Kindshofer: KI-VO als datenschutzrechtliche Herausforderung

ZD 2024, 673

KI-VO als datenschutzrechtliche Herausforderung

Theoretische und praktische Einführung mit Lösungsmöglichkeiten für DSB

Aurea Kindshofer ist Datenschutz- und IT-Sicherheitsbeauftragte im Ingenieurbüro Dr. Plesnik GmbH in Herzogenrath.

■ Mit Veröffentlichung der KI-Verordnung (KI-VO) im EU-Amtsblatt beginnt die Umsetzung einer der umfangreichsten Verordnungen der EU-Kommission. Der Umfang der Verordnung ergibt sich aus der Komplexität der Technologie selbst sowie dem Bestreben, diese mit bestehenden Gesetzen, Normen und Verordnungen zu harmonisieren. Aus datenschutzrechtlicher Perspektive stellen neue Technologien wie die Künstliche Intelligenz (KI) eine besondere Herausforderung dar, da die datengestützte Natur der KI im Widerspruch zu den Prinzipien der DS-GVO steht, deren Fokus auf der Datenminimierung gem. Art. 5 Abs. 1 lit. c DS-GVO liegt. Dieser Beitrag bietet einen Überblick über die KI-VO, erläutert Grundlagen der Risikobewertung und beleuchtet die datenschutzrechtlichen Problemstellungen, die durch das Training, die Implementierung und die Anwendung von KI-Modellen und KI-Systemen entstehen. Anhand eines Use-Case wird die Relevanz geeigneter Compliance- und Governance-Maßnahmen verdeutlicht und praxisorientierte Lösungen für Datenschutzbeauftragte vorgestellt.

■ With the publication of the Artificial Intelligence Act (AI Act) in the Official Journal of the EU, the implementation of one of the EU Commission's most comprehensive regulations begins. The scope of the regulation is a result of the complexity of the technology itself and the desire to harmonise it with existing laws, standards and regulations. From a data protection perspective, new technologies such as artificial intelligence (AI) pose a particular challenge, as the data-based nature of AI is at odds with the principles of the GDPR, which focuses on data minimisation in accordance with Art. 5 para. 1 lit. c GDPR. This article provides an overview of the GDPR, explains the basics of risk assessment and highlights the data protection issues that arise from the training, implementation and use of AI models and AI systems. The relevance of suitable compliance and governance measures is illustrated based on a use case and practical solutions for data protection officers are presented.

Lesedauer: 22 Minuten

I. Einleitung

Der Verabschiedung der KI-VO ging ein Prozess voraus, der bereits 2018 angestoßen wurde. ISO-Normen wie die ISO/IEC TR 24028:2020[1], ISO/IEC 22989:2022[2] oder ISO/IEC 23053:2022[3] bilden nicht nur die Grundlage der Definitionen der OECD und später der KI-VO, sondern beinhalten bereits sehr präzise Informationen über eine vertrauenswürdige KI.

Eine von der EU-Kommission eingesetzte Hochrangige Expertengruppe für KI (HEG-KI) veröffentlichte 2019 eine „Ethik-Leitlinie für eine vertrauenswürdige KI"[4], in der sieben Kernanforderungen definiert wurden, die sicherstellen sollen, dass KI als vertrauenswürdige Technologie entwickelt und genutzt werden kann:

- Vorrang menschlichen Handelns und menschliche Aufsicht
- technische Robustheit und Sicherheit
- Schutz der Privatsphäre und Datenqualitätsmanagement
- Transparenz
- Vielfalt, Nichtdiskriminierung und Fairness

- gesellschaftliches und ökologisches Wohlergehen
- Rechenschaftspflicht

Aus diesen Grundsätzen leiten sich entsprechende Maßnahmen ab, deren Ziel die kontinuierliche Kontrolle und Bewertung eines KI-Systems während seines gesamten Lebenszyklus ist. In Erwägungsgrund 27 KI-VO wird auf den Ethik-Leitfaden und die Kernanforderungen Bezug genommen.

II. Historische Entwicklung der Begriffe „KI" und „KI-System"

KI ist als Überbegriff von KI-Systemen oder KI-Modellen abzugrenzen. Noch bevor der Begriff „KI" Anfang 1940 Einzug in die wissenschaftliche Welt nahm,[5] schrieb Samuel Butler 1872 in seinem Roman „Erewhon" von Maschinen mit menschenähnlicher Intelligenz. Auf der Darthmouth-Konferenz[6] 1956 wurde der Begriff von John McCarthy geprägt – sie gilt als die offizielle Geburtsstunde der KI als Forschungsdisziplin. In der ISO-Norm 22989:2022 wird KI unter dem Abschnitt 3.1.3. ebenfalls als Disziplin beschrieben, die sich mit der Forschung und Entwicklung von Mechanismen und Anwendungen von KI-Systemen befasst. Die Definition in Erwägungsgrund 4 KI-VO unterscheidet sich von der Beschreibung als Disziplin und bezeichnet KI stattdessen als eine Reihe schnell fortschreitender Technologien, die Vorteile für Umwelt, Gesellschaft und Wirtschaft bringen sollen. Eine weitere Definition findet sich in der Norm ISO/IEC TR 24028:2020 unter dem Abschnitt 3.4. Demnach ist KI die „Fähigkeit eines technischen Systems (3.38), Wissen und Fertigkeiten zu erwerben, zu verarbeiten und anzuwenden."

