FAQ
Künstliche Intelligenz

(Stand: März 2026)

Systeme Künstlicher Intelligenz (KI) gehören inzwischen weitgehend zum Alltag. In diesen FAQ klären wir über die (datenschutz-)rechtlichen Aspekte bei der Nutzung von Künstlicher Intelligenz auf. Ein Glossar erläutert zudem einige zentrale Begriffe.

Inhaltsverzeichnis

1. Was regelt die KI-Verordnung? Welche Überschneidungen gibt es mit bestehenden Datenschutzgesetzen?

Am 1. August 2024 ist die europäische Verordnung über **Künstliche Intelligenz** (KI-Verordnung) in Kraft getreten. Sie zielt darauf ab, die verantwortungsvolle Entwicklung und Verwendung künstlicher Intelligenz in der EU zu fördern, indem sie den Einsatz von **KI-Modellen** und **KI-Systemen** in Europa regelt. Zielsetzung dieser Regulierung ist es, auch im KI-Zeitalter den Schutz der Grundrechte der Menschen zu gewährleisten. Zugleich soll die Einführung von menschenzentrierten und vertrauenswürdigen KI-Modellen und KI-Systemen gefördert werden. Diese Ziele stehen grundsätzlich gleichberechtigt nebeneinander.

Die KI-VO enthält bis auf zwei Ausnahmen (Art. 10 Abs. 5 und Art. 59 KI-VO) keine Regelungen zum Schutz personenbezogener Daten. Es gelten gemäß Art. 2 Abs. 7 KI-VO für die Verarbeitung personenbezogener Daten bei der Entwicklung und dem Betrieb von KI-Modellen und KI-Systemen die DSGVO, die JI-Richtlinie und die ePrivacy-Richtlinie sowie die entsprechenden nationalen Umsetzungen im Bundesdatenschutzgesetz und im Telekommunikation-Digitale-Dienste-Datenschutz-Gesetz (TDDDG) (Art. 2 Abs. 7 KI-VO).

2. Ab welchem Zeitpunkt gelten die Vorschriften der KI-VO? Gibt es Übergangsfristen?

Die KI-VO ist bereits im August 2024 in Kraft getreten und als europäische Verordnung unmittelbar anwendbar. Es gibt allerdings abgestufte Fristen für den Geltungsbeginn:

- 2. Februar 2025: allgemeine Bestimmungen wie insbesondere die KI-Kompetenz (Art. 4 KI-VO) und verbotene KI-Praktiken (Art. 5 KI-VO)
- 2. August 2025: Notifizierung von Hochrisiko-KI-Systemen, Vorschriften für **KI-Modelle mit allgemeinem Verwendungszweck**, KI-Governance, Sanktionen, Vertraulichkeitspflicht zuständiger Behörden
- 2. August 2026: alle weiteren Regelungen außer die Einstufung von Hochrisiko-KI-Systemen (Art. 6 Abs. 1 KI-VO) sowie die entsprechenden Pflichten
- 2. August 2027: Uneingeschränkte Geltung aller Vorschriften.

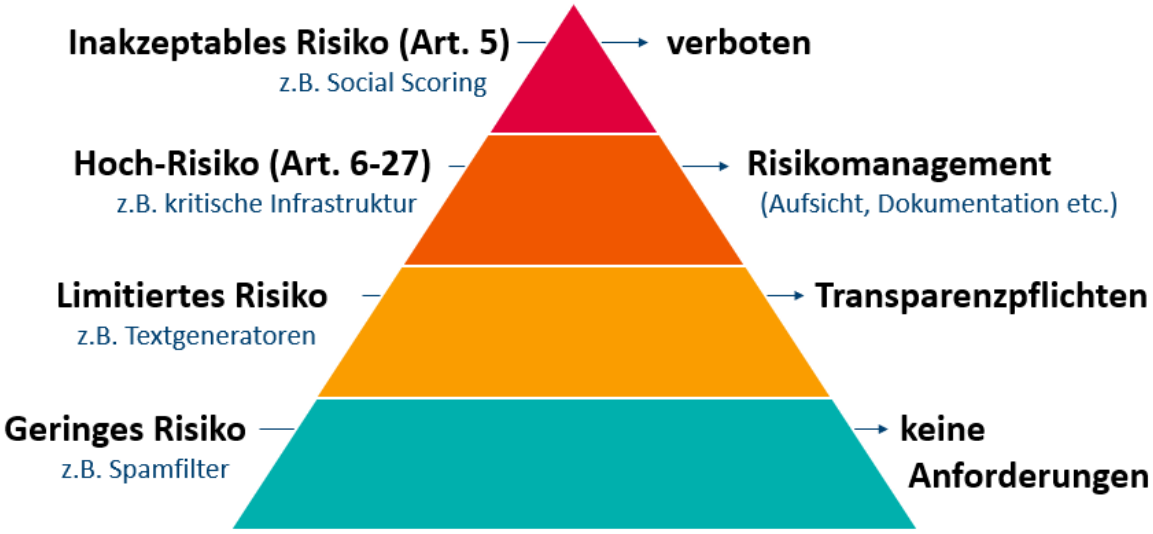
3. Wer hat die KI-VO zu beachten – für wen gilt sie?

