Sven Venzke-Caprarese

Interne Regelungen zum Einsatz von Künstlicher Intelligenz

Eine praktische Orientierungshilfe für eine KI-Richtlinie

Fragen rund um den Einsatz von Künstlicher Intelligenz (KI) sind in vielen Organisationen derzeit ein Thema. Die ungeregelte Einführung von KI stellt allerdings ein Risiko im Hinblick auf den Datenschutz, die Informationssicherheit und weitere Compliance-Anforderungen dar. Dieser Beitrag soll mögliche Regelungspunkte einer KI-Richtlinie darstellen und eine Hilfestellung für Organisationen geben, die Einführung von KI bedarfsgerecht zu regeln.

1 Vorbemerkungen

Der Einsatz von KI-Systemen birgt neben zahlreichen Chancen auch einige Risiken. Sofern personenbezogene Daten mithilfe eines KI-Systems verarbeitet werden, müssen die datenschutzrechtlichen Anforderungen beachtet werden. Die ungeregelte Einführung von KI-Systemen würde ansonsten ein Risiko für die Rechte und Freiheiten der betroffenen Personen darstellen. Auch die Verarbeitung sonstiger schützenswerter Unternehmensinformationen würde bei einem ungeregelten KI-Einsatz insbesondere zu einem Risiko für die Vertraulichkeit dieser Daten führen. Hinzu kommt, dass eine Fehlbedienung von KI-Systemen, bzw. deren Einsatz in nicht dafür vorgesehenen Situationen, weitere Risiken mit sich bringen kann. Zudem müssen die Anforderungen der KI-Verordnung (KI-VO) eingehalten werden. Dies bedeutet unter anderem, dass die verantwortliche Organisation dafür Sorge tragen muss, dass keine verbotenen KI-Praktiken angewendet werden (Art. 5 KI-VO), dass etwaige Transparenzpflichten für den Betrieb von KI-Systemen eingehalten werden (Art. 50 KI-VO) und dass der Einsatz von Hochrisiko-KI-Systemen (Art. 6 KI-VO) erkannt und die damit zusammenhängenden Vorgaben eingehalten werden.

Vor diesem Hintergrund sollten Organisationen klare Regelungen im Hinblick auf die Einführung und den Einsatz von KI-Systemen treffen. Auf diese Weise kann vermieden werden, dass KI-Systeme unbemerkt und ohne die notwendigen Prozesse eingeführt werden.[1] Eine KI-Richtlinie kann zudem der erste Schritt in Richtung der Umsetzung einer KI-Strategie der Organisation sein, die Einführung unterstützen und den Beschäftigten sowie projektverantwortlichen Personen einen Handlungsrahmen vorgeben.

2 Grundsätzliche Weichenstellung

Zu Beginn einer KI-Richtline steht die Frage, welche KI-Strategie die Organisation verfolgt und damit einhergehend, welcher Ansatz für eine KI-Richtlinie gewählt werden soll. Dabei sind im Grundsatz vier Regelungsmöglichkeiten gegeben, von denen aber nicht alle sinnvoll sein werden.

So ist es denkbar, den Einsatz von KI-Systemen über eine entsprechende Richtlinie vollständig zu verbieten. Damit wäre zwar ein geregeltes Umfeld geschaffen, allerdings würde ein solcher Einsatz sowohl der Organisation als auch den Beschäftigten den Aufbau von KI-Kompetenzen grundsätzlich verwehren.

Auf der anderen Seite wäre es denkbar, den Einsatz von KI ungeregelt zu erlauben. Eine solche Regelung wäre aber nicht sinnvoll, denn die Organisation würde dann einen Kontrollverlust in einem hoch reglementierten Bereich erleiden und den eingangs erwähnten Risiken könnte nicht angemessen begegnet werden.

Sinnvoller erscheint an dieser Stelle ein differenzierterer Ansatz. In diesem Zusammenhang ist es denkbar, den Einsatz von KI-Systemen entweder grundsätzlich zu erlauben und Ausnahmen zu regeln (Blacklist-Ansatz) oder den Einsatz von KI-Systemen grundsätzlich zu verbieten und Ausnahmen zu regeln (Whitelist-Ansatz).



Sven Venzke-Caprarese

ist Geschäftsführer der datenschutz nord GmbH.

E-Mail: svenzke@datenschutz-nord.de

1 Vgl. hierzu auch Orientierungshilfe der Konferenz der unabhängigen Datenschutzaufsichtsbehörden des Bundes und der Länder vom 6. Mai 2024, Künstliche Intelligenz und Datenschutz Version 1.0, Ziffer 2.2.

2.1 Verarbeitung personenbezogener Daten

Sofern personenbezogene Daten in einem KI-System verarbeitet werden sollen, ist ein Whitelist-Ansatz erforderlich. Denn die Verarbeitung personenbezogener Daten setzt eine Prüfung der datenschutzrechtlichen Rahmenbedingungen des konkreten KI-Systems voraus, das freigegeben werden soll. Zudem muss stets sichergestellt sein, dass die Verarbeitung personenbezogener Daten durch eine Rechtsgrundlage abgedeckt ist.