Auch anhand der Evolution des Begriffs „KI-System" kann gut nachvollzogen werden, wie rasant sich neue Technologien weiterentwickeln und Definitionen entsprechend angepasst wer-

den müssen. Während das KI-System 2022 noch als „Technisches System [beschrieben wurde], das Outputs wie Inhalte, Prognosen, Empfehlungen oder Entscheidungen für eine bestimmte, vom Menschen definierte Zielsetzung generiert"[7], findet sich in Art. 3 Ziff. 1 KI-VO das „KI-System als 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."

Im Mai 2019 haben OECD-Mitglieder und sechs Partnerländer erstmals grundlegende Prinzipien für den Umgang mit KI festgelegt. Diese sollen gewährleisten, dass KI-Systeme stabil, sicher, fair und vertrauenswürdig funktionieren und als Richtlinien für staatliche Maßnahmen dienen. In Kooperation mit über 50 Fachleuten aus den Bereichen Politik, Wissenschaft, Wirtschaft, Zivilgesellschaft, Technologie, Gewerkschaften und internationale Organisationen wurden fünf Prinzipien und Handlungsempfehlungen für Regierungen für den verantwortungsvollen Einsatz von KI erarbeitet. Auf der Webseite der OECD findet sich die Originalversion der Definition von KI-Systemen und der besseren Nachvollziehbarkeit halber die ursprüngliche Definition nebst Änderungen.[8]

Die Definitionsfindung der Begriffe in der KI-VO entwickelte sich also u.a. aus vorhergehenden Frameworks und deren Terminologien.

III. KI-Systeme und KI-Modelle

Die KI-VO unterscheidet zwischen einem (bestimmten) KI-System (Art. 3 Ziff. 1 KI-VO) und einem KI-System mit allgemeinem Verwendungszweck (Art. 3 Ziff. 66 KI-VO), dem ein KI-Modell mit allgemeinem Verwendungszweck (Art. 3 Ziff. 63 KI-VO) zugrunde liegt. IdR werden KI-Modelle als AI-as-a-Service (AIaaS)[9] zur Verfügung gestellt, damit Unternehmen keine eigene Infrastruktur aufbauen müssen. Dabei wird unterschieden zwischen einem KI-Modell mit allgemeinem Verwendungszweck und einem KI-Modell mit allgemeinem Verwendungszweck, das durch seine „Fähigkeiten mit hoher Wirkkraft" (Art. 3 Ziff. 64 KI-VO) ein systemisches Risiko (Art. 3 Ziff. 65 KI-VO) darstellen kann. Von einer hohen Wirkkraft wird

ausgegangen, wenn die kumulierte Menge der für das Training inkl. Feintuning verwendeten Rechenleistung 10^{25} FLOPS beträgt. Es ist davon auszugehen, dass die Rechenleistung aller bekannten großen Sprachmodelle über diesem Schwellenwert liegt.

FLOPS (Gleitkommaoperationen pro Sekunde) sind eine Maßeinheit, die verwendet wird, um die Rechenleistung eines Computers zu beschreiben, insbesondere wie viele Berechnungen er pro Sekunde durchführen kann.

Welche Fähigkeiten mit hoher Wirkkraft ein KI-Modell tatsächlich besitzt, hängt nicht allein von der Rechenleistung ab, sondern von vielen weiteren Faktoren wie der Prozessorarchitektur, der Art der verwendeten Algorithmen, der Modellarchitektur, der Datenübertragung, dem Speicherzugriff, der Speicherbandbreite, der Compiler-Optimierung oder der Datenqualität, die maßgeblich Einfluss auf die Effizienz der Rechenleistung haben. Ein kleineres Modell kann also uU eine ähnlich hohe Wirkkraft haben wie ein KI-Modell mit über 10^{25} FLOPS.

In Erwägungsgrund 111 KI-VO wird zwar darauf hingewiesen, dass die Schwellenwerte noch angepasst werden können, dies ändert jedoch nichts an der Tatsache, dass FLOPS als alleinige Metrik für die Bewertung der Wirkkraft eines KI-Modells nur bedingt aussagefähig sind. Für KI-Modelle mit systemischem Risiko sind in Annex XIII KI-VO weitere Kriterien vorgesehen, wie zB die Anzahl der Parameter, die Größe und Qualität des Datensatzes sowie die Anzahl der registrierten Endnutzer.

IV. Risikobasierte Klassifizierung

Während KI-Modelle nach ihrer Wirkkraft bewertet werden, sind sowohl bestimmte KI-Systeme als auch KI-Systeme mit allgemeinem Verwendungszweck nach einem risikobasierten Ansatz in verschiedene Kategorien eingeordnet worden. „risikobasiert" bedeutet, dass KI-Systeme entsprechend dem Grad ihrer Risiken für Gesundheit, Sicherheit und Grundrechte, mit denen sie einhergehen, klassifiziert werden. Auch die DS-GVO folgt einem risikobasierten Ansatz, der jedoch nicht – wie in der KI-VO vorgesehen – auf spezifischen Kategorien beruht, sondern Risiken auf einer kontinuierlichen Skala bewertet. Der Begriff des Risikos erscheint in der DS-GVO an vielen Stellen, obwohl er nicht explizit definiert wird. So soll bei der Auswahl technischer und organisatorischer Maßnahmen (TOMs) die Schwere des Risikos für die Rechte und Freiheiten natürlicher Personen berücksichtigt werden, Art. 32 DS-GVO. Darüber hinaus verlangt die DS-GVO die Durchführung bestimmter Maßnahmen, wenn ein „voraussichtlich hohes Risiko" für die persönlichen Rechte und Freiheiten natürlicher Personen besteht. Beispiele hierfür sind die Datenschutz-Folgenabschätzung (DSFA) gem. Art. 35 DS-GVO sowie die Benachrichtigung der betroffenen Personen bei Datenschutzverletzungen gem. Art. 34 DS-GVO. Nach Art. 3 Ziff. 2 KI-VO ist ein Risiko „die Kombination aus der Wahrscheinlichkeit des Auftretens eines Schadens und der Schwere dieses Schadens". Durch die Einteilung von KI-Systemen in Risikoklassen ergeben sich konkrete Maßnahmen, die sich möglicherweise mit den Pflichten der DS-GVO überschneiden oder diese ergänzen. Ein Beispiel ist die Grundrechte-Folgenabschätzung, die laut EU-Kommission zusammen mit einer DSFA durchgeführt werden sollte[10].

