisch erzeugten Protokollen** (Art. 12 Abs. 1 KI-VO) ihres Systems gewähren, sofern sie diese Protokolle kontrollieren (Art. 21 Abs. 2 KI-VO).

Alle von den Behörden erhaltenen Informationen werden vertraulich behandelt (Art. 21 Abs. 3 KI-VO). Die Vertraulichkeitspflichten regelt Art. 78 KI-VO. Hiervon erfasst sind z. B. die Rechte des geistigen Eigentums, vertrauliche Geschäftsinformationen und Geschäftsgeheimnisse sowie öffentliche und nationale Sicherheitsinteressen.

Zwischenergebnis

Wenn alle Anforderungen an Hochrisiko-KI-Systeme durch den Anbieter erfüllt worden sind, ist mit **Schritt 6 (Konformitätsbewertungsverfahren)** fortzufahren.

Schritt 4.2: Welche Pflichten muss ich als Betreiber eines Hochrisiko-KI-Systems erfüllen?

Dr. Benedikt Rohrßen (Taylor Wessing Partnerschaftsgesellschaft mbB)

Nachdem festgestellt worden ist, dass es sich um ein Hochrisiko-KI-System i.S.d. KI-VO handelt und Sie Betreiber sind, ist im nächsten Schritt zu prüfen, welche Anforderungen vom Betreiber zu erfüllen sind.

Die KI-VO sieht im Abschnitt 3, also in den **Art. 16 bis 27 KI-VO**, eine Reihe von Pflichten vor, die vom Anbieter zu erfüllen sind.

Relevante(r) Artikel:

16, 25, 26, 27, 50

Relevante(r) ErwG:

26, 91, 93, 96, 132, 133, 134, 135, 137

Konkretisierungsbedürftig:

Ja (Praxisleitfäden zur Umsetzung der Pflichten zur Feststellung und Kennzeichnung künstlich erzeugter und manipulierter Inhalte

Ggf. Durchführungsrechtsakte zur Genehmigung der Praxisleitfäden oder für die Umsetzung der Pflichten

Musterfragebogen des Büros für Künstliche Intelligenz zur erleichterten Erfüllung der Pflicht der Betreiber zur Grundrechte-Folgeabschätzung für Hochrisiko-KI-Systeme)

Einleitung

Die Einführung der neuen KI-Verordnung („KI-VO“) bringt umfassende Veränderungen und neue Pflichten für zahlreiche Akteure im Bereich der Künstlichen Intelligenz mit sich. Insbesondere Anbieter von KI-Systemen stehen einer Vielzahl neuer Anforderungen und Verantwortlichkeiten gegenüber. Diese Pflichten reichen von der Gewährleistung der Transparenz und Nachvollziehbarkeit ihrer KI-Systeme bis zur Einhaltung strenger Sicherheitsstandards. Doch nicht nur Anbieter, sondern auch Betreiber von KI-Systemen, die KI in ihrer täglichen Praxis nutzen, sind von den neuen Regelungen betroffen. Denn mit der KI-VO gehen spezifische Betreiberpflichten einher, die entscheidend für die sichere KI-Anwendung in verschiedenen Branchen sind.

Anwendungsbereich

Die KI-VO betrifft gem. Art. 2 Abs. 1 KI-VO in persönlicher Hinsicht Anbieter, Betreiber, Einführer und Händler, Produkthersteller von KI-Systemen sowie Bevollmächtigte Vertreter. Gegenüber der Allgemeinen Produktsicherheitsverordnung (auf Englisch abgekürzt: GPSR) erweitert die KI-VO ihren persönlichen Anwendungsbereich um Anbieter und Betreiber. Betreiber meint jede natürliche oder juristische Person, Einrichtung oder sonstige Stelle mit Sitz in der EU oder die sich in der EU befindet, die ein KI-System in eigener Verantwortung verwendet – es sei denn, das KI-System wird im Rahmen einer persönlichen und nicht beruflichen Tätigkeit verwendet, Art. 3 Nr. 4 KI-VO. Insofern gilt die KI-VO nicht für Betreiber, die natürliche Personen sind und KI-Systeme ausschließlich zu persönlichen Zwecken verwenden, Art. 2 Abs. 10 KI-VO.

Zentrale Norm für die spezifischen Betreiberpflichten ist **Art. 26 KI-VO**. Darüber hinaus gelten ggf. die an den Anbieter gestellten Anforderungen und Pflichten auch für Betreiber, **Art. 25 Abs. 1 KI-VO.**

Schritt 4.2.1: Wann muss ich Anbieterpflichten erfüllen?

An Hochrisiko-KI-Systeme stellt die KI-VO strenge Anforderungen – gerade, weil sie ein hohes Risiko für die Gesundheit und Sicherheit oder für die Grundrechte natürlicher Personen darstellen können. Wesentliche Verpflichtete sind primär die **Anbieter** von Hochrisiko-KI-Systemen, welche umfassenden Pflichten nach Art. 16 KI-VO unterliegen. Darunter beispielsweise

- die Einrichtung eines **Risikomanagementsystems**, das eine angemessene Risikobewertung und Risikominimierung oder -beseitigung gewährleistet (Art. 9 KI-VO);
- die Sicherstellung der **Datenqualität**, vor allem durch geeignete Daten-Governance- und Datenverwaltungsverfahren (Art. 10 KI-VO);
- die **Technische Dokumentation** über das Hochrisiko-KI-System (Art.11 KI-VO);
- die automatische **Protokollierung** von Vorgängen und Ereignissen im „Hochrisiko-KI-System“ (Art. 12 KI-VO);
- **Transparenz- und Bereitstellungspflichten** gegenüber den Nutzern (Art.13 KI-VO);
- die Pflicht, das Hochrisiko-KI-System so zu entwickeln, dass eine **menschliche, wirksame Aufsicht** für die Dauer der Verwendung gewährleistet ist (Art. 14 KI-VO);
- die Einhaltung eines angemessenen Maßes an **Robustheit, Sicherheit und Genauigkeit** des jeweiligen Hochrisiko-KI-Systems (Art. 15 KI-VO);
- die Einrichtung eines **Qualitätsmanagementsystems** (Art. 17 KI-VO).

Diese Pflichten können allerdings **auch für Betreiber** von Hochrisiko-KI-Systemen Geltung entfalten, wenn diese als Anbieter eines Hochrisiko-KI-Systems zu klassifizieren sind. **Den Anbieterpflichten unterliegen Betreiber dann, wenn**

- sie ein bereits in Verkehr gebrachtes oder in Betrieb genommenes Hochrisiko-KI-System m**it ihrem Namen oder ihrer Handelsmarke** versehen, unbeschadet vertraglicher Vereinbarungen, die eine andere Aufteilung der Pflichten vorsehen;
- wenn sie eine **wesentliche Änderung** an einem Hochrisiko-KI-System, das bereits in Verkehr gebracht oder in Betrieb genommen wurde, so vornehmen, dass es weiterhin ein Hochrisiko-KI-System im Sinne von Artikel 6 bleibt;
- wenn sie die **Zweckbestimmung** eines KI-Systems, einschließlich eines KI-Systems mit allgemeinem Verwendungszweck, das nicht als hochriskant eingestuft wurde und bereits in Verkehr gebracht oder in Betrieb genommen wurde, so **verändern**, dass das betreffende KI-System zu einem Hochrisiko-KI-System im Sinne von Artikel 6 wird.

Schritt 4.2.2: Welche spezifischen Betreiberpflichten sieht Art. 26 KI-VO vor?

Zusätzlich zu den Anbieterpflichten, die Wirkung auf Betreiber entfalten können, unterliegen Betreiber betreiberspezifischen Pflichten.

Bei Betreibern, die Finanzinstitute sind und gemäß den Rechtsvorschriften der Union über Finanzdienstleistungen Anforderungen in Bezug auf ihre Regelungen oder Verfahren der internen Unternehmensführung, unterliegen, gilt die in Unterabsatz festgelegte Überwachungspflicht als erfüllt, wenn die Vorschriften über Regelungen, Verfahren oder Mechanismen der internen Unternehmensführung gemäß den einschlägigen Rechtsvorschriften über Finanzdienstleistungen eingehalten werden.

(1) Die Betreiber von Hochrisiko-KI-Systemen treffen geeignete technische und organisatorische Maßnahmen, um sicherzustellen, dass sie solche Systeme entsprechend der den Systemen beigefügten Gebrauchsanweisungen und gemäß den Absätzen 3 und 6 verwenden.

(2) Die Betreiber übertragen natürlichen Personen, die über die erforderliche Kompetenz, Ausbildung und Befugnis verfügen, die menschliche Aufsicht und lassen ihnen die erforderliche Unterstützung zukommen.

(3) Die Pflichten nach den Absätzen 1 und 2 lassen sonstige Pflichten der Betreiber nach Unionsrecht oder nationalem Recht sowie die Freiheit der Betreiber bei der Organisation ihrer eigenen Ressourcen und Tätigkeiten zur Wahrnehmung der vom Anbieter angegebenen Maßnahmen der menschlichen Aufsicht unberührt.

(4) Unbeschadet der Absätze 1 und 2, und soweit die Eingabedaten ihrer Kontrolle unterliegen, sorgen die Betreiber dafür, dass die Eingabedaten der Zweckbestimmung des Hochrisiko-KI-Systems entsprechen und ausreichend repräsentativ sind.

(5) Die Betreiber überwachen den Betrieb des Hochrisiko-KI-Systems anhand der Gebrauchsanweisung und informieren gegebenenfalls die Anbieter gemäß Artikel 72. Haben Betreiber Grund zu der Annahme, dass die Verwendung gemäß der Gebrauchsanweisung dazu führen kann, dass das Hochrisiko-KI-System ein Risiko im Sinne des Artikels 79 Absatz 1 birgt, so informieren sie unverzüglich den Anbieter oder Händler und die zuständige Marktüberwachungsbehörde und setzen die Verwendung des Systems aus. Haben die Betreiber einen schwerwiegenden Vorfall festgestellt, informieren sie auch unverzüglich zuerst den Anbieter und dann den Einführer oder Händler und die zuständigen Marktüberwachungsbehörden über den Vorfall. Kann der Betreiber den Anbieter nicht erreichen, so gilt Artikel 73 entsprechend. Diese Pflicht gilt nicht für sensible operative Daten von Betreibern von KI-Systemen, die Strafverfolgungsbehörden sind.

(6) Betreiber von Hochrisiko-KI-Systemen bewahren die von ihrem Hochrisiko-KI-System automatisch erzeugten Protokolle, soweit diese Protokolle ihrer Kontrolle unterliegen, für einen der Zweckbestimmung des Hochrisiko-KI-Systems angemessenen Zeitraum von mindestens sechs Monaten auf, sofern im geltenden Unionsrecht, insbesondere in den Rechtsvorschriften der Union über den Schutz personenbezogener Daten, oder im geltenden nationalem Recht nichts anderes bestimmt ist. Betreiber, die Finanzinstitute sind und gemäß den Rechtsvorschriften der Union über Finanzdienstleistungen Anforderungen in Bezug auf ihre Regelungen oder Verfahren der internen Unternehmensführung unterliegen, bewahren die Protokolle als Teil der gemäß den einschlägigen Rechtsvorschriften der Union über Finanzdienstleistungen aufzubewahrenden Dokumentation auf.

(7) Vor der Inbetriebnahme oder Verwendung eines Hochrisiko-KI-Systems am Arbeitsplatz informieren Betreiber, die Arbeitgeber sind, die Arbeitnehmervertreter und die betroffenen Arbeitnehmer darüber, dass sie Gegenstand des Einsatzes des Hochrisiko- KI-Systems sein werden. Diese Informationen werden gegebenenfalls im Einklang mit den Vorschriften und Gepflogenheiten auf Unionsebene und nationaler Ebene in Bezug auf die Unterrichtung der Arbeitnehmer und ihrer Vertreter bereitgestellt.

(8) Betreiber von Hochrisiko-KI-Systemen, bei denen es sich um Organe, Einrichtungen und sonstige Stellen der Union handelt, müssen den Registrierungspflichten gemäß Artikel 49 nachkommen. Stellen diese Betreiber fest, dass das Hochrisiko-IT-System, dessen Verwendung sie planen, nicht in der in Artikel 71 genannten EU-Datenbank registriert wurde, sehen sie von der Verwendung dieses Systems ab und informieren den Anbieter oder den Händler.

(10) Unbeschadet der Richtlinie (EU) 2016/680 beantragt der Betreiber eines Hochrisiko-KI-Systems zur nachträglichen biometrischen Fernfernidentifizierung im Rahmen von Ermittlungen zur gezielten Suche einer Person, die der Begehung einer Straftat verdächtigt wird oder aufgrund einer solchen verurteilt wurde, vorab oder unverzüglich, spätestens jedoch binnen 48 Stunden bei einer Justizbehörde oder einer Verwaltungsbehörde, deren Entscheidung bindend ist und einer gerichtlichen Überprüfung unterliegt, die Genehmigung für die Nutzung dieses Systems, es sei denn, es wird zur erstmaligen Identifizierung eines potenziellen Verdächtigen auf der Grundlage objektiver und nachprüfbarer Tatsachen, die in unmittelbarem Zusammenhang mit der Straftat stehen, verwendet. Jede Verwendung ist auf das für die Ermittlung einer bestimmten Straftat unbedingt erforderliche Maß zu beschränken.

Wird die beantragte Genehmigung gemäß Unterabsatz 1 abgelehnt, so wird die Verwendung des mit dieser beantragten Genehmigung verbundenen Systems zur nachträglichen biometrischen Fernidentifizierung mit sofortiger Wirkung eingestellt und werden die personenbezogenen Daten im Zusammenhang mit der Verwendung des Hochrisiko-KI-Systems, für das die Genehmigung beantragt wurde, gelöscht. In keinem Fall darf ein solches Hochrisiko-KI-System zur nachträglichen biometrischen Fernidentifizierung zu Strafverfolgungszwecken in nicht zielgerichteter Weise und ohne jeglichen Zusammenhang mit einer Straftat, einem Strafverfahren, einer tatsächlichen und bestehenden oder tatsächlichen und vorhersehbaren Gefahr einer Straftat oder der Suche nach einer bestimmten vermissten Person verwendet werden. Es muss sichergestellt werden, dass die Strafverfolgungsbehörden keine ausschließlich auf der Grundlage des Ergebnisses solcher Systeme zur nachträglichen biometrischen Fernidentifizierung beruhende Entscheidung, aus der sich eine nachteilige Rechtsfolge für eine Person ergibt, treffen.

Dieser Absatz gilt unbeschadet des Artikels 9 der Verordnung (EU) 2016/679 und des Artikels 10 der Richtlinie (EU) 2016/680 für die Verarbeitung biometrischer Daten.

Unabhängig vom Zweck oder Betreiber wird jede Verwendung solcher Hochrisiko-KI-Systeme in der einschlägigen Polizeiakte dokumentiert und der zuständigen Marktüberwachungsbehörde und der nationalen Datenschutzbehörde auf Anfrage zur Verfügung gestellt, wovon die Offenlegung sensibler operativer Daten im Zusammenhang mit der Strafverfolgung ausgenommen ist. Dieser Unterabsatz berührt nicht die den Aufsichtsbehörden durch die Richtlinie (EU) 2016/680 übertragenen Befugnisse. Die Betreiber legen den zuständigen Marktüberwachungsbehörden und den nationalen Datenschutzbehörden Jahresberichte über ihre Verwendung von Systemen zur nachträglichen biometrischen Fernidentifizierung vor, wovon die Offenlegung sensibler operativer Daten im Zusammenhang mit der Strafverfolgung ausgenommen ist. Die Berichte können eine Zusammenfassung sein, damit sie mehr als einen Einsatz abdecken. Die Mitgliedstaaten können im Einklang mit dem Unionsrecht strengere Rechtsvorschriften für die Verwendung von Systemen zur nachträglichen biometrischen Fernidentifizierung erlassen.

(11) Unbeschadet des Artikels 50 der vorliegenden Verordnung informieren die Betreiber der in Anhang III aufgeführten Hochrisiko-KI-Systeme, die natürliche Personen betreffende Entscheidungen treffen oder bei solchen Entscheidungen Unterstützung leisten, die natürlichen Personen darüber, dass sie der Verwendung des Hochrisiko-KI-Systems unterliegen. Für Hochrisiko-KI-Systeme, die zu Strafverfolgungszwecken verwendet werden, gilt Artikel 13 der Richtlinie (EU) 2016/680.

(12) Die Betreiber arbeiten mit den zuständigen Behörden bei allen Maßnahmen zusammen, die diese Behörden im Zusammenhang mit dem Hochrisiko-KI-System zur Umsetzung dieser Verordnung ergreifen.

Betreiber von **Hochrisiko-KI-Systemen** müssen unter anderem

- geeignete technische und organisatorische Maßnahmen treffen, um sicherzustellen, dass sie Hochrisiko-KI-Systeme gemäß der Gebrauchsanweisung nutzen (Art. 26 Abs. 1 KI-VO),
- menschliche Aufsicht über das Hochrisiko-KI-System ermöglichen und dabei durch erforderliche Unterstützung sicherstellen, dass diese kompetent und ausreichend qualifiziert sind (Art. 26 Abs. 2 KI-VO),
- die Funktionsweise des Hochrisiko-KI-Systems überwachen, d. h.
- den Betrieb des Hochrisiko-KI-Systems anhand der Gebrauchsanweisung überwachen und falls nötig, Anbieter gem. Art. 72 KI-VO informieren (Art. 26 Abs. 5 KI-VO),
- die Nutzung des Hochrisiko-KI-Systems aussetzen und den Anbieter, Händler und die zuständige Behörde informieren, wenn der Verdacht besteht, dass die Nutzung des Hochrisiko-KI-Systems gemäß der Gebrauchsanweisung ein Risiko gem. Art. 79 Abs. 1 darstellt (Art. 26 Abs. 5 KI-VO),
- mit Ausnahme von sensiblen operativen Daten von Strafverfolgungsbehörden, die Betreiber von KI-Systemen sind, bei Feststellung eines schwerwiegenden Vorfalls zunächst den Anbieter, dann den Einführer oder Händler und die zuständigen Behörden informieren (Art. 26 Abs. 5 KI-VO),
- die automatisch erzeugten Protokolle ihrer Hochrisiko-KI-Systeme mindestens sechs Monate aufbewahren, wenn sie unter ihrer Kontrolle stehen (Art. 26 Abs. 6 KI-VO),
- vor einer Inbetriebnahme oder Nutzung des Hochrisiko-KI-Systems am Arbeitsplatz die Arbeitnehmervertreter und die betroffenen Arbeitnehmer darüber informieren (Art. 26 Abs. 7 KI-VO). Diese Information muss entsprechend den EU- und nationalen Vorschriften und Gepflogenheiten zur Information der Arbeitnehmer und ihrer Vertreter erfolgen.
- der Registrierungspflicht gem. Art. 49 KI-VO nachkommen, wenn es sich bei dem Betreiber um ein Organ, eine Einrichtung oder sonstige Stelle der Union handelt. Zudem müssen solche Betreiber von der Verwendung des Hochrisiko-KI-Systems absehen, wenn es nicht in der EU-Datenbank gem. Art. 71 KI-VO registriert wurde, und den Anbieter oder Händler diesbezüglich informieren (Art. 26 Abs. 8 KI-VO).
- ihrer Pflicht zur Durchführung einer Datenschutz-Folgenabschätzung gemäß Artikel 35 der Verordnung (EU) 2016/679 oder Artikel 27 der Richtlinie (EU) 2016/680 nachkommen, indem sie die nach Art. 13 KI-VO bereitgestellten Informationen verwenden (Art. 26 Abs. 9 KI-VO),
- als Betreiber eines Hochrisiko-KI-Systems zur nachträglichen biometrischen Fernidentifizierung im Rahmen von Ermittlungen vor der Nutzung oder spätestens innerhalb von 48 Stunden eine Genehmigung von einer Justiz- oder Verwaltungsbehörde einholen, es sei denn, das System wird erstmals zur Identifizierung eines potenziellen Verdächtigen basierend auf objektiven und überprüfbaren Tatsachen verwendet, die im unmittelbaren Zusammenhang mit der Straftat stehen. Wenn die Genehmigung abgelehnt wird, muss die Nutzung des Systems sofort eingestellt und alle damit verbundenen personenbezogenen Daten gelöscht werden. Zudem ist jede solche Verwendung in der jeweiligen Polizeiakte zu dokumentieren und der zuständigen Marktüberwachungsbehörde und nationalen Datenschutzbehörde auf Anfrage zur

Verfügung zu stellen, wobei strafverfolgungsbezogene sensible operative Daten davon ausgenommen sind (Art. 26 Abs. 10 KI-VO).

- mit Ausnahme von sensiblen operativen Daten im Zusammenhang mit der Strafverfolgung den zuständigen Marktüberwachungsbehörden und den nationalen Datenschutzbehörden Jahresberichte über die Nutzung von Systemen zur nachträglichen biometrischen Fernidentifizierung vorlegen (Art. 26 Abs. 10 KI-VO).
- unabhängig von den Transparenzpflichten gem. Art. 50 KI-VO, betroffene natürliche Personen darüber informieren, wenn es sich um ein Hochrisiko-KI-System handelt, das in Anhang III aufgeführt ist und das Entscheidungen über natürliche Personen trifft oder bei Entscheidungen unterstützt, dass das Hochrisiko-KI-System bei ihrem Fall eingesetzt wird (Art. 26 Abs. 11 KI-VO),
- mit den zuständigen Behörden im Zusammenhang mit etwaigen Maßnahmen, die die Behörden zur Umsetzung der KI-VO ergreifen, kooperieren (Art. 26 Abs. 12 KI-VO),

Schritt 4.2.3: Betreiberpflicht zur Grundrechte-Folgenabschätzung nach Art. 27

(1) Vor der Inbetriebnahme eines Hochrisiko-KI-Systems gemäß Artikel 6 Absatz 2 — mit Ausnahme von Hochrisiko-KI-Systemen, die in dem in Anhang III Nummer 2 aufgeführten Bereich verwendet werden sollen — führen Betreiber, bei denen es sich um Einrichtungen des öffentlichen Rechts oder private Einrichtungen, die öffentliche Dienste erbringen, handelt, und Betreiber von Hochrisiko-KI-Systemen gemäß Anhang III Nummer 5 Buchstaben b und c eine Abschätzung der Auswirkungen, die die Verwendung eines solchen Systems auf die Grundrechte haben kann, durch. Zu diesem Zweck führen die Betreiber eine Abschätzung durch, die Folgendes umfasst:

a) eine Beschreibung der Verfahren des Betreibers, bei denen das Hochrisiko-KI-System im Einklang mit seiner Zweckbestimmung verwendet wird;

b) eine Beschreibung des Zeitraums und der Häufigkeit, innerhalb dessen bzw. mit der jedes Hochrisiko-KI-System verwendet werden soll;

c) die Kategorien der natürlichen Personen und Personengruppen, die von seiner Verwendung im spezifischen Kontext betroffen sein könnten;

d) die spezifischen Schadensrisiken, die sich auf die gemäß Buchstabe c dieses Absatzes ermittelten Kategorien natürlicher Personen oder Personengruppen auswirken könnten, unter Berücksichtigung der vom Anbieter gemäß Artikel 13 bereitgestellten Informationen;

e) eine Beschreibung der Umsetzung von Maßnahmen der menschlichen Aufsicht entsprechend den Betriebsanleitungen;

f) die Maßnahmen, die im Falle des Eintretens dieser Risiken zu ergreifen sind, einschließlich der Regelungen für die interne Unternehmensführung und Beschwerdemechanismen.

(2) Die in Absatz 1 festgelegte Pflicht gilt für die erste Verwendung eines Hochrisiko-KI-Systems. Der Betreiber kann sich in ähnlichen Fällen auf zuvor durchgeführte Grundrechte-Folgenabschätzungen oder bereits vorhandene Folgenabschätzungen, die vom Anbieter durchgeführt wurden, stützen. Gelangt der Betreiber während der Verwendung des Hochrisiko-KI-Systems zur Auffassung, dass sich eines der in Absatz 1 aufgeführten Elemente geändert hat oder nicht mehr auf dem neuesten Stand ist, so unternimmt der Betreiber die erforderlichen Schritte, um die Informationen zu aktualisieren.

(3) Sobald die Abschätzung gemäß Absatz 1 des vorliegenden Artikels durchgeführt wurde, teilt der Betreiber der Marktüberwachungsbehörde ihre Ergebnisse mit, indem er das ausgefüllte, in Absatz 5 des vorliegenden Artikels genannte Muster als Teil der Mitteilung übermittelt. In dem in Artikel 46 Absatz 1 genannten Fall können die Betreiber von der Mitteilungspflicht befreit werden.

(4) Wird eine der in diesem Artikel festgelegten Pflichten bereits infolge einer gemäß Artikel 35 der Verordnung (EU) 2016/679 oder Artikel 27 der Richtlinie (EU) 2016/680 durchgeführten Datenschutz-Folgenabschätzung erfüllt, so ergänzt die Grundrechte-Folgenabschätzung gemäß Absatz 1 des vorliegenden Artikels diese Datenschutz-Folgenabschätzung.

(5) Das Büro für Künstliche Intelligenz arbeitet ein Muster für einen Fragebogen — auch mithilfe eines automatisierten Instruments — aus, um die Betreiber in die Lage zu versetzen, ihren Pflichten gemäß diesem Artikel in vereinfachter Weise nachzukommen.

Betreiber, bei denen es sich um Einrichtungen des öffentlichen Rechts oder private Einrichtung, die öffentliche Dienste erbringt, handelt, haben außerdem eine Überprüfung im Hinblick auf Auswirkungen des Hochrisiko-KI-Systems auf Grundrechte durchführen, wenn es Hochrisiko-KI-Systeme gemäß Artikel 6 Absatz 2 — mit Ausnahme von Hochrisiko-KI-Systemen, die in dem in Anhang III Nummer 2 aufgeführten Bereich verwendet werden sollen.

Dies hat vor der Inbetriebnahme eines Hochrisiko-KI-Systems gemäß Art. 6 Abs. 2 KI-VO zu erfolgen. Die gleiche Verpflichtung gilt für Betreiber von Hochrisiko-KI-Systemen, die

- bestimmungsgemäß für die Kreditwürdigkeitsprüfung und die Bonitätsbewertung natürlicher Personen verwendet werden sollen, mit Ausnahme von KI-Systemen, die zur Aufdeckung von Finanzbetrug verwendet werden (Art. 27 Abs. 1 i.V.m. Anhang III Nr. 5 lit. b KI-VO);
- bestimmungsgemäß für die Risikobewertung und Preisbildung in Bezug auf natürliche Personen im Fall von Lebens- und Krankenversicherungen verwendet werden sollen (Art. 27 Abs. 1 i.V.m. Anhang III Nr. 5 lit. c KI-VO).