2.1.1 Allgemeine Rahmenbedingungen

Sofern das KI-System nicht in der eigenen Infrastruktur (on premise) betrieben wird,[2] sondern in externen Rechenzentren oder, was derzeit häufig der Fall sein dürfte, als Software as a Service, muss der Einsatz der entsprechenden Dienstleister datenschutzrechtlich betrachtet werden.

Hierzu gehört die Prüfung der Verträge zur Auftragsverarbeitung (Art. 28 DSGVO), bei der sich grundsätzlich keine Besonderheiten ergeben. Allerdings sollte darauf geachtet werden, dass sich die eingesetzten Dienstleister nicht die Verarbeitung der personenbezogenen Daten zu eigenen Zwecken vorbehalten. Denn die Daten dürfen vom Dienstleister insbesondere nicht zum Training der zugrundeliegenden KI-Modelle verwendet werden.

Sofern Dienstleister aus Drittländern beteiligt sind, bzw. die Datenverarbeitung in Drittländern erfolgt, muss die Prüfung der Angemessenheit des Datenschutzniveaus erfolgen (Art. 44 ff. DSGVO).

Je nach Organisation oder Datenarten können weitere spezifische Anforderungen gelten, so etwa im Bereich der Berufsgeheimnisse die Geheimhaltungsverpflichtung der mitwirkenden Personen (§ 203 Abs. 4 StGB) oder im Bereich der Sozialleistungsträger weitere, besondere Anforderungen (§ 80 SGB X).

Selbst wenn die allgemeinen datenschutzrechtlichen Rahmenbedingungen eingehalten werden, stellen sich weitere datenschutzrechtliche Anforderungen. Wenn die Organisation verpflichtet ist, ein Verzeichnis der Verarbeitungstätigkeiten zu führen (Art. 30 DSGVO), muss geprüft werden, ob der Einsatz des KI-Systems entsprechend aufgenommen werden muss. Darüber hinaus muss geprüft werden, ob eine Datenschutz-Folgenabschätzung (Art. 35 DSGVO) vorzunehmen ist. Sofern mithilfe eines KI-Systems erstmalig Daten erhoben werden, ist eine Datenschutzinformation erforderlich (Art. 13 DSGVO).[3] Sofern die Verarbeitung im KI-System eine Zweckänderung darstellt oder personenbezogene Daten durch das KI-System erstmalig bei Dritten erhoben werden, müssen ebenfalls Informationspflichten geprüft werden (Art. 14 DSGVO). Im Zusammenhang mit KI-Systemen ist es oftmals auch relevant, das Vorhandensein automatisierter Entscheidungen im Einzelfall (Art. 22 DSGVO) zu überprüfen. Und selbst wenn die eingegebenen Daten nicht dauerhaft im KI-System gespeichert und nicht zu Trainingszwecken genutzt werden sollen, ist es doch sinnvoll, zu prüfen, ob Daten innerhalb des KI-Systems gespeichert werden und Speicherfristen bzw. Löschpflichten in den Blick genommen werden müssen.[4]

Je nach Schutzbedarf der Daten können darüber hinaus weitere Anforderungen an die technischen und organisatorischen Maßnahmen rund um das KI-System greifen (Art. 32 DSGVO), wie zum Beispiel die Einrichtung einer Zwei-Faktor-Authentisierung für den Zugang zum KI-System und die Einhaltung eines angemessenen Rechte- und Rollenkonzepts für den Zugriff auf etwaig gespeicherte Daten.

2.1.2. Rechtsgrundlage für die Datenverarbeitung

Schließlich muss gewährleistet werden, dass die Datenverarbeitung innerhalb des KI-Systems auch im Einklang mit einer passenden Rechtsgrundlage erfolgt (Art. 6 ff. DSGVO). Hierbei macht es durchaus einen Unterschied, ob es sich bei dem eingesetzten KI-System um eines mit allgemeinem oder spezifischem Verwendungszweck handelt.

KI-Systeme mit allgemeinem Verwendungszweck

Die Festlegung und Prüfung der konkreten Rechtsgrundlage ist bei KI-Systemen mit allgemeinem Verwendungszweck (Art. 3 Nr. 66 KI-VO) im Vorfeld nicht immer möglich. Denn ein solches KI-System kann für eine Vielzahl von Zwecken eingesetzt werden. Ein Beispiel für ein solches System ist ein KI-Chatbot wie ChatGPT oder Microsoft Copilot. Hier ist der Zweck der Datenverarbeitung allgemein gehalten und je nach Einsatz des KI-Chatbots kann die Rechtsgrundlage für die Verarbeitung variieren.