1. Verbotene KI-Systeme

Zu verbotenen KI-Systemen gehören nach Art. 5 KI-VO u.a. solche, die Personen bewusst manipulieren oder täuschen und Schwächen ausnutzen. Auch Social Scoring ist wegen der Gefahr von Diskriminierung und Stigmatisierung verboten. Weiter sind KI-Systeme, die zur Risikobewertung iVm Profiling im Hinblick auf Straffälligkeit Anwendung finden, nicht erlaubt. Ebenfalls verboten sind das Ableiten von Emotionen durch Überwachung am Arbeitsplatz und biometrische Kategorisierungssysteme, durch die sich Rasse, politische Meinung, Gewerkschaftszugehörigkeit, religiöse oder philosophische Überzeugungen, Sexualleben oder sexuelle Orientierung ableiten lassen, sowie Echtzeit-Fernidentifizierungssysteme in öffentlichen Räumen, deren Anwendung nur unter strengen Auflagen zur Strafverfolgung legal ist.

2. Hochrisiko-KI-Systeme

Die KI-VO unterscheidet zwischen zwei Kategorien von Hochrisiko-KI-Systemen:

Die erste Kategorie umfasst KI-Systeme, die selbst Produkte oder Sicherheitskomponenten von Produkten sind, die unter die im Anhang II KI-VO aufgeführten EU-Harmonisierungsrechtsvorschriften fallen. Diese Produkte oder Sicherheitskomponenten müssen im Hinblick auf das Inverkehrbringen oder die Inbetriebnahme einer Konformitätsbewertung durch Dritte hinsicht-

lich Gesundheit und Sicherheit unterzogen werden. Beispiele hierfür sind Spielzeuge, Flugzeuge, zwei- oder dreirädrige Fahrzeuge, Autos, medizinische Geräte, Eisenbahnsysteme und Aufzüge. Bei generativen KI-Modellen zählen u.a. virtuelle Assistenten oder personalisierte Empfehlungen als Bestandteile solcher Produkte.

Die zweite Kategorie umfasst KI-Systeme, die in einen oder mehrere der in Anhang III KI-VO aufgeführten Bereiche fallen, Art. 6 Abs. 2 KI-VO. Beispiele hierfür sind:

■ biometrische Identifizierungs- oder Kategorisierungssysteme: KI-Systeme, die biometrische Daten wie Gesichtserkennung, Stimmerkennung, Emotionserkennung oder Verhaltensanalyse verwenden, bergen ein hohes Risiko durch mögliche Datenschutz- und Privatsphärenverstöße. Generative KI-Komponenten werden in diesen Systemen jedoch selten eingesetzt.

■ KI-Systeme in sicherheitskritischen Bereichen: Systeme, die in sicherheitsrelevanten Bereichen wie der digitalen Infrastruktur, dem Verkehrssektor oder der Energiewirtschaft (zB Wasser-, Gas-, Wärme- und Stromversorgung) eingesetzt werden, können bei Fehlern oder Fehlfunktionen schwere Schäden oder Verletzungen verursachen und gelten daher als Hochrisiko-KI-Systeme.

■ bewertende KI-Systeme in bestimmten Bereichen: KI-Systeme, die zur Bewertung von Leistungen, zur Einstufung von Personen oder zum Zugang zu privaten und öffentlichen Dienstleistungen eingesetzt werden, können bei Fehlern oder Vorurteilen erhebliche Auswirkungen auf die Karriereaussichten, die gesellschaftliche Teilhabe, den Lebensstandard und die Lebensgrundlagen der betroffenen Personen haben. Besonders hochriskant sind Systeme in folgenden Bereichen:

■ Bildungsbereich (zB Bewertung von Schülerleistungen)

■ automatisierte Einstufung von Bewerbern im Personalmanagement und Zugang zur Selbstständigkeit

■ Bewertung der Kreditwürdigkeit natürlicher Personen (Ausnahme: Eigenbedarf von kleinen Anbietern)

■ Abschluss von Kranken- und Lebensversicherungen

■ KI-Systeme im Strafverfolgungs- und Justizbereich: KI-Anwendungen, die von Strafverfolgungsbehörden oder EU-Behörden zur Unterstützung von Gerichtsverfahren eingesetzt werden (zB Lügendetektoren oder Bewertung der Verlässlichkeit von Beweisen), gelten als hochriskant.

■ KI-Systeme im Zusammenhang mit Migration, Asyl und Grenzkontrolle: KI-Systeme, die in Einreise- oder Asylverfahren eingesetzt werden (zB zur Vorhersage von Sicherheits- oder Gesundheitsrisiken oder zur Überprüfung von Reisedokumenten), sind ebenfalls Hochrisiko-KI-Systeme.

■ KI-Systeme im Zusammenhang mit Wahlen: Systeme, die zur Beeinflussung von Wahlergebnissen oder Wahlverhalten eingesetzt werden, gefährden die Wahlfreiheit und stellen daher ein hohes Risiko dar.

