Blickpunkt Brüssel

EUROPE DIRECT
Duisburg-Niederrhein

KI-Regulierung für Europa

Der Verordnungsentwurf der Europäischen Kommission zur Festlegung harmonisierter Vorschriften für Künstliche Intelligenz vom 21.04.2021

Laura Katharina Pauli

Mai

2022

Inhaltsverzeichnis

I. Einführung

Künstliche Intelligenz (abgekürzt KI) hat inzwischen alle Lebensbereiche – darunter Mobilität, Handel und Gesundheit – erreicht. Die Europäische Kommission hat am 21.04.2021 den weltweit ersten Entwurf für einen Rechtsrahmen zur Regulierung von KI-Systemen vorgelegt.[1] Der Verordnungsentwurf (im Folgenden „KI-VO-E“) ist Teil der europäischen KI-Strategie und soll maßgeblich dazu beitragen, dass die EU bei der Entwicklung einer sicheren, vertrauenswürdigen und ethisch vertretbaren Künstlichen Intelligenz weltweit eine Führungsrolle einnimmt. So sollen nur solche KI-Systeme in den Verkehr gebracht und genutzt werden, die im Einklang mit den bestehenden Grundrechten und Werten der EU stehen. Auf diese Weise soll einerseits das Vertrauen der EU-Bürgerinnen und Bürger in den Einsatz von KI-Systemen gestärkt und andererseits eine werteorientierte Nutzung der Technik gewährleistet werden.[2]

II. Was ist KI?

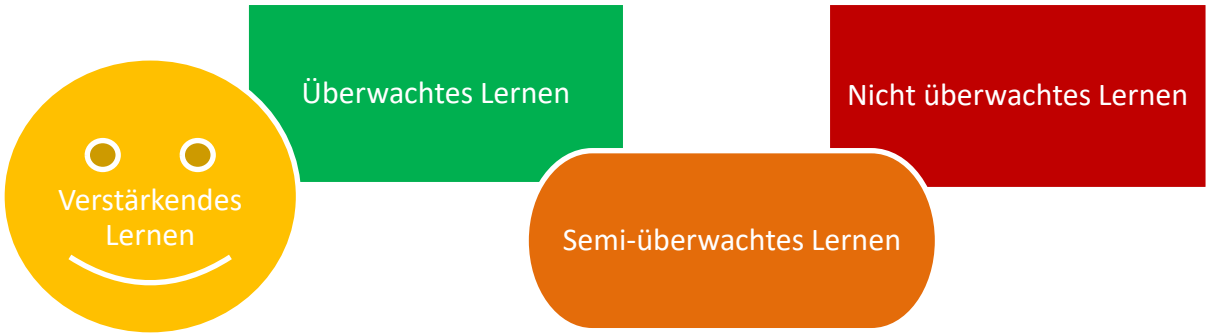
KI ist in aller Munde und doch wissen die wenigsten, was KI eigentlich sein soll. Der Begriff fiel erstmals im Jahr 1955.[3] Damals trafen sich renommierte Wissenschaftler am Darthmouth College im amerikanischen Hanover, New Hampshire, zu einem Workshop, bei dem sie über die Entwicklung von intelligenten Systemen diskutierten, die Probleme bewältigen sollten, deren Lösung dem Menschen besser gelingt. Dieses Ziel verfolgt auch heute noch die KI. Die KI soll in vielerlei Hinsicht die menschlichen Fähigkeiten nachahmen. Dies zeigt sich insbesondere in den KI-Bereichen des maschinellen Sehvermögens, der Spracherkennung, der Verarbeitung natürlicher Sprache sowie der Robotik.

Kennzeichen von KI-Systemen sind insbesondere ihre Lernfähigkeit und Eigenständigkeit. Auf Grundlage von enormen Datenmengen wird das KI-System trainiert, eine bestimmte Aufga-

[1] Der Verordnungsentwurf ist abrufbar unter: https://eur-lex.europa.eu/legal-content/DE/TXT/?uri=CELEX%3A52021PC0206 (zuletzt aufgerufen am 05.05.2022).
[2] *Valta/Vasel*, ZRP 2021, 142.
[3] *McCarthy/Minsky/Rochester/Shannon*, „A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence“, 1955, www-formal.stanford.edu/jmc/history/dartmouth/dartmouth.html (zuletzt aufgerufen am 23.05.2022); siehe ferner *Russell/Norvig*, Künstliche Intelligenz, 3. Aufl. 2012, S. 40 f.

be zu lösen. Durch die Feststellung von Zusammenhängen erkennt das KI-System nach und nach in den Daten ein gewisses Muster. Dieses Muster ermöglicht es dem KI-System regelmäßig, die Aufgabe zu lösen. Je aussagekräftiger die zugrunde liegenden Daten sind, desto zielgenauer kann sich das KI-System weiterentwickeln.

Einen besonderen Stellenwert in der KI-Forschung nimmt das *machine learning* (maschinelles Lernen) ein. Beim maschinellen Lernen werden Algorithmen durch die Anwendung statistischer Methoden trainiert, um Vorhersagen oder Klassifikationen zu machen.[4] Differenziert wird zwischen unterschiedlichen Lernverfahren:



Dem *überwachten Lernen* liegt ein gekennzeichneter Datensatz zugrunde. Dem Algorithmus wird durch die gekennzeichneten Daten das Verhältnis zwischen Input und gewünschtem Output vorgegeben. Auf diese Weise kann der Lernprozess kontrolliert werden.

Im Rahmen des *unüberwachten Lernens* ist der Datensatz demgegenüber nicht gekennzeichnet. Dieses Lernverfahren eignet sich, wenn unbekannte Zusammenhänge aufgedeckt werden sollen. Dem Algorithmus wird es selbst überlassen, ein Muster in den Daten zu entdecken.