Die Abschätzung der Auswirkung des Hochrisiko-KI-Systems auf Grundrechte umfasst

- eine Beschreibung der Verfahren des Betreibers, bei denen das Hochrisiko-KI-System im Einklang mit seiner Zweckbestimmung verwendet wird,
- eine Beschreibung des Zeitraums und der Häufigkeit, innerhalb dessen bzw. mit der jedes Hochrisiko-KI-System verwendet werden soll,

- die Kategorien der natürlichen Personen und Personengruppen, die von seiner Verwendung im spezifischen Kontext betroffen sein könnten,
- die spezifischen Schadensrisiken, dieser Kategorien natürlicher Personen oder Personengruppen, wobei hierbei auch die vom Betreiber bereitgestellten Informationen nach Art. 13 KI-VO berücksichtigt werden,
- eine Beschreibung der Umsetzung von Maßnahmen der menschlichen Aufsicht entsprechend den Betriebsanleitungen,
- die Maßnahmen, die im Falle des Eintretens dieser Risiken zu ergreifen sind, einschließlich der Regelungen für die interne Unternehmensführung und Beschwerdemechanismen.

Die Pflicht zur Grundrechte-Folgenabschätzung gilt bereits für die erste Verwendung eines Hochrisiko-KI-Systems. In ähnlich gelagerten Fällen kann sich der Betreiber auf bestehende Folgenabschätzungen berufen. Sofern der Betreiber während der Verwendung des Hochrisiko-KI-Systems der Auffassung ist, dass sich eines der Elemente der Abschätzung geändert hat, hat er die erforderlichen Schritte zu unternehmen, um die Information zu aktualisieren (Art. 27 Abs. 2 KI-VO).

Nach der Durchführung der Abschätzung hat der Betreiber die Ergebnisse dieser Abschätzung der Marktüberwachungsbehörde mitzuteilen. Hierfür hat er den Musterfragebogen des Büros für Künstliche Intelligenz (Art. 27 Abs. 5 KI-VO) auszufüllen und der Marktüberwachungsbehörde zu übermitteln. Sofern die Marktüberwache eine Ausnahme vom Konformitätsbewertungsverfahren gem. Art. 46 Abs. 1 KI-VO macht, können die Betreiber auch von der Mitteilungspflicht befreit werden (Art. 27 Abs. 3 KI-VO).
Mit dem Fragebogenmuster sollen Betreiber, die Erfüllung ihrer Grundrechte-Folgenabschätzung für Hochrisiko-KI-Systeme erleichtert werden (Art. 27 Abs. 5 KI-VO).

Wenn bereits eine Datenschutz-Folgenabschätzung infolge von Artikel 35 der Verordnung (EU) 2016/679 oder Artikel 27 der Richtlinie (EU) 2016/680 vorliegt, ergänzt die Grundrechte-Folgenabschätzung die Datenschutz-Folgenabschätzung (Art. 27 Abs. 4 KI-VO).

Schritt 4.2.4: Transparenzpflichten des Betreibers, Art. 50 Abs. 3 und 4

(3) Die Betreiber eines Emotionserkennungssystems oder eines Systems zur biometrischen Kategorisierung informieren die davon betroffenen natürlichen Personen über den Betrieb des Systems und verarbeiten personenbezogene Daten gemäß den Verordnungen (EU) 2016/679 und (EU) 2018/1725 und der Richtlinie (EU) 2016/680. Diese Pflicht gilt nicht für gesetzlich zur Aufdeckung, Verhütung oder Ermittlung von Straftaten zugelassene KI-Systeme, die zur biometrischen Kategorisierung und Emotionserkennung im Einklang mit dem Unionsrecht verwendet werden, sofern geeignete Schutzvorkehrungen für die Rechte und Freiheiten Dritter bestehen.

(4) Betreiber eines KI-Systems, das Bild-, Ton- oder Videoinhalte erzeugt oder manipuliert, die ein Deepfake sind, müssen offenlegen, dass die Inhalte künstlich erzeugt oder manipuliert wurden. Diese Pflicht gilt nicht, wenn die Verwendung zur Aufdeckung, Verhütung, Ermittlung oder Verfolgung von Straftaten gesetzlich zugelassen ist. Ist der Inhalt Teil eines offensichtlich künstlerischen, kreativen, satirischen, fiktionalen oder analogen Werks oder Programms, so beschränken sich die in diesem Absatz festgelegten Transparenzpflichten darauf, das Vorhandensein solcher erzeugten oder manipulierten Inhalte in geeigneter Weise offenzulegen, die die Darstellung oder den Genuss des Werks nicht beeinträchtigt. Betreiber eines KI-Systems, das Text erzeugt oder manipuliert, der veröffentlicht wird, um die Öffentlichkeit über Angelegenheiten von öffentlichem Interesse zu informieren, müssen offenlegen, dass der Text künstlich erzeugt oder manipuliert wurde. Diese Pflicht gilt nicht, wenn die Verwendung zur Aufdeckung, Verhütung, Ermittlung oder Verfolgung von Straftaten gesetzlich zugelassen ist oder wenn die durch KI erzeugten Inhalte einem Verfahren der menschlichen Überprüfung oder redaktionellen Kontrolle unterzogen wurden und wenn eine natürliche oder juristische Person die redaktionelle Verantwortung für die Veröffentlichung der Inhalte trägt.

Betreiber von **Emotionserkennungssystemen** oder **Systemen zur biometrischen Kategorisierung** müssen spätestens zum Zeitpunkt der Erstinteraktion oder -aussetzung in klar verständlicher Weise

- betroffene Personen über den Betrieb eines solchen Systems informieren (Art. 50 Abs. 3 KI-VO),

- bei der Verarbeitung personenbezogener Daten die Verordnung (EU) 2016/679, Verordnung (EU) 2016/1725 bzw. die Richtlinie (EU) 2016/280 einhalten (Art. 50 Abs. 3 KI-VO),
- ausgenommen hiervon sind jedoch Emotionserkennungssysteme und Systeme zur biometrischen Kategorisierung, die zur Aufdeckung, Verhütung und Ermittlung von Straftaten zugelassen sind und im Einklang mit dem Unionsrecht verwendet werden, wenn die Rechte und Freiheiten Dritter geschützt sind (Art. 50 Abs. 3 KI-VO).

Betreiber von KI-Systemen, die sogenannte **Deepfakes** erstellen, müssen spätestens zum Zeitpunkt der Erstinteraktion oder -aussetzung in klar verständlicher Weise

- offenlegen, dass die Inhalte künstlich erzeugt oder manipuliert wurden, außer das KI-System ist zur Aufdeckung, Verhütung, Ermittlung und Verfolgung von Straftaten gesetzlich zugelassen (Art. 50 Abs. 4 KI-VO).

Betreiber von KI-Systemen, die **Texte erzeugen oder manipulieren**, die veröffentlicht werden, um die Öffentlichkeit zu informieren, müssen spätestens zum Zeitpunkt der Erstinteraktion oder -aussetzung in klar verständlicher Weise

- angeben, dass diese Texte künstlich erzeugt oder manipuliert wurden.
- Diese Pflicht entfällt, wenn
- die Nutzung des KI-Systems gesetzlich erlaubt ist, um Straftaten aufzudecken, zu verhindern, zu ermitteln und zu verfolgen oder
- die KI-erzeugten Inhalte von Menschen überprüft oder redaktionell kontrolliert werden und eine Person oder Organisation die redaktionelle Verantwortung für die Veröffentlichung übernimmt (Art. 50 Abs. 4 KI-VO).

Um die wirksame Umsetzung der Pflichten bezüglich künstlich erzeugter oder manipulierter Inhalte zu erleichtern, unterstützt das Amt für künstliche Intelligenz die Erstellung von Verhaltenskodizes auf EU-Ebene. Die EU-Kommission hat die Befugnis, diese Kodizes durch Durchführungsrechtsakte zu genehmigen. Wenn die Kommission einen Kodex für unangemessen hält, kann durch Durchführungsrechtsakt eigene verbindliche Vorschriften für die Umsetzung dieser Pflichten erlassen (Art. 50 Abs. 7 KI-VO).

Parallel zu den in Art. 50 Abs. 3 bis 4 festgelegten Anforderungen und Pflichten haben Betreiber solcher KI-Systeme andere einschlägige unionsrechtliche oder nationalrechtliche Vorschriften einzuhalten (Art. 50 Abs. 6 KI-VO).

Zwischenergebnis

Wenn die Betreibereigenschaft gegeben ist und die Betreiberpflichten erfüllt sind, ist als Nächstes mit **Schritt 5.1 (KI-Kompetenz)** fortzufahren.

7 Compliance-Anforderungen für KI-Systeme mit geringem Risiko

Zusätzlich zu den in den vorangegangenen Abschnitten aufgeführten speziellen Pflichten für Hochrisiko-KI-Systeme sieht die KI-VO weitere allgemeine Transparenz- und Compliance-Pflichten mit einem breiteren Adressatenkreis vor. Diese Pflichten, die in diesem Abschnitt vorgestellt werden, gelten je nach Anwendungsbereich für eine oder mehrere der folgenden Gruppen von KI:

- **KI-Systeme unterhalb der Hochrisiko-Schwelle** – für Anbieter und Betreiber: Neben den Hochrisiko-KI-Systemen reguliert die KI-VO auch bestimmte KI-Systeme, von denen ein geringeres, unterhalb der Hochrisiko-Schwelle zu verortendes Risiko ausgeht. Sollte eine Prüfung des Schritts 2 ergeben haben, dass kein hohes Risiko im Sinne der KI-VO gegeben ist, ist die Prüfung hier fortzusetzen.
- **Hochrisiko-KI-Systeme** – für Anbieter und Betreiber: Sollte die Prüfung ergeben haben, dass es sich bei dem in Rede stehenden KI-System um ein Hochrisiko-KI-System handelt, ist die Prüfung hier ebenfalls fortzusetzen. Compliance-Anforderungen für KI-Systeme mit geringerem Risiko gelten (erst recht) auch für Hochrisiko-KI-Systeme.
- **KI-Modelle mit allgemeinem Verwendungszweck (mit systemischem Risiko)** – für Anbieter: Schließlich stellt die KI-VO auch spezielle Pflichten für Anbieter von KI-Modellen mit allgemeinem Verwendungszweck auf. Sollte die Prüfung ergeben haben, dass es sich bei der in Rede stehenden KI um ein solches KI-Modell mit allgemeinem Verwendungszweck handelt, ggf. sogar mit systemischem Risiko, dann ist die Prüfung hier fortzusetzen.

Wichtig: Richtige Einschätzung der Qualifikation als Betreiber und/oder Anbieter

Für die korrekte Erfassung des individuellen Pflichtenkatalogs ist es unerlässlich, eine zutreffende Einschätzung über die eigene Einordnung als Anbieter und/oder Betreiber eines erfassten (Hochrisiko-)KI-Systems bzw. als Anbieter eines erfassten KI-Modells vorzunehmen (s. hierzu oben Schritt 3.3.1 und 3.3.2).

Schritt 5.1: Wie baue ich KI-Kompetenz auf?

Tim Sauerhammer (Reed Smith LLP), Alexander Schmalenberger (Taylor Wessing Partnerschaftsgesellschaft mbB)

Art. 4 KI-VO sieht vor, dass im Hinblick auf alle KI-Systeme eine hinreichende **KI-Kompetenz** aufzubauen ist.

Relevante(r) Artikel:

Art.3 lit. 56, 4

Relevante(r) ErwG:

20, 21, 73

Konkretisierungsbedürftig:

Ja durch, juristische Auslegungsmethoden.

> Die Anbieter und Betreiber von KI-Systemen ergreifen Maßnahmen, um nach besten Kräften sicherzustellen, dass ihr Personal und andere Personen, die in ihrem Auftrag mit dem Betrieb und der Nutzung von KI-Systemen befasst sind, über ein ausreichendes Maß an KI-Kompetenz verfügen, wobei ihre technischen Kenntnisse, ihre Erfahrung, ihre Ausbildung und Schulung und der Kontext, in dem die KI-Systeme eingesetzt werden sollen, sowie die Personen oder Personengruppen, bei denen die KI-Systeme eingesetzt werden sollen, zu berücksichtigen sind.

KI-Kompetenz ist **in Art. 3 lit. 56 KI-VO legaldefiniert** als:

> die Fähigkeiten, die Kenntnisse und das Verständnis, die es Anbietern, Betreibern und Betroffenen unter Berücksichtigung ihrer jeweiligen Rechte und Pflichten im Rahmen dieser Verordnung ermöglichen, KI-Systeme sachkundig einzusetzen sowie sich der Chancen und Risiken von KI und möglicher Schäden, die sie verursachen kann, bewusst zu werden.

Überblick und Allgemeines

Artikel 4 der KI-Verordnung (KI-VO) legt fest, dass Anbieter und Betreiber von KI-Systemen Maßnahmen ergreifen müssen, um sicherzustellen, dass ihr Personal und andere Personen, die in ihrem Auftrag mit dem Betrieb und der Nutzung von KI-Systemen befasst sind, über ausreichende KI-Kompetenz verfügen. Diese Kompetenz umfasst technische Kenntnisse, Erfahrung, Ausbildung und Schulung sowie den Kontext, in dem die KI-Systeme eingesetzt werden sollen, und die Personen oder Personengruppen, bei denen die KI-Systeme eingesetzt werden sollen.

Definition von Personal und im Auftrag befasste andere Personen

Welche Personengruppen genau zu schulen sind, ist leider nicht eindeutig definiert. Es ist daher wichtig zu betonen, dass die folgenden Definitionen und Interpretationen auf unseren eigenen Analysen basieren und nicht direkt aus dem Gesetz abgeleitet werden können. Personal umfasst in der Regel alle Personen, die in einem Arbeitsverhältnis zum Unternehmen stehen und somit dem Weisungsrecht des Arbeitgebers unterliegen. Dies schließt sowohl festangestellte Mitarbeiter als auch Leiharbeitnehmer ein, sofern letztere in die Betriebsorganisation des Unternehmens integriert sind und dessen Weisungen folgen. Im Auftrag befasste andere Personen können sowohl freie Dienstleister als auch Subunternehmer umfassen, die nicht direkt dem Weisungsrecht des Unternehmens unterliegen, sondern ihre Aufgaben eigenverantwortlich und selbstständig ausführen. Diese Personen sind in der Regel nicht in die Betriebsorganisation des Auftraggebers integriert und unterliegen nicht dessen Weisungen.

Verpflichtungen der Anbieter und Betreiber

Die Verordnung betont die Notwendigkeit der „digitalen Alphabetisierung“ des Personals und der anderen Personen, was bedeutet, dass die Mitarbeiter über grundlegende Kenntnisse und Fähigkeiten im Umgang mit KI-Systemen verfügen müssen. Diese Maßnahmen sind Teil eines umfassenderen Rahmens, der sicherstellen soll, dass KI-Systeme sicher und verantwortungsvoll betrieben werden. Dazu gehören auch technische und organisatorische Maßnahmen sowie Transparenzpflichten, insbesondere bei Hochrisiko-KI-Systemen. Zusätzlich zu den Anforderungen an die Kompetenz des Personals müssen Betreiber von Hochrisiko-KI-Systemen vor der Inbetriebnahme eine Grundrechte-Folgenabschätzung durchführen, um die Auswirkungen auf die Grundrechte Dritter zu bewerten. Dies zeigt, dass die Verordnung nicht nur technische Fähigkeiten, sondern auch ethische und rechtliche Aspekte berücksichtigt. Insgesamt zielt Artikel 4 der KI-VO darauf ab, ein hohes Maß an Kompetenz und Verantwortungsbewusstsein bei allen Beteiligten sicherzustellen, um die sichere und ethische Nutzung von KI-Systemen zu gewährleisten.

Allgemeiner Lernpfad mit Zertifikaten

Ein allgemeiner Lernpfad mit Zertifikaten kann als unverbindlicher Vorschlag formuliert werden, der auf den jeweiligen Einsatz der KI und das Unternehmen angepasst werden muss. Dieser Lernpfad soll als Orientierungshilfe dienen und kann je nach spezifischen Anforderungen und Kontext des Unternehmens modifiziert werden. Der Lernpfad könnte folgende Stufen umfassen:

Level 1: Grundlegende KI-Kompetenz

Der erste Level fördert das selbstgesteuerte Lernen und führt zur Zertifizierung in KI-Grundkenntnissen. Ziel ist es, die Lernenden in die Welt des maschinellen Lernens

einzuführen und das Thema Ethik in der KI zu erkunden. Die Module umfassen Grundlagen der KI, Ethik in der KI und praktische Übungen. Die geschätzte Dauer beträgt etwa fünf Stunden. Die Lernressourcen bestehen aus selbstgesteuertem Online-Lernen, Zugang zu Community-Gruppen zur Unterstützung des Lernens sowie Zugang zu Ask-Me-Anything-Sessions zur Unterstützung des Lernens, Netzwerkens und zur Karriereentwicklung.

Level 2: Grundlagen der KI

Der zweite Level fördert ebenfalls das selbstgesteuerte Lernen und führt zur Zertifizierung in den Grundlagen der KI. Ziel ist es, den Lernenden zu helfen, KI angemessen zu verstehen und zu nutzen sowie grundlegende KI- und Datenanwendungen in Python zu programmieren. Die Module umfassen unter anderem die Einführung in Python, Bibliotheken und Datenmanipulation, explorative Datenanalyse, statistisches Denken, überwachtes und unüberwachtes Lernen, Deep Learning sowie andere Programmiersprachen und Tools. Die geschätzte Dauer beträgt etwa 140 Stunden. Die Lernressourcen bestehen aus selbstgesteuertem Online-Lernen, Zugang zu Community-Gruppen zur Unterstützung des Lernens sowie Zugang zu monatlichen Ask-Me-Anything-Sessions zur Unterstützung des Lernens, Netzwerkens und zur Karriereentwicklung. Voraussetzung für diesen Level ist die Zertifizierung in KI-Grundkenntnissen (Level 1).

Level 3: Maschinelles lernen anwenden

Der dritte Level fördert das selbstgesteuerte Lernen und führt zur selbstständigen Entwicklung von ML-Methoden. Ziel dieser Zertifizierung ist es, die Fähigkeit der Lernenden zu bewerten, ML-Probleme zu formulieren, ML-Modelle zu entwickeln, ML-Lösungen zu entwerfen, ML-Pipelines zu automatisieren und zu orchestrieren, Systeme zur Datenvorbereitung und -verarbeitung zu gestalten sowie ML-Lösungen zu überwachen, zu optimieren und zu warten. Die Lernressourcen umfassen Zugang zu Online-Lernplattformen und Community-Gruppen.

Vertiefte Diskussion

Die technische Kompetenz ist das Fundament der KI-Kompetenz. Sie umfasst ein tiefes Verständnis der Algorithmen und Modelle, die KI-Systeme antreiben, sowie die Fähigkeit, diese Modelle zu entwickeln und zu implementieren. Dies erfordert Kenntnisse in Bereichen wie maschinelles Lernen, Datenanalyse und Programmierung. Ein tiefes Verständnis der verschiedenen Algorithmen des maschinellen Lernens, einschließlich überwachten und unüberwachten Lernens, ist unerlässlich. Dies umfasst auch das Wissen über neuronale Netze, Entscheidungsbäume und andere gängige Modelle. Die Fähigkeit, große Datensätze zu analysieren und zu interpretieren, ist entscheidend. Dies umfasst Kenntnisse in Statistik, Datenvisualisierung und Datenvorverarbeitung. Grundlegende Programmierkenntnisse in Sprachen wie Python, R oder Java sind notwendig, um KI-Modelle zu entwickeln und zu implementieren.

Neben den technischen Fähigkeiten ist die praktische Anwendung von KI-Systemen ein weiterer wichtiger Aspekt der KI-Kompetenz. Dies umfasst die Fähigkeit, KI-Tools und -Plattformen zur Lösung von Problemen zu nutzen und die Ergebnisse dieser Systeme zu interpretieren und zu bewerten. Die Fähigkeit, verschiedene KI-Tools und -Plattformen zu nutzen, ist entscheidend. Dies umfasst auch die Fähigkeit, diese Tools in bestehende Arbeitsprozesse zu integrieren. Die Fähigkeit, die Ergebnisse von KI-Systemen kritisch zu bewerten und zu interpretieren, ist wichtig. Dies umfasst auch das Verständnis der Grenzen

und möglichen Fehlerquellen der Systeme. Die Fähigkeit, KI-Systeme in bestehende Arbeitsprozesse zu integrieren und deren Nutzung zu optimieren, ist entscheidend für die effektive Anwendung der Technologie.

Ein weiterer wichtiger Aspekt der KI-Kompetenz ist das Verständnis der ethischen und gesellschaftlichen Implikationen der Nutzung von KI. Dies umfasst die Berücksichtigung von Fairness, Transparenz und Verantwortlichkeit bei der Entwicklung und Anwendung von KI-Systemen. Die Fähigkeit, ethische Überlegungen in die Entwicklung und Anwendung von KI-Systemen einzubeziehen, ist entscheidend. Dies umfasst die Berücksichtigung von Fragen der Fairness, Transparenz und Verantwortlichkeit. Das Verständnis der möglichen gesellschaftlichen Auswirkungen der Nutzung von KI ist wichtig. Dies umfasst die Berücksichtigung von Fragen der Arbeitsplatzsicherheit, des Datenschutzes und der sozialen Gerechtigkeit. Die Fähigkeit, Verantwortung für die Auswirkungen der Nutzung von KI-Systemen zu übernehmen und Maßnahmen zu ergreifen, um mögliche negative Auswirkungen zu minimieren, ist entscheidend.

Die kontinuierliche Schulung und Weiterbildung ist ein zentraler Aspekt der Förderung der KI-Kompetenz. Anbieter und Betreiber von KI-Systemen müssen sicherstellen, dass ihr Personal regelmäßig geschult und weitergebildet wird, um mit den neuesten Entwicklungen und Best Practices im Bereich der KI schrittzuhalten. Regelmäßige Schulungen und Fortbildungen sind notwendig, um sicherzustellen, dass das Personal stets auf dem neuesten Stand ist. Dies umfasst sowohl technische als auch ethische Schulungen. Schulungen sollten interdisziplinär angelegt sein und sowohl technische als auch ethische Aspekte der KI abdecken. Dies fördert ein umfassendes Verständnis der Technologie und ihrer Auswirkungen. Praktische Übungen und Fallstudien sind wichtig, um das theoretische Wissen in die Praxis umzusetzen. Dies kann durch Workshops, Simulationen und andere praxisorientierte Lernmethoden erreicht werden.

Folgen und Anforderungen an Schulungen

Die Anbieter und Betreiber von KI-Systemen müssen sicherstellen, dass ihr Personal und andere beauftragte Personen über ausreichende KI-Kompetenz verfügen. Dies kann durch verschiedene Maßnahmen wie Schulungen, Weiterbildungen und Trainings erfolgen. Ein einfaches PDF zum Selbststudium dürfte in den meisten Fällen nicht ausreichen, da eine vertiefte Schulung erforderlich ist, um die notwendigen technischen Kenntnisse und Fähigkeiten zu vermitteln. Wenn die erforderlichen Schulungen nicht angeboten werden, könnten Personal und andere beauftragte Personen unter bestimmten Umständen ein Leistungsverweigerungsrecht haben. Dies ist vergleichbar mit Regelungen im Arbeitsschutz, wo Arbeitnehmer das Recht haben, die Arbeit zu verweigern, wenn die Sicherheitsvorschriften nicht eingehalten werden. Regelungen zur Schulung von gefährlichen Arbeitsmitteln und die Folgen bei Unterlassung der Schulung finden sich im Arbeitsschutzrecht. Hier sind Arbeitgeber verpflichtet, ihre Mitarbeiter umfassend zu unterweisen und zu schulen, um Unfälle und Gesundheitsgefahren zu vermeiden. Bei Nichteinhaltung dieser Pflichten können Arbeitnehmer ihre Arbeitsleistung verweigern und der Arbeitgeber haftet für eventuelle Schäden. Diese Überlegungen könnten auch auf den Bereich der KI-Kompetenz übertragen werden.

Zusammenfassend lässt sich sagen, dass sowohl festangestellte Mitarbeiter als auch freie Dienstleister und Subunternehmer unter den Begriffen „Personal“ und „im Auftrag befasste andere Personen“ fallen können, je nach ihrem Grad der Weisungsgebundenheit und Integration in die Betriebsorganisation. Die Schulungsmaßnahmen müssen umfassend und

vertieft sein, um den Anforderungen gerecht zu werden, und es bestehen Leistungsverweigerungsrechte, wenn diese Schulungen nicht angeboten werden.

Beispiele und Besprechung aus einem weiten Anwendungsfeld von KI-Anwendungen

Beispiel 1: KI in der medizinischen Diagnose

Ein KI-System zur Unterstützung bei der Diagnose muss die Grenzen und Möglichkeiten des Systems verstehen. Die Schulung umfasst die Interpretation von KI-Ausgaben und die Erkennung von Anomalien. Erforderliche Schulungen könnten Online-Kurse, Workshops mit Experten und praktische Übungen umfassen.

Beispiel 2: KI in der industriellen Produktion

Ein KI-gesteuertes Produktionssystem erfordert Schulungen zur Überwachung des Systems und zum Eingriff bei Fehlfunktionen. Der Fokus liegt auf den technischen Aspekten und möglichen Risiken. Schulungen könnten durch formale Programme, Simulationen und regelmäßige Fortbildungen erfolgen.

Beispiel 3: KI in der Kundenbetreuung

Ein KI-Tool zur Beantwortung von Kundenanfragen benötigt Schulungen zur Funktionsweise des Tools und zur Erkennung von fehlerhaften Antworten. Schulungen könnten Online-Kurse, interaktive Workshops und praxisorientierte Übungen umfassen.

Beispiel 4: KI in der industriellen Anwendung

Ein KI-Dienstleistungskatalog umfasst Fähigkeiten wie Computer Vision für die visuelle Qualitätskontrolle, Mustererkennung zur Erkennung von Anomalien, Verarbeitung natürlicher Sprache zur Klassifizierung von Daten, Optimierung zur Verbesserung der Planung und hybride Modellierung zur Erstellung von Ersatzmodellen physikalischer Systeme. Schulungen könnten durch Expertenvorträge, praktische Workshops und Online-Ressourcen erfolgen.