3. KI-Systeme mit begrenztem Risiko

KI-Systeme mit begrenztem Risiko können Chatbots sein, die direkt mit Menschen interagieren, die Bild-, Audio-, Text- oder Videoinhalte erzeugen oder manipulieren, oder auch biometrische Kategorisierungs- und Emotionserkennungssysteme (davon zu unterscheiden sind KI-Systeme, die verboten sind). Die Risikoklassifizierung in ein KI-System mit begrenztem Risiko kann auch für KI-Systeme nach Art. 6 Abs. 3 KI-VO gelten, sofern nicht mindestens eine der in Erwägungsgrund 53 KI-VO aufgeführten Bedingungen erfüllt ist.

4. KI-Systeme mit minimalem oder keinem Risiko

Zu KI-Systemen mit geringem oder keinem Risiko gehören KI-Systeme, deren Aufgabe die Durchführung eng gefasster Verfahrensaufgaben ist (Art. 6 Abs. 3 lit. a KI-VO), die Optimierung von vorheriger abgeschlossener menschlicher Tätigkeit (Art. 6 Abs. 3 lit. b KI-VO), die Erkennung von Abweichungen von bekannten Entscheidungsmustern (Art. 6 Abs. 3 lit. c KI-VO) oder die Vorbereitung einer Bewertung, die durch menschliche Entscheidung vollendet wird (Art. 6 Abs. 3 lit. d KI-VO). KI-Systeme dieser Kategorie unterliegen keinen spezifischen Anforderungen gemäß der KI-VO. Die Einhaltung von Verhaltenskodizes (Code of Practices) wird zwar gefördert, ist aber freiwillig, Art. 95, Erwägungsgrund 165 KI-VO. Dabei ist aber zu beachten, dass immer eine individuelle Bewertung iVm einer technischen Dokumentation notwendig ist, um nachzuweisen, dass KI-Systeme die Bedingungen erfüllen.

V. Anforderungen an Hochrisiko-Systeme

Hochriskante KI-Systeme müssen grundlegende Anforderungen erfüllen, um eine EU-Konformitätsbewertung zu erhalten. Hierzu gehören:

- Einrichtung und Erhalt eines Risikomanagementsystems (Art. 9 KI-VO)
- Daten und Daten-Governance (Art. 10 KI-VO)
- technische Dokumentation (Art. 11 KI-VO)
- Aufzeichnungspflichten (Art. 12 KI-VO)
- Transparenz und Bereitstellung von Informationen für Betreiber (Art. 13 KI-VO)
- menschliche Aufsicht (Art. 14 KI-VO)
- Genauigkeit, Robustheit, Cybersicherheit (Art. 15 KI-VO)

Darüber hinaus muss der Anbieter ein Qualitätsmanagement einrichten und unterhalten, Art. 17 KI-VO. Es müssen sämtliche Dokumentationen für mindestens 10 Jahre aufbewahrt werden, Art. 18 KI-VO. Auch automatisch erstellte Protokolle (Art. 19 KI-VO) sind verpflichtend. Es besteht sowohl eine Informationspflicht gegenüber den Marktüberwachungsbehörden als auch die Pflicht zu Korrekturmaßnahmen, sollte der Anbieter den Verdacht haben, dass das KI-System nicht KI-VO-konform ist, Art. 20 KI-VO. Um Konformität nachzuweisen, müssen auf Anforderung der Behörden iRe Zusammenarbeit Informationen und Dokumentationen vollständig vorgelegt werden, Art. 21 KI-VO. Anbieter von hochriskanten KI-Systemen müssen zusätzlich eine maschinenlesbare EU-Konformitätserklärung ausstellen, Art. 47 KI-VO.

VI. Use-Case: KI-basierte Hiring-Software

Neben dem Einsatz von Chatbots gehört Hiring-Software zu den häufig verwendeten KI-Systemen. Der folgende Use-Case demonstriert die Notwendigkeit von Transparenzpflichten aus Sicht von DS-GVO und KI-VO.

Beispiel:

Firma A führt eine KI-basierte Hiring-Software ein, die den Bewerbungsprozess und das Personalmanagement automatisieren soll.

Zweckbestimmung (Art. 3 Ziff. 12 KI-VO) des Hiring-Systems ist das Klassifizieren und Vorsortieren von Bewerbenden sowie das Führen eines Personalverzeichnisses. Hierbei werden Profile angelegt. Da das KI-System auch Profiling gem. Art. 4 Abs. 2 DS-GVO durchführt, wird es nach Annex III Abs. 4 lit. a KI-VO als hochriskant klassifiziert. Überdies werden besondere Kategorien personenbezogener Daten gem. Art. 9 DS-GVO verarbeitet. Um Bewerbungsprozesse zu erleichtern, werden auf Plattformen oder in sozialen Medien Formulare bereitgestellt, in die personenbezogene Daten (strukturierte Daten) sowie Anschreiben, Lebenslauf, Zeugnisse und Zertifikate (unstrukturierte Daten) eingegeben und hochgeladen werden können.

Verarbeitungstätigkeit 1: Der Bewerber A gibt in das Formular seine personenbezogenen Daten ein und lädt Bewerbungsunterlagen hoch. Die Angestellte B ist für das Personalmanage-

ment zuständig und bedient das KI-System. Gem. Art. 26 Abs. 11 KI-VO muss eine natürliche Person vor der Eingabe darüber informiert werden, dass sie mit einem Hochrisiko-KI-System interagiert und gem. Art. 12 ff. DS-GVO aufgeklärt werden, inwieweit ihre Daten (weiter-)verarbeitet werden. Wird ein hochriskantes KI-System genutzt, so muss der Betreiber gem. Art. 26 Abs. 7 KI-VO auch die Angestellte B darüber informieren, dass sie ein KI-System verwendet. Überdies muss die Angestellte B gem. Art. 4 KI-VO KI-Kompetenz erworben haben.

