Entwicklung der EU-Verordnung für künstliche Intelligenz (KI)

Entwicklungsstand und aktuelle Einwirkung durch ChatGPT

Von Prof. Dr. Christoph Graf von Bernstorff, Bremen

Seit etwa zwei Jahren wird im EU-Parlament bereits über den Vorschlag der EU-Kommission für eine Verordnung für Künstliche Intelligenz (KI) debattiert und es wurden über 3.300 Änderungsanträge vorgelegt. Seit Anfang 2023 besteht der aktuelle Hype um den Sprachroboter ChatGPT, der die Überlegungen zur Weiterentwicklung der EU-Gesetzgebung in eine völlig neue Richtung bringt, da sich für KI-Systeme vollkommen neue Einsatzzwecke auftun. So wird überlegt, ob es für diese neuesten KI-Entwicklungen spezieller Regulierungen bedarf oder ob von vornherein eine Einstufung als Hochrisiko-Technologie (beispielsweise für die Frage, ob die Chatbots ebenso wie eine automatische Gesichtserkennung als hochriskante Technik eingestuft werden müssten) mit besonders hohen Anforderungen hinsichtlich Transparenz, Datenschutz usw. angebracht sein könnte.

INHALT

Entwurf einer KI-Verordnung

Am 21.4.2021 legte die EU-Kommission einen „Vorschlag für eine Verordnung des Europäischen Parlaments und des Rates zur Festlegung harmonisierter Vorschriften für künstliche Intelligenz (*Gesetz über Künstliche Intelligenz*)" vor, KOM (2021) 206 final 2021/0106 (COD). Dieser Vorschlag geht auf die politischen Leitlinien für die Kommission (2019–2024) – „Eine Union, die mehr erreichen will" zurück, in denen ein Legislativvorschlag der EU-Kommission für ein Konzept zu menschlichen und ethischen Aspekten der KI angekündigt wurde. Im Nachgang zu dieser Ankündigung veröffentlichte die Kommission am 19.2.2020 ihr Weißbuch zur KI „Ein europäisches Konzept für Exzellenz und Vertrauen" [KOM (2020) 65 final], worin sie die politischen Optionen darlegte, wie die Nutzung von KI gefördert und gleichzeitig die mit bestimmten Anwendungen dieser Technologie verbundenen Risiken eingedämmt werden können. Der Vorschlag beruht auf den Werten und Grundrechten der EU und will erreichen, dass Privatpersonen und andere Nutzer KI-gestützten Lösungen vertrauen und gleichzeitig Unternehmen Anreize erhalten, diese zu entwickeln.

Zielsetzung

Der Vorschlag für eine neue KI-Verordnung aus dem Jahr 2021 enthält harmonisierte Vorschriften für die Entwicklung, das Inverkehrbringen und die Verwendung von KI-Systemen in der Union, die im Verhältnis zu den Risiken stehen. Während einige besonders schädliche KI-Praktiken, die gegen die Werte der Union verstoßen, verboten werden, werden für die Zwecke der Strafverfolgung für bestimmte Anwendungen biometrischer Fernidentifizierungssysteme konkrete Beschränkungen und Sicherheitsmaßnahmen vorgeschlagen. Der Vorschlag enthält eine Risiko-Methodik zur Einstufung von Hochrisiko-KI-Systemen, also Systemen, die erhebliche Risiken für die Gesundheit und Sicherheit oder die Grundrechte von Personen bergen. Solche KI-Systeme müssen horizontalen Auflagen für vertrauenswürdige KI genügen und Konformitätsbewertungsverfahren unterzogen werden, bevor sie in der Union in Verkehr gebracht werden dürfen. Damit die Sicherheit und die Einhaltung bestehender Rechtsvorschriften zum Schutz der Grundrechte über den gesamten Lebenszyklus von KI-Systemen hinweg gewahrt bleiben, werden Anbietern und Nutzern dieser Systeme berechenbare, verhältnismäßige und klare Pflichten auferlegt. Für einige KI-Systeme werden nur minimale Transparenzpflichten vorgeschlagen, insbesondere für den Einsatz von Chatbots (KI-basierte Text-Dialogsysteme) oder „Deepfakes" (durch KI verfälschte Medieninhalte).