Das *semi-überwachte Lernen* stellt – wie der Name schon sagt – einen Mittelweg zwischen den vorgenannten Lernverfahren dar. Hier wird ein kleinerer Datensatz mit Kennzeichnung verwendet, um einem größeren Datensatz ohne Kennzeichnung anzuleiten. Das in dem kleineren Datensatz erkennbare Muster lässt sich somit auf den größeren Datensatz übertragen.

[4] Siehe hierzu und zum Folgenden: https://www.ibm.com/de-de/cloud/learn/machine-learning#toc-methoden-d-_2x9Z3-t (zuletzt aufgerufen am 23.05.2022).

Das *verstärkende Lernen* zeichnet sich dadurch aus, dass der der Algorithmus mithilfe eines positiven oder negativen Feedbacks trainiert wird. Das Modell lernt mithilfe von Versuch und Irrtum. Wenn mehrere Ergebnisse nacheinander richtig sind, erhält das Modell einen positiven Impuls. Insoweit kann auf das gewünschte Lernergebnis Einfluss genommen werden.

Eine besondere Methode des machine learning stellt das *deep learning* (mehrschichtiges bzw. tiefgehendes Lernen) dar. Beim tiefgehenden Lernen werden mehrschichtige *künstliche neuronale Netze* eingesetzt. Diese basieren auf Neuronen und Verknüpfungen, die durch die Überschreitung bestimmter Schwellenwerte aktiviert werden. Mithilfe von gewichteten Synapsen werden die Neuronen zu einem Netzwerk verbunden. Die Neuronen sind dabei über Schichten miteinander verknüpft. Im Gegensatz zum klassischen machine learning ist das deep learning nicht auf die menschliche Datenaufbereitung angewiesen. Vielmehr kann das deep learning dort zum Tragen kommen, wo in einem großen unbearbeiteten Datensatz unbekannte Strukturen aufgedeckt werden sollen.

Welches Lernverfahren und welche Methodik ein Programmierer auswählt, um ein KI-System zu entwickeln, ist demnach davon abhängig, welche Daten ihm zur Verfügung stehen und wie offen die für das System zu bewältigende Aufgabenstellung gefasst ist.

III. Notwendigkeit einer Regulierung

Die Notwendigkeit, das Inverkehrbringen und die Nutzung von KI-Systemen bereits zum jetzigen Zeitpunkt zu reglementieren, ergibt sich aus den KI-spezifischen Risiken sowie den damit einhergehenden grundrechtlich relevanten Gefährdungslagen.

1. KI-spezifische Risiken

Durch ihr eigenständiges Lernverhalten bergen KI-Systeme auch besondere Risiken in sich, die es regulatorisch einzudämmen gilt. So besteht insbesondere die Gefahr der Diskriminierung sowie der fehlenden menschlichen Kontrollmöglichkeit.

Aus dem Umstand, dass KI-Systeme auf Grundlage von Trainingsdaten versuchen, Korrelationen zu entdecken und Individuen lediglich auf einer Metaebene betrachtet werden,[5] folgt zwangsläufig ein Diskriminierungsrisiko. Denn der Algorithmus kann in den Trainingsdaten auch Zusammenhänge zwischen einem zulässigen und einem diskriminierungsträchtigen Merkmal erkennen, die der Programmierer nicht antizipieren konnte. In den USA wurde bspw. Künstliche Intelligenz – namentlich die Software COMPAS – zur Ermittlung der Rückfallanfälligkeit von verurteilten Straftätern eingesetzt.[6] Die Ergebnisse der Software wurden bei der Urteilsfindung, der Ermittlung des Strafmaßes sowie der Bescheidung eines Antrags auf vorzeitige Haftentlassung zugrunde gelegt. Aus der statistischen Analyse der durch COMPAS entschiedenen Fälle ergab sich, dass Afroamerikaner im Vergleich zu Weißen diskriminiert wurden, da sie von dem System eher als kriminell eingestuft wurden. Grund hierfür war das in den zugrunde liegenden Trainingsdaten angelegte Diskriminierungspotenzial („*bias*“).

Darüber hinaus besteht die Gefahr, dass die Entscheidung durch ein KI-System nicht oder nur unzureichend von einem Menschen überprüft und nachvollzogen werden kann. Entscheidungen durch KI-Systeme sind aufgrund ihres eigenständigen Lernvorgangs aus ex-ante-Sicht nur schwer vorhersehbar. Dies hat zur Folge, dass die möglicherweise tangierten Grundrechte sowie die Wahrscheinlichkeit eines Schadenseintritts vorab nur schwer vorausgesagt werden können. Aus ex-post-Perspektive stellt sich spiegelbildlich hierzu das Problem, dass die Entscheidung nur schwerlich rekonstruiert werden kann, da das innere Prozessieren des KI-Systems regelmäßig im Dunkeln bleibt (sog. *black box*[7]). Selbst ein Programmierer kann häufig nur im Nachhinein erläutern, dass es zu einer bestimmten Entscheidung gekommen ist, nicht jedoch aus welchen Beweggründen.[8] Dieses Transparenzdefizit auf ex-post-Ebene führt dazu, dass das Risiko besteht, die Entscheidung durch das KI-System nicht überprüfen zu können. Dieses Risiko ist besonders hoch im Falle des Einsatzes eines mehrschichtigen neu-

[5] *Ernst*, JZ 2017, 1026, 1028; *Martini*, JZ 2017, 1017, 1018.
[6] Siehe hierzu und zum Folgenden: https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing (zuletzt aufgerufen am 09.05.2022).
[7] *Herberger*, NJW 2018, 2825, 2828.
[8] *Greco*, RW 2020, 29, 42; *Pfeil*, InTeR 2020, 17, 20.

ronalen Netzwerks. Aufgrund seines komplexen Entscheidungsmodells gilt dessen Entscheidung als nicht oder kaum mehr rekonstruierbar.[9] Um das Vertrauen von Menschen in den Einsatz von KI-Systemen zu gewinnen und zu stärken, muss das intransparente Entscheidungsverhalten transparenter gemacht werden. Diesem Anliegen widmet sich die sog. *explainable AI*[10], die versucht, die Entscheidung eines KI-Systems für einen Menschen erklärbar zu machen.