Verarbeitungstätigkeit 2: Anschließend werden die Daten in einem Algorithmic Decision Making(ADM)-System tokenisiert, vektorisiert und ausgewertet. Damit das ADM-System möglichst präzisen Output generiert, wird es uU mit folgenden Daten trainiert:

- durch Eingabe strukturierter und unstrukturierter Informationen durch Bewerbende
- evtl. Kontextinformation für Datenanreicherung
- evtl. Sourcing über öffentlich verfügbare Informationen (Social Media etc)
- historische personenbezogene Daten der Mitarbeitenden

Verarbeitungstätigkeit 3: Das KI-System hat einen Abgleich und eine abschließende Bewertung der Bewerbenden vorgenommen und gibt entsprechende Inhalte aus. Hier muss festgestellt werden, inwieweit eine vertragliche Ausgestaltung zwischen Anbieter und Betreiber hinsichtlich der Weiterverarbeitung der Daten als mögliche Trainingsdaten und nicht mehr nur zum Zweck der vorvertraglichen oder vertraglichen Erfüllung besteht. In diesem Fall muss die betroffene Person gem. § 32 Nr. 1 Abs. 1 BDSG darüber informiert werden.

Der Transparenzbegriff im Kontext beider Verordnungen:

Während sich die Transparenz in der DS-GVO auf die Verarbeitung personenbezogener Daten bezieht, gehen die Transparenzpflichten der KI-VO darüber hinaus, da auch die Technologie selbst transparent gestaltet werden muss. Dies kann zu einem datenschutzrechtlichen Konflikt führen, da die gemäß der KI-VO verpflichtende Offenlegung von Trainingsdaten ggf. personenbezogene Daten enthalten kann, die nach der DS-GVO geschützt werden müssen. Zur Transparenzpflicht nach Art. 26 und 50 KI-VO, Erwägungsgrund 132 KI-VO und Art. 12 ff. DS-GVO gehört es, betroffenen Personen barrierefrei Informationen zur Verfügung zu stellen, die sie hinreichend über Verarbeitungstätigkeiten und Funktionsweisen aufklären, damit eine informierte Einwilligung gegeben werden kann. Transparenz gehört gem. Art. 5 Abs. 1 lit. a DS-GVO zu den Grundsätzen der DS-GVO und ist eng mit der Informationspflicht verbunden, Art. 13 und 14 DS-GVO. Gem. Art. 12. Abs. 1 S. 1 DS-GVO muss der Verantwortliche geeignete Maßnahmen treffen, um seinen Transparenzpflichten hinreichend nachzukommen. Hierbei ist die Information „in präziser, transparenter, verständlicher und leicht zugänglicher Form in einer klaren und einfachen Sprache zu übermitteln“.

Ergänzende Angaben zur Informationspflicht werden in Erwägungsgrund 60 DS-GVO gemacht.

In Erwägungsgrund 58 S. 3 DS-GVO wird angegeben, dass Transparenz insbesondere dort nötig ist, „wo die große Zahl der Beteiligten und die Komplexität der dazu benötigten Technik es der betroffenen Person schwer machen, zu erkennen und nachzuvollziehen, ob, von wem und zu welchem Zweck sie betreffende personenbezogene Daten erfasst werden, wie etwa bei der Werbung im Internet.“

Nach Erwägungsgrund 39 S. 2 DS-GVO „sollte für natürliche Personen Transparenz dahingehend bestehen, dass sie betreffende personenbezogene Daten erhoben, verwendet, eingesehen oder anderweitig

verarbeitet werden und in welchem Umfang die personenbezogenen Daten verarbeitet werden und künftig noch verarbeitet werden."

In Erwägungsgrund 78 S. 3 DS-GVO soll Transparenz in Bezug auf die Funktionen und die Verarbeitung personenbezogener Daten hergestellt werden, die es den betroffenen Person ermöglicht, die Verarbeitung personenbezogener Daten zu überwachen.

Die KI-VO sieht keine Definition der Transparenz in Art. 3 KI-VO vor. In Erwägungsgrund 27 KI-VO wird jedoch auf den Ethik-Leitfaden der HEG-KI 2019 verwiesen und Transparenz wie folgt beschrieben: „'Transparenz' bedeutet, dass KI-Systeme so entwickelt und verwendet werden, dass sie angemessen nachvollziehbar und erklärbar sind, wobei den Menschen bewusst gemacht werden muss, dass sie mit einem KI-System kommunizieren oder interagieren, und dass die Betreiber ordnungsgemäß über die Fähigkeiten und Grenzen des KI-Systems informieren und die betroffenen Personen über ihre Rechte in Kenntnis setzen müssen."

Der Begriff „Transparenz" iSd Art. 50, 26 Ziff. 7 und 11 KI-VO, Erwägungsgrund 132, 133 und 134 KI-VO bezieht sich hier insbesondere auf die Kenntlichmachung, dass mit (Hochrisiko-)KI-Systemen interagiert und ein synthetischer Output generiert wird.

In Erwägungsgrund 67 S. 3 KI-VO wird auf die Einhaltung der Pflichten nach der DS-GVO hingewiesen. Demnach sollen Daten-Governance- und Datenverwaltungsverfahren bei personenbezogenen Daten Transparenz hinsichtlich des ursprünglichen Zwecks der Datenerhebung sicherstellen.