Häufig wird es aber dennoch keiner neuen Rechtsgrundlage bedürfen. Denn sofern ein KI-System lediglich als Werkzeug eingesetzt wird, um einen Zweck zu erfüllen, für den es im Vorfeld schon eine Rechtsgrundlage gab, kann die Verarbeitung der personenbezogenen Daten im Zusammenhang mit dem KI-System als Fortführung des ursprünglichen Erhebungszwecks angesehen werden. Sofern keine Zweckänderung eintritt, bedarf es dann keiner neuen Rechtsgrundlage. Der Einsatz muss insofern aber auch erforderlich sein, z.B. um dem Zweck innerhalb der angemessenen Zeit und mit angemessener Effizienz nachkommen zu können.[5]

Würde man hingegen vertreten, dass derartige Verarbeitungen in KI-Systemen nicht basierend auf der ursprünglichen Rechtsgrundlage durchgeführt werden könnten, sondern alleine für den Einsatz des KI-Systems als Werkzeug schon eine neue Rechtsgrundlage fordern (z.B. Art. 6 Abs. 1 lit. f DSGVO), so würde hier eine unnötige Prüfung erfolgen müssen und ggf. erneute Transparenzpflichten nach Art. 14 DSGVO bzw. Widerspruchsrechte (Art. 21 DSGVO) ausgelöst werden. Ein praxisgerechter Einsatz von KI-Systemen wäre dann kaum möglich.

KI-Systeme mit spezifischem Verwendungszweck

Sofern KI-Systeme mit spezifischem Verwendungszweck eingesetzt werden sollen, ist die Prüfung der Rechtsgrundlage im konkreten Einzelfall möglich. Hier wäre es sinnvoll, wenn die KI-Richtline deutlich macht, dass entsprechende KI-Systeme nur für die vorgesehenen Zwecke verwendet werden dürfen.

2 Im Gegensatz zu den häufig anzutreffenden KI-Systemen, die als Software as a Service angeboten werden (z.B. ChatGPT von OpenAI), ist es durchaus möglich, sich KI-Modelle (z.B. Llama 4 von Meta) lokal herunterzuladen und über ein selbst gehostetes Interface (z.B. OpenWebUI) anzusprechen. Auf diese Weise ist es möglich, ein KI-System (z.B. einen KI-Chatbot) lokal zu betreiben. Hierzu ist neben Speicherplatz insbesondere auch die Rechenleistung von speziellen Grafik- bzw. KI-Prozessoren erforderlich.

3 Dies kann zum Beispiel im Rahmen der Bereitstellung eines Kunden-Chatbots der Fall sein, der direkt mit Kunden interagiert und in diesem Rahmen erstmals Daten erhebt.

4 Dies kann zum Beispiel auch die Speicherung von Chathistorien betreffen, die durch eine Interaktion von Beschäftigten mit KI-Chatbots entstehen.

5 Vgl. zum Begriff der Erforderlichkeit auch Gola/Heckmann/Schulz, DS-GVO, Art. 6 RN 20 mit weiteren Nachweisen.

2.1.3 Weitere datenschutzrechtliche Herausforderungen

Fraglich ist, inwiefern der Betreiber von KI-Systemen für das zugrundeliegende Training der KI-Modelle verantwortlich gemacht werden kann und ob eine KI-Richtlinie hierzu etwas regeln muss.

Sofern die Organisation selbst KI-Systeme bzw. KI-Modelle mit personenbezogenen Trainingsdaten anlernt, stellt dies eine Datenverarbeitung dar und bedarf einer eindeutigen Rechtsgrundlage. Hier kommt sowohl die Einwilligung der betroffenen Person, eine Interessenabwägung oder eine vertragliche Ausgestaltung als Rechtsgrundlage für das Training mit personenbezogenen Daten in Betracht. Selbst dann bleiben jedoch im Einzelfall die Umsetzung von Transparenzpflichten und Betroffenenrechten als Herausforderung bestehen. Eine KI-Richtlinie sollte daher das Training mit personenbezogenen Daten ausdrücklich untersagen und sich gesonderte Regelungen vorbehalten, falls doch ein Training mit personenbezogenen Daten geplant ist.

Es stellt sich darüber hinaus die Frage, ob bzw. inwieweit eine Organisation für KI-Modelle verantwortlich gemacht werden kann, die sie selbst nicht trainiert hat, jedoch im Rahmen von KI-Systemen nutzt. Konkret könnte hier die Frage gestellt werden, ob insofern proaktive Prüfpflichten bestehen bzw. sich die Organisation ab einem gewissen Punkt ein rechtswidrig erfolgtes Training mit personenbezogenen Daten durch den Anbieter des KI-Modells zurechnen lassen muss. Ob und welche Zurechnungsmaßstäbe hier gelten, ist umstritten.