Beispiel 5: KI im Bankwesen

KI-Anwendungen im Bankwesen umfassen die Automatisierung interner Prozesse, die Verbesserung des Kundenerlebnisses, die Informationsbeschaffung für Pitches, die automatisierte Dokumentenverarbeitung, die Personalisierung der Kundenkommunikation und die Analyse von unstrukturierten Daten. Erforderliche Schulungen könnten Online-Kurse, praxisorientierte Workshops und kontinuierliche Weiterbildung umfassen.

Zusammenfassung

Die Förderung der KI-Kompetenz gemäß Art. 4 KI-VO ist entscheidend für den sicheren und effektiven Einsatz von KI-Systemen. Anbieter und Betreiber müssen sicherstellen, dass ihr Personal über die notwendigen Fähigkeiten und Kenntnisse verfügt, um die Chancen und Risiken von KI zu verstehen und sachkundig mit diesen Systemen umzugehen. Die Anforderungen an die menschliche Aufsicht nach Art. 14 und 26 der KI-VO ergänzen diese Kompetenzanforderungen, indem sie spezifische Maßnahmen zur Überwachung und Kontrolle von Hochrisiko-KI-Systemen festlegen.

Zwischenergebnis

Ist ein ausreichendes Maß an KI-Kompetenz sichergestellt, geht es weiter mit **Schritt 5.2 (Transparenzpflichten).**

Schritt 5.2: Wie gewährleiste ich Transparenz für bestimmte KI-Systeme?

Susan Bischoff (Morrison & Foerster LLP), Christiane Stützle (Morrison & Foerster LLP), Filipp Revinzon (Vay Technology GmbH), Dr. Anastasia Linnik (Retresco GmbH), Tim Sauerhammer (Reed Smith LLP),

Die KI-VO sieht in Art. 50 unterschiedliche Transparenzpflichten für Anbieter und Betreiber von KI-Systemen vor. Je nach Rolle sind die Anforderungen von einem anderen Blickwinkel zu betrachten und unterschiedliche Maßnahmen zu ergreifen. Ziel ist dabei immer, natürliche Personen, die mit KI-Systemen interagieren, vor den mit der neuen Technologie einhergehenden Risiken zu schützen und das Vertrauen in die Integrität des Informationsökosystems zu erhalten. Fehlinformationen, Manipulation in großem Maßstab, Betrug, Identitätsbetrug und Täuschung der Verbraucher sollen effektiv eingeschränkt werden. Dies bedeutet im Regelfall, dass interagierenden Personen und insbesondere Endverbrauchern mitgeteilt werden muss, dass sie es mit einem KI-System zu tun haben.

Dabei kann die Art und Weise, wie diese Mitteilung zu erfolgen hat, Anbieter und Betreiber vor erhebliche Herausforderungen stellen. Es gibt zwar eine Vielzahl von Technologien, die zur Übermittlung der entsprechenden Information nutzbar sind. Es ist jedoch weitgehend ungeklärt, ob diese auch die Vorgaben der KI-VO erfüllen. Es ist zu erwarten, dass die Kommission – wie gemäß Art. 96 Abs. 1 lit. d KI-VO vorgesehen – Leitfäden erarbeiten wird. Eine wichtige Rolle werden auch die Praxisleitfäden einnehmen, die bis zum Mai 2025 unter der Schirmherrschaft des Büros für KI und unter Mitwirkung von Industrie und Interessenträgern wie (potenziellen) Anbietern ausgearbeitet werden, Art. 56, 50 Abs. 7 KI-VO.

Die Pflichten gelten ab dem 2. August 2026, Art. 113 S. 2 KI-VO. Bei vorsätzlicher oder fahrlässiger Verletzung können Geldbußen in Höhe von bis zu 3 % des gesamten weltweiten Jahresumsatzes des Akteurs im vorangegangenen Geschäftsjahr oder bis zu 15 Mio. EUR verhängt werden, je nachdem, welcher Betrag höher ist, Art. 99 Abs. 4 lit. g KI-VO.

Im Folgenden werden zunächst allgemeingültig die verschiedenen Technologien zur Schaffung von Transparenz vorgestellt (5.2.1). Anschließend werden die konkreten sich aus Art. 50 KI-VO ergebenden Transparenzpflichten erläutert, sowohl für Anbieter (5.2.2), als auch für Betreiber (5.2.3) erfasster KI-Systeme sowie für Anbieter von KI-Modellen mit allgemeinem Verwendungszweck (mit systemischem Risiko) (5.2.4).

Technologien zur Schaffung von Transparenz

Grundsätzlich stehen Anbietern und Betreibern von KI-Systemen zwei Möglichkeiten zur Verfügung, wie einer Person mitgeteilt werden kann, dass sie mit einem KI-System oder mit von KI erzeugten oder modifizierten Inhalten interagieren:

- **Direkte Kennzeichnung**: Ein Kennzeichen, welches der interagierenden Person deutlich und erkennbar vermittelt, dass es sich um ein KI-System handelt oder ein bestimmter Inhalt künstlich erzeugt oder manipuliert ist.
- Beispiele: Beschriftungen, Overlays, Banner, Pop-ups, sichtbare Wasserzeichen, auditive Hinweise.

- **Indirekte Kennzeichnung**: Ein Kennzeichen, welches an dem durch das KI-System erzeugten oder manipulierten Inhalt angebracht wird und von der interagierenden Person in der Regel nur unter Zuhilfenahme von zusätzlichen Hilfsmitteln als solches erkannt werden kann.

  Beispiele: Metadaten, unsichtbare Wasserzeichen, Hashing.

Im Folgenden wird eine detaillierte Betrachtung der verschiedenen Technologien vorgenommen. Dabei wird auf die jeweiligen Funktionsweisen und Anwendungsbereiche eingegangen, um ein besseres Verständnis zu vermitteln:

- Beschriftungen, Overlays, Banner, Pop-ups, auditive Hinweise

  Eine einfache Möglichkeit, Informationen zu übermitteln sind Beschriftungen, Overlays, Banner, Pop-ups und auditive Hinweise. Overlays sind grafische Elemente, die temporär über und um das Darstellungsfeld eines Inhalts gelegt und häufig für wichtige Hinweise genutzt werden. Banner sind längliche Grafikflächen, die in der Regel am oberen oder unteren Rand einer Webseite platziert sind. Pop-ups sind kleine Fenster, die im Darstellungsfeld erscheinen. Auditive Hinweise sind akustische Signale bei einer Sprachausgabe oder gesprochene Elemente, die dem Benutzer Informationen vermitteln. Allen dieser Kennzeichen ist gemein, dass sie das Nutzererlebnis beeinflussen, indem sie aktiv Aufmerksamkeit erregen und bestimmte Informationen angeben.

- Wasserzeichen

  - Sichtbare Wasserzeichen

  - Sichtbare Wasserzeichen sind klar für die interagierende Person erkennbar. Diese Art von Wasserzeichen wird häufig in Bildern und Videos eingesetzt und beinhalten oft Text oder Logos, die auf die Herkunft hinweisen – z. B., dass der Inhalt von einem KI-System generiert wurde.

  - Unsichtbare Wasserzeichen

  - Demgegenüber sind unsichtbare Wasserzeichen in die Daten direkt eingebettet und für die interagierende Person nicht erkennbar. Sie können jedoch durch spezielle Algorithmen identifiziert werden, die das Wasserzeichen und die eingebetteten Informationen auslesen. Diese Wasserzeichen werden in verschiedenen Medientypen wie Text, Audio, Bild und Video verwendet. Sie ermöglichen es, die Herkunft und Authentizität von Inhalten zu überprüfen, ohne das Benutzererlebnis zu beeinträchtigen. Die Implementierung ist allerdings technisch anspruchsvoll. Insbesondere in Texten und Audiodateien ist es oft nicht möglich ein unsichtbares Wasserzeichen zu implementieren, ohne das Risiko einzugehen, dass die inhaltliche Bedeutung des Textes oder der Audiodatei verändert wird.

- Hashing

  Hashing ist ein Verfahren, bei dem mittels eines Algorithmus auf Basis des per KI generierten oder modifizierten Inhalts eine einzigartige ID generiert wird, um diesen Inhalt zu einem späteren Zeitpunkt identifizieren zu können. Die Hashes müssen in einer Datenbank gespeichert werden, um zukünftige Inhalte mit dem Original vergleichen zu können. Im Gegensatz zum Wasserzeichen ist dieser Hash nicht in die Inhaltsdatei selbst eingebettet.

Pflichten für Anbieter

Transparenzpflichten für Anbieter von KI-Systemen ergeben sich in erster Linie aus Art. 50 Abs. 1 und Abs. 2 KI-VO. Die KI-VO sieht darin besondere Anforderungen vor, für

a) KI-Systeme, die für die direkte Interaktion mit natürlichen Personen bestimmt sind (Abs. 1); und

b) KI-Systeme, die synthetische Audio-, Bild-, Video- oder Textinhalte erzeugen (Abs. 2).

Der Anbieter muss prüfen, ob das betreffende KI-System in ein oder mehrere der o.g. Kategorien fällt. Die Regelungen der Nummer 1–2 können kumulativ vorliegen. Der Anbieter sollte also für jeden der folgenden Abschnitte prüfen, ob es für das betreffende KI-System relevant ist.

Schritt 5.2.1: Transparenzpflicht für KI-Systeme, die für die direkte Interaktion mit natürlichen Personen bestimmt sind – Art. 50 Abs. 1 KI-VO

Die Anbieter stellen sicher, dass KI-Systeme, die für die direkte Interaktion mit natürlichen Personen bestimmt sind, so konzipiert und entwickelt werden, dass die betreffenden natürlichen Personen informiert werden, dass sie mit einem KI-System interagieren, es sei denn, dies ist aus Sicht einer angemessen informierten, aufmerksamen und verständigen natürlichen Person aufgrund der Umstände und des Kontexts der Nutzung offensichtlich. Diese Pflicht gilt nicht für gesetzlich zur Aufdeckung, Verhütung, Ermittlung oder Verfolgung von Straftaten zugelassene KI-Systeme, wenn geeignete Schutzvorkehrungen für die Rechte und Freiheiten Dritter bestehen, es sei denn, diese Systeme stehen der Öffentlichkeit zur Anzeige einer Straftat zur Verfügung.

Relevante(r) Artikel:

50 Abs.1, Abs. 5

Relevante(r) ErwG:

132

Konkretisierungsbedürftig:

Kasuistik, wann die Interaktion mit einem KI-System aus Sicht einer angemessenen informierten, aufmerksamen und verständigen natürlichen Person aufgrund der Umstände und des Kontexts der Nutzung offensichtlich sein soll.

Anwendungsbereich:

Art. 50 Abs. 1 KI-VO bezieht sich spezifisch auf KI-Systeme, die für die direkte Interaktion mit Menschen vorgesehen sind. Daher sind KI-Systeme, die nicht für eine direkte Interaktion mit natürlichen Personen bestimmt sind, von dieser Regelung nicht erfasst.

Zu diesen nicht erfassten KI-Systemen können z. B. industrielle Automatisierungssysteme gehören, die Produktionsprozesse steuern und optimieren, oder etwa Roboter, die autonom in Fertigungsstraßen arbeiten. Von hoher Relevanz sind auch Backend-Analyse-Tools, die große Datenmengen analysieren, um Muster und Trends zu erkennen, ohne direkt mit menschlichen Nutzern zu kommunizieren. Dies können u. a. Systeme sein, die Finanzmärkte analysieren und automatisiert Investitionsstrategien entwickeln oder Personalisierungssysteme, die im Hintergrund ablaufen und basierend auf Nutzerdaten personalisierte Werbung indizieren. Die möglichen Anwendungsfälle sind zahlreich. Entscheidend für die Prüfung, ob Absatz 1 einschlägig ist, ist zunächst stets, ob das KI-System für die direkte Interaktion vorgesehen ist.

Kommt man zu dem Ergebnis, dass das KI-System für die direkte Interaktion mit dem Menschen bestimmt ist, muss in einem zweiten Schritt geprüft werden, ob die Transparenzpflicht des Art. 50 Abs. 1 KI-VO ausnahmsweise trotzdem entfällt: Wenn es bei der Interaktion mit dem System offensichtlich ist, dass es sich um ein KI-System handelt, bedarf es keiner zusätzlichen Maßnahmen, die der Anbieter vorzunehmen hätte. Maßstab für die Prüfung ist, ob es für eine angemessen informierte, aufmerksame und verständige natürlichen Person aufgrund der Umstände und des Kontexts der Nutzung offensichtlich ist, dass sie es mit einem KI-System zu tun hat. Dabei muss der Anbieter allerdings berücksichtigen, dass er die Art und Weise des Einsatzes des KI-Systems durch einen Betreiber oftmals nicht beeinflussen kann. Er darf sich nicht darauf verlassen, dass ein Betreiber bestimmte Kennzeichnungsanforderungen einhält. Art. 50 Abs. 1 KI-VO sieht entsprechend vor, dass bereits bei Konzeption und Entwicklung des KI-Systems die potenzielle Informationspflicht berücksichtigt werden muss. Im Ergebnis dürften Anbieter aus Gründen der Risikominimierung daher immer dazu neigen, die Informationspflichten zu erfüllen.

Beispiele für KI-Systeme, für die eine Transparenzpflicht entfallen könnte:

- Text- oder Codevorschläge, Autovervollständigung in Smartphones, Texteditoren, Notiz-Apps, IDEs;
- KI-basierte Anwendungen und Komponenten für alltägliche Aufgaben, wie etwa Google Search, Optical Character Recognition (OCR) in Adobe Acrobat, und Spracherkennungskomponenten (ASR) in Smartphones;
- Prozessautomatisierung-Software, z. B. RPA-Tools;
- Empfehlungssysteme z. B. bei Amazon, Spotify oder Netflix.

Die Pflicht gilt gemäß Art. 50 Abs. 1 S. 2 KI-VO außerdem nicht für hinreichend drittschützende KI-Systeme, die gesetzlich im Bereich der Strafverfolgung zugelassen sind, außer solche, die der Öffentlichkeit zur Anzeige einer Straftat zur Verfügung gestellt werden.

Transparenzpflicht:

Der Anbieter muss KI-Systeme so konzipieren und entwickeln, dass interagierende Personen informiert werden, dass es sich um ein KI-System handelt.

Praktische Umsetzung:

Beschriftungen, Overlays, Banner, Pop-ups, auditive Hinweise.

Erläuternde Erklärung zur Umsetzung:

Die Kennzeichnungen müssen in dem KI-System an sich angebracht werden und nicht etwa nur in der Umgebung, in die das KI-System eingebettet ist. Zu diesem Zweck sollten entsprechende Kennzeichnungen über die Benutzeroberfläche eines KI-Systems implementiert werden (z. B. das Eingabefenster eines KI-Chatbots). Allerdings verfügen nicht alle KI-Systeme über eine Benutzeroberfläche, weil sie z. B. in andere (komplexe) Systeme integriert werden. In diesem Fall ist es ggf. möglich, die interagierende Person über die API-Response oder einen anderen Kanal (z. B. akustisch) zu informieren.

Eine indirekte Kennzeichnung ist im Hinblick auf Art. 50 Abs. 5 der KI-VO nicht ausreichend. Demnach muss die Information spätestens im Zeitpunkt der ersten Interaktion, zudem in klar und eindeutiger Weise und hinreichend barrierefrei, bereitgestellt werden.

Schritt 5.2.2: Transparenzpflicht für KI-Systeme, die synthetische Audio-, Bild-, Video- oder Textinhalte erzeugen – Art. 50 Abs. 2 KI-VO

Relevante(r) Artikel:

50 Abs. 2, Abs. 5

Relevante(r) ErwG:

133

Konkretisierungsbedürftig:

Einzelheiten konkret zulässiger technischer Lösungen zur Kennzeichnung

Umfang der Standardbearbeitungen, die von der Pflicht ausgenommen sind, z.B. in der Postproduktion im Film oder der Bildbearbeitung

(2) Anbieter von KI-Systemen, einschließlich KI-Systemen mit allgemeinem Verwendungszweck, die synthetische Audio-, Bild-, Video- oder Textinhalte erzeugen, stellen sicher, dass die Ausgaben des KI-Systems in einem maschinenlesbaren Format gekennzeichnet und als künstlich erzeugt oder manipuliert erkennbar sind. Die Anbieter sorgen dafür, dass — soweit technisch möglich — ihre technischen Lösungen wirksam, interoperabel, belastbar und zuverlässig sind und berücksichtigen dabei 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. Diese Pflicht gilt nicht, soweit die KI-Systeme eine unterstützende Funktion für die Standardbearbeitung ausführen oder die vom Betreiber bereitgestellten Eingabedaten oder deren Semantik nicht wesentlich verändern oder wenn sie zur Aufdeckung, Verhütung, Ermittlung oder Verfolgung von Straftaten gesetzlich zugelassen sind.

Anwendungsbereich:

Art. 50 Abs. 2 KI-VO bezieht sich ausschließlich auf KI-Systeme, die synthetische Audio-, Text-, Bild- und Videoinhalte erzeugen.

Hiervon ausgenommen sind jedoch KI-Systeme, die entweder eine unterstützende Funktion für die Standardbearbeitung haben oder die Eingabedaten oder deren Semantik nicht wesentlich verändern. Dies bedeutet, dass die Kennzeichnungspflichten nicht gelten, wenn das KI-System bei der Erzeugung oder Manipulation des Inhalts nur eine untergeordnete Rolle spielt. Das KI-System „unterstützt“ nur oder verändert nur „unwesentlich“. Generell handelt es sich dabei um grundlegende Anpassungen oder Optimierungen, die routinemäßig

durchgeführt werden. Beispielsweise umfasst die Standardbearbeitung einer Audiodatei oft die Normalisierung der Lautstärke und die Reduktion von Hintergrundrauschen. Bei Texten können Korrekturen von Rechtschreib- und Grammatikfehlern sowie das Formatieren des Textes für ein einheitliches Erscheinungsbild vorgenommen werden, ohne, dass die Semantik der Eingabedaten unwesentlich verändert wird. Bei Bildern beinhaltet dies typischerweise Anpassungen von Helligkeit, Kontrast, Farbbalance und Schärfe, während es bei Videos um das Schneiden von Clips, Hinzufügen von Übergängen und die Anpassung der Farbgebung geht. Im Film- und sonstigen audiovisuellen Bereich ist diese Ausnahme für Arbeiten in der Postproduktion (Stichwort „post enhancement“) besonders relevant; zu den Standardbearbeitungen, die keiner Kennzeichnung unter Art. 50 Abs. 2 KI-VO bedürfen, wenn sie von einem KI-System durchgeführt oder unterstützt werden, dürften etwa solche im Bereich Kostüm, Geräuschreduzierung und zusätzliche Soundeffekte, Timing und Geschwindigkeit, Kontinuität, Ton, sowie visuelle Klarheit und Filter gehören. Auch in Bezug auf KI-Anpassungen von Dialogen und ggf. zusätzlich den Gesichtsbewegungen („AI-Dubbing“), etwa zur Anpassung an Jugendschutzvorgaben oder an Besonderheiten unterschiedlicher Vertriebsmärkte, könnte die Ausnahme greifen – denn jedenfalls im fiktionalen Bereich greift der Zweck der Kennzeichnungspflicht (Fehlinformationen, Betrug, Verbrauchertäuschung, Authentizität von Inhalten) so nicht.

Allerdings ist die genaue Grenze, wann noch von einer Standardbearbeitung und außerdem von einer bloß untergeordneten Rolle ausgegangen werden kann, derzeit noch unklar und durch Rechtsprechung und Leitlinien der Europäischen Kommission (Art. 96 Abs. 1 lit. d KI-VO) auszuformen. Zeitnähere Klarstellungen könnten durch Praxisleitfäden erfolgen, die bis Mai 2025 durch Interessenträger und die Industrie unter der Schirmherrschaft des Büros für KI ausgearbeitet werden sollen (Art. 50 Abs. 7, Art. 56 KI-VO).

Transparenzpflicht:

Audio-, Text-, Bild- und Videoinhalte, die mithilfe eines KI-Systems generiert oder modifiziert wurden, müssen explizit als KI-erzeugte Inhalte in einem maschinenlesbaren Format gekennzeichnet und erkennbar sein.

Praktische Umsetzung:

Beschriftungen, sichtbare Wasserzeichen, auditive Hinweise, Metadaten.

Erläuternde Erklärung zur Umsetzung:

Die Transparenzpflicht betrifft die Ausgabe selbst. Die Kennzeichnung muss an der erzeugten oder manipulierten Datei selbst angebracht werden. Dabei sind nach der KI-VO explizit Techniken, wie Wasserzeichen, Metadatenidentifizierungen, kryptografische Methoden zum Nachweis der Herkunft und Authentizität des Inhalts, Protokollierungsmethoden, Fingerabdrücke oder andere Techniken, oder eine Kombination solcher Techniken zu berücksichtigen. Anbieter haben hier einen Ermessensspielraum und können Besonderheiten, sowie die einschlägigen technologischen Entwicklungen und Marktentwicklungen, einbeziehen.

Die Kennzeichnung muss zum Zeitpunkt der ersten Aussetzung in klarer und eindeutiger sowie hinreichend barrierefreier Weise bereitgestellt werden, Art. 50 Abs. 5 KI-VO.

Zwischenergebnis

Sind die allgemeinen Transparenzpflichten erfüllt, ist für **reine Anbieter von KI-Systemen die Prüfung beendet**. Kommt (auch) eine Akteursstellung als **Betreiber in Betracht, ist mit Schritt 5.2.3** fortzufahren.

Pflichten für Betreiber

Transparenzpflichten für Betreiber von KI-Systemen ergeben sich aus Art. 50 Abs. 3 bis 4 KI-VO. Die KI-VO sieht darin besondere Anforderungen vor, für

a) KI-Systeme, die Emotionen erkennen oder biometrischen Kategorisierung vornehmen (Abs. 3);

b) KI-Systeme, die Bild-, Ton- oder Videoinhalte erzeugen oder manipulieren, die ein Deepfake sind (Abs. 4 UAbs. 1);

c) KI-Systeme, die Text erzeugen oder manipulieren, der veröffentlicht wird, um die Öffentlichkeit über Angelegenheiten von öffentlichem Interesse zu informieren (Abs. 4 UAbs. 2).

Der Betreiber muss prüfen, ob das betreffende KI-System in ein oder mehrere der o.g. Kategorien fällt. Die KI-VO sieht für Betreiber von KI-Systemen im Kontext von Art. 50 Abs. 3–4 Ausnahmen bzw. Beschränkungen vor, insbesondere in Hinblick auf die präventive und repressive Kriminalitätsbekämpfung (Abs. 3, Abs. 4), den Einsatz im offensichtlich künstlerisch-kreativen Rahmen (Abs. 4 UAbs. 1) oder bei menschlicher Abschlusskontrolle und -Verantwortung (Abs. 4 UAbs. 2). Der Betreiber sollte daher für jeden der bezeichneten Abschnitte prüfen, ob und in welchem konkreten Umfang die Transparenzpflicht jeweils besteht.

Schritt 5.2.3: Transparenzpflicht für Emotionserkennungssysteme und Systeme zur biometrischen Kategorisierung – Art. 50 Abs. 3 KI-VO

Relevante(r) Artikel:

Art. 50 Abs. 3, Abs. 5

Relevante(r) ErwG:

18, 30, 132

Konkretisierungsbedürftig:

/

(3) Die Betreiber eines Emotionserkennungssystems oder eines Systems zur biometrischen Kategorisierung informieren die davon betroffenen natürlichen Personen über den Betrieb des Systems und verarbeiten personenbezogene Daten gemäß den Verordnungen (EU) 2016/679 und (EU) 2018/1725 und der Richtlinie (EU) 2016/680. Diese Pflicht gilt nicht für gesetzlich zur Aufdeckung, Verhütung oder Ermittlung von Straftaten zugelassene KI-Systeme, die zur biometrischen Kategorisierung und Emotionserkennung im Einklang mit dem Unionsrecht verwendet werden, sofern geeignete Schutzvorkehrungen für die Rechte und Freiheiten Dritter bestehen.

Anwendungsbereich:

Art. 50 Abs. 3 KI-VO enthält eine besondere Informationspflicht für Betreiber von Emotionserkennungssystemen (Alt. 1) und Systemen zur biometrischen Kategorisierung (Alt. 2).

Nach der KI-VO handelt es sich bei einem „Emotionserkennungssystem“ um ein KI-System, das dem Zweck dient, Emotionen oder Absichten natürlicher Personen auf der Grundlage ihrer biometrischen Daten festzustellen oder daraus abzuleiten (Art. 3 Nr. 39 KI-VO). Dabei geht es ausweislich der Erwägungsgründe um Emotionen oder Absichten wie Glück, Trauer, Wut, Überraschung, Ekel, Verlegenheit, Aufregung, Scham, Verachtung, Zufriedenheit und Vergnügen. Nicht umfasst sind demgegenüber physische Zustände wie Schmerz oder Ermüdung. Dies bedeutet, dass beispielsweise Systeme, die zur Erkennung des Zustands der Ermüdung von Berufspiloten oder -fahrern eingesetzt werden, um Unfälle zu verhindern, nicht unter Art. 50 Abs. 3 KI-VO fallen. Ebenso wenig richtet sich Art. 50 Abs. 3 KI-VO an KI-Systeme, mit welchen offensichtliche Ausdrucksformen, Gesten und Bewegungen erkannt werden können, es sei denn, sie werden zum Erkennen oder Ableiten von Emotionen verwendet.

Auch Betreiber eines „Systems zur biometrischen Kategorisierung“ sollen betroffene Personen informieren. Darunter versteht die KI-VO ein KI-System, das dem Zweck dient, natürliche Personen auf der Grundlage ihrer biometrischen Daten (z. B. Gesicht oder Fingerabdruck) bestimmten Kategorien zuzuordnen. Eine Ausnahme besteht jedoch, sofern es sich dabei um eine bloße Nebenfunktion eines anderen kommerziellen Dienstes handelt und aus objektiven technischen Gründen unbedingt erforderlich ist (Art. 3 Nr. 40 KI-VO).

Die Pflicht gilt auch nicht für gesetzlich zur Aufdeckung, Verhütung oder Ermittlung von Straftaten zugelassene KI-Systeme, die zur biometrischen Kategorisierung und Emotionserkennung im Einklang mit dem Unionsrecht verwendet werden, sofern geeignete Schutzvorkehrungen für die Rechte und Freiheiten Dritter bestehen.

Transparenzpflicht:

Betreiber müssen betroffenen natürlichen Personen über den Betrieb des Systems informieren.

Praktische Umsetzung:

Beschriftungen, Overlays, Banner, Pop-ups, auditive Hinweise.