2. Grundrechtsrelevanz

Die Verordnungsentwurf muss einerseits den unternehmerischen Freiheiten Rechnung tragen und andererseits die mit dem KI-Einsatz verbundenen Gefahren für die Grundrechte Dritter und die öffentliche Sicherheit auf ein angemessenes Maß reduzieren. Insoweit dient die Regulierung einem Interessenausgleich.

Nicht nur die Verbotstatbestände, sondern auch die diversen zwingenden Vorgaben greifen in die grundrechtlich geschützte unternehmerische Freiheit ein (Art. 16 EU-GrCh). Aus unternehmerischer Sicht wird durch die Vielzahl an zwingenden Vorgaben der Innovationseifer eingeengt. So muss sich der Unternehmer bereits bei Entwicklung seines KI-Systems die Frage stellen, ob dieses den Anforderungen genügt und ein Inverkehrbringen seines KI-Produktes im europäischen Markt möglich sein wird. Die KI-Verordnung führt mithin dazu, dass von Unternehmern, die in der EU ihr KI-Produkt vertreiben wollen, von vornherein solche KI-Techniken nicht entwickelt werden, die den Verbotstatbeständen unterfallen. Die EU beabsichtigt jedoch genau diesen Steuerungseffekt, damit nur solche KI-Systeme entwickelt werden, die im Einklang mit den europäischen Werten stehen und damit den Anforderungen an eine vertrauenswürdige KI genügen. Aufgrund ihrer besonderen Risiken (z. B. Intransparenz, Komplexität, Datenabhängigkeit, autonomes Verhalten) kann die Verwendung von KI dazu führen, dass einige der in der EU-Grundrechtecharta verankerten Grundrechte verletzt werden.[11] Ziel der Verordnung ist es, diese Grundrechte – insbesondere die Würde des Men-

[9] *Zech*, ZfPW 2019, 198, 217.

[10] *Deeks*, Columbia Law Review 119 (2019), 1829 ff.; *Waltl/Vogl*, DuD 42 (2018), 613 ff.; siehe ferner *Beuth*, Die rätselhafte Gedankenwelt eines Computers, abrufbar unter: https://www.zeit.de/digital/internet/2017-03/kuenstliche-intelligenz-black-box-transparenz-fraunhofer-hhi-darpa (zuletzt aufgerufen am 06.05.2022).

[11] Die folgenden Ausführungen beruhen auf 3.5. der Vorbemerkungen zum Verordnungsentwurf, abrufbar

schen (Art. 1 EU-GrCh), die Achtung des Privatlebens und der Schutz personenbezogener Daten (Artt. 7 und 8 EU-GrCh), die Nichtdiskriminierung (Art. 21 EU-GrCh), die Gleichheit von Frauen und Männern (Art. 23 EU-GrCh) sowie Meinungs-, Versammlungs- und Vereinigungsfreiheit (Artt. 11 und 12 EU-GrCh) – zu schützen. Sollte es trotzdem zu einer Grundrechtsverletzung kommen, erhalten die betroffenen Personen die Möglichkeit, Rechtsmittel einzulegen, da für Transparenz und Rückverfolgbarkeit der KI-Systeme im Verbund mit starken Kontrollen auf ex-post-Ebene gesorgt ist.

IV. Darf die EU das regeln?

Durch die einheitlichen Regelungen für die Inbetriebnahme, das Inverkehrbringen und die Verwendung von KI-Systemen möchte die EU einen rechtssicheren Rahmen schaffen, der in der gesamten EU zwingende Geltung entfaltet. Wie auch die Datenschutzgrundverordnung (DS-GVO) soll die KI-Verordnung bei deren Inkrafttreten unmittelbar in jedem Mitgliedstaat gelten (Art. 288 Abs. 2 AEUV), d.h., dass den Mitgliedstaaten keinerlei Spielraum bei der Umsetzung verbleibt. Dabei stellt sich die Frage, ob die EU überhaupt zu solch weitreichenden Regelungen berechtigt ist. Nach dem Prinzip der begrenzten Einzelermächtigung (Art. 5 Abs. 1 S. 1, Abs. 2 EUV) darf die EU nur in den Bereichen tätig werden, die die Mitgliedstaaten ihr übertragen haben. Der Vorschlag wird insbesondere auf die allgemeine Binnenmarktkompetenz des Art. 114 Abs. 1 AEUV gestützt. Danach sind das Europäische Parlament und der Rat berechtigt, Maßnahmen zur Rechtsangleichung zu erlassen, um den Binnenmarkt zu verwirklichen. Problematisch ist insoweit, dass bisher gar keine unterschiedlichen nationalen Regelungen zum Inverkehrbringen und Einsatz von KI-Systemen existieren, die eine Rechtsangleichung gebieten könnten. Vielmehr wird die EU proaktiv tätig und will nationalen Regelungen vorgreifen. Der Verordnungsentwurf steht daher auf recht wackeligen Füßen.

unter: https://eur-lex.europa.eu/legal-content/DE/TXT/?uri=CELEX%3A52021PC0206 (zuletzt aufgerufen am 05.05.2022).

V. Anwendungsbereich der Verordnung

Der persönliche und sachliche Anwendungsbereich der Verordnung geht sehr weit.

Adressaten der Verordnung sind nicht nur private, sondern auch staatliche Akteure innerhalb und außerhalb der EU. Entscheidend ist allein, dass das KI-System in der Union in den Verkehr gebracht wird oder Menschen in der EU von seiner Verwendung betroffen sind (vgl. Art. 2 KI-VO-E). Der Verordnungsentwurf adressiert insoweit nicht nur die Anbieter, sondern auch die Nutzer eines KI-Systems.