ISd Art. 13 und 53 KI-VO, Erwägungsgrund 72 KI-VO soll Transparenz durch detaillierte technische Erläuterungen des KI-Modells oder KI-Systems gewährleistet werden. Gemäß Erwägungsgrund 101 KI-VO sollten „verhältnismäßige Transparenzmaßnahmen festgelegt werden, einschließlich der Erstellung und Aktualisierung von Dokumentation und der Bereitstellung von Informationen über das KI-Modell mit allgemeinem Verwendungszweck für dessen Nutzung durch die nachgelagerten Anbieter."

In Annex XI KI-VO finden sich überdies weiterführende Informationen zur technischen Dokumentation von KI-Modellen, die einen maßgeblichen Bestandteil der Transparenz darstellt.

Gem. Art. 26 Abs. 6 KI-VO müssen Betreiber eine automatisierte Protokollführung gewährleisten und die Ergebnisse bis zu sechs Monate aufbewahren.

Es muss im Einzelfall nicht nur geprüft werden, ob und wie KI-Systeme und -Modelle personenbezogene (Trainings-)Daten verarbeiten, sondern auch, inwieweit personenbezogene Daten im Zuge der Offenlegung von Trainingsdaten für eine angestrebte EU-Konformitätsbewertung prüfenden Dritten zur Verfügung gestellt werden.

VII. Rechtsgrundlagen

1. Einwilligung

Aktuell ist für Bewerbende noch nicht ersichtlich, dass sie mit einem KI-basierten System interagieren. Eine Einwilligung gem. Art. 6 Abs. 1 lit. a DS-GVO ist nur dann möglich, wenn für den Betroffenen ersichtlich ist, wofür er seine Einwilligung erteilt.

Bei KI-Systemen wie dem Hiring-System der Firma A stellt dies aufgrund der fehlenden Transparenz der Verarbeitungstätigkeiten eine Herausforderung dar. Zumindest müssen aber gemäß Erwägungsgrund 42 S. 4 DS-GVO der Zweck der Verarbeitung personenbezogener Daten und der Verantwortliche angegeben werden. Auch im Hinblick auf die Widerruflichkeit der Einwilli-

gung gem. Art. 7 Abs. 3 S. 1 DS-GVO stellt sich die Frage nach einer sinnvollen Umsetzung. Nutzt die betroffene Person ihr Widerrufsrecht, muss der Verantwortliche ihre personenbezogenen Daten gem. Art.

17 Abs. 1 lit. b DS-GVO unverzüglich löschen, sofern keine andere Rechtsgrundlage für die Datenverarbeitung besteht. Dies könnte uU mit einem Verlust der Funktionalität des KI-Systems verbunden sein, sofern dieses mit den zu widerrufenden personenbezogenen Daten trainiert wurde. Das KI-System oder KI-Modell müsste ganz neu trainiert werden. Damit würden auch alle anderen Trainingsdaten, für die eine Einwilligung zur Verarbeitung vorliegt, verloren gehen und es müssten neue Einwilligungen eingeholt werden.

2. Vertragserfüllung und Zweckänderung

Eine Vertragserfüllung kann gem. Art. 6 Abs. 2 lit. b DS-GVO insbesondere bei KI-Systemen, die für eine bestimmte Funktion entwickelt wurden, als Rechtfertigungsgrund gelten. Die Zweckänderung und Weiterverarbeitung personenbezogener Daten gem. Art. 6 Abs. 4 DS-GVO stellt für KI-Modell- oder KI-System-Nutzende aber auch hier eine datenschutzrechtliche Herausforderung dar. Wenn eine betroffene Person die Einwilligung zur Verarbeitung ihrer personenbezogenen Daten für eine Vertragserfüllung gegeben hat, muss geklärt werden, ob es eines neuen Zwecks bedarf, wenn die gleichen Daten anschließend für das Training genutzt werden. Gemäß Erwägungsgrund 50 DS-GVO müssen die Zwecke vereinbar sein, damit eine Weiterverarbeitung zulässig ist. Bei einem KI-System, das dazu konzipiert wurde, aus eingegebenen Daten zu lernen, um die Vertragserfüllung des ursprünglich genannten Zwecks gewährleisten zu können, lässt sich darüber diskutieren, inwieweit eine Zweckänderung notwendig oder möglich ist.

In der KI-VO findet sich in der Begriffsbeschreibung von Art. 3 Ziff. 12 KI-VO die Zweckbestimmung. Sie bezieht sich nicht auf personenbezogene Daten, sondern auf das KI-System selbst. In Erwägungsgrund 12 KI-VO wird erwähnt, dass „durch die Bezugnahme auf explizite oder implizite Ziele betont [wird], dass KI-Systeme gemäß explizit festgelegten Zielen oder gemäß impliziten Zielen arbeiten können. Die Ziele des KI-Systems können sich – unter bestimmten Umständen – von der Zweckbestimmung des KI-Systems unterscheiden."

Es ergibt sich also nicht nur die Fragestellung nach einer Verarbeitung personenbezogener Daten für ein konkretes KI-System, sondern es besteht auch die Möglichkeit, dass das KI-System impliziten Zielen und nicht mehr der ursprünglichen Zweckbestimmung folgt. Die Zweckbestimmung legt also fest, wofür ein KI-System (theoretisch) konzipiert ist. In der DS-GVO finden sich die Begriffe „Zweck" (Art. 5 Abs. 1 lit. b DS-GVO) und „Zweckbindung", die sich aus Art. 6 DS-GVO ableiten.