Anwendungsbereich

Der in 12 „Titel" untergliederte Verordnungs-Entwurf nimmt in *Titel I* (Art. 1 bis 4 des VO-Entwurfs) den Gegenstand der Verordnung und den Anwendungsbereich der neuen Vorschriften für das Inverkehrbringen, die Inbetriebnahme und die Verwendung von KI-Systemen. Er enthält auch die Begriffsbestimmungen, die in diesem Rechtsinstrument durchweg verwendet werden. Ziel der in diesem Rechtsrahmen festgelegten Begriffsbestimmung für KI-Systeme ist es, so technologieneutral und zukunftstauglich wie möglich zu sein und den rasanten Entwicklungen in der KI-Technologie und auf dem KI-Markt Rechnung zu tragen. Damit die notwendige Rechtssicherheit gegeben ist, wird Titel I durch Anhang I ergänzt, in welchem Konzepte und Techniken für die KI-Entwicklung detailliert aufgeführt sind und von der Kommission in dem Umfang angepasst werden, wie sich neue technologische Entwicklungen ergeben.

Verbotene Praktiken

Titel II (Art. 5 VO-Entwurf) enthält eine Liste verbotener KI-Praktiken. Die Verordnung verfolgt einen risikobasierten Ansatz, bei dem zwischen Anwendungen von KI unterschieden wird, die ein

- unannehmbares Risiko,
- ein hohes Risiko und
- ein geringes oder minimales Risiko darstellen.

Die Aufstellung der verbotenen Praktiken in Titel II umfasst alle KI-Systeme, die als unannehmbar gelten, weil sie Werte der Union, beispielsweise Grundrechte, verletzen. Die Verbote gelten für Praktiken, die ein erhebliches Potenzial haben, Personen zu manipulieren.

Hohes Risiko

Der sehr umfangreiche *Titel III* (Art. 6 bis 51 VO-Entwurf) enthält spezifische Vorschriften für KI-Systeme, die ein hohes Risiko für die Gesundheit und Sicherheit oder für die Grundrechte natürlicher Personen darstellen. Entsprechend dem risikobasierten Ansatz sind solche Hochrisiko-KI-Systeme auf dem europäischen Markt zugelassen, sofern sie bestimmten zwingend vorgeschriebenen Anforderungen genügen und vorab eine Konformitätsbewertung durchgeführt wird. Die Einstufung als Hochrisiko-KI-System beruht auf der Zweckbestimmung des KI-Systems entsprechend den bestehenden EU-Produktsicherheitsvorschriften. Hier erfolgt also ein Verweis auf die EU-Regelungen zur Produkthaftung, die sich derzeit ebenfalls in einem Neuerungsprozess befinden (vgl. *AW-Prax 2023, 64 ff.*, Neues zur EU-Produkthaftung).

In Titel III Kapitel 1 sind die Einstufungsregeln angegeben und zwei Hauptkategorien für *Hochrisiko-KI-Systeme* festgelegt:

- KI-Systeme, die als Sicherheitskomponenten von Produkten, die einer Vorab-Konformitätsbewertung durch Dritte unterliegen, verwendet werden sollen;
- sonstige eigenständige KI-Systeme, die ausdrücklich in *Anhang III* genannt werden und sich vor allem auf die Grundrechte auswirken.

Hochrisiko-KI-Systeme umfassen Systeme, bei denen eine besonders hohe Gefahr für die Gesundheit und Sicherheit oder die Grundrechte von EU-Bürgern befürchtet wird. Die Einstufung in den Hochrisikobereich beschränkt sich unter anderem auf die in Anhang III des Verordnungsentwurfs gelisteten Systeme. Mensch-Maschine-Interaktionen werden dort bisher nicht erfasst. Würde ChatGPT jedoch als Hochrisiko-KI-System bewertet, würde es einer strengen Regulierung unterliegen, und es müsste zusätzlich den Transparenzanforderungen des Art. 52 des VO-Entwurfs genügen.

Für *KI-Systeme mit einem lediglich geringen Risiko* gelten nach Art. 52 VO-Entwurf grundsätzlich besondere Transparenzpflichten. *KI-Systeme mit einem minimalen Risiko* unterfallen nicht dem Art. 52. Für diese Systeme empfiehlt Art. 69 VO-Entwurf nur die Aufstellung fakultativer Verhaltenskodizes.

Transparenzpflichten

Titel IV (Art. 52 VO-Entwurf) befasst sich mit spezifischen Manipulationsrisiken bestimmter KI-Systeme. Transparenzpflichten gelten für Systeme, die

- mit Menschen interagieren,
- zur Erkennung von Emotionen oder zur Assoziierung (gesellschaftlicher) Kategorien anhand biometrischer Daten eingesetzt werden oder
- Inhalte erzeugen oder manipulieren („*Deepfakes*“).