Auch der sachliche Anwendungsbereich wird in Art. 3 Nr. 1 KI-VO-E sehr weit gefasst. Da KI eine Vielzahl von Modellen und Techniken erfasst, ist es schwierig, den Begriff abschließend zu definieren. Die EU-Kommission hat insoweit versucht, eine Definition zu entwickeln, die einerseits hinreichend bestimmt und andererseits offen für zukünftige Entwicklungen ist. Ein KI-System im Sinne der Verordnung ist „*eine Software, die mit einer oder mehreren der in Anhang I aufgeführten Techniken und Konzepte entwickelt worden ist und im Hinblick auf eine Reihe von Zielen, die vom Menschen festgelegt werden, Ergebnisse wie Inhalte, Vorhersagen, Empfehlungen oder Entscheidungen hervorbringen kann, die das Umfeld beeinflussen, mit dem sie interagieren*“. Die im Anhang I niedergelegten Techniken umfassen ein breites Spektrum. So zählen hierunter

- machine learning einschließlich deep learning,
- wissensbasierte Ansätze inklusive Expertensysteme,
- statistische Ansätze sowie
- Such- und Optimierungsmethoden.

Die Liste ist dabei bewusst nicht statischer Natur. Vielmehr ist die Europäische Kommission nach Artt. 4, 73 KI-VO-E ermächtigt, Aktualisierungen zu veranlassen und neue Techniken darin aufzunehmen. Dies trägt dem Umstand Rechnung, dass sich die Entwicklung von KI-Systemen dynamisch fortentwickeln wird und dabei auch die zugrunde liegenden Techniken neue, zum jetzigen Zeitpunkt noch nicht absehbare Formen annehmen können. Überraschend ist an dem Anwendungsbereich, dass sogar deterministische Systeme wie Experten-

systeme erfasst werden, obwohl von diesen die KI-spezifischen Risiken wie die Unvorhersehbarkeit und Intransparenz des Entscheidungsverhaltens gerade nicht ausgehen.[12]

VI. Risikobasierter Regulierungsansatz

Die KI-Verordnung verfolgt einen risikobasierten Ansatz.[13] Hiernach sind die regulatorischen Anforderungen an das KI-System umso höher, je größer die zu erwartenden Gefahren im jeweiligen Einsatzgebiet sind.[14] Vor diesem Hintergrund differenziert der Verordnungsentwurf zwischen drei verschiedenen Gruppen von KI-Systemen, den

- verbotenen KI-Systemen,
- Hochrisiko-KI-Systemen sowie
- den übrigen KI-Systemen,

die im Folgenden näher beleuchtet werden sollen.

1. Verbotene KI-Systeme

Die erste Gruppe von KI-Systemen umfasst solche, deren Einsatz, Inbetriebnahme und Inverkehrbringen per se verboten wird. Untersagt wird hingegen nicht, solche Systeme zu entwickeln.

Zu den verbotenen KI-Systemen zählen nach Art. 5 Abs. 1 lit. a KI-VO-E Techniken der unterschwelligen Beeinflussung außerhalb des Bewusstseins einer Person, die darauf abzielen, das Verhalten einer Person in einer Weise wesentlich zu beeinflussen, die dieser Person oder einer anderen Person einen physischen oder psychischen Schaden zufügt oder zufügen kann.

In Art. 5 Abs. 1 lit. b KI-VO-E werden ergänzend Systeme aufgeführt, die die Schwächen einer Person aufgrund ihres Alters oder einer Behinderung ausnutzen.

Art. 5 Abs. 1 lit. c KI-VO-E untersagt KI-Systeme, die zur Beurteilung der Vertrauenswürdigkeit von Personen im Zusammenhang mit sozialer Zugehörigkeit oder gesellschaftlichem

[12] So auch *Spindler*, CR 2021, 361, 363.
[13] *Spindler*, CR 2021, 361, 362.
[14] https://www.fieldfisher.com/de-de/insights/die-eu-ki-verordnung-teil-1 (zuletzt aufgerufen am 23.05.2022).

Verhalten genutzt werden können, wie man sie teilweise aus China kennt (social scoring). Zu beachten ist hierbei, dass sich das Verbot nur an staatliche Stellen richtet.

In Art. 5 Abs. 1 lit. d KI-VO-E wird das Verbot statuiert, Echtzeit-Fernidentifizierungssystemen zur biometrischen Identifizierung in öffentlich zugänglichen Räumen zu nutzen. Diese können jedoch im Einzelfall eingesetzt werden, sofern ein deutlich überwiegendes öffentliches Interesse den Einsatz im Hinblick auf näher dargelegte Fallgruppen gebietet (z. B. bei der gezielten Suche nach bestimmten potenziellen Opfern von Straftaten oder nach vermissten Kindern).

2. Hochrisikosysteme

Der Schwerpunkt der Verordnung bildet die zweite Gruppe. In Art. 6 KI-VO-E werden sog. Hochrisiko-KI-Systeme aufgeführt, deren Einsatz nur erlaubt wird, wenn sie erhöhte Anforderungen erfüllen.

a. Definition

Die Einordnung als Hochrisiko-KI-System basiert auf der Zweckbestimmung des KI-Systems entsprechend den bestehenden EU-Produktsicherheitsvorschriften.[15] Damit hängt die Einstufung als Hochrisiko-KI-System nicht nur von der Funktion dieses Systems ab, sondern auch von seinem konkreten Zweck und seinen Anwendungsmodalitäten.[16]

Zu den Hochrisiko-KI-Systemen zählen zum einen solche Systeme, die als Sicherheitskomponenten von Produkten oder als eigenständiges Produkt einer Vorab-Konformitätsbewertung durch Dritte unterliegen. Zum anderen fallen hierunter KI-Systeme, die ausdrücklich in Anhang III genannt werden und sich vor allem auf die Grundrechte auswirken. Hierunter fallen folgende Bereiche:[17]

[15] Siehe 5.2.3. der Vorbemerkungen zum Verordnungsentwurf, abrufbar unter: https://eur-lex.europa.eu/legal-content/DE/TXT/?uri=CELEX%3A52021PC0206 (zuletzt aufgerufen am 05.05.2022).
[16] Siehe 5.2.3. der Vorbemerkungen zum Verordnungsentwurf, abrufbar unter: https://eur-lex.europa.eu/legal-content/DE/TXT/?uri=CELEX%3A52021PC0206 (zuletzt aufgerufen am 05.05.2022).
[17] Die nachfolgende Aufzählung wurde übernommen von: https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/excellence-trust-artificial-intelligence_de (zuletzt aufgerufen am 23.05.2022).