Organisationen, die KI-Modelle nutzen wollen, werden in der Regel gar nicht über die notwendigen Informationen verfügen, um die Rechtmäßigkeit des Trainings durch einen Dritten Anbieter im Detail bewerten zu können. Dies gilt erst recht, solange die Transparenzpflichten für Anbieter von KI-Modellen im Hinblick auf die verwendeten Trainingsdaten noch nicht in Kraft getreten sind (Art. 53 Abs. 1 lit. d KI-VO, Art. 113 KI-VO). In diesem Zusammenhang ist die Ansicht des Hamburgischen Beauftragten für Datenschutz und Informationsfreiheit zu begrüßen, wonach sich *„eine eventuelle Rechtswidrigkeit beim Training eines Modells […] nicht auf die Rechtmäßigkeit des Einsatzes eines solchen Modells aus[wirkt]“*[6]. Denn andernfalls könnte kaum ein KI-Modell rechtssicher genutzt werden. Zudem zieht der Hamburgische Beauftragte für Datenschutz und Informationsfreiheit die Grenzen der Verantwortlichkeit des Betreibers des KI-Systems klar beim Input und beim Output der Daten: Hier muss die das KI-System betreibende Organisation den Anforderungen der DSGVO entsprechen. Das dem KI-System zu Grunde liegende KI-Modell bliebe insoweit also eine „Black-Box“.

Allerdings gibt es auch Stimmen, die vertreten, dass eine Rechtswidrigkeit beim Training eines Modells durchaus zur Rechtswidrigkeit des Einsatzes des Modells führen könne. In diesem Rahmen hält es das European Data Protection Board für erforderlich, dass der Betreiber eines KI-Systems eine Bewertung des zugrundeliegenden KI-Modells durchführen muss. Das Detaillevel der Bewertung sei dabei vom Einzelfall abhängig. Insbesondere, wenn die Rechtswidrigkeit des Trainings von einer Aufsichtsbehörde oder einem Gericht festgestellt worden sei, könne dies aber nicht ignoriert werden.[7] Etwas anderes gelte jedoch dann, wenn das KI-Modell nach einem rechtswidrigen Training vom Anbieter anonymisiert wurde – dann würde sich die Rechtswidrigkeit des Trainings nicht mehr auswirken.[8] Welche Voraussetzungen für eine Anonymisierung genau vorliegen müssen, lässt das European Data Protection Board im Detail jedoch offen. Anders als der Hamburgischen Beauftragte für Datenschutz und Informationsfreiheit, geht das European Data Protection Board aber davon aus, dass ein KI-Modell, selbst wenn es die personenbezogenen Daten nicht mehr in ihrer ursprünglichen Form beinhaltet, diese dennoch in gewisser Weise „absorbiert“ haben und durch mathematische Objekte darstellen könne.[9]

Je nachdem, welcher Ansicht man an dieser Stelle folgt, kann es sinnvoll sein, die Whitelist auch auf KI-Modelle zu erstrecken. Da eine Detail-Prüfung der Rechtmäßigkeit des Trainings von KI-Modellen durch Dritte in der Praxis aber kaum möglich sein wird und es eine Vielzahl von KI-Modellen auf dem Markt gibt, wäre es sinnvoller, der Ansicht des Hamburgischen Beauftragten für Datenschutz und Informationsfreiheit zu folgen. Sofern Gerichte oder Aufsichtsbehörden allerdings die Rechtswidrigkeit des Trainings eines KI-Modells feststellen, ist es aus Gründen der Risikovermeidung dennoch ratsam, die Verwendung der betroffenen KI-Modelle bzw. die damit in Zusammenhang stehenden KI-Systeme dann zu untersagen bzw. nicht zu erlauben.

Insofern kann es zum Beispiel relevant sein, das Verfahren der irischen Datenschutzbehörde im Zusammenhang mit dem Training des KI-Modells „Grok“ durch das Unternehmen „X“ zu beobachten.[10] Hier wird derzeit geprüft, ob zum Training in unzulässigerweise Nutzerdaten der Social Media Plattform verwendet wurden.

Im Hinblick auf DeepSeek könnte das Verfahren der italienischen Datenschutzaufsichtsbehörde relevant werden. Diese hat am 30. Januar 2025 festgestellt, dass die Datenverarbeitung des Anbieters gegen datenschutzrechtliche Vorschriften verstößt und der Dienst deshalb in Italien nicht angeboten werden darf.[11] Allerdings scheint sich das Verbot nur auf das Anbieten des KI-Systems DeepSeek zu beschränken. Unklar ist daher, ob auch der Betrieb der unterschiedlichen DeepSeek-KI-Modelle in einer on premise Infrastruktur hiervon betroffen ist, zumal die Entscheidungsgründe der italienischen Aufsichtsbehörde nicht auf die Unrechtmäßigkeit des Trainings der KI-Modelle abstellen, sondern auf Umstände im Zusammenhang mit dem Anbieten des KI-Systems.