Sowohl der Zweck der Verarbeitung personenbezogener Daten als auch die Zweckbindung ergeben sich aus der Zweckbestimmung, die in der KI-VO Anwendung findet. Folgt das KI-System Zielen, die von seiner Zweckbestimmung abweichen, kann sich daraus ggf. eine neue Rechtsgrundlage ergeben, die geprüft werden muss.

3. Berechtigtes Interesse

Das berechtigte Interesse als Rechtsgrundlage gem. Art. 6 Abs. 1 lit. f DS-GVO kann zur Anwendung kommen, wenn Verantwortliche oder Dritte personenbezogene Daten verarbeiten, um ihr berechtigtes Interesse zu wahren. Hierbei darf das berechtigte Interesse die Interessen, Grundrechte oder Grundfreiheiten der betroffenen Person, die den Schutz personenbezogener Daten erfordern, nicht überwiegen.

Erwägungsgrund 47 DS-GVO geht auf das berechtigte Interesse im Kundenkontext expliziter ein. Nach Erwägungsgrund 48 DS-GVO können Unternehmen innerhalb einer Gruppe oder eines Konzerns ein berechtigtes Interesse daran haben, personenbezogene Daten für interne Verwaltungszwecke, einschließlich Kunden- und Mitarbeiterdaten, zu übermitteln – dies dürfte insbesondere im Hinblick auf KI-basierte Personalmanagementsoftware von Bedeutung sein.

4. Erforderlichkeit

Ein wichtiges Kriterium ist auch die Erforderlichkeit. Müssen personenbezogene Daten zwingend für das Training von KI-Modellen oder KI-Systemen genutzt werden oder kommen auch andere, datenminimierende

Möglichkeiten in Betracht? Wird ein KI-System im Personalwesen eingesetzt, ist eine Verarbeitung personenbezogener Daten für die Funktionalität der KI-basierten Anwendung erforderlich. Ist ein KI-System aber dafür konzipiert, Unregelmäßigkeiten in der Buchhaltung zu entdecken, sind keine personenbezogenen Daten erforderlich. Somit besteht keine Rechtsgrundlage nach Art. 6 Abs. 1 lit. f DS-GVO.

5. Abwägung

Bei der Anwendung von KI-Systemen stellt sich grundsätzlich immer die Frage, ob Betroffene einschätzen können, wie groß die Auswirkungen auf die Interessen, Grundrechte und Grundfreiheiten sind, wenn personenbezogene Daten durch ein KI-System oder KI-Modell (weiter-)verarbeitet werden. Es muss also abgewogen werden, inwieweit die Verarbeitung personenbezogener Daten in einem angemessenen Verhältnis zum berechtigen Interesse steht.

Die Verarbeitung besonderer Kategorien personenbezogener Daten ist gem. Art. 9 Abs. 1 DS-GVO verboten. Ausnahmen finden sich in Art. 9 Abs. 2 lit. a und lit. e DS-GVO: wenn eine ausdrückliche Einwilligung zur Verarbeitung gegeben wurde oder die personenbezogenen Daten von der betreffenden Person öffentlich gemacht wurden.

Mit zu berücksichtigen ist auch hier wieder das Wesen eines KI-Systems oder KI-Modells, weil es durch einseitige oder nicht hochwertige Trainingsdaten zu Verzerrungen und infolgedessen zu Diskriminierungen kommen kann, die direkte Auswirkungen auf die Betroffenen haben. Hier entsteht ggf. ein Konflikt zwischen der DS-GVO und der KI-VO, da in Erwägungsgrund 70 KI-VO eine Ausnahme zum Verbot der Verarbeitung besonderer Kategorien personenbezogener Daten erwähnt wird, um Verzerrung und Diskriminierung zu verhindern.

6. Recht auf Löschen und Vergessenwerden

Das Recht auf Löschung gem. Art. 17 DS-GVO ist schwer umzusetzen, da generative KI-Systeme möglicherweise Schwierigkeiten haben, Daten tatsächlich zu „vergessen". Stattdessen nutzen sie diese Daten zur eigenen Verbesserung, wodurch personenbezogene Daten unwiderruflich in das Daten- und Verhaltensmuster der KI integriert werden können. Gleiches gilt gem. Erwägungsgrund 55 DS-GVO für das Berichtigen von personenbezogenen Daten. Es können zwar aktualisierte personenbezogene Daten eingegeben werden, dies führt jedoch nicht zu einem „Überschreiben" der ursprünglichen Eingabe. Somit ist eine Berichtigung nur bedingt möglich.

7. Drittstaatentransfer

Häufig werden KI-Systeme genutzt, die außerhalb der EU entwickelt wurden. Hier gilt es, auf den Drittstaatentransfer zu achten. Falls Trainingsdaten zB von Dienstleistern außerhalb der EU/des EWR verarbeitet werden, sind die Bestimmungen gem. Art. 44 ff. DS-GVO zu beachten. Datenschutzbeauftragte sollten prüfen, ob die Drittstaaten als sicher gelten. Falls es sich um nicht sichere Drittstaaten handelt, müssen im Falle einer Auf-

tragsverarbeitung sog. Transfer Impact Assessments (TIA) durchgeführt und Standardvertragsklauseln (SCC) abgeschlossen werden.