Erläuternde Erklärung zur Umsetzung:

Vor dem Hintergrund der besonderen Gefahren, die von der Feststellung und Ableitung von Emotionen bzw. der Kategorisierung von biometrischen Daten ausgeht, sollte das Informieren der betroffenen Personen mit einer gesteigerten Transparenz vorgenommen werden. Eine indirekte Kennzeichnung ist im Hinblick auf Art. 50 Abs. 5 der KI-VO nicht ausreichend. Demnach muss die Information spätestens im Zeitpunkt der ersten Interaktion, zudem in klar und eindeutiger Weise und hinreichend barrierefrei, bereitgestellt werden.

Schritt 5.2.4: Transparenzpflicht für KI-Systeme, die Deepfakes oder Texte erzeugen oder manipulieren – Art. 50 Abs. 4 KI-VO

(4) Betreiber eines KI-Systems, das Bild-, Ton- oder Videoinhalte erzeugt oder manipuliert, die ein Deepfake sind, müssen offenlegen, dass die Inhalte künstlich erzeugt oder manipuliert wurden. Diese Pflicht gilt nicht, wenn die Verwendung zur Aufdeckung, Verhütung, Ermittlung oder Verfolgung von Straftaten gesetzlich zugelassen ist. Ist der Inhalt Teil eines offensichtlich künstlerischen, kreativen, satirischen, fiktionalen oder analogen Werks oder Programms, so beschränken sich die in diesem Absatz festgelegten Transparenzpflichten darauf, das Vorhandensein solcher erzeugten oder manipulierten Inhalte in geeigneter Weise offenzulegen, die die Darstellung oder den Genuss des Werks nicht beeinträchtigt.

Relevante(r) Artikel:

Art. 50 Abs. 4, Abs. 5 KI, Art. 3 Nr. 60 VO

Relevante(r) ErwG:

134

Konkretisierungsbedürftig:

Welche Inhalte unter die Kennzeichnungs-erleichterung für „offensichtlich“ künstlerische etc. Werke und Programme fallen und welche Art von Kennzeichnung in diesen Fällen konkret genügt.

Anwendungsbereich:

Der Regelungsumfang von Art. 50 Abs. 4 KI-VO bezieht sich spezifisch auf KI-Systeme, die Bild-, Ton-, oder Videoinhalte erzeugen oder manipulieren, die ein Deepfake sind (Abs. 4 UAbs. 1) bzw. auf KI-Systeme, die Text erzeugen oder manipulieren, der veröffentlicht wird, um die Öffentlichkeit über Angelegenheiten von öffentlichem Interesse zu informieren (Abs. 4 UAbs. 2).

Deepfakes

Art. 50 Abs. 4 KI-VO erlegt Betreibern im Zusammenhang mit der Verwendung von Deepfakes umfangreiche Transparenzpflichten auf. Unter Deepfakes werden gemäß Art. 3 Nr. 60 KI-VO durch KI erzeugte oder manipulierte Bild-, Ton- oder Videoinhalte, die wirklichen Personen, Gegenständen, Orten, Einrichtungen oder Ereignissen ähneln und einer Person fälschlicherweise als echt oder wahrheitsgemäß erscheinen würde, verstanden.

Bei

- echt aussehenden **Bildern**, die Menschen vor einer Landschaft zeigen, wobei Menschen und/oder die Landschaft tatsächlich nicht (so) existieren,
- **Audiodateien**, die beispielsweise im Podcast-Format einen Dialog zwischen real nichtexistierenden Menschen beinhalten,
- **Videos**, deren optische und/oder audiovisuelle Komponente durch KI beispielsweise in Form eines Voice-Over überschrieben wurden,

handelt es sich folglich um Deepfakes im Sinne der KI-VO. Weil ein Personen-Deepfake einem real existierenden Menschen ähneln muss, fallen etwa rein synthetische Darsteller nicht unter diese Kennzeichnungspflicht (können aber von Art. 50 Abs. 2 KI-VO erfasst sein.). Deepfakes verkörpern in dieser Hinsicht unter Einsatz modernster Technologie erzeugte oder manipulierte Werktypen, die vermeintlich reale Situationen und Lebenssachverhalte darstellen, welche jedoch in der Realität zu keinem Zeitpunkt existiert bzw. so stattgefunden

haben. Im Vordergrund steht hierbei das nicht offensichtliche und deshalb besonders gefährliche Potenzial der Irreführung für potenzielle Adressaten.

In Konstellationen, in denen sich das beschriebene Irreführungspotential bereits dem Grunde nach nicht realisieren kann, besteht deshalb auch keine erweiterte Verpflichtung für Betreiber gemäß Art. 50 Abs. 4 KI-VO. Konkret bedeutet dies für Betreiber, dass z. B. Bilder, die aufgrund unterschiedlicher Elemente bereits auf den ersten Blick unnatürlich oder realitätsfremd erscheinen, oder lediglich unter Zuhilfenahme von KI auf richtige Formatmaße zugeschnitten wurden, ohne gesonderte Kennzeichnung verwenden werden können.

Entscheidend für die Einstufung eines Inhalts als Deepfake ist der (inter-)subjektive Eindruck des Rezipienten. Daher ist es im Rahmen der Kennzeichnungspflicht unerheblich, ob beispielsweise eine nachgestellte Person Kenntnis von diesem Vorgang hat respektive ihr ausdrückliches Einverständnis zum Deepfake gegeben hat.

KI-erzeugte oder -manipulierte Texte

Texte können ausweislich der Definition in Art. 3 Nr. 60 KI-VO kein Deepfake im Sinne der KI-VO darstellen. KI-erzeugte oder manipulierte Texte werden aber von der Transparenzpflicht des Art. 50 Abs. 4 UAbs. 2 KI-VO erfasst. Texte, wie beispielsweise solche von Songs und anderen Gestaltungsprozessen, die nicht von echten Menschen, sondern durch KI erstellt wurden, fallen somit in den regulierten Anwendungsbereich der Vorschrift.

Wird ein solcher Text zum Zweck der Information der Öffentlichkeit über Angelegenheiten von öffentlichem Interesse veröffentlicht, dann muss ein solches Zustandekommen des Textes grundsätzlich offengelegt werden.

Transparenzpflicht:

Neben den bereits beschriebenen technischen Lösungen, deren Umsetzung sich primär an Anbieter von KI-Systemen richtet, sollen Betreiber, die Deepfakes bzw. vergleichbar erzeugte oder manipulierte Texte im Sinne des Abs. 4 verwenden, eigenen Transparenzpflichten unterliegen. Die Umsetzung dieser Pflichten muss hierbei gemäß Art. 50 Abs. 5 KI-VO spätestens zum Zeitpunkt der ersten Aussetzung in klarer und eindeutiger Weise erfolgen und geltenden Barrierefreiheitsanforderungen entsprechen.

Die Erstreckung der Verpflichtung zur Negativkennzeichnung synthetischer Inhalte auch auf Betreiber im Sinne der KI-VO bezweckt eine doppelte Absicherung des potenziellen Adressatenkreises. Während Anbietern die initiale Kennzeichnung auf technischer Ebene obliegt, sollen Betreiber potenzielle Lücken schließen, um – unabhängig von einer gesetzlich nicht erforderlichen und unter Umständen auch tatsächlich nicht gegebenen Irreführungsabsicht im Einzelfall – bestmögliche Transparenz zu erzielen. Nur auf diese Weise kann, auch in Übereinstimmung mit den in den Erwägungsgründen Nr. 133ff niedergeschriebenen Zielvorgaben der KI-VO, eine rechtzeitige, dauerhafte und schlussendlich effektive Aufklärung gegenüber allen potenziellen Adressaten erreicht werden.

Die Transparenzpflicht ist in Konstellationen ausgeschlossen, in denen das Ziel präventiver oder repressiver Kriminalitätsbekämpfung durch die Verwendung von Deepfakes verfolgt wird.

Eine abgeschwächte Transparenzpflicht gilt für Betreiber, wenn Deepfakes im Kontext der Meinungs-, Kunst- und Wissenschaftsfreiheit verwendet werden sowie bei KI-

Nachrichtentexten unter menschlicher Aufsicht bzw. Verantwortung. Ist das Deepfake „Teil eines offensichtlich künstlerischen, kreativen, satirischen, fiktionalen oder analogen Werks oder Programms“, sieht die KI-VO eine Abschwächung, aber keine vollständige Befreiung von der Transparenzpflicht vor – die Kennzeichnung muss dann so erfolgen, dass die Darstellung oder der Genuss des Werks, also etwa des Films oder des Hörbuchs, nicht beeinträchtigt wird. Dies dürfte beispielsweise Hinweise am Anfang eines Films, ähnlich der Kennzeichnung von Produktplatzierungen, oder in einem Text neben einem Bild umfassen. Was ein „offensichtlich“ künstlerisches etc. Werk darstellt, ist ebenso wie die konkrete Kennzeichnung vorerst in jedem Einzelfall und vor dem Hintergrund einer möglicherweise unzureichenden Kennzeichnung und damit mangelnden Konformität mit der KI-VO zu bestimmen. Improvisierte (*unscripted*) und nicht-fiktionale Inhalte bedürfen dabei besonderer Aufmerksamkeit, da hier einerseits die Gefahr eines falschen, irregeführten Eindrucks beim Rezipienten höher ist als bei fiktionalen Inhalten und andererseits die Schwelle zu einem offensichtlich künstlerischen o. Ä. Inhalt schwerer zu erreichen sein kann – während z. B. die bloße Hinzufügung eines KI-generierten Talkshow-Gastes diese künstlerische Schwelle möglicherweise nicht erreicht, könnte ein solcher Einsatz in einer satirischen Sendung ausreichen.

Praktische Umsetzung:

Beschriftungen, Overlays, Banner, Pop-ups, sichtbare Wasserzeichen, auditive Hinweise.

Erläuternde Erklärung zur Umsetzung:

Trotz der unterschwelligen Zielvorgaben der KI-VO insbesondere mit Blick auf informierte Entscheidungsfindung potenzieller Adressaten im Einzelfall bzw. der Möglichkeit eines „step back“, eröffnet der stark abschwächte Wortlaut des Art. 50 Abs. 4 KI-VO einen großen Gestaltungsspielraum im Einzelfall.

Der Betreiber kann grundsätzlich in allen Sachverhaltskonstellationen selbst entscheiden, ob und auf welche Art und Weise eine Kennzeichnung von Deepfakes erfolgen soll. Es erscheint naheliegend, dass dies vor allem in Abhängigkeit von dem konkret verwendeten Medium einerseits, d. h. Bild, Ton oder Video, und dem thematischen Rahmen andererseits erfolgen wird.

Mangels strikter Vorgaben bzw. Maßstäbe zur praktischen Umsetzung der Kennzeichnungspflicht im Sinne von Art. 50 Abs. 4 KI-VO kann jedenfalls zu Beginn davon ausgegangen werden, dass die Kennzeichnung von Deepfakes auf sehr unterschiedliche und nicht standardisierte Art und Weise erfolgen wird. Eine eindeutige Kennzeichnung von Deepfakes dürfte hierbei nicht im primären Interesse der Industrie liegen; insbesondere, weil ein optischer „KI-generiert“-Stempel auf einem bildlichen Deepfake mit hoher Wahrscheinlichkeit die Werbebotschaft konterkarieren dürfte.

Vorstellbar ist deshalb, dass die Kennzeichnung nicht unmittelbar beim Deepfake selbst erfolgen wird, sondern – räumlich und zeitlich versetzt – vergleichbar zu aus dem Fernsehen bekannten Produktplatzierungshinweisen bzw. auf Social Media verbreiteten und wettbewerbsrechtlich zwingenden „#Werbung“-Tags in Beschreibungstexten. Ob solche Ausformungen der Kennzeichnungspflicht dem Maßstab einer „klar[en] und deutlich[en]“ Offenlegung im Sinne des Erwägungsgrundes Nr. 134 bzw. Art. 50 Abs. 4 KI-VO genügen werden, wird sich erst noch herauskristallisieren müssen.

Auch mit Blick auf die in Art. 50 Abs. 4 KI-VO statuierten Schranken wird sich ein gültiges Rechtsverständnis herausbilden müssen; ein klarer Maßstab, wann ein Deepfake Teil eines

offensichtlich künstlerischen, kreativen, satirischen, fiktionalen oder analogen Werks ist, existiert zum aktuellen Zeitpunkt jedenfalls noch nicht und wird mit hoher Wahrscheinlichkeit durch gerichtliche Einzelfallentscheidungen oder bestenfalls bis Mai 2025 durch die Praxisleitfäden unter Verantwortung des Büros für KI (Art. 50 Abs. 7, Art. 56 KI-VO) näher konkretisiert werden.

Zum aktuellen Zeitpunkt existieren weder Leitlinien der Kommission im Sinne von Art. 96 Abs. 1 lit. d KI-VO zur praktischen Umsetzung der in Art. 50 Abs. 4 KI-VO normierten Transparenzpflichten für Betreiber noch vergleichbare Praxisleitfäden des Büros für Künstliche Intelligenz im Sinne von Art. 50 Abs. 7 KI-VO. Mangels anderweitiger, rechtsgültiger Standards ist folglich davon auszugehen, dass sich der hinreichende Maßstab einer rechtskonformen Umsetzung der Transparenzpflichten erst noch herausbilden wird – entweder durch zeitlich vorgezogene, in der Industrie geschaffene und selbstauferlegte Kennzeichnungsmaßstäbe oder Gerichte.

Zwischenergebnis

Sind die allgemeinen Transparenzpflichten erfüllt, ist **die Prüfung beendet**. Kommt (auch) eine Anbieterstellung mit Blick auf ein **KI-Modell mit allgemeinem Verwendungszweck** in Betracht, **ist mit Schritt 5.2.5** fortzufahren.

Pflichten für Anbieter von KI-Modellen mit allgemeinem Verwendungszweck

Susan Bischoff (Morrison & Foerster LLP), Christiane Stützle (Morrison & Foerster LLP)

Anbieter von KI-Modellen mit allgemeinem Verwendungszweck unterliegen besonderen Pflichten der Dokumentation, Urheberrechts-Compliance, Transparenz ihrer Trainingsinhalte und der Bestellung von Bevollmächtigten.

Die Pflichten für Anbieter solcher KI-Modelle gelten grundsätzlich ab dem 2. August 2025, Art. 113 S. 3 lit. b KI-VO. Eine Ausnahme besteht für Modelle, die zu diesem Zeitpunkt in der EU bereits in Verkehr gebracht wurden; sie haben bis zum 2. August 2027 Konformität mit der KI-VO sicherzustellen, Art. 111 Abs. 3 KI-VO. Bei vorsätzlicher oder fahrlässiger Verletzung können Geldbußen in Höhe von bis zu 3 % des gesamten weltweiten Jahresumsatzes des Anbieters im vorangegangenen Geschäftsjahr oder bis zu 15 Mio. EUR verhängt werden, je nachdem, welcher Betrag höher ist, Art. 101 Abs. 1 lit. a KI-VO.

Orientierung an Praxisleitfäden

Konkretisierungen für die praktische Umsetzung der Handlungspflichten von Anbieter von KI-Modellen mit allgemeinem Verwendungszweck sind von den Praxisleitfäden zu erwarten, deren zeitnahe Ausarbeitung unter Beteiligung von Anbietern und anderen Interessenträgern vom Büro für KI gefördert wird, Art. 56 KI-VO. Unter regelmäßiger Berichterstattung an das Büro für KI können sich Anbieter auf diese Praxisleitfäden stützen, um die Einhaltung ihrer Pflichten nachzuweisen, Art. 53 Abs. 4 S. 1, Art. 56 Abs. 5 KI-VO.

Die Leitfäden sollen spätestens am 2. Mai 2025 vorliegen, andernfalls kann die Europäische Kommission Umsetzungsvorschriften festlegen, Art. 56 Abs. 9 S. 1 KI-VO. Die Kommission kann Praxisleitfäden genehmigen und ihnen damit allgemeine Gültigkeit verleihen, Art. 53 Abs. 6 S. 4 KI-VO.

Einhaltung harmonisierter Normen

Wird schließlich eine harmonisierte Norm veröffentlicht, begründet deren Einhaltung für die Anbieter die Vermutung der Konformität mit ihren Handlungspflichten, Art. 53 Abs. 4 S. 2 KI-VO.

Alternative Verfahren der Einhaltung

Alternativ zur Befolgung von Praxisleitfäden und der Einhaltung harmonisierter Normen können Anbieter von KI-Modellen mit allgemeinem Verwendungszweck auch geeignete andere Verfahren zur Einhaltung ihrer Handlungspflichten aufzeigen, Art. 53 Abs. 4 S. 3 KI-VO. Diese Verfahren werden von der Kommission bewertet.

Schritt 5.2.5: Pflicht zur technischen Dokumentation des Modells – Art. 53 Abs. 1 lit. a, lit. b KI-VO

Relevante(r) Artikel:

Art. 53 Abs. 1 lit. a, lit. b, Abs. 2

Relevante(r) ErwG:

101

Konkretisierungsbedürftig:

/

Art. 53 Abs. 1 lit. a, lit. b: Anbieter von KI-Modellen mit allgemeinem Verwendungszweck

erstellen und aktualisieren die technische Dokumentation des Modells, einschließlich seines Trainings- und Testverfahrens und der Ergebnisse seiner Bewertung, die mindestens die in Anhang XI aufgeführten Informationen enthält, damit sie dem Büro für Künstliche Intelligenz und den zuständigen nationalen Behörden auf Anfrage zur Verfügung gestellt werden kann;

erstellen und aktualisieren Informationen und die Dokumentation und stellen sie Anbietern von KI-Systemen zur Verfügung, die beabsichtigen, das KI-Modell mit allgemeinem Verwendungszweck in ihre KI-Systeme zu integrieren. Unbeschadet der Notwendigkeit, die Rechte des geistigen Eigentums und vertrauliche Geschäftsinformationen oder Geschäftsgeheimnisse im Einklang mit dem Unionsrecht und dem nationalen Recht zu achten und zu schützen, müssen die Informationen und die Dokumentation

die Anbieter von KI-Systemen in die Lage versetzen, die Fähigkeiten und Grenzen des KI-Modells mit allgemeinem Verwendungszweck gut zu verstehen und ihren Pflichten gemäß dieser Verordnung nachzukommen, und

ii) zumindest die in Anhang XII genannten Elemente enthalten;

Die nach lit. a zu erstellende Dokumentation ist auf Anfrage dem Büro für KI sowie nationalen Behörden zur Verfügung zu stellen. Dagegen sind die nach lit. b zu erstellenden und zu aktualisierenden Informationen und Dokumentation den nachgelagerten Anbietern von solchen KI-Systemen bereitzustellen, die das KI-Modell integrieren. Damit soll sichergestellt werden, dass nachgelagerte Anbieter ein gutes Verständnis des KI-Modells und seiner Fähigkeiten haben, um dessen Integration zu ermöglichen und auch, um ihre eigenen Pflichten im Rahmen der KI-VO erfüllen zu können.

Die Dokumentationspflichten gelten nicht für Anbieter von KI-Modellen, die unter einer freien und quelloffenen Lizenz, d. h. Open Source, bereitgestellt werden, es sei denn, es handelt sich um ein KI-Modell mit systemischen Risiken, Art. 53 Abs. 2 KI-VO.

Die Kommission ist befugt, die Anhänge XI und XII und damit die für die Einhaltung dieser Dokumentationspflicht erforderlichen Informationen und Elemente zu ändern, Art. 53 Abs. 6 KI-VO. Anbieter sollten daher künftige Änderungen im Auge behalten und ihre Dokumentationen entsprechend anpassen.

Schritt 5.2.6: Pflicht zur Einrichtung einer Strategie zur Einhaltung des Urheberrechts, einschließlich der Rechtsvorbehalte unter der Text- und Data Mining-Schranke – Art. 53 Abs. 1 lit. c KI-VO

Anbieter von KI-Modellen mit allgemeinem Verwendungszweck müssen gemäß Art. 53 Abs. 1 lit. c KI-VO eine Strategie zur Einhaltung des Unionsurheberrechts auf den Weg bringen:

> Anbieter von KI-Modellen mit allgemeinem Verwendungszweck
>
> c) bringen eine Strategie zur Einhaltung des Urheberrechts der Union und damit zusammenhängender Rechte und insbesondere zur Ermittlung und Einhaltung eines gemäß Artikel 4 Absatz 3 der Richtlinie (EU) 2019/790 geltend gemachten Rechtsvorbehalts, auch durch modernste Technologien, auf den Weg

Relevante(r) Artikel:

Art. 53 Abs. 1 lit. c

Art. 4 Richtlinie (EU) 2019/790/§ 44b UrhG

Relevante(r) ErwG:

105, 106, 108

Konkretisierungsbedürftig:

Werden spezielle Maßnahmen im Rahmen der Strategie zur Urheberrechts-Compliance erwartet?

(Fehlende) Rückwirkung eines Rechtsvorbehalts des Text und Data Mining; keine nachträgliche Rechtswidrigkeit bereits trainierter KI-Modelle

Ob/wie Anbieter bestehende Sammlungen von Trainingsinhalten vor einem erneuten Training auf zwischenzeitlich erklärte Rechtsvorbehalte prüfen müssen; Dauer rechtmäßiger Speicherung von Trainingsinhalten

Bedeutung und Reichweite der Regelung im ErwG 106, dass Rechtsvorbehalte an den Trainingsinhalten auch beim Training von KI-Modellen außerhalb der EU zu berücksichtigen sind

Allgemeine Urheberrechts-Compliance

Bei der Nutzung urheberrechtlich geschützter Inhalte, die einen hinreichenden EU-Bezug aufweisen (z. B. Vervielfältigung von Trainingsmaterial in der EU, Output mit erkennbaren Drittinhalten wird in der EU vervielfältigt oder öffentlich zugänglich gemacht), muss ein Anbieter das europäische Urheberrecht beachten – dies ergibt sich bereits aus dem Urheberrecht selbst. Andernfalls droht eine urheberrechtliche Haftung und das selbstverständlich auch schon vor Geltung der KI-VO. Insoweit gibt Art. 53 Abs. 1 lit. c KI-VO wieder, was ohnehin schon unter dem Urheberrecht gilt. Es bleibt aber abzuwarten, ob der europäische Gesetzgeber mit Art. 53 Abs. 1 lit. c KI-VO *spezielle* Handlungspflichten für Anbieter zur Urheberrechts-Compliance einführen will. Denkbar wäre auf der Output-Ebene etwa die Blockierung bestimmter Prompts, bei denen ein besonders hohes Risiko urheberrechtsverletzender Outputs besteht, oder die Einrichtung von Maßnahmen zur Moderation der Output-Inhalte, wie etwa ein Meldesystem für Nutzer. Bis zu einer ausdrücklichen Verpflichtung zu bestimmten Maßnahmen (etwa durch weitere Rechtsakte, Klarstellungen durch das Büro für KI oder Auslegung der Norm durch die Rechtsprechung) bleibt es unter Art. 53 Abs. 1 lit. c KI-VO zunächst einmal bei einer Appellfunktion an die Anbieter, auch das Urheberrecht nicht aus den Augen zu verlieren. Dabei müssen Anbieter berücksichtigen, dass das „Urheberrecht der Union“ nur teilweise vollharmonisiert ist. Unterschiede in den nationalen Umsetzungen und Besonderheiten der Urheberrechtsordnungen der Mitgliedstaaten sind deshalb gesondert zu beachten.

Besonders wichtig: Compliance mit Text- und Data Mining-Rechtsvorbehalten

Schon jetzt hat die Pflicht aus Art. 53 Abs. 1 lit. c KI-VO insoweit eine ganz besondere Bedeutung. Die Norm verlangt ausdrücklich, dass insbesondere Rechtsvorbehalte für das Text- und Data Mining zu beachten sind. Das europäische Urheberrecht erlaubt Vervielfältigungen und Entnahmen urheberrechtlich geschützter, rechtmäßig zugänglicher Inhalte für Zwecke des Text- und Data Mining (Art. 4 Richtlinie (EU) 2019/790, umgesetzt in § 44b UrhG). Auf diese Ausnahme können sich auch Anbieter für das Trainieren von KI-Modellen mit allgemeinem Verwendungszweck berufen. Werden Trainingsinhalte nicht

ausschließlich selbst erstellt oder umfassend einlizenziert, ist diese Schranke entscheidend für ein urheberrechtskonformes Trainieren von KI-Modellen.

Eine große technische und organisatorische Herausforderung besteht allerdings darin, dass die Rechteinhaber der Trainingsinhalte einen Rechtsvorbehalt (sog. Opt-out) erklären können (Art. 4 Abs. 3 Richtlinie (EU) 2019/790). Auf diesen Rechtsvorbehalt nimmt Art. 53 Abs. 1 lit. c KI-VO ausdrücklich Bezug und verpflichtet den Anbieter, solche Opt-outs zu ermitteln und einzuhalten. Unterlässt ein Anbieter dies, verletzt er nicht nur die Urheberrechte der Rechteinhaber an den Trainingsinhalten, sondern unterläuft zugleich auch die Konformitätsanforderungen der KI-VO an das Training von KI-Modellen, sodass die so trainierten KI-Modell weder konform mit der KI-VO sind noch zulässig in der EU in Verkehr gebracht werden dürfen – andernfalls drohen Bußgelder und andere Maßnahmen unter der KI-VO, bis zur Zurücknahme des KI-Modells vom europäischen Markt.

Ein Rechtsvorbehalt für online veröffentlichte Inhalte ist allerdings nur dann wirksam und damit von den Anbietern zu beachten, wenn er in maschinenlesbarer Form erklärt wird. Bislang hat sich hierfür kein technischer Standard herausgebildet. Diskutiert und genutzt werden derzeit insbesondere Erklärungen in robots.txt-Dateien oder Metadaten. Es ist zu erwarten, dass das Büro für KI im Anschluss an die Stakeholder-Gespräche konkretisieren wird, wo und in welcher Form Anbieter nach Rechtsvorbehalten Ausschau halten müssen. Opt-out-Erklärungen in reiner Textform, etwa in den AGB oder im Impressum einer Website, stellen keinen wirksamen Rechtsvorbehalt dar und müssen von Anbietern beim Text- und Data Mining insoweit nicht berücksichtigt werden.

Vorteilhaft für Anbieter beim Trainieren von KI-Modellen ist hingegen, dass die Text- und Data Mining Schranke für alle Inhalte (ohne Rechtsvorbehalt) gilt, die für den Anbieter „rechtmäßig zugänglich“ sind. Dies umfasst bereits alles, was im Internet frei zugänglich ist, unabhängig davon, ob es an der konkreten Stelle mit oder ohne Zustimmung des Rechteinhabers online gestellt wurde.