6 Der Hamburgische Beauftragte für Datenschutz und Informationsfreiheit, Diskussionspapier: Large Language Models und personenbezogene Daten, Seite 9 (Ziffer III.1), abrufbar unter https://datenschutz-hamburg.de/fileadmin/user_upload/HmbBfDI/Datenschutz/Informationen/240715_Diskussionspapier_HmbBfDI_KI_Modelle.pdf (zuletzt abgerufen am 15.04.2025).

7 Vgl. hierzu näher: European Data Protection Board, Opinion 28/2024 on certain data protection aspects related to the processing of personal data in the context of AI models, Ziffer 3.4.2, RN 129 https://www.edpb.europa.eu/system/files/2024-12/edpb_opinion_202428_ai-models_en.pdf (zuletzt abgerufen am 15.04.2025).

8 Ebenda, Ziffer 3.4.3., RN 135.

9 Ebenda, Ziffer 3.1., RN 31., vgl. auch LfDI Baden-Württemberg, Diskussionspapier: Rechtsgrundlagen im Datenschutz beim Einsatz Künstlicher Intelligenz (Version 2.0), Seite 8 m.w.N.

10 Data Protection Commission, Data Protection Commission Announces commencement of inquiry into X Internet Unlimited Company (XIUC), https://www.dataprotection.ie/en/news-media/latest-news/data-protection-commission-announces-commencement-inquiry-x-internet-unlimited-company-xiuc (zuletzt abgerufen am 15.04.2025).

11 Garante per la protezione die data personali, Provvedimento del 30 gennaio 2025 [10098477], https://www.garanteprivacy.it/web/guest/home/docweb/-/docweb-display/docweb/10098477 (zuletzt abgerufen am 15.04.2025).

Ein ähnliches Verbot sprach die italienische Datenschutzaufsichtsbehörde bereits im März 2023 in Bezug auf ChatGPT aus, hob dies im April aber wieder auf.[12]

Ein Anbieter, der von Datenschutzaufsichtsbehörden sogar mit einem Bußgeld (auch) für das unrechtmäßige Training mit personenbezogenen Daten belegt wurde, ist Clearview AI.[13]

2.1.4 Zwischenfazit

Sofern personenbezogene Daten verarbeitet werden dürfen, muss die Nutzung der KI-Systeme im Einzelfall geprüft werden. Hier sollte der Datenschutzbeauftragte der Organisation hinzugezogen werden, sofern vorhanden. Die Freigabe von KI-Systemen für die Verarbeitung personenbezogener Daten erfordert die Aufnahme des konkreten KI-Systems in eine Whitelist. Eine entsprechende Regelung könnte wie folgt lauten:

- **„Erlaubt sind nur die KI-Systeme, die ausdrücklich von der Organisation freigebegeben wurden. Eine Liste der freigegebenen KI-Systeme wird in Anlage 1 der Richtlinie dargestellt. Sofern nichts anderes geregelt wurde, dürfen keinerlei personenbezogene Daten in diesen KI-Systemen verarbeitet werden. Die Verarbeitung von Daten mit besonderem Vertrauensschutz (§ 203 StGB) ist ebenfalls nur mit einer Genehmigung erlaubt, die sich ausdrücklich auch auf die Verarbeitung dieser Daten erstreckt."**

Insbesondere wenn KI-Systeme mit allgemeinem Verwendungszweck erlaubt werden, sollte in der Richtlinie folgendes aufgenommen werden:

- **„Sofern die Verarbeitung von personenbezogenen Daten für ein bestimmtes KI-System erlaubt ist, dürfen diese nur verarbeitet werden, wenn eine datenschutzrechtliche Rechtsgrundlage für die Datenverarbeitung vorliegt. KI-Systeme dürfen nicht genutzt werden, um automatisierte Einzelentscheidungen zu treffen, von denen Personen betroffen sind, außer dies wird ausdrücklich in der Anlage genehmigt."**

Sofern KI-Systeme mit spezifischem Verwendungszweck zum Einsatz kommen, sollte jede zweckfremde Nutzung ausgeschlossen werden. Das Training von KI-Modellen sollte geregelt werden, z.B. wie folgt:

- **„Sofern in der Anlage 1 für spezielle KI-Systeme bestimmte Zwecke vorgeschrieben sind, darf nur ein zweckentsprechender Einsatz erfolgen. Das Training von KI-Modellen ist vorab mit der für die KI-Koordination zuständigen Person abzustimmen."**

Im Hinblick auf KI-Modelle, die nach Feststellungen von Aufsichtsbehörden oder Gerichten in datenschutzrechtlich unzulässigerweise Trainingsdaten verwendet haben, sollte der Einsatz im Zweifel untersagt bleiben. Dies gilt unabhängig davon, ob die Verarbeitung von personenbezogenen Daten im Zusammenhang mit dem betroffenen KI-Modell vorgesehen ist oder nicht, da die Rechtswidrigkeit bereits das KI-Modell betrifft.