- Kritische Infrastrukturen (z. B. im Verkehr), in denen das Leben und die Gesundheit der Bürger gefährdet werden könnten
- Schul- oder Berufsausbildung, wenn der Zugang einer Person zur Bildung und zum Berufsleben beeinträchtigt werden könnte (z. B. Bewertung von Prüfungen)
- Sicherheitskomponenten von Produkten (z. B. eine KI-Anwendung für die roboterassistierte Chirurgie)
- Beschäftigung, Personalmanagement und Zugang zu selbstständiger Tätigkeit (z. B. Software zur Auswertung von Lebensläufen für Einstellungsverfahren)
- Zentrale private und öffentliche Dienstleistungen (z. B. Bewertung der Kreditwürdigkeit, wodurch Bürgern Darlehen verwehrt werden)
- Strafverfolgung, die in die Grundrechte der Menschen eingreifen könnte (z. B. Überprüfung der Echtheit von Beweismitteln)
- Migration, Asyl und Grenzkontrolle (z. B. Überprüfung der Echtheit von Reisedokumenten)
- Rechtspflege und demokratische Prozesse (z. B. Anwendung der Rechtsvorschriften auf konkrete Sachverhalte)

Die Europäische Kommission ist nach Artt. 7, 73 KI-VO-E berechtigt, diese Liste auszuweiten, sofern das KI-System unter die bereits bestehenden Bereiche fällt und ein vergleichbares Risiko begründet.

b. Pflichtenprogramm

Für den Einsatz von Hochrisiko-KI-Systemen gelten besondere Pflichten wie Risiko- und Qualitätsmanagementsysteme sowie Transparenz- und Publizitätspflichten (Artt. 8-15 KI-VO-E). So werden konkrete Anforderungen zur Minimierung des Risikos der Diskriminierung durch Algorithmen, insbesondere im Hinblick auf die Qualität der verwendeten Datensätzen, etabliert. Ferner werden Tests, Risikomanagement, Dokumentation und menschliche Aufsicht über die gesamte Lebensdauer von KI-Systemen hinweg verbindlich vorgeschrieben.

Zunächst verlangt Art. 9 KI-VO-E die Implementierung eines Risikomanagementsystems. Das Risikomanagementsystem soll die bekannten und vorhersehbaren Risiken des KI-Systems ermitteln, diese bewerten und bei Bedarf Maßnahmen der Risikominimierung ergreifen. Diese Vorschrift bildet ein zentrales Element der systeminternen Kontrolle und ermöglicht eine fortlaufende effektive Risikoüberwachung.

Die in Art. 10 KI-VO-E verankerten Anforderungen an den zugrunde liegenden Datensatz sollen der Gefahr der Diskriminierung begegnen. So müssen die Trainings-, Validierungs- und Testdatensätze nicht nur relevant, repräsentativ und vollständig, sondern auch fehlerfrei sein.

Um eine effektive Nachkontrolle der Entscheidung des KI-Systems zu gewährleisten, sehen die Art. 12 und Art. 13 KI-VO-E diverse Aufzeichnungs- und Transparenzpflichten vor. Die Vorschriften fördern insoweit den vermehrten Einsatz von sog. *explainable AI*.

Schließlich wird in Art. 14 KI-VO-E das Erfordernis der menschlichen Beaufsichtigung des Systems statuiert. Danach soll eine natürliche Person während der Dauer der Verwendung des KI-Systems deren wirksame Beaufsichtigung übernehmen, damit Dysfunktionalitäten und Anomalien frühzeitig erkannt und behoben werden müssen. Die beaufsichtigende Person soll hierfür in der Lage sein, die Fähigkeiten und Grenzen des Hochrisiko-KI-Systems vollständig zu verstehen (Art. 14 Abs. 4 lit. a KI-VO-E).

Art. 15 KI-VO-E verlangt überdies, dass das Hochrisiko-KI-System ein angemessenes Maß an Genauigkeit, Robustheit und Cybersicherheit an den Tag legen muss.

Flankiert werden die spezifischen Anforderungen an Hochrisiko-KI-Systeme durch konkret normierte Pflichten der Anbieter, Nutzer und anderer Beteiligter ebensolcher Systeme (Artt. 16-29 KI-VO-E).

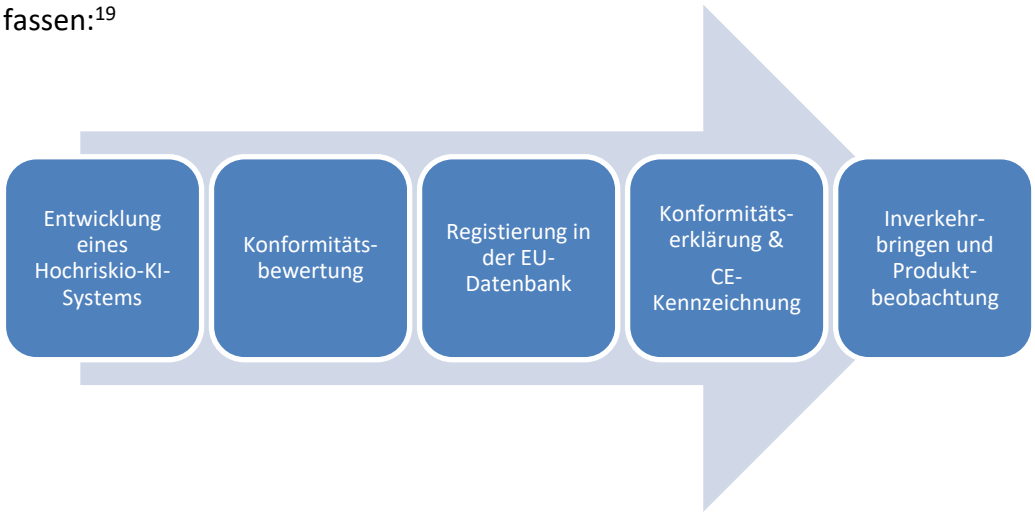
Ein weiterer Baustein zur Gewährleistung der Sicherheit beim Einsatz von Hochrisiko-KI-Systemen bildet die aus Art. 51 KI-VO-E folgende Registrierungspflicht, wonach diese in einer öffentlich zugänglichen EU-Datenbank (Art. 60 KI-VO-E) eingepflegt werden müssen.