Interagieren Personen mit KI-Systemen oder werden deren Emotionen oder Merkmale durch automatisierte Mittel erkannt, müssen die Menschen hierüber informiert werden.

Titel V (Art. 53 bis 55 VO-Entwurf) befasst sich mit geplanten Maßnahmen zur Innovationsförderung (etwa durch Aufbau von „Reallaboren“ mit kontrollierten Testumgebungen). *Titel VI bis VIII* (Art. 56 bis 68 VO-Entwurf) enthalten Grundlagen für die Einrichtung eines Europäischen Ausschusses für künstliche Intelligenz auf Unionsebene, der sich aus Vertretern der Mitgliedstaaten und der Kommission zusammensetzt. Der Ausschuss soll zu einer wirksamen Zusammenarbeit der nationalen Aufsichtsbehörden und der Kommission beitragen und so eine reibungslose, wirksame und harmonisierte Durchführung dieser Verordnung erleichtern und darüber hinaus die Kommission fachlich beraten. Die *Schlusstitel IX bis XII* enthalten Verhaltenskodizes (Art. 69 VO-Entwurf), Vertraulichkeit und Sanktionen (ab Art. 70) und Schlussbestimmungen.

Die Verhandlungen über den Verordnungsentwurf waren zum Jahresende 2022 eigentlich schon abgeschlossen, nachdem der Rat der EU-Mitgliedstaaten seine Verhandlungsposition festgelegt hatte; das EU-Parlament sollte im Frühjahr 2023 über den Entwurf abstimmen. Im Februar 2023 tauchte die neue Problematik auf, wie das Programm ChatGPT einzustufen sei und ob eventuell eine (weitere) spezielle Regulierung für eine solche, eventuell gar als „Hochrisikotechnologie“ einzustufende KI, erforderlich werden könnte.

Neue Anforderungen durch den Chatbot „ChatGPT“

ChatGPT (*Generative Pre-trained Transformer*) ist ein vom US-Unternehmen OpenAI entwickelter und im November 2022 veröffentlichter Prototyp eines Chatbots, also eines textbasierten Dialogsystems als Benutzerschnittstelle, der auf maschinellem Lernen beruht. ChatGPT formuliert Texte, indem Wort um Wort die wahrscheinliche Fortsetzung eines Satzes eingeschätzt wird. Eine Folge des Verfahrens ist aktuell, dass ChatGPT neben korrekten Angaben auch völlig falsche Informationen erfindet, dabei aber für den Nutzer kein Unterschied erkennbar ist.

Als die EU-Kommission ihren Entwurf zur KI-Verordnung im April 2021 vorstellte, spielten Programme wie ChatGPT noch keine große Rolle. Inzwischen jedoch dreht sich die Diskussion um die Frage, ob Chatbots wie ChatGPT ebenso wie eine automatische Gesichtserkennung als hochriskante Technik eingestuft werden müssen.

Die Meinungen zu dieser Technologie sind derzeit streitig und weit auseinander – sie reichen von der Notwendigkeit einer Einstufung des ChatGPT als Hochtechnologie, die einer speziellen Regulierung bedürfe, bis hin zu der Einschätzung, dass ChatGPT oder vergleichbare KI keiner besonderen Regulierung bedürfe, um Innovationen in Europa nicht zu behindern.

Der Verordnungsentwurf sieht in Titel III bereits Regelungen zu Hochrisiko-KI-Systemen vor. Sollten Anwendungen wie Microsofts ChatGPT (oder auch Googles Bard) als hochriskant eingestuft werden, hätte dies höhere Auflagen für die Anbieter zur Folge. Dazu zählen unter anderem hohe Anforderungen an angemessene Risikobewertungs- und Risikominderungssysteme, eine hohe Qualität der Datensätze, die in das System eingespeist werden, um Risiken und diskriminierende Ergebnisse so gering wie möglich zu halten, sowie eine zusätzliche Protokollierung der Vorgänge, um die Rückverfolgbarkeit von Ergebnissen zu ermöglichen.

Jüngste Entwicklung

Am 28.4.2023 entschied das EU-Parlament, dass generative KI-Systeme nicht grundsätzlich als Hochtechnologierisiko eingestuft werden sollen. Vorausgesetzt werde aber, dass die Daten, von und mit denen die KI-Anwendungen lernen, so ausgewählt sind, dass niemand benachteiligt wird und ein Mensch stets die letzte Kontrolle haben kann.