12 Garante per la protezione die data personali, Provvedimento del 30 marzo 2023 [9870832], https://www.garanteprivacy.it/home/docweb/-/docweb-display/docweb/9870832 (zuletzt abgerufen am 15.04.2025).

13 Autoriteit Persoonsgegevens, Dutch DPA imposes a fine on Clearview because of illegal data collection for facial recognition, https://autoriteitpersoonsgegevens.nl/en/current/dutch-dpa-imposes-a-fine-on-clearview-because-of-illegal-data-collection-for-facial-recognition, (zuletzt abgerufen am 15.04.2025).

2.2 Verarbeitung schützenswerter Unternehmensinformationen

Eine KI-Richtline muss neben den datenschutzrechtlichen Anforderungen auch sonstige, nicht personenbezogene Daten in den Blick nehmen. Sollen schützenswerte Unternehmensinformationen verarbeitet werden, empfiehlt sich wiederum ein Whitelist-Ansatz. Denn in diesem Fall muss erneut im Einzelfall geprüft werden, ob etwaige Dienstleister die Anforderungen der Organisation erfüllen.

Hier ist insbesondere das Schutzziel der Vertraulichkeit zu beachten. In diesem Zusammenhang ist es erforderlich, die vertraglichen Regelungen mit etwaigen Dienstleistern zu überprüfen. Dabei kommt es insbesondere darauf an, dass der Anbieter die zweckfremde Verwendung der eingegebenen Daten ausschließt und diese nicht an Dritte weitergibt, oder für eigene Trainingszwecke benutzt. Entsprechende Regelungen finden sich häufig in den Allgemeinen Geschäftsbedingungen bzw. den Terms of Service des Dienstleisters bzw. dem zugrundeliegenden Leistungsvertrag.

Auch in Bezug auf vertrauliche Unternehmensinformationen können weitere Sicherheitsmaßnahmen eine Rolle spielen, die bereits in Ziffer 2.1.1 auszugsweise dargestellt wurden.

Der Umgang mit vertraulichen Unternehmensinformationen könnte in einer KI-Richtlinie wie folgt geregelt werden:

- **„Sofern nichts anderes geregelt wurde, dürfen keinerlei vertrauliche Unternehmensdaten in den in der Anlage 1 genannten KI-Systemen verarbeitet werden."**

Insgesamt erscheint es sinnvoll, vor der Aufnahme von KI-Systemen auf die Whitelist auch den Informationssicherheitsbeauftragten der Organisation zu involvieren, sofern vorhanden. Sofern die Organisation bereits im Rahmen eines ISMS eine Informationsklassifizierung durchgeführt hat, könnten einzelne KI-Systeme auch nur für bestimmte Schutzstufen freigegeben werden (z.B. für vertrauliche Daten, aber nicht für streng vertrauliche Daten).

2.3 Hochrisiko-KI-Systeme

Sofern KI-Systeme genutzt werden sollen, die nach Art. 6 KI-VO als Hochrisiko-KI-Systeme einzustufen sind, ist erneut ein Whitelist-Ansatz erforderlich. Denn im Rahmen des Betriebs solcher KI-Systeme müssen nach dem 2. August 2026 zahlreiche Pflichten eingehalten werden (Art. 8 ff. KI-VO), deren Umsetzung bereits jetzt geplant werden sollte. Sofern Betreiber planen, sich stattdessen auf die Bestandsschutzregeln für bereits eingeführte KI-Systeme zu berufen, muss sichergestellt werden, dass das KI-System nach dem 2. August 2026 keine wesentlichen Änderungen erfährt bzw. sonstige Gründe für die Anwendbarkeit von bereits eingeführten Hochrisiko-KI-Systemen greifen (Art. 111 KI-VO).

Aus diesen Gründen ist die Aufnahme der folgenden Regelung in eine KI-Richtlinie sinnvoll:

- **„Sofern nichts anderes geregelt wurde, dürfen keine Zwecke verfolgt werden, die als hochriskant im Sinne von Art. 6 KI-VO eingestuft werden."**

Eine entsprechende Regelung innerhalb der KI-Richtlinie sollte jedoch näher erläutern, welche Umstände zu einer Klassifizierung als Hochrisiko-KI-System führen können. Hierbei können je nach Organisation branchenspezifische Schwerpunkte gesetzt werden. Branchenübergreifend sollte jedoch immer ein Schwer-

punkt auch auf den Einsatz von KI im Bereich des Arbeitsverhältnisses gesetzt werden und auf die in Anhang III Ziffer 4 KI-VO dargestellten Situationen. Möglicherweise entscheiden sich Organisationen auch bewusst dafür, die Verarbeitung von Personaldaten durch die Personalabteilung unter ein gesondertes Genehmigungserfordernis zu stellen. Insgesamt sollten die Beschäftigten geschult werden, um auch in der Praxis hochriskante KI-Praktiken identifizieren zu können und diese nicht aus Versehen bei KI-Systemen mit allgemeinem Verwendungszweck anzuwenden. [14]

3 Weitere allgemeine Regelungen

Im Rahmen einer KI-Richtlinie können weitere allgemeine Regelungen aufgenommen werden. Diese gelten dann für alle KI-Systeme auf der Whitelist, unabhängig davon, ob diese für die Verarbeitung personenbezogener Daten oder sonstiger schützenswerter Unternehmensdaten zugelassen sind und unabhängig von ihrer Klassifizierung als Hochrisiko-KI-System.