Bislang fehlt es an praktischen Hinweisen des Gesetzgebers und des Büros für KI, wie der Anbieter mit nachträglich erklärten Rechtevorbehalten umzugehen hat. Wurde das KI-Modell bereits mit dem betreffenden Material trainiert, kann die KI dies nicht mehr „verlernen“, sodass der später erklärte Opt-out keine Auswirkungen auf die Rechtmäßigkeit des Modells haben kann. Eine entsprechende Klarstellung wäre aus Gründen der Rechtssicherheit wünschenswert. Ebenso ist zu klären, ob ein Anbieter vor der erneuten Verwendung eines Trainingsinhalts prüfen muss, ob zwischenzeitlich ein Rechtsvorbehalt erklärt wurde. Es ist nicht ersichtlich, wie dies angesichts der schieren Menge an Trainingsinhalten organisatorisch oder technisch möglich sein soll. Für die durch Text- und Data Mining gewonnenen Trainingsinhalte ist jedenfalls zu beachten, dass diese im Rahmen der Schranke nur so lange aufbewahrt werden dürfen, wie dies für die Zwecke des Text- und Data Mining notwendig ist (Art. 4 Abs. 2 Richtlinie (EU) 2019/790). Für ein rechtssicheres Training von KI-Modellen bedarf es auch hier weiterer Konkretisierungen, etwa ob eine Speicherung für ein weiteres oder erneutes Training (z. B. zur Vermeidung von Modelldrift) noch von der Schranke erfasst und damit Urheberrechts-Compliance nach Art. 53 Abs. 1 lit. c KI-VO gewährleistet ist.

Rechtsvorbehalte sollen auch beim Trainieren außerhalb der EU beachtet werden

Auch Anbieter, die ihr KI-Modell außerhalb der EU trainieren, aber in der EU in Verkehr bringen wollen, müssen wohl beachten, dass ihr Modell nur dann mit der KI-VO konform ist, wenn wirksam erklärte Vorbehalte zum Text- und Data Mining berücksichtigt wurden. Dies ist insoweit ungewöhnlich, als das europäische Urheberrecht – und damit auch die Text- und Data Mining-Schranke und ihre Grenze des Rechtsvorbehalts – nur für Nutzungshandlungen in der EU oder mit hinreichendem EU-Bezug gilt. Für Vervielfältigungen beim Training eines KI-Modells, die außerhalb der EU stattfinden, ist das europäische Urheberrecht daher eigentlich nicht anwendbar. ErwG 106 S. 3 KI-VO stellt aber im Hinblick auf die Pflicht zur Beachtung von Rechtevorbehalten ausdrücklich fest:

> Jeder Anbieter, der ein KI-Modell mit allgemeinem Verwendungszweck in der Union in Verkehr bringt, sollte diese Pflicht erfüllen, unabhängig davon, in welchem Hoheitsgebiet die urheberrechtlich relevanten Handlungen, die dem Training dieser KI-Modelle mit allgemeinem Verwendungszweck zugrunde liegen, stattfinden.

Erwägungsgründe sind nicht bindend. Es bleibt daher abzuwarten, wie diese Feststellung, die sich so nicht in den Artikeln der KI-VO findet, in der Praxis vom Büro von KI, von der Kommission und auch von den Gerichten gehandhabt wird. Möglicherweise will der europäische Gesetzgeber hiermit die Einhaltung wirksamer Opt-outs beim Training zu einer Produktanforderung für jedes KI-Modell mit allgemeinem Verwendungszweck machen, das in der EU in Verkehr gebracht wird – unabhängig davon, wo es trainiert wurde. Wird diese Anforderung nicht erfüllt, so hätte dies in Bezug auf das KI-Modell dann nach der KI-VO Geldbußen und Durchsetzungsmaßnahmen zur Folge. Dass mit ErwG 106 S. 3 KI-VO gar der territoriale Geltungsbereich des europäischen Urheberrechts ausgeweitet werden soll, erscheint weit weniger wahrscheinlich, da dies im Widerspruch zu internationalem Recht stünde. Für die Anbieter von KI-Modellen mit allgemeinem Verwendungszweck hätte ErwG 106 S. 3 KI-VO – so er denn in der Praxis angewendet wird – zur Folge, dass Rechtsvorbehalte auch beim Training außerhalb der EU zu beachten wären.

Schritt 5.2.7 Pflicht zur Veröffentlichung einer Zusammenfassung der Trainingsinhalte – Art. 53 Abs. 1 lit. d KI-VO

Anbieter von KI-Modellen mit allgemeinem Verwendungszweck

d) erstellen und veröffentlichen eine hinreichend detaillierte Zusammenfassung der für das Training des KI-Modells mit allgemeinem Verwendungszweck verwendeten Inhalte nach einer vom Büro für Künstliche Intelligenz bereitgestellten Vorlage

Relevante(r) Artikel:

Art. 53 Abs. 1 lit. d

Relevante(r) ErwG:

107, 108, 109

Konkretisierungsbedürftig:

Erforderliche der Aktualisierung der Zusammenfassung

Zulässige Platzierung der Zusammenfassungen

Konkreter Inhalt der Zusammenfassungen, insbesondere wenn bereits die Herkunft/der bereitstellende Drittanbieter der Datensätze eine vertrauliche oder schutzbedürftige Information darstellen

Anbieter von KI-Modellen mit allgemeinem Verwendungszweck müssen gem. Art. 53 Abs. 1 lit. d KI-VO öffentlich eine Zusammenfassung der Trainingsinhalte bereitstellen. Damit soll zum einen im Interesse der Allgemeinheit und der Forschung die Transparenz hinsichtlich der für das Training solcher KI-Modelle verwendeten Daten erhöht werden, zum anderen sollen die Rechteinhaber der Trainingsinhalte in die Kenntnislage versetzt werden, ihre Rechte wirksam wahrzunehmen (ErwG 107 KI-VO).

Die Transparenzpflicht umfasst alle Formen von Trainingsinhalten, nicht nur solche, die urheberrechtlich geschützt sind. Hinsichtlich des erforderlichen Detaillierungsgrades stellt ErwG 107 S. 2 KI-VO klar, dass für eine „hinreichend detaillierte Zusammenfassung“ nicht jeder einzelne Trainingsinhalt aufgelistet werden muss:

Unter gebührender Berücksichtigung der Notwendigkeit, Geschäftsgeheimnisse und vertrauliche Geschäftsinformationen zu schützen, sollte der Umfang dieser Zusammenfassung allgemein weitreichend und nicht technisch detailliert sein, um Parteien mit berechtigtem Interesse, einschließlich der Inhaber von Urheberrechten, die Ausübung und Durchsetzung ihrer Rechte nach dem Unionsrecht zu erleichtern, beispielsweise indem die wichtigsten Datenerhebungen oder Datensätze aufgeführt werden, die beim Training des Modells verwendet wurden, etwa große private oder öffentliche Datenbanken oder Datenarchive, und indem eine beschreibende Erläuterung anderer verwendeter Datenquellen bereitgestellt wird. Es ist angebracht, dass das Büro für Künstliche Intelligenz eine Vorlage für die Zusammenfassung bereitstellt, die einfach und wirksam sein sollte und es dem Anbieter ermöglichen sollte, die erforderliche Zusammenfassung in beschreibender Form bereitzustellen

Die KI-VO erkennt ausdrücklich an, dass die Transparenzverpflichtung mit dem Bedürfnis des Anbieters, seine Geschäftsgeheimnisse und vertraulichen Geschäftsinformationen zu schützen, kollidieren kann. In der Praxis ist eine sorgfältige Abwägung zwischen diesem Schutz und der Erfüllung der Transparenzverpflichtung vorzunehmen. So kann etwa schon die Information, dass ein bestimmter Datensatz von einem bestimmten Datenanbieter für das Training des KI-Modells verwendet wird, an sich schutzbedürftig sein. Es ist aber davon auszugehen, dass sich das Büro für KI allein durch den Hinweis auf eine solche Vertraulichkeit oder Schutzbedürftigkeit nicht grundsätzlich von einer Veröffentlichungspflicht hinreichend detaillierter Angaben zu den Trainingsinhalten abhalten lassen wird.

Begrüßenswert ist, dass das Büro für KI eine Vorlage für eine solche Zusammenstellung von Trainingsinhalten bereitstellen wird. Konkretisierungen zum angemessenen Detaillierungsgrad der Zusammenfassungen sind von den Praxisleitfäden zu erwarten.

Schritt 5.2.8: Pflicht zur Benennung eines Bevollmächtigten – Art. 54 Abs. 1 KI-VO

Relevante(r) Artikel:

Art. 54

Relevante(r) ErwG:

/

Konkretisierungsbedürftig:

/

Art. 54:

(1) Anbieter, die in Drittländern niedergelassen sind, benennen vor dem Inverkehrbringen eines KI-Modells mit allgemeinem Verwendungszweck auf dem Unionsmarkt schriftlich einen in der Union niedergelassenen Bevollmächtigten.

(2) Der Anbieter muss seinem Bevollmächtigten ermöglichen, die Aufgaben wahrzunehmen, die im vom Anbieter erhaltenen Auftrag festgelegt sind.

Die KI-VO gilt auch für Anbieter, die KI-Modelle mit allgemeinem Verwendungszweck in der Union in Verkehr bringen unabhängig davon, ob der Anbieter in der Union oder in einem Drittland niedergelassen ist, Art. 3 Abs. 1 lit. a KI-VO. Anbieter mit Niederlassung in einem Drittland haben einen in der Union niedergelassenen Bevollmächtigten zu erlassen:

Die Pflicht gilt nicht für Anbieter von KI-Modellen, die im Rahmen einer freien und quelloffenen Lizenz bereitgestellt werden, es sei denn, es handelt sich dabei um ein KI-Modell mit allgemeinem Verwendungszweck mit systematischen Risiken, Art. 54 Abs. 6 KI-VO.

Exkurs: Pflichten für Anbieter von KI-Modellen mit allgemeinem Verwendungszweck mit systemischem Risiko

Relevante(r) Artikel:

Art. 55

Relevante(r) ErwG:

114, 115

Konkretisierungsbedürftig:

/

Für Anbieter von KI-Modellen mit allgemeinem Verwendungszweck treten zusätzlich zu den in Schritt **7.2.4** aufgeführten Pflichten weitere Pflichten hinzu, wenn es sich um ein Modell mit systemischem Risiko im Sinne der Artikel 51, 52 KI-VO handelt (zur Kategorisierung solcher Modelle oben in Schritt **3.2**).

Art. 55 Abs. 1: Zusätzlich zu den in den Artikeln 53 und 54 aufgeführten Pflichten müssen Anbieter von KI-Modellen mit allgemeinem Verwendungszweck mit systemischem Risiko

a) eine Modellbewertung mit standardisierten Protokollen und Instrumenten, die dem Stand der Technik entsprechen, durchführen, wozu auch die Durchführung und Dokumentation von Angriffstests beim Modell gehören, um systemische Risiken zu ermitteln und zu mindern,

b) mögliche systemische Risiken auf Unionsebene – einschließlich ihrer Ursachen –, die sich aus der Entwicklung, dem Inverkehrbringen oder der Verwendung von KI-Modellen mit allgemeinem Verwendungszweck mit systemischem Risiko ergeben können, bewerten und mindern,

c) einschlägige Informationen über schwerwiegende Vorfälle und mögliche Abhilfemaßnahmen erfassen und dokumentieren und das Büro für Künstliche Intelligenz und gegebenenfalls die zuständigen nationalen Behörden unverzüglich darüber unterrichten,

d) ein angemessenes Maß an Cybersicherheit für die KI-Modelle mit allgemeinem Verwendungszweck mit systemischem Risiko und die physische Infrastruktur des Modells gewährleisten.

Die zusätzlichen Pflichten umfassen damit:

- Eine **Modellbewertung**, um die systemischen Risiken zu ermitteln und zu mindern (lit. a),
- **Bewertung und Minderung** möglicher systemischer Risiken (lit. b),
- **Erfassung und Dokumentation** schwerwiegender Vorfälle und möglicher Abhilfemaßnahmen für die Behörden (lit. c), und
- Angemessene **Cybersicherheit** (lit. d). Dabei sind insbesondere unbeabsichtigter Modelldatenverlust, die unerlaubte Bereitstellung, die Umgehung von Sicherheitsmaßnahmen und der Schutz vor Cyberangriffen, unbefugtem Zugriff oder

Modelldiebstahl zu beachten. Als mögliche Schutzmaßnahmen listet ErwG 115 KI-VO die Sicherung von Modellgewichten, Algorithmen, Servern und Datensätzen auf.

Anbieter können eine Konformität mit diesen Pflichten durch die Einhaltung von Praxisleitfäden nachweisen, die bis Mai 2025 durch Interessenträger und die Industrie unter Verantwortung des Büros für KI zu erstellen sind, Art. 55 Abs. 2, Art. 56 KI-VO. Dies gilt bis zur Veröffentlichung einer harmonisierten Norm – danach begründet die Einhaltung dieser Norm die Vermutung der Konformität.

Zwischenergebnis

Sind alle Transparenzpflichten erfüllt, ist die Prüfung hier beendet.

8 Das Konformitäts-bewertungsverfahren

Vasilios Danos (TÜV Informationstechnik GmbH), Stephan Kress (Morrison & Foerster LLP)

Für Anbieter von **Hochrisiko-KI-Systemen und KI-Modellen mit allgemeinem Verwendungszweck** regeln die **Art. 40 ff. KI-VO**, wie die Konformität solcher Systeme mit den Anforderungen der Verordnung bewertet und nachgewiesen werden kann.

Relevante(r) Artikel:

Art. 40-49

Relevante(r) ErwG:

78, 123, 124, 125, 126, 127, 128, 129, 130, 131

Konkretisierungsbedürftig:

-

(1) Bei Hochrisiko-KI-Systemen oder KI-Modellen mit allgemeinem Verwendungszweck, die mit harmonisierten Normen oder Teilen davon, deren Fundstellen gemäß der Verordnung (EU) Nr. 1025/2012 im Amtsblatt der Europäischen Union veröffentlicht wurden, übereinstimmen, wird eine Konformität mit den Anforderungen gemäß Abschnitt 2 des vorliegenden Kapitels oder gegebenenfalls mit den Pflichten gemäß Kapitel V Abschnitte 2 und 3 der vorliegenden Verordnung vermutet, soweit diese Anforderungen oder Verpflichtungen von den Normen abgedeckt sind.

(2) Gemäß Artikel 10 der Verordnung (EU) Nr. 1025/2012 erteilt die Kommission unverzüglich Normungsaufträge, die alle Anforderungen gemäß Abschnitt 2 des vorliegenden Kapitels abdecken und gegebenenfalls Normungsaufträge, die Pflichten gemäß Kapitel V Abschnitte 2 und 3 der vorliegenden Verordnung abdecken. In dem Normungsauftrag werden auch Dokumente zu den Berichterstattungs- und Dokumentationsverfahren im Hinblick auf die Verbesserung der Ressourcenleistung von KI-Systemen z. B. durch die Verringerung des Energie- und sonstigen Ressourcenverbrauchs des Hochrisiko-KI-Systems während seines gesamten Lebenszyklus und zu der energieeffizienten Entwicklung von KI-Modellen mit allgemeinem Verwendungszweck verlangt. Bei der Ausarbeitung des Normungsauftrags konsultiert die Kommission das KI-Gremium und die einschlägigen Interessenträger, darunter das Beratungsforum.

Bei der Erteilung eines Normungsauftrags an die europäischen Normungsorganisationen gibt die Kommission an, dass die Normen klar und — u. a. mit den Normen, die in den verschiedenen Sektoren für Produkte entwickelt wurden, die unter die in Anhang I aufgeführten geltenden Harmonisierungsrechtsvorschriften der Union fallen — konsistent sein müssen und sicherstellen sollen, dass die in der Union in Verkehr gebrachten oder in Betrieb genommenen Hochrisiko-KI-Systeme oder KI-Modelle mit allgemeinem Verwendungszweck die in dieser Verordnung festgelegten einschlägigen Anforderungen oder Pflichten erfüllen. Die Kommission fordert die europäischen Normungsorganisationen auf, Nachweise dafür vorzulegen, dass sie sich nach besten Kräften bemühen, die in den Unterabsätzen 1 und 2 dieses Absatzes genannten Ziele im Einklang mit Artikel 24 der Verordnung (EU) Nr. 1025/2012 zu erreichen.

(3) Die am Normungsprozess Beteiligten bemühen sich, Investitionen und Innovationen im Bereich der KI, u. a. durch Erhöhung der Rechtssicherheit, sowie der Wettbewerbsfähigkeit und des Wachstums des Unionsmarktes zu fördern und zur Stärkung der weltweiten Zusammenarbeit bei der Normung und zur Berücksichtigung bestehender internationaler Normen im Bereich der KI, die mit den Werten, Grundrechten und Interessen der Union im Einklang stehen, beizutragen und die Multi-Stakeholder-Governance zu verbessern, indem eine ausgewogene Vertretung der Interessen und eine wirksame Beteiligung aller relevanten Interessenträger gemäß den Artikeln 5, 6 und 7 der Verordnung (EU) Nr. 1025/2012 sichergestellt werden.

Schritt 6.1: Gibt es harmonisierte Normen oder gemeinsame Spezifikationen, die die Anforderungen aus Kapitel III Abschnitt 2 abdecken?

Harmonisierte Normen

Bei Übereinstimmung des KI-Systems oder KI-Modells mit allgemeinem Verwendungszweck mit relevanten harmonisierten Normen wird eine Konformität mit den Anforderungen gemäß Abschnitt 2 von Kapitel III oder gegebenenfalls mit den Pflichten gemäß Kapitel V Abschnitte 2 und 3 der KI-Verordnung vermutet, soweit diese Anforderungen oder Verpflichtungen von den Normen abgedeckt sind. Die relevanten harmonisierten Normen müssen solche sein, deren Fundstellen gemäß der Verordnung (EU) Nr. 1025/2012 im Amtsblatt der Europäischen Union veröffentlicht wurden. Sie werden von anerkannten europäischen Normungsorganisation entwickelt, nämlich von der CEN (Europäisches Komitee für Normung), der CENELEC (Europäisches Komitee für elektrotechnische Normung) oder der ETSI (Europäisches Institut für Telekommunikationsnormen).

Das Gemeinsame Technische Komitee 21 von CEN and CENELEC ist derzeit damit befasst, bestehende ISO/IEC Normen anzupassen und neue zu entwickeln. Dafür hat das Komitee ein Arbeitsprogramm veröffentlicht, anhand dessen die Standardsetzung verfolgt werden kann: CEN – CEN/CLC/JTC 21 (cencenelec.eu). Bisher wurden keine Fundstellen für harmonisierte Normen mit Bezugnahme auf die KI-Verordnung im Amtsblatt der Europäischen Union veröffentlicht.

Gemeinsame Spezifikationen

Wenn es keine harmonisierten Normen gibt, oder diese dem Auftrag der Kommission nicht entsprechen, kann die Kommission Durchführungsrechtsakte zur Festlegung gemeinsamer Spezifikationen erlassen (Art. 41 KI-VO). Das entsprechende Prüfverfahren ist in Art. 5 der Verordnung (EU) Nr. 182/2011 geregelt. Auch hier gilt, dass bei Hochrisiko-KI-Systemen oder KI-Modellen mit allgemeinem Verwendungszweck, die mit gemeinsamen Spezifikationen oder Teilen dieser Spezifikationen übereinstimmen, eine Konformität mit den Anforderungen in Abschnitt 2 von Kapitel III oder gegebenenfalls die Einhaltung der in Kapitel V Abschnitte 2 und 3 genannten Pflichten vermutet wird, soweit diese Anforderungen oder diese Pflichten von den gemeinsamen Spezifikationen abgedeckt sind.

Andere technische Lösungen

Den Anbietern von KI-Systemen steht es frei, andere technische Lösung zu wählen, um die Einhaltung der verbindlichen gesetzlichen Anforderungen nachzuweisen. Die Vermutungswirkungen wie bei den harmonisierten Normen und gemeinsamen Spezifikationen entfällt dann allerdings. Soweit harmonisierte Normen bestehen, empfiehlt es sich, diese auch heranzuziehen.

Schritt 6.2: Durchführung des erforderlichen Konformitätsbewertungsverfahrens

Art. 43 (1) Hat ein Anbieter zum Nachweis, dass ein in Anhang III Nummer 1 aufgeführtes Hochrisiko-KI-System die in Abschnitt 2 festgelegten Anforderungen erfüllt, harmonisierte Normen gemäß Artikel 40 oder gegebenenfalls gemeinsame Spezifikationen gemäß Artikel 41 angewandt, so entscheidet er sich für eines der folgenden Konformitätsbewertungsverfahren auf der Grundlage

a) der internen Kontrolle gemäß Anhang VI oder

b) der Bewertung des Qualitätsmanagementsystems und der Bewertung der technischen Dokumentation unter Beteiligung einer notifizierten Stelle gemäß Anhang VII.

Zum Nachweis, dass sein Hochrisiko-KI-System die in Abschnitt 2 festgelegten Anforderungen erfüllt, befolgt der Anbieter das Konformitätsbewertungsverfahren gemäß Anhang VII, wenn

a) es harmonisierte Normen gemäß Artikel 40 nicht gibt und keine gemeinsamen Spezifikationen gemäß Artikel 41 vorliegen,

b) der Anbieter die harmonisierte Norm nicht oder nur teilweise angewandt hat;

c) die unter Buchstabe a genannten gemeinsamen Spezifikationen zwar vorliegen, der Anbieter sie jedoch nicht angewandt hat;

d) eine oder mehrere der unter Buchstabe a genannten harmonisierten Normen mit einer Einschränkung und nur für den eingeschränkten Teil der Norm veröffentlicht wurden.

Für die Zwecke des Konformitätsbewertungsverfahrens gemäß Anhang VII kann der Anbieter eine der notifizierten Stellen auswählen. Soll das Hochrisiko-KI-System jedoch von Strafverfolgungs-, Einwanderungs- oder Asylbehörden oder von Organen, Einrichtungen oder sonstigen Stellen der Union in Betrieb genommen werden, so übernimmt die in Artikel 74 Absatz 8 bzw. 9 genannte Marktüberwachungsbehörde die Funktion der notifizierten Stelle. (2) Bei den in Anhang III Nummern 2 bis 8 aufgeführten Hochrisiko-KI-Systemen befolgen die Anbieter das Konformitätsbewertungsverfahren auf der Grundlage einer internen Kontrolle gemäß Anhang VI, das keine Beteiligung einer notifizierten Stelle vorsieht.

(3) Bei den Hochrisiko-KI-Systemen, die unter die in Anhang I Abschnitt A aufgeführten Harmonisierungsrechtsakte der Union fallen, befolgt der Anbieter die einschlägigen Konformitätsbewertungsverfahren, die nach diesen Rechtsakten erforderlich sind. Die in Abschnitt 2 dieses Kapitels festgelegten Anforderungen gelten für diese Hochrisiko-KI-Systeme und werden in diese Bewertung einbezogen. Anhang VII Nummern 4.3, 4.4 und 4.5 sowie Nummer 4.6 Absatz 5 finden ebenfalls Anwendung.

Wenn ein in Anhang I Abschnitt A aufgeführter Rechtsakte es dem Hersteller des Produkts ermöglicht, auf eine Konformitätsbewertung durch Dritte zu verzichten, sofern dieser Hersteller alle harmonisierten Normen, die alle einschlägigen Anforderungen abdecken, angewandt hat, so darf dieser Hersteller nur dann von dieser Möglichkeit Gebrauch machen, wenn er auch harmonisierte Normen oder gegebenenfalls gemeinsame Spezifikationen gemäß Artikel 41, die alle in Abschnitt 2 dieses Kapitels festgelegten Anforderungen abdecken, angewandt hat.

(4) Hochrisiko-KI-Systeme, die bereits Gegenstand eines Konformitätsbewertungsverfahren gewesen sind, werden im Falle einer wesentlichen Änderung einem neuen Konformitätsbewertungsverfahren unterzogen, unabhängig davon, ob das geänderte System noch weiter in Verkehr gebracht oder vom derzeitigen Betreiber weitergenutzt werden soll. Bei Hochrisiko-KI-Systemen, die nach dem Inverkehrbringen oder der Inbetriebnahme weiterhin dazulernen, gelten Änderungen des Hochrisiko-KI-Systems und seiner Leistung, die vom Anbieter zum Zeitpunkt der ursprünglichen Konformitätsbewertung vorab festgelegt wurden und in den Informationen der technischen Dokumentation gemäß Anhang IV Nummer 2 Buchstabe f enthalten sind, nicht als wesentliche Veränderung;

(5) Die Kommission ist befugt, gemäß Artikel 97 delegierte Rechtsakte zu erlassen, um die Anhänge VI und VII zu ändern, indem sie sie angesichts des technischen Fortschritts aktualisiert

(6) Die Kommission ist befugt, gemäß Artikel 97 delegierte Rechtsakte zur Änderung der Absätze 1 und 2 des vorliegenden Artikels zu erlassen, um die in Anhang III Nummern 2 bis 8 genannten Hochrisiko-KI-Systeme dem Konformitätsbewertungsverfahren gemäß Anhang VII oder Teilen davon zu unterwerfen. Die Kommission erlässt solche delegierten Rechtsakte unter Berücksichtigung der Wirksamkeit des Konformitätsbewertungsverfahrens auf der Grundlage einer internen Kontrolle gemäß Anhang VI hinsichtlich der Vermeidung oder Minimierung der von solchen Systemen ausgehenden Risiken für die Gesundheit und Sicherheit und den Schutz der Grundrechte sowie hinsichtlich der Verfügbarkeit angemessener Kapazitäten und Ressourcen in den notifizierten Stellen.

Welches Konformitätsbewertungsverfahren ist einschlägig?

Die Frage, welches Konformitätsverfahren vor Inverkehrbringen durchgeführt werden muss, richtet sich sowohl nach Art und Zweck des Hochrisiko-KI-Systems als auch danach, ob auf harmonisierte Normen bzw. gemeinsame Spezifikationen zurückgegriffen wird.