8. Exkurs: Notwendige Maßnahmen für Betreiber von HR-Software

■ Qualitätsmanagementsystem gem. Art. 17 KI-VO, ggf. Erwägungsgrund 146 KI-VO

■ DSFA gem. Art. 26 Abs. 9 KI-VO, Art. 35 DS-GVO, Erwägungsgrund 91 DS-GVO und § 67 BDSG mit vorangehender Schwellenwertanalyse

■ ggf. Grundrechte-Folgenabschätzung gem. Art. 27 KI-VO, Erwägungsgrund 96 KI-VO

■ Sicherstellung der KI-Kompetenz gem. Art. 4 KI-VO, Erwägungsgrund 91 KI-VO

■ Benennung und Unterstützung von Fachpersonen gem. Art. 26 Abs. 2 KI-VO

■ Verarbeitungsverzeichnis gem. Art. 30 DS-GVO, Erwägungsgrund 82 DS-GVO, § 70 BDSG

■ TOMs gem. Art. 11 und 15 KI-VO, Art. 25 und 32 DS-GVO und § 64 BDSG

■ Transparenz- und Informationspflicht gem. Art. 12 ff. DS-GVO, Art. 26 Abs. 11 KI-VO und § 32 BDSG

■ vertragliche Absicherung zwischen Anbieter und Betreiber

Sinnvoll bei der Umsetzung von Compliance-Maßnahmen[11] sind Hilfsmittel wie die Checkliste des HmbBfDI[12] oder der DSK[13] sowie die ISO/IEC 42001:2023 13[14] als KI-Risikomanagement-Framework.

VIII. Fazit

Datenschutzbeauftragte werden vor die Aufgabe gestellt, sich technisches Wissen anzueignen und KI-Systeme und KI-Modelle aus unterschiedlichen rechtlichen Perspektiven zu betrachten. Nur weil ein KI-System oder KI-Modell von der KI-VO ausgenommen ist, bedeutet dies nicht automatisch, dass Anbieter, Betreiber oder sonstige Personen von jeglichen Governance- und Compliance-Maßnahmen befreit sind. Es muss gemäß Erwägungsgrund 53 KI-VO trotzdem geprüft werden, ob die KI-Modelle oder KI-Systeme unter die KI-VO fallen und eine technische Dokumentation geführt werden. IRe Schwellenwertanalyse kann überdies auch festgestellt werden, ob eine DSFA und weitere Maßnahmen notwendig sind.

Schnell gelesen …

■ Die KI-VO gehört durch die Komplexität der Technologie und die weitreichenden Folgen, die durch fehlerhaftes oder missbräuchliches Nutzen der KI-Systeme und KI-Modelle entstehen können, zu den umfangreichsten EU-Verordnungen.

■ Der risikobasierte Ansatz der KI-VO erfordert die Kategorisierung von KI-Systemen und KI-Modellen.

■ Die Harmonisierung von KI-VO und DS-GVO stellt eine regulatorische Herausforderung für beratende Personen und Unternehmen dar.

■ Insbesondere Anbieter und Betreiber von Hochrisiko-KI-Systemen oder KI-Modellen mit allgemeinem Verwendungszweck mit systemischem Risiko müssen umfangreiche Governance- und Compliance-Anforderungen erfüllen.

Aurea Kindshofer
ist Datenschutz- und IT-Sicherheitsbeauftragte im Ingenieurbüro Dr. Plesnik GmbH in Herzogenrath.

1 International Organization for Standardization, 2020, Information technology – Artificial intelligence – Management System (ISO/IEC 24028:2020), abrufbar unter: www.iso.org/standard/77608.html.

2 International Organization for Standardization, 2022, Information Technology – Artificial Intelligence – Artificial intelligence concepts and terminology (ISO/IEC 22989:2022), abrufbar unter: www.iso.org/standard/74296.html.

3 International Organization for Standardization, 2022, Framework for Artificial Intelligence (AI) Systems Using Machine Learning (ML) (ISO/IEC 23053:2022), abrufbar unter: www.iso.org/standard/74438.html.

4

EU-Kommission, Generaldirektion Kommunikationsnetze, Inhalte und Technologien, 2019, Ethik-leitlinien für eine vertrauenswürdige KI, abrufbar unter: https://data.europa.eu/doi/10.2759/22710.

5 McCulloch, W.S., & Pitts, W. (1943), The Bulletin of Mathematical Biophysics, 5(4), 115–133. doi:10.1007/bf02478259; Piccinini, G. (2004). The first computational theory of mind and brain: A close look at Mcculloch and pitts's "logical calculus of ideas immanent in nervous activity". Synthese, 141(2), S. 175–215. doi: 10.1023/b:synt.0000043018.52445.3e.

6 Artificial intelligence (AI) coined at Dartmouth, abrufbar unter: https://home.dartmouth.edu/about/artificial-intelligence-ai-coined-dartmouth.

7 ISO 22989:2022.

8 What is AI? Can you make a clear distinction between AI and non-AI systems?, abrufbar unter: https://oecd.ai/en/wonk/definition.

9 Ebers/Heinze/Krügel/Steinrötter, Künstliche Intelligenz und Robotik/Berberich/Conrad, 2020, § 30 Rn. 31–33.

10 https://ec.europa.eu/commission/presscorner/detail/de/QANDA_21_1683.

11 Le/Treibel ZD 2024, 370.

12 https://datenschutz-hamburg.de/fileadmin/user_upload/HmbBfDI/Datenschutz/Informationen/20231113_Checkliste.

13 Checkliste mit Prüfkriterien nach DS-GVO, abrufbar unter: https://www.lda.bayern.de/media/ki_checkliste.pdf.

14 International Organization for Standardization, 2023, Information Technology – Artificial Intelligence – Artificial intelligence concepts and terminology (ISO/IEC 22989:2022), abrufbar unter: www.iso.org/standard/74296.html.