14 So wäre es ansonsten denkbar, dass Beschäftigte ein KI-System mit allgemeinem Verwendungszweck (z.B. ChatGPT) für hochriskante KI-Praktiken verwenden (z.B. im Rahmen der Bewerberauswahl).

3.1 Verbotene KI-Praktiken nach Art. 5 KI-VO

Grundsätzlich ist es empfehlenswert in der KI-Richtlinie noch einmal ausdrücklich auf die in Art. 5 KI-VO geregelten verbotenen KI-Praktiken hinzuweisen. Hierzu könnte das Folgende aufgenommen werden:

- **„Es dürfen keine verbotenen KI-Praktiken im Sinne von Art. 5 KI-VO ausgeführt werden."**

Wie bereits im Rahmen der Regelungen zu Hochrisiko-KI-Systemen (Ziffer 2.3) wird es auch bei Aufnahme dieser Regelung sinnvoll sein, die entsprechenden verbotenen KI-Praktiken noch einmal näher zu erläutern und diese auch zum Gegenstand einer Schulung zu machen.

3.2 Transparenzpflichten nach Art. 50 KI-VO

Im Hinblick auf die Transparenzpflichten kann es für Betreiber von KI-Systemen sinnvoll sein, folgende Regelung in die KI-Richtline aufzunehmen:

- **„Ab dem 2. August 2026 gelten bestimmte Transparenzpflichten. Insbesondere wenn Texte erzeugt werden, die ohne weitere Redaktion veröffentlicht werden sollen, um die Allgemeinheit zu informieren, können Transparenzpflichten entstehen. Transparenzpflichten können auch bei der Erzeugung**

Springer Vieweg springer.com/informatik

Künstliche Intelligenz



U. Barthelmeß, U. Furbach
Künstliche Intelligenz aus ungewohnten Perspektiven
Ein Rundgang mit Bergson, Proust und Nabokov
2019, X, 190 S. 18 Abb., 10 Abb. in Farbe. Brosch.
€ (D) 29,99 | € (A) 30,83 | *CHF 33.50
ISBN 978-3-658-24569-6
€ 22,99 | *CHF 26.50
ISBN 978-3-658-24570-2 (eBook)



C. Beierle, G. Kern-Isberner
Methoden wissensbasierter Systeme
Grundlagen, Algorithmen, Anwendungen
6., überarb. Aufl. 2019, XVIII, 564 S. 165 Abb.
Mit Online-Extras. Brosch.
€ (D) 39,99 | € (A) 41,11 | *CHF 44.50
ISBN 978-3-658-27083-4
€ 29,99 | *CHF 35.50
ISBN 978-3-658-27084-1 (eBook)

Ihre Vorteile in unserem Online Shop:
Über 280.000 Titel aus allen Fachgebieten | eBooks sind auf allen Endgeräten nutzbar | Kostenloser Versand für Printbücher weltweit

€ (D): gebundener Ladenpreis in Deutschland, € (A): in Österreich. * : unverbindliche Preisempfehlung. Alle Preise inkl. MwSt.

Jetzt bestellen auf springer.com/informatik oder in der Buchhandlung

Part of SPRINGER NATURE

A85542

oder der Manipulation von Bild-, Ton- oder Videoinhalten entstehen, nicht nur in Bezug auf Personen, sondern auch in Bezug auf Gegenstände, Orte, Einrichtungen oder Ereignisse (sog. Deepfakes).“

Die entsprechenden Anforderungen folgen aus Art. 50 Abs. 4 KI-VO. Auch hier ist eine nähere Erläuterung in der Richtlinie bzw. die Aufnahme der entsprechenden Themen in eine Schulung sinnvoll. Dabei sollte den Beschäftigten auch erklärt werden, was Deepfakes sind (Art. 3 Nr. 60 KI-VO).