c. Akteurspezifische Konkretisierung des Pflichtenprogramms

Die Artt. 16 ff. KI-VO-E legen die Verhaltenspflichten für die verschiedenen Akteure – wie Anbieter, Einführer, Händler und Nutzer – fest, die im Zusammenhang mit dem Einsatz eines Hochrisiko-KI-Systems beachtet werden müssen. Sinn und Zweck dieser Vorschriften ist es, dass die zuvor normierten Anforderungen durch die Akteure umgesetzt werden.

aa. Anbieter

Nach Art. 16 KI-VO-E müssen die Anbieter (vgl. Art. 3 Nr. 2 KI-VO-E) eines Hochrisiko-KI-Systems nicht nur dafür Sorge tragen, dass die zuvor genannten Anforderungen erfüllt werden, sondern auch sicherstellen, dass das System – bevor es in den Verkehr gebracht oder in Betrieb genommen wird – das betreffende Konformitätsbewertungsverfahren durchlaufen hat (Art. 16 lit. e i.V.m. Art. 19 i.V.m. Art. 43 KI-VO-E) und eine CE-Kennzeichnung angebracht worden ist (Art. 16 lit. i i.V.m. Art. 49 KI-VO-E).[18] Darüber hinaus sind Hochrisiko-KI-Systeme i.S.d. Art. 6 Abs. 2 KI-VO-E vor deren Inverkehrbringen oder Inbetriebnahme in eine von der Kommission verwaltete EU-Datenbank einzutragen. Letzteres dient der Gewährleistung der Aufsicht und Überwachung durch die zuständigen Behörden. Schließlich sind Anbieter dazu verpflichtet, Hochrisiko-KI-Systeme nach deren Inverkehrbringen zu beobachten und etwaige schwerwiegende Vorfälle und Fehlfunktionen den Marktüberwachungsbehörden anzuzeigen (Artt. 61-62 KI-VO-E). Die den Anbieter treffenden Pflichten lassen wie folgt zusammenfassen:[19]



[18] Zur Vertiefung *Roos/Weitz*, MMR 2021, 844, 846 ff.
[19] Die Grafik ist angelehnt an https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/excellence-trust-artificial-intelligence_de (zuletzt aufgerufen am 23.05.2022).

bb. Einführer

Die formelle Prüfpflicht des Einführers (vgl. Art. 3 Nr. 7 KI-VO-E) eines Hochrisiko-KI-Systems ist eng mit dem Pflichtenkreis des Anbieters verwoben.[20] Einführer haben zu überprüfen, ob der jeweilige Anbieter seinen Pflichten nachgekommen ist, bevor sie das Hochrisiko-KI-System in den Verkehr bringen (Art. 26 Abs. 1 KI-VO-E). Sind sie der Auffassung oder haben Grund zur Annahme, dass das KI-System nicht den zwingenden Vorgaben entspricht, sind sie verpflichtet, die Konformität vor dem Inverkehrbringen herzustellen (Art. 26 Abs. 2 S. 1 KI-VO-E). Sofern dem KI-System ein Risiko i.S.d. Art. 65 Abs. 1 KI-VO-E immanent ist, hat der Einführer sowohl den Anbieter als auch die Marktüberwachungsbehörden hiervon in Kenntnis zu setzen (Art. 26 Abs. 2 S. 2 KI-VO-E). Auf diese Weise wird eine effektive Kontrolle auf verschiedenen Akteursebenen sowie deren Interaktion untereinander gefördert. Zuletzt treffen den Einführer Pflichten bei der Verpackung, Lagerung und dem Transport (Art. 26 Abs. 3-4 KI-VO-E).

cc. Händler

Die Pflichten des Händlers ähneln denjenigen des Einführers. Auch den Händler (vgl. Art. 3 Nr. 8 KI-VO-E) eines Hochrisiko-KI-Systems trifft zunächst eine Prüfpflicht, bevor er das System auf dem Markt bereitstellt (Art. 27 Abs. 1 KI-VO-E). Dabei hat der Händler die Einhaltung sämtlicher Pflichten der Anbieter und Einführer zu kontrollieren. Anders als der Einführer hat der Händler nicht nur die formellen Voraussetzungen zu überprüfen, sondern auch die Einhaltung der materiellen Anforderungen. Insoweit setzt der Verordnungsentwurf ein hohes Maß an Fachkunde des Händlers voraus.[21]

Art. 27 Abs. 2 KI-VO-E ist inhaltlich Art. 26 Abs. 2 KI-VO-E nachgebildet. Im Übrigen treffen den Händler auch weitgehende Pflichten nach Bereitstellung des KI-Systems. Sofern sich erst nach Bereitstellung des KI-Systems auf dem Markt herausstellt, dass dieses nicht den zwingenden Anforderungen entspricht, hat der Händler umgehend die erforderlichen Korrekturmaßnahmen zu ergreifen (Art. 27 Abs. 4 S. 1 KI-VO-E). Dies kann zur Folge haben, dass er