Art des Hochrisiko-KI-Systems	Erforderliches Konformitätsbewertungsverfahren
Biometrische KI-Systeme gem. Anhang III Nummer 1, d. h.: ■ Biometrische Fernidentifizierungssysteme, ■ KI-Systeme, die bestimmungsgemäß zur Emotionserkennung verwendet werden sollen, sowie ■ KI-Systeme, die bestimmungsgemäß für die biometrische Kategorisierung nach sensiblen oder geschützten Attributen oder Merkmalen auf der Grundlage von Rückschlüssen auf diese Attribute oder Merkmale verwendet werden sollen.	Der Anbieter kann das Konformitätsbewertungsverfahren der internen Kontrolle gemäß Anhang VI wählen, wenn er alle Anforderungen vollständig durch Anwendung harmonisierter Normen gemäß Artikel 40 oder gemeinsamer Spezifikationen gemäß Artikel 41 umgesetzt hat. Falls dies nicht der Fall ist, kommt nur das Konformitätsbewertungsverfahren unter Beteiligung einer notifizierten Stelle gemäß Anhang VII infrage.
Hochrisiko-KI-Systemen gem. Anhang III Nummern 2 bis 8.	Konformitätsbewertungsverfahren der internen Kontrolle gemäß Anhang VI.
Hochrisiko-KI-Systemen, die unter die in Anhang I Abschnitt A aufgeführten Harmonisierungsrechtsakte der Union fallen, also z. B.: ■ Spielzeug ■ Medizinprodukte ■ Sportboote ■ Funkanlagen	Einschlägig ist das Konformitätsbewertungsverfahren, das nach dem jeweiligen Rechtsakt erforderlich ist. Die in Abschnitt 2 des Kapitels III festgelegten Anforderungen werden in diese Bewertung einbezogen. Anhang VII Nummern 4.3, 4.4 und 4.5 sowie Nummer 4.6 Absatz 5 finden ebenfalls Anwendung.

Für die Zwecke des Konformitätsbewertungsverfahrens gemäß Anhang VII kann der Anbieter eine der notifizierten Stellen auswählen. Soll das Hochrisiko-KI-System jedoch von Strafverfolgungs-, Einwanderungs- oder Asylbehörden oder von Organen, Einrichtungen oder sonstigen Stellen der Union in Betrieb genommen werden, so übernimmt die in Artikel 74 Absatz 8 bzw. 9 genannte Marktüberwachungsbehörde die Funktion der notifizierten Stelle.

Ist eine Ausnahme von der Erforderlichkeit eines Konformitätsbewertungsverfahrens einschlägig?

Auf ein hinreichend begründetes Ersuchen hin kann eine Marktüberwachungsbehörde das Inverkehrbringen oder die Inbetriebnahme bestimmter Hochrisiko-KI-Systeme aus außergewöhnlichen Gründen der öffentlichen Sicherheit, des Schutzes des Lebens und der Gesundheit von Personen, des Umweltschutzes oder des Schutzes wichtiger Industrie- und Infrastrukturanlagen genehmigen. Diese Genehmigung wird auf die Dauer der erforderlichen Konformitätsbewertungsverfahren befristet, wobei den außergewöhnlichen Gründen für die Ausnahme Rechnung getragen wird. Die Genehmigung wird nur erteilt, wenn die Marktüberwachungsbehörde zu dem Schluss gelangt, dass das Hochrisiko-KI-System die Anforderungen des Abschnitts 2 des Kapitels III erfüllt.

Beispiele für das Eingreifen dieser Ausnahmen wegen des Bedürfnisses einer raschen

(1) Der Anbieter stellt für jedes Hochrisiko-KI-System eine schriftliche maschinenlesbare, physische oder elektronisch unterzeichnete EU-Konformitätserklärung aus und hält sie für einen Zeitraum von 10 Jahren ab dem Inverkehrbringen oder der Inbetriebnahme des Hochrisiko-KI-Systems für die zuständigen nationalen Behörden bereit. Aus der EU-Konformitätserklärung geht hervor, für welches Hochrisiko-KI-System sie ausgestellt wurde. Eine Kopie der EU-Konformitätserklärung wird den zuständigen nationalen Behörden auf Anfrage übermittelt.

(2) Die EU-Konformitätserklärung muss feststellen, dass das betreffende Hochrisiko-KI-System die in Abschnitt 2 festgelegten Anforderungen erfüllt. Die EU-Konformitätserklärung enthält die in Anhang V festgelegten Informationen und wird in eine Sprache übersetzt, die für die zuständigen nationalen Behörden der Mitgliedstaaten, in denen das Hochrisiko-KI-System in Verkehr gebracht oder bereitgestellt wird, leicht verständlich ist.

(3) Unterliegen Hochrisiko-KI-Systeme anderen Harmonisierungsrechtsvorschriften der Union, die ebenfalls eine EU-Konformitätserklärung vorschreiben, so wird eine einzige EU-Konformitätserklärung ausgestellt, die sich auf alle für das Hochrisiko-KI-System geltenden Rechtsvorschriften der Union bezieht. Die Erklärung enthält alle erforderlichen Informationen zur Feststellung der Harmonisierungsrechtsvorschriften der Union, auf die sich die Erklärung bezieht.

Verfügbarkeit könnten etwa sein:

- Eine medizinische App wird zur Verringerung der schädlichen Auswirkung einer sich rasch ausbreitenden ansteckenden Krankheit benötigt.
- Eine Sicherheitskomponente für kritische Infrastruktur muss zur Abwehr einer drohenden Gefahr unverzüglich ausgetauscht werden.

Schritt 6.3: Was ist nach dem Konformitätsbewertungsverfahren zu tun?

Die Artikel 47 bis 37 KI-VO regeln, was nach dem Konformitätsbewertungsverfahren zu tun ist.

Konformitätserklärung

(4) Mit der Ausstellung der EU-Konformitätserklärung übernimmt der Anbieter die Verantwortung für die Erfüllung der in Abschnitt 2 festgelegten Anforderungen. Der Anbieter hält die EU-Konformitätserklärung gegebenenfalls auf dem neuesten Stand.

(5) Der Kommission ist befugt, gemäß Artikel 97 delegierte Rechtsakte zur Aktualisierung des in Anhang V festgelegten Inhalts der EU-Konformitätserklärung zu erlassen, um den genannten Anhang durch die Einführung von Elementen zu ändern, die angesichts des technischen Fortschritts erforderlich werden.

Nach dem Konformitätsbewertungsverfahren stellt der Anbieter für jedes Hochrisiko-KI-System eine schriftliche maschinenlesbare, physische oder elektronisch unterzeichnete Konformitätserklärung aus und hält sie für einen Zeitraum von 10 Jahren ab dem Inverkehrbringen oder der Inbetriebnahme des Hochrisiko-KI-Systems für die zuständigen nationalen Behörden bereit (Art. 47 Abs. 1 KI-VO). Mit der Konformitätserklärung erklärt der Anbieter verbindlich, dass alle einschlägigen Pflichten in Bezug auf das Hochrisiko-KI-System befolgt wurden. Die EU-Konformitätserklärung hat die in Anhang V der KI-VO festgelegten Informationen zu enthalten. Dies umfasst etwa Name und Typ des KI-Systems, Name und Anschrift des Anbieters, einen Verweis auf einschlägige harmonisierte Normen oder andere gemeinsame Spezifikationen, für die die Konformität erklärt wird, und gegebenenfalls Name und Kennnummer der benannten Stelle, eine Beschreibung des durchgeführten Konformitätsbewertungsverfahrens und die Bezeichnung der ausgestellten Bescheinigung.

Registrierung in einer Datenbank

(1) Vor dem Inverkehrbringen oder der Inbetriebnahme eines in Anhang III aufgeführten Hochrisiko-KI-Systems — mit Ausnahme der in Anhang III Nummer 2 genannten Hochrisiko-KI-Systeme — registriert der Anbieter oder gegebenenfalls sein Bevollmächtigter sich und sein System in der in Artikel 71 genannten EU-Datenbank.

(2) Vor dem Inverkehrbringen oder der Inbetriebnahme eines Hochrisiko-KI-Systems, bei dem der Anbieter zu dem Schluss gelangt ist, dass es nicht hochriskant gemäß Artikel 6 Absatz 3 ist, registriert dieser Anbieter oder gegebenenfalls sein Bevollmächtigter sich und dieses System in der in Artikel 71 genannten EU-Datenbank.

(3) Vor der Inbetriebnahme oder Verwendung eines in Anhang III aufgeführten Hochrisiko-KI-Systems — mit Ausnahme der in Anhang III Nummer 2 aufgeführten Hochrisiko-KI-Systeme — registrieren sich Betreiber, bei denen es sich um Behörden oder Organe, Einrichtungen oder sonstige Stellen der Union oder in ihrem Namen handelnde Personen handelt, in der in Artikel 71 genannten EU-Datenbank, wählen das System aus und registrieren es dort.

(4) Bei den in Anhang III Nummern 1, 6 und 7 genannten Hochrisiko-KI-Systemen erfolgt in den Bereichen Strafverfolgung, Migration, Asyl und Grenzkontrolle die Registrierung gemäß den Absätzen 1, 2 und 3 des vorliegenden Artikels in einem sicheren nicht öffentlichen Teil der in Artikel 71 genannten EU-Datenbank und enthält, soweit zutreffend, lediglich die Informationen gemäß a) Anhang VIII Abschnitt A Nummern 1 bis 10 mit Ausnahme der Nummern 6, 8 und 9, b) Anhang VIII Abschnitt B Nummern 1 bis 5 sowie Nummern 8 und 9, c) Anhang VIII Abschnitt C Nummern 1 bis 3, d) Anhang IX Nummern 1, 2, 3 und Nummer 5. Nur die Kommission und die in Artikel 74 Absatz 8 genannten nationalen Behörden haben Zugang zu den jeweiligen beschränkten Teilen der EU-Datenbank gemäß Unterabsatz 1 dieses Absatzes.

(5) Die in Anhang III Nummer 2 genannten Hochrisiko-KI-Systeme werden auf nationaler Ebene registriert.

Vor dem Inverkehrbringen oder der Inbetriebnahme eines in Anhang III aufgeführten Hochrisiko-KI-Systeme – mit Ausnahme von Hochrisiko-KI-Systeme der Kritischen Infrastruktur (d. h. KI-Systeme, die bestimmungsgemäß als Sicherheitsbauteile im Rahmen der Verwaltung und des Betriebs kritischer digitaler Infrastruktur, des Straßenverkehrs oder der Wasser-, Gas-, Wärme- oder Stromversorgung verwendet werden sollen) – registriert der Anbieter oder gegebenenfalls sein Bevollmächtigter sich und sein System in einer von der Kommission betriebenen EU-Datenbank (Art. 16 lit. i i.V.m. Art. 49 KI-VO). Die Datenbank

ist noch im Aufbau begriffen und wird wahrscheinlich über den Internetauftritt der Kommission erreichbar sein.

Anbringung eines CE-Kennzeichens

Hochrisiko-KI-Systeme müssen grundsätzlich mit der CE-Kennzeichnung versehen sein, aus der ihre Konformität mit der KI-Verordnung hervorgeht (Art. 48 KI-VO).

(1) Für die CE-Kennzeichnung gelten die in Artikel 30 der Verordnung (EG) Nr. 765/2008 festgelegten allgemeinen Grundsätze.

(2) Bei digital bereitgestellten Hochrisiko-KI-Systemen wird eine digitale CE-Kennzeichnung nur dann verwendet, wenn sie über die Schnittstelle, von der aus auf dieses System zugegriffen wird, oder über einen leicht zugänglichen maschinenlesbaren Code oder andere elektronische Mittel leicht zugänglich ist.

(3) Die CE-Kennzeichnung wird gut sichtbar, leserlich und dauerhaft an Hochrisiko-KI-Systemen angebracht. Falls die Art des Hochrisiko-KI-Systems dies nicht zulässt oder nicht rechtfertigt, wird sie auf der Verpackung bzw. der beigefügten Dokumentation angebracht.

(4) Gegebenenfalls wird der CE-Kennzeichnung die Identifizierungsnummer der für die in Artikel 43 festgelegten Konformitätsbewertungsverfahren zuständigen notifizierten Stelle hinzugefügt. Die Identifizierungsnummer der notifizierten Stelle ist entweder von der Stelle selbst oder nach ihren Anweisungen durch den Anbieter oder den Bevollmächtigten des Anbieters anzubringen. Diese Identifizierungsnummer wird auch auf jeglichem Werbematerial angegeben, in dem darauf hingewiesen wird, dass das Hochrisiko-KI-System die Anforderungen für die CE-Kennzeichnung erfüllt.

(5) Falls Hochrisiko-KI-Systeme ferner unter andere Rechtsvorschriften der Union fallen, in denen die CE-Kennzeichnung auch vorgesehen ist, bedeutet die CE-Kennzeichnung, dass das Hochrisiko-KI-System auch die Anforderungen dieser anderen Rechtsvorschriften erfüllt.

- Bei in ein Produkt integrierten Hochrisiko-KI-Systemen sollte eine physische CE-Kennzeichnung angebracht werden. Die CE-Kennzeichnung ist gut sichtbar, leserlich und dauerhaft an das Hochrisiko-KI-System anzubringen (Art. 48 Abs. 3 KI-VO). Falls die Art des Hochrisiko-KI-Systems dies nicht zulässt oder nicht rechtfertigt, wird sie auf der Verpackung bzw. der beigefügten Dokumentation angebracht. Bei in ein Produkt

integrierten Hochrisiko-KI-Systemen kann die physische CE-Kennzeichnung durch eine digitale CE-Kennzeichnung ergänzt werden.

- Bei Hochrisiko-KI-Systemen, die digital bereitgestellt werden, soll eine digitale CE-Kennzeichnung verwendet werden. Eine digitale CE-Kennzeichnung soll allerdings nur dann verwendet werden, wenn sie über die Schnittstelle, von der aus auf dieses System zugegriffen wird, oder über einen leicht zugänglichen maschinenlesbaren Code oder andere elektronische Mittel leicht zugänglich ist (Art. 48 Abs. 2 KI-VO). Sollte dies nicht der Fall sein, ist die CE-Kennzeichnung etwa auf der beigefügten Dokumentation anzubringen.

(4) Hochrisiko-KI-Systeme, die bereits Gegenstand eines Konformitätsbewertungsverfahren gewesen sind, werden im Falle einer wesentlichen Änderung einem neuen Konformitätsbewertungsverfahren unterzogen, unabhängig davon, ob das geänderte System noch weiter in Verkehr gebracht oder vom derzeitigen Betreiber weitergenutzt werden soll. Bei Hochrisiko-KI-Systemen, die nach dem Inverkehrbringen oder der Inbetriebnahme weiterhin dazulernen, gelten Änderungen des Hochrisiko-KI-Systems und seiner Leistung, die vom Anbieter zum Zeitpunkt der ursprünglichen Konformitätsbewertung vorab festgelegt wurden und in den Informationen der technischen Dokumentation gemäß Anhang IV Nummer 2 Buchstabe f enthalten sind, nicht als wesentliche Veränderung;

Schritt 6.4: Vorgehen bei wesentlichen Änderungen

Art. 43 Abs. 4 KI-VO regelt, was bei wesentlichen Änderungen zu tun ist.

Pflichten bei einer wesentlichen Änderung.

Hochrisiko-KI-Systeme, die bereits Gegenstand eines Konformitätsbewertungsverfahren gewesen sind, müssen im Falle einer wesentlichen Änderung einem neuen Konformitätsbewertungsverfahren unterzogen werden. Bei der Frage, welches Konformitätsverfahren zu wählen ist, gilt das oben Gesagte entsprechend.

Wann liegt eine wesentliche Änderung vor?

Nach Art. 3 Nr. 23 der KI-VO ist eine wesentliche Veränderung eine Veränderung eines KI-Systems nach dessen Inverkehrbringen oder Inbetriebnahme, die

- in der vom Anbieter durchgeführten ursprünglichen Konformitätsbewertung nicht vorgesehen oder geplant war und durch welche die Konformität des KI-Systems mit den Anforderungen in Kapitel III Abschnitt 2 beeinträchtigt wird oder werden könnte (z. B. Änderung des Betriebssystems oder der Softwarearchitektur) oder
- die zu einer Änderung der Zweckbestimmung führt, für die das KI-System bewertet wurde.

Zwar stellt die KI-VO klar, dass Änderungen, die den Algorithmus und die Leistung von Hochrisiko-KI-Systemen betreffen, die nach dem Inverkehrbringen oder der Inbetriebnahme weiterhin dazulernen — d. h., sie passen automatisch an, wie die Funktionen ausgeführt werden —, keine wesentliche Veränderung darstellen, sofern diese Änderungen vom Anbieter vorab festgelegt, zum Zeitpunkt der Konformitätsbewertung bewertet wurden und in den Informationen der technischen Dokumentation gemäß Anhang IV Nummer 2 Buchstabe f enthalten sind.

Aus technischer Sicht kann jedoch selbst eine geringfügige Änderung des Lernalgorithmus oder der Modellparameter zu unvorhersehbaren Veränderungen im Verhalten eines KI-Systems führen. Streng genommen müsste daher die in Kapitel III, Abschnitt 2, insbesondere in Artikel 15 vorgesehene Evaluierung hinsichtlich Genauigkeit, Robustheit und Cybersicherheit vollständig wiederholt werden. Die genannte Ausnahme ist in der Praxis also kaum praktikabel. Besonders die sog. selbstlernenden bzw. kontinuierlich lernenden Systemen sind dadurch kaum konformitätsfähig.

Zwischenergebnis

Ist das Konformitätsbewertungsverfahren erfolgreich abgeschlossen worden, ist mit **Schritt 7** (fortlaufende Pflichten) fortzufahren.

9 Fortlaufende Pflichten

Prof. Dr. Heinz-Uwe Dettling Ernst & Young GmbH Wirtschaftsprüfungsgesellschaft), Michael Schemel (UnternehmerTUM GmbH)

Fortlaufende Pflichten bzw. Anbieterpflichten sind gesetzliche Vorgaben, die sich an Unternehmen richten, die KI-Systeme entwickeln, vertreiben oder einsetzen. Diese Pflichten dienen dazu, sicherzustellen, dass KI-Systeme während es gesamten Lebezyklus sicher, transparent und ethisch entwickelt und eingesetzt werden.

Schritt 7: Welche Pflichten sind nach dem Inverkehrbringen zu erfüllen?

Schritt 7.1.: Allgemeine Pflichten der Anbieter (Art. 16 KI-VO)

Art. 16 beschreibt in einer Übersicht, vergleichbar zu Art. 8 für die Anforderungen an Hochrisiko-KI-Systeme, welche Pflichten die Anbieter von Hochrisiko-KI-Systemen erfüllen werden müssen:

- Buchst. a: Erfüllen der Anforderungen für Hochrisiko-KI-Systeme (2. Teil 1. Kap. Rn. 266ff.);
- Buchst. b: Angabe von Namen, eingetragenem Handelsnamen/eingetragene Handelsmarke und Kontaktanschrift;
- Buchst. c: Vorhandensein eines Qualitätsmanagements gem. Art. 17 (2. Teil 1. Kap. Rn. 368 ff.);
- Buchst. d: Aufbewahren der Dokumentation gem. Art 18 (2. Teil 1. Kap. Rn. 373 ff.);
- Buchst. e: Aufbewahren der Protokolle gem. Art. 19 (2. Teil 1. Kap. Rn. 378 ff.);
- Buchst. f: Durchführen des Konformitätsbewertungsverfahren gem. Art 43;
- Buchst. g: Ausstellen der CE-Kennzeichnung zwecks Konformität mit Art. 48;
- Buchst. h: Anbringen einer CE-Kennzeichnung zwecks Konformität mit Art. 48;
- Buchst. i: Erfüllen der Registrierungspflichten gem. Art. 49;
- Buchst. j: Ergreifen erforderlicher Korrekturmaßnahmen und Bereitstellen erforderlicher Informationen gem. Art. 20 (2. Teil 1. Kap. Rn. 381 ff.);
- Buchst. k: auf begründete Anfrage einer nationalen Aufsichtsbehörde gem. Art. 21 nachweisen, dass Anforderungen aus Kapitel III, Abschnitt 2 erfüllt sind (2. Teil 1. Kap. Rn. 388 ff.); und
- Buchst. l: Erfüllen der Barrierefreiheitsanforderungen gem. Richtlinie 2016/2102 und Richtlinie 2019/882, im Idealfall mit Blick auf ErwG 80 durch Voreinstellungen, die bereits bei der Konzeption in das Hochrisiko-KI-System integriert werden.

Der „Verschiebebahnhof“ des Art. 16 ist hier im Kontext anderer Vorschriften der KI-VO zu betrachten und zu erläutern. So muss der Anbieter eines Hochrisiko-KI-Systems nach Art. 16 Buchst. i eine Registrierung gem. Art. 49 Abs. 1 in der von der Kommission nach Art. 71 Abs. 1 zu errichtenden Datenbank vornehmen. Die dabei anzugebenden Informationen sind Anhang VIII Abschnitt A zu entnehmen. Es gibt Ausnahmeregelungen hinsichtlich Anhang III Punkt 1, 2, 6 und 7. Der Pflichtenkatalog ist zudem nicht abschließend: Weitere Pflichten finden sich etwa in Art. 72. (Beobachtung nach dem Inverkehrbringen) und Art. 73 (Austausch von Informationen schwerwiegende Vorfälle).

Schritt 7.2: Die Pflichten nach Art. 72 und 73 KI-VO

Die Verpflichtung zur Beobachtung nach dem Inverkehrbringen gemäß Art. 72 ähnelt vergleichbaren Vigilanz-Pflichten bei anderen hochregulierten Produkten wie etwa Arzneimitteln nach der Richtlinie 2001/83/EG oder Medizinprodukten nach der Medizinprodukte-Verordnung (EU) 2017/745. Anbieter von Hochrisiko-KI-Systemen müssen ein System zur Beobachtung nach dem Inverkehrbringen, Einrichten und Dokumentieren. Das Beobachtungssystem (KI-Vigilanz-System) muss auf einem Plan für die Beobachtung nach dem Inverkehrbringen beruhen, der Teil der in Anhang IV genannten technischen Dokumentation ist. Die Kommission wird bis zum 2.2.2026 detaillierte Bestimmungen für die Erstellung eines Musters des Plans für die Beobachtung nach dem Inverkehrbringen sowie die Liste der in den Plan aufzunehmenden Elemente detailliert festlegen.

Die Vigilanz-Pflicht nach Art. 72 steht in engem Zusammenhang mit der Verpflichtung nach Art. 73 zur Meldung schwerwiegender Vorfälle an die zuständige Marktüberwachungsbehörde, wie sie ebenfalls etwa aus dem Bereich von Arzneimitteln oder Medizinprodukten bekannt ist. „Schwerwiegender Vorfall“ ist nach Art. 3 Nr. 49 einen Vorfall oder eine Fehlfunktion bezüglich eines KI-Systems, das bzw. die direkt oder indirekt eine der nachstehenden Folgen hat: a) den Tod oder die schwere gesundheitliche Schädigung einer Person; b) eine schwere und unumkehrbare Störung der Verwaltung oder des Betriebs kritischer Infrastrukturen; c) die Verletzung von Pflichten aus den Unionsrechtsvorschriften zum Schutz der Grundrechte; d) schwere Sach- oder Umweltschäden.

Entsprechende Meldungen an die Behörden müssen spätestens innerhalb von 15 Tagen erfolgen, nachdem der Anbieter oder gegebenenfalls der Betreiber Kenntnis von dem schwerwiegenden Vorfall erlangt hat. Im Falle eines weitverbreiteten Verstoßes oder eines schwerwiegenden Vorfalls i.S.d. Art. 3 Nr. 49 Buchstabe b muss die Meldung sogar spätestens innerhalb von zwei Tagen erfolgen, nachdem der Anbieter oder gegebenenfalls der Betreiber von diesem Vorfall Kenntnis erlangt hat. Im Falle des Todes einer Person hat die Meldung spätestens innerhalb von zehn Tagen zu erfolgen, nachdem der Anbieter oder gegebenenfalls der Betreiber von dem schwerwiegenden Vorfall Kenntnis erlangt hat.

Solche Meldungen lösen bei den Marktüberwachungsbehörden neben sonstigen sachgerechten Maßnahmen u. a. die unionsweiten elektronischen Meldeverfahren über das Schnellinformationssystem („ICSMS“) gemäß Art. 20 und 34 der Marktüberwachungs-Verordnung (EU) 2019/1020 und gegebenenfalls über das Schnellwarnsystem (Rapid Information Exchange System, RAPEX) gemäß Art. 12 der Richtlinie 2001/95/EG über die allgemeine Produktsicherheit bzw. ab dem 13.12.2024 gegebenenfalls über das Schnellwarnsystem Safety Gate nach Art. 26 der Verordnung über die allgemeine Produktsicherheit VO (EU) 2023/988 aus.

Zwischenergebnis

Wenn auch die fortlaufenden Pflichten erfüllt sind, ist die **Prüfung hier beendet**. Alle Schritte, die für die Compliance nach der KI-VO zu erledigen sind, sind erledigt.

Für Anbieter (und ggf. auch für Betreiber) von Hochrisiko-KI-Systemen, können jedoch die folgenden Kapitel von Interesse sein. **Kapitel 10 befasst sich mit Ko- und Selbstregulierungsmechanismen der KI-VO, Kapitel 11 mit Standardisierungsfragen und Kapitel 12 mit KI-Reallaboren.**

10 Ko- und Selbstregulierung

Dr. Frank Beer (INFODAS Gesellschaft für Systementwicklung und Informationsverarbeitung mbH), Hung Pham (INFODAS Gesellschaft für Systementwicklung und Informationsverarbeitung mbH)

Allgemeines zu Ko- und Selbstregulierung

Relevante(r) Artikel:

56, 95

Relevante(r) ErwG:

2, 5, 8

Konkretisierungsbedürftig:

-

Anbieter von Systemen mit KI müssen den regulatorischen Vorgaben der KI-VO der EU gerecht werden, damit eine sichere und ethische Nutzung von KI gewährleistet werden kann. Dabei sind die Anforderungen aus der KI-VO für die vier Risikoklassen von KI-Systemen unterschiedlich verpflichtend. Die Systeme mit inakzeptablem Risiko werden verboten (vgl. Kapitel 4 und 5). Für Hochrisiko-KI-Systeme und ihre Anbieter sind alle Anforderungskerngruppen aus der KI-VO bindend (vgl. Kapitel 6). Dabei handelt es sich um die Anforderungen an das Risikomanagementsystem, das Daten-Governance, die Protokollierung sowie die Aufzeichnungspflicht, die technische Dokumentation, die Transparenz, die menschliche Aufsicht (vgl. Human-In-The-Loop) und die Aspekte der Robustheit, Sicherheit und Genauigkeit der KI-Systeme. Im Gegensatz dazu gelten für KI-Systeme mit begrenztem Risiko nur spezifische Transparenzpflichten, und für Systeme mit gar keinem oder nur minimalem Risiko erlaubt die KI-VO sogar die freie Nutzung, das heißt für letztere gelten die Anforderungskerngruppen nicht. Für Anbieter von KI-Systemen ohne Risiko bis maximal begrenztem Risiko sieht die KI-VO stattdessen die **Ko- und Selbstregulierung** mithilfe von **Verhaltenskodize**s vor, die die Anbieter dazu motivieren sollen, sich dennoch an die Anforderungen der KI-VO auf freiwilliger Basis zu halten. Ein Spezialfall ist die Gruppe der GPAI-Systeme. Ihre Anbieter sind immer dazu verpflichtet, eine ausführliche technische Dokumentation zu führen. Ansonsten können GPAI wiederum in die Risikoklassen eingestuft werden, und die restlichen Anforderungen würden dann dementsprechend wieder unterschiedlich greifen.