Die weiteren Transparenzanforderungen des Art. 50 KI-VO müssen nicht zwingend im Rahmen der allgemeinen Regelungen einer KI-Richtlinie aufgenommen werden, denn sie gelten entweder nur für die Anbieter von KI-Systemen (also nicht für Betreiber) oder setzen die Verarbeitung personenbezogener Daten voraus. Im letzten Fall muss ohnehin eine Einzelfallprüfung erfolgen. Im Hinblick auf die Gewährleistung der Erfüllung der Anbieterpflichten kann folgende Regelung aufgenommen werden:

- **„Die Entwicklung von KI-System ist vorab mit der für die KI-Koordination zuständigen Person abzustimmen.“**

3.3 KI-Kompetenz

Art. 4 KI-VO schreibt vor, dass Betreiber von KI-Systemen sicherstellen müssen, dass unter anderem alle Beschäftigten, denen KI-Systeme zur Verfügung gestellt werden sollen, über ein angemessenes Maß an KI-Kompetenz verfügen müssen. In diesem Rahmen wird eine bedarfsgerechte Schulungsplanung in der Regel unumgänglich sein. Da die entsprechende Pflicht bereits seit dem 2. Februar 2025 gilt, empfiehlt sich die Aufnahme folgender Regelung sowie die Planung der entsprechenden Schulungsprozesse:

- **„KI-Systeme dürfen nur von Beschäftigten genutzt werden, die vorher zumindest an einer allgemeinen KI-Grundkompetenzschulung teilgenommen haben.“**

Je nach eingesetztem KI-System können neben der Grundkompetenzschulung weitere Schulungsmaßnahmen erforderlich sein.

3.4 Ansprechpartner

Eine KI-Richtline sollte immer einen Ansprechpartner benennen, der für Fragen bereitsteht und sich um die Einhaltung der erforderlichen Prozesse kümmert. Eine entsprechende Regelung könnte wie folgt lauten:

- **„Bei allen Fragen rund um KI oder zu dieser Richtlinie steht die für die KI-Koordination zuständige Person zur Verfügung. Eine Erweiterung bzw. Veränderung der Anlage 1 erfolgt ebenfalls stets durch die KI-Koordination zuständige Person, die ggf. weitere Personen oder relevante Funktionsträger beteiligt.“**

Wer an dieser Stelle von der Organisation benannt wird, ist gesetzlich nicht vorgeschrieben. Insbesondere eine Regelung, die die Benennung eines „KI-Beauftragten“ vorschreiben würde, gibt es nicht. Gleichwohl ist es sinnvoll, abseits der KI-Richtlinie, eine Person als ersten Ansprechpartner für alle Fragen rund um den Einsatz von KI zu benennen und sich in diesem Rahmen auch um die weitere Aufgabenverteilung und Prozessgestaltung Gedanken zu machen.

3.5 Sonstige Punkte

In eine KI-Richtlinie können weitere Punkte aufgenommen werden, zum Beispiel die Ziele der Organisation, der Geltungsbereich, Regelungen zum Umgang mit Verstößen, eine verständliche Definition von KI in Abgrenzung zu klassischer Software mit herkömmlichen Programmierungsansätzen, weitere Regelungen zu anderen Rechtsgebieten (wie dem Umgang mit Urheberrechten) oder eine Regelung zur Zulässigkeit bzw. Unzulässigkeit der Privatnutzung bestimmter KI-Systeme.

Geprüft werden kann auch, inwieweit die Mitarbeitervertretung zu involvieren ist und, je nach Organisation, ob eine Ausgestaltung als Dienstanweisung oder als Vereinbarung Sinn macht.

3.6 Umgang mit der Whitelist

In Abhängigkeit der KI-Strategie der Organisation kann der Umgang mit der Whitelist der beispielhaft vorgeschlagenen KI-Richtlinie variieren.

Sofern eine Organisation zum Beispiel allgemein den Umgang mit KI-Bildgeneratoren freigegeben hat, diese jedoch nicht in jedem Einzelfall prüfen möchte, könnte der Eintrag auf der Whitelist wie folgt aussehen:

- *Genehmigtes KI-System: KI-Systeme, die Bild- oder Videoinhalte erzeugen oder manipulieren.*
- *Personenbezogene Daten: Nicht zugelassen*
- *Schützenswerte Unternehmensinformationen: Nicht zugelassen*
- *Bemerkungen: Sofern Deepfakes erzeugt werden, muss dies ab dem 2. August 2026 kenntlich gemacht werden.*

Sofern ein Chatbot geprüft wurde und alle Anforderungen der Organisation erfüllt werden, könnte dieser beispielhaft wie folgt aufgenommen werden:

- *Genehmigtes KI-System: ChatGPT Enterprise*
- *Personenbezogene Daten: Zugelassen*
- *Schützenswerte Unternehmensinformationen: Zugelassen bis einschließlich Schutzstufe „vertraulich“*
- *Bemerkungen: Bitte nur nach Anmeldung mit den von der Organisation zur Verfügung gestellten Anmeldedaten nutzen (nicht die kostenlose Version ohne Anmeldung).*

4 Fazit

Um einen Kontrollverlust zu vermeiden, sollten Organisationen interne Regelungen zum Einsatz von Künstlicher Intelligenz treffen. Dabei gehen die Regelungsbedarfe weit über den Datenschutz hinaus. Der vorliegende Beitrag gibt eine erste Hilfestellung für einen möglichen Regelungsansatz und stellt Punkte dar, die beachtet werden können.