[20] Hierzu und zum Folgenden siehe *Roos/Weitz*, MMR 2021, 844, 849.
[21] *Roos/Weitz*, MMR 2021, 844, 849.

verpflichtet ist, eine umfassende Rückrufaktion unter Einbeziehung der bis dahin involvierten Akteure einzuleiten.

dd. Nutzer

Um eine sichere Nutzung des Hochrisiko-KI-Systems zu gewährleisten, hat auch der Nutzer (vgl. Art. 2 Nr. 4 KI-VO-E) diverse Pflichten zu erfüllen (Art. 29 KI-VO-E). Die Nutzer haben hiernach die Vorgaben entsprechend der beigefügten Gebrauchsanweisung einzuhalten und den Betrieb des Systems zu überwachen. Sofern sich ein Risiko i.S.d. Art. 65 Abs. 1 KI-VO-E abzeichnet, sind sie verpflichtet, den Betrieb umgehend einzustellen und den Anbieter oder Händler hierüber zu informieren. Dies gilt auch im Falle der Offenbarung eines schwerwiegenden Vorfalls oder einer Fehlfunktion. Überdies sind die Nutzer zur Aufbewahrung der von dem KI-System erzeugten Protokolle verpflichtet. Durch die Einbeziehung des Nutzers in den Kreis der handlungspflichtigen Personen wird somit eine umfassende Kontrolle des Einsatzes auf allen Akteursebenen gewährleistet.

3. Die übrigen KI-Systeme

KI-Systeme, die im Sinne der Verordnung kein hohes Risiko darstellen, unterliegen allgemeinen Regelungen wie insbesondere der Richtlinie 2001/95/EG über die allgemeine Produktsicherheit.[22] Nichtsdestotrotz sollen Anbieter dazu angehalten werden, Verhaltenskodizes zu erstellen, um eine freiwillige Unterwerfung unter die für Hochrisiko-KI-Systeme vorgegebenen Anforderungen zu fördern (Art. 69 KI-VO-E).[23]

VII. Informationspflichten für bestimmte KI-Systeme

Art. 52 KI-VO-E sieht Informations- und Kennzeichnungspflichten für bestimmte KI-Systeme vor. Erfasst werden solche KI-Systeme, die besondere Risiken der Manipulation aufweisen, sei es durch Interaktion mit Menschen (Abs. 1), oder durch Gefühlserkennungssysteme oder biometrische Kategorisierung (Abs. 2), oder schließlich durch die Erzeugung oder Manipula-

[22] *Günther-Burmeister*, DB 2021, 1858, 1862.
[23] *Günther-Burmeister*, DB 2021, 1858, 1862.

tion von Inhalten, insbesondere Bildern oder Videos, die real existierenden Personen oder Begebenheiten ähneln (sog. „deep fakes“, Abs. 3).[24]

VIII. Haftungsrechtliche Dimension

Der Verordnungsentwurf selbst beinhaltet zwar keine ausdrücklichen Haftungsregeln, weil diese einem separaten Gesetzesentwurf[25] vorbehalten bleiben. Gleichwohl entfaltet er eine haftungsrechtliche Dimension im Bereich des nationalen Vertrags- und Deliktsrechts.[26] Wenngleich die Verpflichtungen der Akteure nach dem Verordnungsentwurf dem öffentlichen Recht angehören, können diese in das Privatrecht hineinwirken.[27] Da die Verordnung bei deren Inkrafttreten unmittelbar in jedem Mitgliedstaat gilt (Art. 288 Abs. 2 AEUV) und die Verhaltenspflichten im Umgang mit KI-Systemen regelmäßig nicht nur die Allgemeinheit, sondern auch den Einzelnen schützen sollen, handelt es sich bei der Mehrzahl der in der KI-VO-E niedergelegten Pflichten um Schutzgesetze i.S.d. § 823 Abs. 2 BGB.[28] Vor diesem Hintergrund lösen schuldhafte Pflichtverstöße – insbesondere gegen Artt. 18 ff. KI-VO-E – eine Schadensersatzverpflichtung des verantwortlichen Akteurs aus. Gleichzeitig können in den Verhaltenspflichten auch Verkehrssicherungspflichten i.S.d. § 823 Abs. 1 BGB erblickt werden, sodass auch nach dieser Vorschrift eine Schadensersatzhaftung begründet werden kann. Letztlich konkretisieren die in der Verordnung niedergelegten Pflichten auch den vertraglichen Bereich, da sich das vertragliche Pflichtenprogramm u.a. an Branchenstandards orientiert. Der separate Regelungsentwurf zur zivilrechtlichen Haftung beim Einsatz von KI-Systemen sollte daher die hier bereits vorliegende haftungsrechtliche Dimension berücksichtigen.

[24] *Spindler*, CR 2021, 361, 368.
[25] Siehe dazu den Entwurf einer Entschließung des Europäischen Parlaments mit Empfehlungen an die Kommission für eine Regelung der zivilrechtlichen Haftung beim Einsatz künstlicher Intelligenz (2020/2014(INL)) vom 20.10.2020, abrufbar unter: https://www.europarl.europa.eu/doceo/document/TA-9-2020-0276_DE.html#title1 (zuletzt aufgerufen am 06.05.2022).
[26] Zum Folgenden siehe *Linardatos*, GPR 2022, 58, 60; *Grützmacher*, CR 2021, 433 ff.
[27] *Roos/Weitz*, MMR 2021, 844, 849 f.
[28] *Roos/Weitz*, MMR 2021, 844, 850.

IX. Experimentierfelder

Der Verordnungsentwurf sieht ferner Maßnahmen zur Innovationsförderung vor (Artt. 53-54 KI-VO-E). Danach werden KI-Reallabore eingerichtet, damit eine kontrollierte Umgebung geschaffen wird, um die Entwicklung, Erprobung und Validierung innovativer KI-Systeme für einen begrenzten Zeitraum vor ihrem Inverkehrbringen oder ihrer Inbetriebnahme nach einem spezifischen Plan zu erleichtern. Dies wird von den zuständigen Behörden beaufsichtigt.