Unter dem Begriff der **Selbstregulierung** ist in diesem Kontext eine freiwillige Selbstverpflichtung durch Anbieter von KI-Systemen an die Erfüllung der Anforderungen der KI-VO gemeint, meist in Form von Verhaltenskodizes. Eine **Ko-Regulierung** ist das gegenseitige Unterstützen bei der Erstellung von Verhaltenskodizes zur Selbstregulierung. Insgesamt können eine Vielzahl von Stakeholdern, die ein Interesse an einer tiefgründigeren Regulierung von KI-Systemen tragen oder vertreten, die über die gesetzlichen Mindestanforderungen hinausgehen, an der Ko-Regulierung beteiligt sein. Darunter sind unter anderem Industrieverbände, Unternehmen, Regulierungsbehörden, Standardisierungsorganisationen, zivilgesellschaftliche Organisationen, Nichtregierungsorganisationen (NGOs) und wissenschaftliche Institutionen. Es wurde beispielsweise das Europäische Büro für Künstliche Intelligenz ins Leben gerufen, was laut den beiden **Artikeln 56 und 95 der KI-VO** dabei unterstützen soll, Verhaltenskodizes für die Selbstregulierung bis zum 2. Mai 2025 zu etablieren.

Art. 95

(1) Das Büro für Künstliche Intelligenz und die Mitgliedstaaten fördern und erleichtern die Aufstellung von Verhaltenskodizes, einschließlich damit zusammenhängender Governance-Mechanismen, mit denen die freiwillige Anwendung einiger oder aller der in Kapitel III Abschnitt 2 genannten Anforderungen auf KI-Systeme, die kein hohes Risiko bergen, gefördert werden soll, wobei den verfügbaren technischen Lösungen und bewährten Verfahren der Branche, die die Anwendung dieser Anforderungen ermöglichen, Rechnung zu tragen ist.

(2) Das Büro für Künstliche Intelligenz und die Mitgliedstaaten erleichtern die Aufstellung von Verhaltenskodizes in Bezug auf die freiwillige Anwendung spezifischer Anforderungen auf alle KI-Systeme, einschließlich durch Betreiber, auf der Grundlage klarer Zielsetzungen sowie wesentlicher Leistungsindikatoren zur Messung der Erfüllung dieser Zielsetzungen, einschließlich unter anderem folgender Elemente:

a) in den Ethik-Leitlinien der Union für eine vertrauenswürdige KI enthaltene anwendbare Elemente;

b) Beurteilung und Minimierung der Auswirkungen von KI-Systemen auf die ökologische Nachhaltigkeit, einschließlich im Hinblick auf energieeffizientes Programmieren, und Techniken, um KI effizient zu gestalten, zu trainieren und zu nutzen;

c) Förderung der KI-Kompetenz, insbesondere der von Personen, die mit der Entwicklung, dem Betrieb und der Nutzung von KI befasst sind;

d) Erleichterung einer inklusiven und vielfältigen Gestaltung von KI-Systemen, unter anderem durch die Einsetzung inklusiver und vielfältiger Entwicklungsteams und die Förderung der Beteiligung der Interessenträger an diesem Prozess;

e) Bewertung und Verhinderung der negativen Auswirkungen von KI-Systemen auf schutzbedürftige Personen oder Gruppen schutzbedürftiger Personen, einschließlich im Hinblick auf die Barrierefreiheit für Personen mit Behinderungen, sowie auf die Gleichstellung der Geschlechter.

(3) Verhaltenskodizes können von einzelnen KI-System-Anbietern oder -Betreibern oder von Interessenvertretungen dieser Anbieter oder Betreiber oder von beiden aufgestellt werden, auch unter Einbeziehung von Interessenträgern sowie deren Interessenvertretungen einschließlich Organisationen der Zivilgesellschaft und Hochschulen. Verhaltenskodizes können sich auf ein oder mehrere KI-Systeme erstrecken, um ähnlichen Zweckbestimmungen der jeweiligen Systeme Rechnung zu tragen.

(4) Das Büro für Künstliche Intelligenz und die Mitgliedstaaten berücksichtigen die besonderen Interessen und Bedürfnisse von KMU, einschließlich Startups, bei der Förderung und Erleichterung der Aufstellung von Verhaltenskodizes.

Insbesondere der Artikel 95 gibt mehr Aufschluss über den Prozess zur Erstellung solcher Verhaltenskodizes. Letztere können von einzelnen KI-System-Anbietern, -Betreibern bzw. ihren Interessenvertretungen aufgestellt werden. Die Einbeziehung von Organisationen der Zivilgesellschaft und Hochschulen ist auch zulässig. Dabei werden die spezifischen Interessen und Bedürfnisse von KMUs, einschließlich Startups, bei der Erstellung dieser Kodizes mitberücksichtigt (vgl. ErwG 8). All diese Stakeholder sind (mit)verantwortlich dafür, die Kodizes zu initiieren, auszuarbeiten, zu überprüfen und zu implementieren, wobei sie auf der Grundlage klarer Zielsetzungen, wie sie von der EU vorgegeben werden, aufbauen müssen. Diese Ziele und ihre wesentlichen Leistungsindikatoren zur Messung des Erfüllungsgrads für KI-Systeme sind laut Artikel 95 die ethischen Leitlinien der EU, die Erhaltung der ökologischen Nachhaltigkeit, die Förderung der KI-Kompetenz von Stakeholdern, die Erleichterung einer vielfältigen Gestaltung von KI-Systemen und die Verhinderung der negativen Auswirkungen durch KI-Systeme auf schutzbedürftige Personen (vgl. ErwG 2 und 5).

Exkurs: Codes of Practice für GPAI

Auf internationaler Ebene gibt es seit dem Jahr 2023 mit dem G7-Gipfel in Hiroshima einen „Code of Conduct for Organizations Developing Advanced AI Systems“ – ein Verhaltenskodex mit einem risikobasierten Ansatz, der darauf abzielt, einen freiwillig einzuhaltenden Leitfaden darzustellen, um die Entwicklung von sicherer und vertrauenswürdiger KI-Systeme zu fördern. Der Hiroshima Verhaltenskodex enthält unter anderem die folgenden wichtigen Themenpunkte, die wir in diesem Abschnitt um internationale Normen und Standards ergänzen, um einen verdichteten Überblick zu geben:

1. Anbieter von KI-Systemen von fortschrittlichen KI-Systemen (darunter z. B. generative KI) sollten schon während des Designs darauf achten, dass die Potenziale und Risiken ihrer KI-Systeme bezüglich der Anwendbarkeit auf die Bereiche des Cyber-, chemischen, biologischen und atomaren Krieges minimiert werden. Zusätzlich dazu sind die negativen Auswirkungen auf die Gesellschaft als Ganzes, insbesondere Menschenrechte, Gesundheit und Sicherheit, untersagt. Auch die Gefahr, dass sich ein KI-System selbst vervielfältigt oder andere KI-Systeme trainiert, sollen sie berücksichtigen. Dazu empfiehlt der Kodex die Implementierung eines **Risikomanagementsystems**, das diese Risiken, die sich aus der Nutzung ergeben, überwachen und gegebenenfalls eindämmen soll. Anbieter können hier regelmäßige Risikoanalysen bzw. geeignete Testmaßnahmen durchführen, um potenzielle Risiken zu identifizieren und ihre Eintrittswahrscheinlichkeiten sowie Schweregrad zu bewerten. Standards für die Prozesse eines Risikomanagements können beispielsweise aus der ISO 31000:2018 (Norm für Risikomanagement) und der ISO/IEC 27005 (Standards für Informationssicherheitsrisikomanagement) bezogen werden. Für eine bessere Nachvollziehbarkeit sollen Vorfälle sowie sämtliche Entscheidungen in Form von **Protokollen** und **technischen Dokumentationen** festgehalten werden. Vorgaben dafür können beispielsweise aus der IEC/IEEE 82079-1 (Erstellung von Nutzungsinformation (Gebrauchsanleitungen) für Produkte – Teil 1 Grundsätze und allgemeine Anforderungen) und der ISO/IEC 15289 (Inhaltsanforderungen an System- und Softwarelebenszyklusprozesse) entnommen werden.

2. Im Sinne der **Transparenz** wird im Verhaltenskodex zu Mitteilungen bzw. Veröffentlichungen über die Fähigkeitsgrenzen der eigenen KI-Systeme sowie über die Bereiche der Anwendbarkeit geraten. Diese Informationen können wiederum in einer Dokumentation festgehalten werden. Dieser soll Informationen über die Bewertungen bezüglich der Tauglichkeit der KI-Systeme im Markt, sprich Funktionsweise und Einschränkungen (unter anderem die Maße an Genauigkeit, Robustheit und Cybersicherheit), und ihre potenziellen Risiken auf die Gesellschaft enthalten. Die KI-Systeme und der Betrieb müssen so transparent konzipiert sein, dass Nutzer die produzierten Ergebnisse angemessen interpretieren und verwenden können. Neben der detaillierten technischen Dokumentation zu den KI-Systemen können Anbieter und Händler Anleitungen oder weiterführende Informationsmaterialien zur Nutzung entwickeln und bereitstellen. Auch Schulungen oder Support für Nutzer können angeboten werden. Die Standards dafür sind beispielsweise aus dem Katalog ISO/IEC 25000 zu finden.

3. Auf das Thema **IT- und Cybersicherheit** wird im Verhaltenskodex auch sehr großen Wert gelegt. Leitlinien und Maßnahmen sollen definiert und Sicherheitskontrollen gegen äußere und innere Gefahren etabliert werden. Ein weiteres wichtiges Thema ist

die **Datensicherheit** und der **Datenschutz**. Anbieter von KI-Systemen sollen ihre Trainings-, Validierungs- und Testdatensätze mit Hinblick auf diese Aspekte so auswählen, dass kein Bias entstehen kann und die vertraulichen Daten vor unberechtigtem Zugriff geschützt werden. Entsprechend sind ihre KI-Systeme mit Hinblick auf Genauigkeit, Robustheit und Cybersicherheit so zu konzipieren, dass diese die Sicherheit und den Schutz personenbezogener Daten gewährleisten. Insbesondere sind Datenschutz-Folgeabschätzungen (engl. Data Protection Impact Assessments) nach den Vorgaben der DSGVO durchzuführen. Auch sind regelmäßige Überprüfungen der eigenen Datensicherheitsmaßnahmen sinnvoll. Als Leitfaden können hier die Standards aus der ISO/IEC 27001 (Informationssicherheits-Managementsysteme) sowie der ISO/IEC 27701 (Erweiterung des ISO/IEC 27001) genommen werden.

Neben dem Hiroshima Verhaltenskodex und dem bald zu erscheinenden praktischen Leitregeln für GPAI in Europa (unter der Leitung des Büros für KI) existieren momentan weitere Initiativen, die für eine effiziente Ko- und Selbstregulierung herangezogen werden können. Namhaft sind beispielsweise die Arbeiten des Joint Technical Committee 21 (JTC 21[7]) des Europäischen Komitees für elektronische Normung (vgl. Kapitel 11) oder des TÜV AI.Lab[8], die sich mit vertrauenswürdigen KI-Systemen beschäftigen. Es handelt sich zwar hierbei um Normen und Standards, die eher einen Fokus auf KI-Zertifizierungen und Konformitätsbewertungen nach KI-VO legen. Dennoch können sie als Vorlage genutzt werden, um auch auf freiwilliger Basis die Pflichten von KI-Systemen mit höherem Risiko zu erfüllen.

[7] https://www.cencenelec.eu/areas-of-work/cen-cenelec-topics/artificial-intelligence/
[8] https://www.tuev-lab.ai/

11 Standardisierung

Valentino Halim (Oppenhoff & Partner Rechtsanwälte Steuerberater mbB), Frank Wisselink (Deutsche Telekom AG)

Relevante(r) Artikel:

40, 41, 42, 43, 47

Relevante(r) ErwG:

77, 78, 121, 77, 122, 123, 124, 125, 126, 128, 147

Konkretisierungsbedürftig:

/

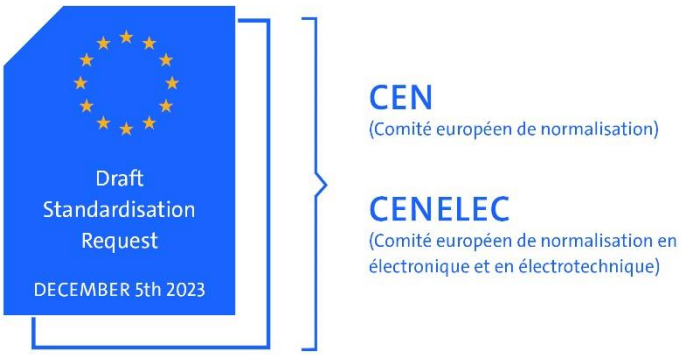
Standardisierung wird für die Umsetzung der KI-VO eine herausragende Rolle spielen. Anders als z. B. Richtlinien wird die KI-VO nicht in nationales Recht der Mitgliedstaaten umgesetzt, sondern durch Normen.

Ausgangspunkt hierfür ist Artikel 40(1) KI-VO. Das New Legislative Framework (NLF) bietet der EU-Kommission die Möglichkeit, eine Verordnung durch die Industrie in technische Anforderungen umzusetzen.

> (1) Bei Hochrisiko-KI-Systemen oder KI-Modellen mit allgemeinem Verwendungszweck, die mit harmonisierten Normen oder Teilen davon, deren Fundstellen gemäß der Verordnung (EU) Nr. 1025/2012 im Amtsblatt der Europäischen Union veröffentlicht wurden, übereinstimmen, wird eine Konformität mit den Anforderungen gemäß Abschnitt 2 des vorliegenden Kapitels oder gegebenenfalls mit den Pflichten gemäß Kapitel V Abschnitte 2 und 3 der vorliegenden Verordnung vermutet, soweit diese Anforderungen oder Verpflichtungen von den Normen abgedeckt sind.

Die EU-Kommission hat CEN-CENELEC um die Erstellung harmonisierter Normen gebeten

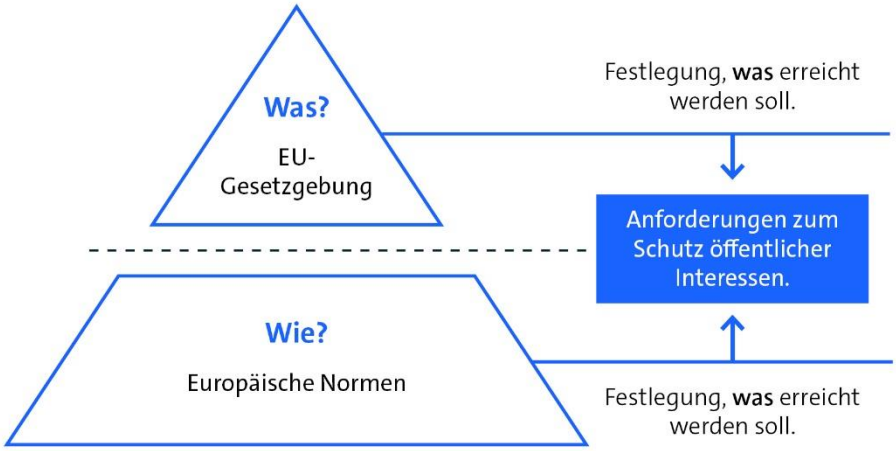
Die EU-Kommission kann durch einen Normungsauftrag die europäischen Standardisierungsorganisationen (in diesem Fall CEN-CENELEC) ersuchen, Normen zu erstellen. Da CEN-CENELEC vor der Veröffentlichung der KI-VO beauftragt worden ist, ist ein Entwurf eines Normungsauftrags (*draft standardization request*) für die KI-VO erteilt worden [Link zum Normungsauftrag].



Der Normungsauftrag ist auf die Etablierung einer harmonisierten Norm gerichtet. Gemäß Artikel 2(1)(c) der Verordnung (EU) 1025/2012 ist eine „harmonisierte Norm“: eine europäische Norm, die auf der Grundlage eines Auftrags der Kommission zur Durchführung von Harmonisierungsrechtsvorschriften der Union angenommen wurde.

Diese Normen beschreiben, wie die KI-VO umgesetzt werden kann

Die KI-VO enthält nur sehr allgemein formulierte Anforderungen. Beispielsweise verpflichtet Artikel 9 KI-VO Anbieter/Betreiber von Hochrisiko-KI-Systemen dazu, ein Risikomanagementsystem zu implementieren (das „Was“). Die KI-VO bestimmt aber nicht, welche konkreten Maßnahmen Unternehmen umsetzen müssen, um ein Risikomanagementsystem angemessen umzusetzen. Dies übernehmen technische Normen, die konkrete Maßnahmen einschließlich Prozesse und Verfahren formulieren, wie Unternehmen eine gesetzliche Anforderung erfüllen können (das „Wie“).



Die EU-Kommission kann technische Normen akzeptieren, indem sie diese annimmt. Die Annahme endet mit der Veröffentlichung der Norm im Amtsblatt der EU (Official Journal) veröffentlicht. Damit ist der Harmonisierungsprozess abgeschlossen.

In der EU haben Normen für den Binnenmarkt eine ganz wesentliche Bedeutung, insbesondere aufgrund der Verwendung harmonisierter Normen, verbunden mit der Vermutung der Konformität von Produkten, die auf dem Markt angeboten werden sollen.

Hochrisiko-KI-Systeme müssen nach der KI-VO strenge Anforderungen erfüllen ([Verweis Bitkom Leitfaden]), damit sie auf dem EU-Markt in Verkehr gebracht und eingesetzt werden dürfen. Um diese Konformität nachzuweisen, können Unternehmen harmonisierte Normen implementieren. In diesem Fall greift die gesetzliche Konformitätsvermutung (Artikel 40(1) KI-VO). Alternativ können Unternehmen auch eigene Maßnahmen definieren und umsetzen, um die Anforderungen der KI-VO zu erfüllen. Die oben genannte Konformitätsvermutung greift in diesem Fall allerdings nicht. Dies bedeutet regelmäßig einen erheblichen Mehraufwand und Rechtsunsicherheit für Unternehmen. Insofern besteht ein starker Anreiz, harmonisierte Normen umzusetzen.

10 Themen werden durch CEN-CENELEC normiert

Der Normungsauftrag enthält 10 Themen, welche durch CEN-CENELEC bearbeitet werden.

Folgende 3 Beispiele erklären eine Auswahl der Themen näher.

Das Risikomanagementsystem ist ein zentrales Thema, da die KI-VO selbst einem risikobasierten Ansatz folgt. Nach dem Normungsauftrag wird von der harmonisierten Norm unter anderem erwartet, dass diese zum Thema Risikomanagement-System „Spezifikationen für ein Risikomanagementsystem für KI-Systeme festlegt." Dabei folgt das Risikomanagement-System einem *lifecycle approach,* der aus dem Produktsicherheitsrecht stammt. D. h. „[d]as Risikomanagement ist als kontinuierlicher iterativer Prozess zu verstehen, der über den gesamten Lebenszyklus des KI-Systems abläuft und darauf abzielt, die einschlägigen Risiken für Gesundheit, Sicherheit oder Grundrechte zu verhindern oder zu minimieren." Artikel 9 der KI-VO beschreibt die meisten der gesetzlichen Anforderungen an diese(r) Norm(en) womit diese harmoniert werden. Diese Anforderungen gehen wie in der Normungsanfrage über technische Risiken hinaus. Vielmehr müssen auch Risiken für Gesundheit, Sicherheit und Grundrechte müssen aktiv berücksichtigt werden. Zudem muss auch eine „Abschätzung und Bewertung der Risiken […] im Rahmen einer vernünftigerweise vorhersehbaren Fehlanwendung verwendet wird" stattfinden.

Daten und Daten-Governance sind für eine vertrauenswürdige KI essenziell. Im Rahmen des Normungsauftrags müssen hier Normen entwickelt werden, welche beschreiben, wie die KI-Modelle mit Trainings-, Validierungs- und Testdatensätzen entwickelt werden und verschiedenen Qualitätskriterien entsprechen (Artikel 10 KI-VO). Zum Beispiel muss sichergestellt werden, dass „mögliche Verzerrungen, die die Gesundheit und Sicherheit von Personen beeinträchtigen, sich negativ auf die Grundrechte auswirken oder zu einer nach den Rechtsvorschriften der Union verbotenen Diskriminierung führen könnten, insbesondere wenn die Datenausgaben die Eingaben für künftige Operationen beeinflussen"

ausgeschlossen werden. Nach Art. 10 KI-VO müssen die Trainings-, Validierungs- und Testdaten

- relevant
- hinreichend repräsentativ, und
- so weit wie möglich, frei von Fehlern, und
- vollständig sein.

Zur Beurteilung dieser Kriterien erfolgt auf Basis der Zweckbestimmung des konkreten Anwendungsfalls, für den das KI-System eingesetzt werden soll. Berücksichtigt werden:

- Merkmale oder Elemente, die dem spezifischen geografischen, kontextuellen, verhaltensbezogenen oder funktionalen Umfeld eigen sind
- Verarbeitung besonderer Kategorien personenbezogener Daten: Es müssen angemessene Garantien für die Grundrechte und -freiheiten natürlicher Personen vorgesehen werden
- Bewerbungsverfahren für Führungspositionen: Datensätze, bei denen Frauenquoten nicht berücksichtigt wurden
- System zur Zugangskontrolle: wenn das System mit „falschen“ Daten oder Statistiken über die Strafbarkeit bestimmter Personengruppen trainiert wird und deshalb eine Warnmeldung anzeigt, wenn bestimmte äußere Erscheinungsmerkmale vorhanden sind

Menschliche Aufsicht ist bei Hoch Risiko KI eine wesentliche Komponente und dient (Artikel 14) „der Verhinderung oder Minimierung der Risiken für Gesundheit, Sicherheit oder Grundrechte, die entstehen können, wenn ein Hochrisiko-KI-System im Einklang mit seiner Zweckbestimmung oder im Rahmen einer vernünftigerweise vorhersehbaren Fehlanwendung verwendet wird…“.

„Hochrisiko-KI-Systeme werden so konzipiert und entwickelt, dass sie während der Dauer ihrer Verwendung — auch mit geeigneten Instrumenten einer Mensch-Maschine-Schnittstelle — von natürlichen Personen wirksam beaufsichtigt werden können“.

- Verarbeitung besonderer Kategorien personenbezogener Daten: Es müssen angemessene Garantien für die Grundrechte und -freiheiten natürlicher Personen vorgesehen werden
- Bewerbungsverfahren für Führungspositionen: Datensätze, bei denen Frauenquoten nicht berücksichtigt wurden
- System zur Zugangskontrolle: wenn das System mit „falschen“ Daten oder Statistiken über die Strafbarkeit bestimmter Personengruppen trainiert wird und deshalb eine Warnmeldung anzeigt, wenn bestimmte äußere Erscheinungsmerkmale vorhanden sind. Menschliche Aufsicht ist bei Hoch Risiko KI eine wesentliche Komponente und dient (Artikel 14) „der Verhinderung oder Minimierung der Risiken für Gesundheit, Sicherheit oder Grundrechte, die entstehen können, wenn ein Hochrisiko-KI-System im Einklang mit seiner Zweckbestimmung oder im Rahmen einer vernünftigerweise vorhersehbaren Fehlanwendung verwendet wird…“.

„Hochrisiko-KI-Systeme werden so konzipiert und entwickelt, dass sie während der Dauer ihrer Verwendung — auch mit geeigneten Instrumenten einer Mensch-Maschine-Schnittstelle — von natürlichen Personen wirksam beaufsichtigt werden können“.

12 KI-Reallabore

Jan-Dierk Schaal (SKW Schwarz Rechtsanwälte*)*

Die **Art. 57ff.** der KI-VO adressieren unter der Überschrift „Maßnahmen zur Innovationsförderung“ KI-Reallabore (engl. „regulatory sandboxes“) als innovationsförderndes Mittel. Art. 57 KI-Verordnung regelt dabei, wie derartige KI-Reallabore durch die EU-Mitgliedstaaten einzurichten sind.

Relevante(r) Artikel:

57-63

Relevante(r) ErwG:

138, 139

Konkretisierungsbedürftig:

- Benennung bzw. Errichtung der zuständigen Aufsichtsbehörde

- Durchführungsrechtsakte für detaillierte Regelungen für KI-Reallabore und deren Funktionsweise, insb. zu den Voraussetzungen und Auswahlkriterien für eine Beteiligung am KI-Reallabor

Art. 3 KI-Verordnung: Begriffsbestimmungen

54. „Plan für das Reallabor“ ein zwischen dem teilnehmenden Anbieter und der zuständigen Behörde vereinbartes Dokument, in dem die Ziele, die Bedingungen, der Zeitrahmen, die Methodik und die Anforderungen für die im Reallabor durchgeführten Tätigkeiten beschrieben werden;

55. „KI-Reallabor“ einen kontrollierten Rahmen, der von einer zuständigen Behörde geschaffen wird und den Anbieter oder zukünftige Anbieter von KI-Systemen nach einem Plan für das Reallabor einen begrenzten Zeitraum und unter regulatorischer Aufsicht nutzen können, um ein innovatives KI-System zu entwickeln, zu trainieren, zu validieren und – gegebenenfalls unter Realbedingungen – zu testen.

Sinn und Zweck von KI-Reallaboren

Eine der zentralen Herausforderungen der KI-Verordnung liegt für den europäischen Gesetzgeber darin, Innovationsanreize für künstliche Intelligenz in einen angemessenen Ausgleich zu ihrer Regulierungsbedürftigkeit zu setzen. Die KI-Verordnung zielt deshalb darauf ab, die Entwicklung von KI-Systemen innerhalb der Europäischen Union gerade unter Berücksichtigung des hohen regulatorischen Rahmens zu fördern, um so den europäischen Markt für hochinnovative Unternehmen attraktiv zu gestalten und nicht zu versperren.

In ErwG 138 wird das Verhältnis zwischen risikominimierenden und innovationsfördernden Vorgaben wie folgt erläutert:

„KI ist eine sich rasch entwickelnde Technologiefamilie, die eine regulatorische Aufsicht und einen sicheren und kontrollierten Raum für Experimente erfordert, wobei gleichzeitig eine verantwortungsvolle Innovation und die Integration geeigneter Schutzmaßnahmen und Maßnahmen zur Risikominderung gewährleistet werden müssen.“

Ein Mittel, um dieses Ziel zu erreichen, sind sogenannte KI-Reallabore. Der englische Begriff „Sandboxes“ beschreibt die Idee der KI-Reallabore treffend: Es soll gleich einem Sandkasten für Kinder ein sicheres Environment geschaffen werden, in dem KI entwickelt und getestet werden kann. Dabei stammt der Begriff der Sandbox, der mit dem Begriff des Reallabors synonym verwendet wird, ursprünglich aus der Informatik. Dort dienen derartige

Reallabore dazu, neue Programmierungen, deren Risikopotential noch nicht abzuschätzen ist, in einer kontrollierten und sicheren Umgebung zu testen, um so eine Marktreife herzustellen und etwaige Risiken für die Anwender zu minimieren. Daran angelehnt werden heute Reallabore als regulatorisches Instrument dazu genutzt, neue und innovative Produkte unter der Aufsicht einer Regulierungsbehörde für einen begrenzten Zeitraum zu testen. Zudem besteht die Möglichkeit für den Gesetzgeber, auf Grundlage der in den Reallaboren gewonnenen Erkenntnisse den Rechtsrahmen zu optimieren.

Die Art. 57ff. KI-Verordnung adressieren KI-Reallabore explizit als Maßnahme zur Innovationsförderung. Nach ErwG 139 der KI-Verordnung sollen diese ein kontrolliertes Experimentier- und Testumfeld schaffen, um die Entwicklung, Erprobung und Validierung innovativer KI-Systeme vor ihrem Inverkehrbringen oder ihrer Inbetriebnahme zu erleichtern und gleichzeitig die Vereinbarkeit mit der KI-Verordnung und sonstigen Rechtsvorschriften zu gewährleisten. Hierdurch soll die Entwicklung der Europäischen Union im Rahmen eines innovationsfreundlichen Klimas zu einem global wettbewerbsfähigen Akteur auf dem Gebiet der künstlichen Intelligenz gefördert und zudem Rechtssicherheit für Innovatoren geschaffen werden. Das KI-Reallabor unterstützt damit, den Widerspruch aufzulösen, der sich daraus begründet, dass auf der einen Seite ein hochregulierter Rechtsrahmen geschaffen wird, der die Wirtschaft mit erheblichen Kosten und Bürokratieaufwand belastet und damit Innovationshemmnisse schafft, was zu einer potenziellen Schwächung des europäischen Standortes im internationalen Vergleich führt, auf der anderen Seite jedoch die Sicherheit gerade von Hochrisiko-KI-Systemen gewährleistet und die Entwicklung und Nutzung bestimmter KI-Systeme untersagt.

Art. 57 Abs. 9 KI-Verordnung listet die verschiedenen Ziele im Einzelnen auf, die mit der Einrichtung von KI-Reallaboren verfolgt werden:

a) Verbesserung der Rechtssicherheit, um für die Einhaltung der Regulierungsvorschriften dieser Verordnung oder, ggf. anderem geltenden Unionsrecht und nationalem Recht zu sorgen

b) Förderung des Austauschs bewährter Verfahren durch Zusammenarbeit mit den am KI-Reallabor beteiligten Behörden

c) Förderung von Innovation und Wettbewerbsfähigkeit sowie Erleichterung der Entwicklung eines KI-Ökosystems

d) Leisten eines Beitrags zum evidenzbasierten regulatorischen Lernen

e) Erleichterung und Beschleunigung des Zugangs von KI-Systemen zum Unionsmarkt, insbesondere wenn sie von KMU – einschließlich Startup-Unternehmen – angeboten werden

Während in anderen Bereichen das Ziel der Ermöglichung kontrollierter Ausnahmen von regulatorischen Vorgaben und Verboten für die Erprobung neuer Rechtsrahmen oder innovativer Produkte dadurch umgesetzt worden ist, dass Experimentierklauseln (z. Bsp. in § 2 Abs. 7 Personenbeförderungsgesetz) oder Ausnahmeregelungen (z. Bsp. § 21i LuftVO) geschaffen worden sind, wurde in der KI-Verordnung ein anderer Weg beschritten.

Regelungsüberblick

Verpflichtung der Mitgliedsstaaten zur Errichtung von KI-Reallaboren und Zuständigkeiten
Die KI-Verordnung enthält in Art. 57 grundsätzliche Regelungen für die Einrichtung von KI-

> Art. 57 KI-Verordnung
>
> (1) Die Mitgliedstaaten sorgen dafür, dass ihre zuständigen Behörden mindestens ein KI-Reallabor auf nationaler Ebene einrichten, das bis zum 2. August 2026 einsatzbereit sein muss. Dieses Reallabor kann auch gemeinsam mit den zuständigen Behörden anderer Mitgliedstaaten eingerichtet werden. Die Kommission kann technische Unterstützung, Beratung und Instrumente für die Einrichtung und den Betrieb von KI-Reallaboren bereitstellen.
>
> Die Verpflichtung nach Unterabsatz 1 kann auch durch Beteiligung an einem bestehenden Reallabor erfüllt werden, sofern eine solche Beteiligung die nationale Abdeckung der teilnehmenden Mitgliedstaaten in gleichwertigem Maße gewährleistet.
>
> (2) Es können auch zusätzliche KI-Reallabore auf regionaler oder lokaler Ebene oder gemeinsam mit den zuständigen Behörden anderer Mitgliedstaaten eingerichtet werden;
> (3) Der Europäische Datenschutzbeauftragte kann auch ein KI-Reallabor für Organe, Einrichtungen und sonstige Stellen der Union einrichten und die Rollen und Aufgaben der zuständigen nationalen Behörden im Einklang mit diesem Kapitel wahrnehmen.

Reallaboren als Maßnahme zur Innovationsförderung. Art. 57 Abs. 1 KI-VO sieht dabei nicht vor, dass ein einheitliches KI-Reallabor als europäisches Werkzeug für die gesamte Europäische Union geschaffen wird, sondern verlagert die Einrichtung der KI-Reallabore jeweils auf nationale Ebene. Jeder Mitgliedsstaat ist verpflichtet, bis zum 2. August 2026 ein KI-Reallabor bereitzustellen, wobei dieses auch gemeinsam mit anderen Mitgliedstaaten eingerichtet werden kann, etwa durch Beteiligung an einem bereits bestehenden KI-Reallabor (Art. 57 Abs. 1 KI-Verordnung). Voraussetzung hierfür ist lediglich, dass die nationale Abdeckung der jeweils teilnehmenden Mitgliedstaaten durch die Beteiligung an einem bestehenden KI-Reallabor in gleichem Maße gewährleistet ist, wie bei einer Einrichtung eines eigenständigen KI-Reallabors.

Dabei sind die Mitgliedstaaten gem. Art. 57 Abs. 2 KI-Verordnung nicht darauf beschränkt, lediglich ein KI-Reallabor zu errichten, sondern können auch mehrere KI-Reallabore auf regionaler oder lokaler Ebene oder gemeinsam mit anderen Mitgliedstaaten oder auch für verschiedene Anwendungsbereiche einrichten.

Die Befugnis, KI-Reallabore einzurichten, haben gem. Art. 57 Abs. 3 KI-Verordnung der Europäische Datenschutzbeauftragte sowie die zuständigen Behörden eines oder mehrerer Mitgliedstaaten, Art. 57 Abs. 1 und 2 KI-Verordnung.

Lokal zuständige Behörde ist gem. Art. 3 Nr. 48 KI-Verordnung eine notifizierende Behörde oder eine Marktüberwachungsbehörde, die von dem jeweiligen Mitgliedstaat zum Zwecke der Durchführung und Anwendung der KI-Verordnung zu errichten oder zu benennen ist, Art. 70 KI-Verordnung. Nach Art. 57 Abs. 10 KI-VO sind die nationalen Datenschutzbehörden entsprechend ihrer Zuständigkeit sowie ihrer Aufgaben und Befugnisse am Betrieb der Reallabore zu beteiligen und in die Aufsicht einzubeziehen, soweit die innovativen KI-Systeme personenbezogene Daten verarbeiten. Inwieweit dies zu eigenen Kompetenzverlagerungen auf die Datenschutzbehörden führt oder lediglich eine Konsultationspflicht beinhaltet, ist in der KI-Verordnung nicht eindeutig geregelt und wird noch zu klären sein.

Die Benennung bzw. Errichtung einer zuständigen Aufsichtsbehörde hat in Deutschland bis zum 2. August 2025 zu erfolgen. Vorreiter unter den EU-Mitgliedstaaten in diesem Bereich ist Spanien. Bereits im Juni 2022 präsentierte die spanische Regierung ein Pilotprojekt für ein erstes KI-Reallabor, im August 2023 richtete Spanien als erster EU-Mitgliedstaat eine nationale KI-Aufsichtsbehörde (AESIA) ein.

Aufgaben der zuständigen Behörden

Die Aufgaben der zuständigen Behörden umfassen gem. Art. 57 KI-Verordnung insbesondere folgende Aspekte:

- Vereinbarung eines Reallabor-Plans mit den jeweiligen Anbietern zur Erleichterung der Entwicklung, des Trainings und der Validierung der KI-Systeme (Art. 57 Abs. 5 KI-Verordnung),
- Gegebenenfalls Anleitung, Aufsicht und Unterstützung der Anbieter, um Risiken, insbesondere im Hinblick auf Grundrechte, Gesundheit und Sicherheit sowie die Einhaltung der Vorschriften der KI-Verordnung zu ermitteln (Art. 57 Abs. 6 KI-Verordnung),
- Bereitstellung von Leitfäden zu regulatorischen Erwartungen und zur Erfüllung der Anforderungen und Pflichten der KI-Verordnung (Art. 57 Abs. 7 KI-Verordnung),
- Überwachung der KI-Reallabore, ob während der Erprobung erhebliche Risiken für die Gesundheit und Sicherheit sowie die Grundrechte zu erwarten sind und sofern dies der Fall ist, Durchführung von Risikominimierungsmaßnahmen (Art. 57 Abs. 11 KI-Verordnung),
- Befugnis, das Testverfahren oder die Beteiligung am Reallabor vorübergehend oder dauerhaft auszusetzen, sofern keine Risikominimierung möglich ist (Art. 57 Abs. 11 KI-Verordnung),
- Übermittlung jährlicher Berichte sowie eines Abschlussberichts über das Reallabor an das Büro für künstliche Intelligenz und das KI-Gremium (Art. 57 Abs. 16 KI-Verordnung).

Art. 57 KI-Verordnung ermöglicht demnach insbesondere einen engen Behördenkontakt und ein Beratungsangebot der Behörden an die jeweiligen (potenziellen) Anbieter von KI-Systemen. Durch eine möglichst enge Zusammenarbeit zwischen Behörde und Anbietern sollen mit den KI-Systemen verbundene Risiken durch die Vereinbarung eines Reallabor-Plans, die Anleitung, die Bereitstellung von Leitfäden und die Überwachung der Reallabore

früh erkannt und minimiert sowie die Einhaltung der Vorschriften der KI-Verordnung sichergestellt werden. Unsicherheit besteht, inwieweit Art. 57 Abs. 6 KI-Verordnung eine tatsächliche Verpflichtung zur Konsultation und Aufsicht der zuständigen Behörde beinhaltet, da der dortige Wortlaut durch den Begriff „gegebenenfalls“ den Eindruck erwecken mag, dass die entsprechende Verpflichtung im Ermessen des Mitgliedsstaates bei Ausgestaltung der zuständigen Behörde mit deren Rechten und Pflichten läge. Dies stände allerdings nach diesseitiger Auffassung im Widerspruch zum Sinn und Zweck der KI-Reallabore, die Entwicklung sicherer KI zu gewährleisten, weshalb die Formulierung des Verordnungsgebers dahingehend zu verstehen ist, dass „bei Bedarf“ die zuständige Behörde zur Anleitung, Aufsicht und Unterstützung der Anbieter verpflichtet ist.

Haftung innerhalb der KI-Reallabore

Die wichtige Frage der Haftung der am KI-Reallabor beteiligten Anbieter wird in Art. 57 Abs. 12 KI-Verordnung adressiert.

> Art. 57 KI-Verordnung
>
> (12) Die am KI-Reallabor beteiligten Anbieter und zukünftigen Anbieter bleiben nach geltendem Recht der Union und nationalem Haftungsrecht für Schäden haftbar, die Dritten infolge der Erprobung im Reallabor entstehen. Sofern die zukünftigen Anbieter den spezifischen Plan und die Bedingungen für ihre Beteiligung beachten und der Anleitung durch die zuständigen nationalen Behörden in gutem Glauben folgen, werden jedoch von den Behörden keine Geldbußen für Verstöße gegen diese Verordnung verhängt. In Fällen, in denen andere zuständige Behörden, die für anderes Unionsrecht und nationales Recht zuständig sind, aktiv an der Beaufsichtigung des KI-Systems im Reallabor beteiligt waren und Anleitung für die Einhaltung gegeben haben, werden im Hinblick auf dieses Recht keine Geldbußen verhängt.

Danach bleiben diese nach geltendem nationalen Haftungsrecht und Unionsrecht für Schäden haftbar, die Dritten infolge der Erprobung im Reallabor entstehen. Die maßgeblichen Haftungsregelungen ergeben sich nicht aus der KI-Verordnung selbst, sondern aus der geplanten KI-Haftungsrichtlinie, deren Entwurfsfassung die EU-Kommission im September 2022 veröffentlicht hat (https://eur-lex.europa.eu/legal-content/DE/TXT/?uri=CELEX%3A52022PC0496). Diese Richtlinie soll außervertragliche Haftungsregelungen für Schäden durch KI im mitgliedsstaatlichen Recht harmonisieren. KI-Reallabore werden innerhalb dieser Richtlinie nicht berücksichtigt oder gar einer Haftungsprivilegierung unterworfen. In den KI-Reallaboren gelten mithin keine zivilrechtlichen Haftungserleichterungen.

Soweit die Anbieter sich jedoch an die „Spielregeln“ innerhalb der KI-Reallabore halten, sind die zuständigen Behörden nach Art. 57 Abs. 12 KI-VO daran gehindert, Geldbußen zu

verhängen. Dies betrifft auch etwaige Verstöße gegen andere Rechtsvorschriften, wie etwa datenschutzrechtliche Bestimmungen, soweit die entsprechenden Behörden in die Entwicklung eingebunden waren.

Einrichtung und Betrieb der KI-Reallabore

Aus Art. 58 Abs. 1 KI-Verordnung ergibt sich, dass detaillierte Regelungen für die Einrichtung, Entwicklung, Umsetzung, den Betrieb sowie die Beaufsichtigung von KI-Reallaboren durch von der Kommission noch zu erlassende Durchführungsrechtsakte festgelegt werden sollen, um eine Rechtszersplitterung zu vermeiden. Demnach sind die genauen Voraussetzungen für die Teilnahme an den Reallaboren sowie die in den Reallaboren geltenden Regelungen in der KI-Verordnung selbst noch nicht festgesetzt.

(1) Um eine Zersplitterung in der Union zu vermeiden, erlässt die Kommission Durchführungsrechtsakte, in denen detaillierte Regelungen für die Einrichtung, Entwicklung, Umsetzung, den Betrieb und die Beaufsichtigung der KI-Reallabore enthalten sind. In den Durchführungsrechtsakten sind gemeinsame Grundsätze zu den folgenden Aspekten festgelegt:

g) Voraussetzungen und Auswahlkriterien für eine Beteiligung am KI-Reallabor;

h) Verfahren für Antragstellung, Beteiligung, Überwachung, Ausstieg und Beendigung bezüglich des KI-Reallabors, einschließlich Plan und Abschlussbericht für das Reallabor;

i) für Beteiligte geltende Anforderungen und Bedingungen. Diese Durchführungsrechtsakte werden gemäß dem in Artikel 98 Absatz 2 genannten Prüfverfahren erlassen

j) für Beteiligte geltende Anforderungen und Bedingungen.

Die Durchführungsrechtsakte sollen Grundsätze zu folgenden Aspekten festlegen:

a) Voraussetzungen und Auswahlkriterien für eine Beteiligung am KI-Reallabor;

b) Verfahren für Antragstellung, Beteiligung, Überwachung, Ausstieg und Beendigung bezüglich des KI-Reallabors, einschließlich Plan und Abschlussbericht für das Reallabor;

c) für Beteiligte geltende Anforderungen und Bedingungen.

Insoweit sollen die Durchführungsakte gewährleisten, dass die KI-Reallabore allen Anbietern eines KI-Systems offenstehen, und grundsätzlich einen gleichberechtigten Zugang ermöglichen.

Datenschutzrechtliche Bestimmungen

Rechtmäßig für andere Zwecke erhobene personenbezogene Daten dürfen im KI-Reallabor gemäß Art. 59 KI-Verordnung ausschließlich für die Zwecke der Entwicklung, des Trainings und des Testens bestimmter KI-Systeme im Reallabor verarbeitet werden, wenn alle Voraussetzungen des Art. 59 Abs. 1 a)-j) KI-Verordnung erfüllt sind. Insbesondere muss das KI-System zur Wahrung eines erheblichen öffentlichen Interesses in einem in Art. 59 Abs. 1 a) i)-v) KI-Verordnung genannten Bereich wie der öffentlichen Sicherheit oder öffentlichen Gesundheit entwickelt werden. Weiter müssen die verarbeiteten Daten für die Erfüllung einer oder mehrerer Anforderungen an Hochrisiko-KI nach Kapitel III Abschnitt 2 der KI-Verordnung erforderlich sein, sofern diese Anforderungen durch die Verarbeitung nicht personenbezogener Daten nicht wirksam erfüllt werden können. Zudem ist besonders zu beachten, dass die im Rahmen des Reallabors verarbeiteten personenbezogenen Daten durch geeignete technische und organisatorische Maßnahmen zu schützen sind und gelöscht werden müssen, sobald die Beteiligung an dem Reallabor endet oder das Ende der Speicherfrist für die personenbezogenen Daten erreicht ist.

Insoweit stellt diese Regelung nach diesseitiger Auffassung allerdings keine Ermächtigungsgrundlage im Sinne des Art. 6 Abs. 3 DSGVO dar, da die Verarbeitung der personenbezogenen Daten im Rahmen eines KI-Reallabors regelmäßig nicht im öffentlichen Interesse oder in Ausübung der öffentlichen Gewalt erfolgen wird. Insoweit bedarf es mithin weiterhin einer gesonderten datenschutzrechtlichen Ermächtigungsgrundlage.

Privilegierung von KMU

Das europäische KI-Ökosystem besteht bislang zu einem großen Teil aus kleinen und mittleren Unternehmen (im Folgenden KMU), die als wesentliche Innovationstreiber fungieren. Da KMU in der Regel über weniger Ressourcen verfügen als große Unternehmen gilt es, die diese stärker treffenden Einschränkungen aufgrund der zu beachtenden Regulatorik durch besondere Förderung hinsichtlich Zugang zu und Nutzung von KI-Systemen auszugleichen.

KMU sind nach der Empfehlung der Kommission vom 6. Mai 2003 betreffend die Definition der Kleinstunternehmen sowie der kleinen und mittleren Unternehmen (ABl. der EU L 124/36 vom 20.05.2003) solche Unternehmen, die weniger als 250 Personen beschäftigen und einen Jahresumsatz von höchstens 50 Mio. Euro erzielen, oder deren Jahresbilanzsumme sich auf höchstens 43 Mio. Euro beläuft. Start-ups werden häufig unter den Begriff der KMU subsumiert, da in der europäischen und nationalen Rechtsordnung bislang keine einheitliche Definition für diese existiert. Diese Einordnung findet sich auch in

der KI-Verordnung durch die verwendete Formulierung „KMU, einschließlich Start-up-Unternehmen“ wieder.

ErwG 139 führt aus, dass die Förderung der KI-Innovation durch KI-Reallabore auch durch die Beseitigung von Hindernissen für KMU, einschließlich Start-ups, erreicht werden und darüber hinaus darauf geachtet werden soll, dass die KI-Reallabore für KMU zugänglich sind. Dies wird in Art. 57 Abs. 9 e) KI-Verordnung normativ festgehalten, indem die Erleichterung und Beschleunigung des Zugangs von KI-Systemen zum Unionsmarkt, insbesondere wenn sie von KMU angeboten werden, ausdrücklich als Ziel genannt werden. Darüber hinaus enthält Art. 58 KI-Verordnung einige Regelungen, die KMU privilegieren. So müssen gem. Art. 58 Abs. 2 d) KI-Verordnung die Durchführungsakte für die detaillierten Regelungen zu KI-Reallaboren unter anderem gewährleisten, dass der Zugang zu diesen für KMU, einschließlich Start-up-Unternehmen, kostenlos ist. Ausgenommen hiervon sind lediglich außergewöhnliche Kosten, die die zuständigen nationalen Behörden in einer fairen und verhältnismäßigen Weise einfordern können. Zudem sollen die Verfahren für die Antragstellung, Beteiligung, Auswahl und den Ausstieg aus dem KI-Reallabor einfach und leicht verständlich gehalten werden, um die Beteiligung von KMU mit begrenzten rechtlichen und administrativen Ressourcen zu erleichtern, Art. 58 Abs. 2 g) KI-Verordnung. Weiter sind insbesondere KMU vor der Einrichtung gegebenenfalls an Dienste zu verweisen, die eine Anleitung zur Umsetzung der KI-Verordnung bereitstellen.

Schließlich normiert Art. 62 Abs. 1 a) KI-Verordnung, dass die Mitgliedsstaaten KMU soweit diese die noch durch Durchführungsrechtsakte zu konkretisierenden Voraussetzungen und Auswahlkriterien erfüllen vorrangigen Zugang zu den KI-Reallaboren zu gewähren haben. Dies schließt allerdings nicht grundsätzlich aus, dass auch Unternehmen, die nicht unter die KMU-Definition fallen, Zugang zu den KI-Reallaboren erhalten, soweit die jeweils zuständige Behörde aufgrund ihrer tatsächlichen und organisatorischen Möglichkeiten verbleibende Kapazitäten für Erprobungen im Reallabor vorweisen kann.

Zusammenfassung

Die KI-Verordnung definiert in den Art. 57 bis 63 zwar unter anderem die Ziele, die mit den KI-Reallaboren verfolgt werden, und verankert die Einrichtung der KI-Reallabore rechtlich, verlagert jedoch den exakten Ablauf und ihre Funktionsweise auf die noch zu erlassenden Durchführungsakte. Aus diesem Grunde kann bislang noch keine exakte Aussage darüber getroffen werden, welche Voraussetzungen für die Beteiligung an KI-Reallaboren für Unternehmen gelten, wie weitreichend die regulatorischen Erleichterungen im KI-Reallabor sind und ob insbesondere (partielle) Ausnahmen von den allgemeinen regulatorischen Anforderungen für die Entwicklungsarbeit in den KI-Reallaboren der KI-Verordnung bestehen werden. Denn die KI-Verordnung enthält bisher keine Ausnahmeregelungen des allgemeinen KI-Verordnungsregimes. Somit sind mit der Teilnahme an einem KI-Reallabor nicht zwingend geringere Anforderungen an die Einhaltung geltenden Rechts verbunden; die regulatorischen Erleichterungen beschränken sich vielmehr bislang auf einen engen Behördenkontakt und ein umfassendes Beratungsangebot.

Herausgeber
Bitkom e.V.
Albrechtstr. 10 | 10117 Berlin

Ansprechpartner/in
Lea Ossmann-Magiera | Wissenschaftliche Mitarbeiterin
T +49 30 27576-181 | l.ossmann@bitkom.org

Janis Hecker | Referent Künstliche Intelligenz
T +49 30 27576-239 | j.hecker@bitkom.org

Verantwortliches Bitkom-Gremium
AK AI

Autorinnen und Autoren
Sandra Baum (Bundesdruckerei GmbH), Dr. Frank Beer INFODAS Gesellschaft für Systementwicklung und Informationsverarbeitung mbH), Eric Behrendt (TÜV Informationstechnik GmbH), Susan Bischoff (Morrison & Foerster LLP), Arnd Böken (GvW Graf von Westphalen Rechtsanwälte Steuerberater Partnerschaft mbB), Jan Breuer (Detecon International GmbH), Camilla Dalerci (Bundesdruckerei GmbH), Vasilios Danos (TÜV Informationstechnik GmbH), Prof. Dr. Heinz-Uwe Dettling Ernst & Young GmbH Wirtschaftsprüfungsgesellschaft), Marius Drabiniok (SKW Schwarz Rechtsanwälte), Markus Frowein (RWE AG), Dr. Axel Grätz (Oppenhoff & Partner Rechtsanwälte Steuerberater mbB), Valentino Halim (Oppenhoff & Partner Rechtsanwälte Steuerberater mbB), Janis Hecker (Bitkom), Dr. Rachel Hegemann (Deutsche Bahn AG), Benedict Huyeng (RWE AG), Sven Jacobs (Cisco Systems GmbH), Ali-Reza Khalaji (R+V Versicherung AG), Stephan Kress (Morrison & Foerster LLP), Dr. Christoph Krück (SKW Schwarz Rechtsanwälte), Malte Lange (Finanz Informatik GmbH & Co. KG), Dr. Kim Lauenroth (Fachhochschule Dortmund), Dr. Anastasia Linnik (Retresco GmbH), Stefan Mangold (Datev eG), Martin Meyer (Siemens Healthcare GmbH), Dilan Mienert (GÖRG Partnerschaft von Rechtsanwälten mbB), Lea Ludmilla Ossmann-Magiera (LL.M. Leiden) (Bitkom), Hung Pham (INFODAS Gesellschaft für Systementwicklung und Informationsverarbeitung mbH), Filipp Revinzon (Vay Technology GmbH), Lys Riemenschneider (Holisticon AG), Dr. Benedikt Rohrßen (Taylor Wessing Partnerschaftsgesellschaft mbB), Tim Sauerhammer (Reed Smith LLP), Jan-Dierk Schaal, LL.M (University of Melbourne) (SKW Schwarz Rechtsanwälte), Michael Schemel (UnternehmerTUM GmbH), Alexander Schmalenberger (Taylor Wessing Partnerschaftsgesellschaft mbB), Ferdinand Schwarz (SKW Schwarz Rechtsanwälte), Maria Stammwitz (Bundesdruckerei GmbH), Christiane Stützle (Morrison & Foerster LLP), Frank Wisselink (Deutsche Telekom AG)