X. Aufsicht und Überwachung

Ein besonderes Augenmerk legt der Vorschlag der Kommission auch auf die Aufsicht und Überwachung der Einhaltung der vorgenannten Vorschriften.

Nach Art. 30 Abs. 1 KI-VO-E hat jeder Mitgliedstaat dafür zu sorgen, dass eine notifizierende Behörde geschaffen wird, die für die Einrichtung und Durchführung der erforderlichen Verfahren zur Bewertung, Benennung und Notifizierung von Konformitätsbewertungsstellen und für deren Überwachung zuständig ist. Dabei muss die Behörde so ausgerichtet sein, dass die Objektivität und Unparteilichkeit ihrer Tätigkeiten gewährleistet sind.

Artt. 56-59 KI-VO-E sehen weiterhin vor, dass ein „Europäischer Ausschuss für künstliche Intelligenz“ eingerichtet wird, der insbesondere die einheitliche Anwendung dieser Verordnung durch die nationalen Aufsichtsbehörden überwachen soll. Dem Ausschuss kommt insoweit eine beratende und koordinierende Funktion zu.

Schließlich ordnet Art. 63 KI-VO-E an, dass die europäische Marktüberwachungsverordnung entsprechende Anwendung findet. Hierdurch soll gewährleistet werden, dass auch nach Inverkehrbringen von Hochrisiko-KI-Systemen die in der Verordnung niedergelegten Vorgaben eingehalten werden.

XI. Sanktionen

Abgerundet wird der Verordnungsentwurf durch harte Sanktionen in Form von hohen Bußgeldern (Art. 71 KI-VO-E). So drohen bei Verstößen gegen Art. 5 und Art. 10 KI-VO-E (grober Verstoß) Bußgelder bis zu 30 Mio. Euro oder 6 % des gesamten weltweiten Jahresumsatzes,

je nachdem, welcher Betrag höher ist (Art. 71 Abs. 3 KI-VO-E). Bei einem mittelschweren Verstoß i.S.d. Art. 71 Abs. 4 KI-VO-E drohen Bußgelder bis zu 20 Mio. Euro oder 4 % des gesamten weltweiten Jahresumsatzes. Bei kleineren Verstößen i.S.d. Art 71 Abs. 5 KI-VO-E können Bußgelder bis zu 10 Mio. Euro oder 2 % des gesamten weltweiten Jahresumsatzes verhängt werden. Die eklatante Höhe der zu verhängenden Bußgelder soll in effektiver Weise die Anbieter und anderen Akteure dazu anhalten, die Vorschriften der Verordnung einzuhalten.

XII. Kritik

Der Verordnungsentwurf schafft einen effektiven Sicherungsmechanismus im Umgang mit risikoträchtigen KI-Systemen. So werden auf allen Akteursebenen weitreichende Handlungspflichten begründet, die über den gesamten Lebenszyklus des KI-Systems andauern. Durch die gegenseitige Kontrolle der Akteure können besondere Risiken und Fehlfunktionen des Systems zügig aufgedeckt und behoben werden. Gleichwohl müssen diese Handlungspflichten im Lichte der hohen Bußgelder dem Verhältnismäßigkeitsgrundsatz sowie Bestimmtheitsgrundsatz genügen. Insoweit besteht derzeit noch Handlungsbedarf.

Einerseits ist bereits fragwürdig, weshalb der Verordnungsentwurf auch auf solche Systeme Anwendung finden soll, die nicht die skizzierten KI-spezifischen Risiken in sich tragen. Dies widerspricht dem risikobasierten Regulierungsansatz.

Andererseits bedarf die Definition der Hochrisiko-KI-Systeme der Überarbeitung. Durch die zahlreichen Verweisungen ist sie für den Normadressaten zu unübersichtlich. Vielmehr sollten Hochrisiko-KI-Systeme anhand der zugrunde liegenden Technik, dem Einsatzzweck und -gebiet definiert werden. Denn auch im Rahmen der dem KI-System zugrunde liegenden Lerntechniken bestehen große Diskrepanzen mit Blick auf ihren Gefährdungsgrad. So sind bspw. KI-Systeme, denen die Methodik des überwachten oder verstärkenden Lernens zugrunde liegen, in ihrem Entwicklungsprozess deutlich beherrschbarer als solche, die auf mehrschichtigen künstlichen neuronalen Netzen beruhen. Im Übrigen ist die bereits vorzufindende An-

knüpfung an das Einsatzgebiet mit einem ergänzungsfähigen Katalog von Anwendungsfeldern ein probates Mittel, um dem Bestimmtheitserfordernis Rechnung zu tragen.

XIII. Fazit

Der Vorstoß der Europäischen Kommission, den weltweit ersten Rechtsrahmen für KI schaffen zu wollen, ist zu begrüßen. Denn nur auf diese Weise lässt sich die gebotene Rechtssicherheit im Umgang mit KI-Systemen erreichen. Die risikobasierte und werteorientierte Ausrichtung des Verordnungsentwurfs überzeugt in seiner grundlegenden Konzeption. So versucht der Entwurf einen angemessenen Ausgleich zwischen der Beachtung der Grundrechte und der Ermöglichung von Innovation zu finden. Dies gelingt in vielen Bereichen, wenngleich die technische Realisierbarkeit der Vorgaben zuweilen noch zweifelhaft erscheint. Europa erhält durch den Rechtsrahmen den Stempel einer vertrauenswürdigen KI, die sich dadurch auszeichnet, dass europäische Werte gewahrt werden. Nichtsdestotrotz führt der Entwurf in seiner jetzigen Fassung tendenziell zu einer Überregulierung, die eine investitionshemmende Wirkung zur Folge haben kann.