Die KI-VO gilt vor allem für alle Entwickler und Betreiber von KI-Modellen und -Systemen, die in Europa zum Einsatz kommen. Darüber hinaus adressiert sie Anbieter, Einführer und Händler von KI-Systemen, Produkthersteller, die KI-Systeme zusammen mit ihrem Produkt unter ihrem eigenen Namen oder ihrer Handelsmarke in Verkehr bringen sowie Bevollmächtigte von Anbietern (Art. 2 Abs. 1 KI-VO). Die KI-VO gilt nicht, wenn natürliche Personen KI-Systeme im Rahmen einer ausschließlich persönlichen und nicht beruflichen Tätigkeit verwenden (Art. 2 Abs. 10 KI-VO). Werden hierbei personenbezogene Daten verarbeitet, findet sich in der DSGVO eine vergleichbare Ausnahmeregelung. Die DSGVO ist nicht anwendbar, wenn natürliche Personen zur Ausübung ausschließlich persönlicher oder familiärer Tätigkeiten personenbezogene Daten verarbeiten (Art. 2 Abs. 2 lit. c DSGVO).

4. Welche Risikostufen unterscheidet die KI-VO für KI-Systeme und welche Pflichten ergeben sich daraus?

Die KI-VO verfolgt einen risikobasierten Ansatz. Je höher das Risiko, desto strengere Anforderungen müssen erfüllt werden. Unter Risiko wird gem. Art. 3 Nr. 2 KI-VO die Kombination aus der Wahrscheinlichkeit des Auftretens eines Schadens und der Schwere dieses Schadens verstanden. Die in Art. 5 KI-VO aufgezählten Praktiken sind mit einem so hohen Risiko verbunden, dass sie verboten sind. Auf der nächsten Stufe stehen Hochrisiko-KI-Systeme im Sinne von Art. 6 KI-VO. Vor deren Einsatz hat der Anbieter u.a. ein Konformitätsbewertungsverfahren gemäß Art. 43 KI-VO durchzuführen.

Da beim Einsatz von KI-Systemen, die für die direkte Interaktion mit natürlichen Personen bestimmt sind, z.B. Chatbots, Transparenzrisiken bestehen, haben Anbieter solcher Systeme Transparenzpflichten gem. Art. 50 KI-VO zu erfüllen. Bezüglich KI-Systemen, die nicht in einer dieser drei Risikostufen fallen, bestehen keine besonderen Verpflichtungen. KI-Modelle mit allgemeinem Verwendungszweck unterliegen besonderen Pflichten gem. Art. 53 KI-VO. Diese bestehen unabhängig von ggf. bestehenden Pflichten aus der Eigenschaft als Hochrisiko KI-System oder Transparenzpflichten gem. Art. 50 KI-VO.



5. Welche Praktiken sind im KI-Bereich verboten?

Art. 5 KI-VO definiert eine Reihe von Praktiken als „verbotene Praktiken" im KI-Bereich. Diese Praktiken sind mit einem so hohen Risiko für die Betroffenen verbunden, dass diese untersagt sind. Dazu zählt beispielsweise der Einsatz von KI-Systemen zur manipulativen oder täuschenden Verhaltensbeeinflussung oder die Verwendung von KI-Systemen, die Datenbanken zur Gesichtserkennung durch das ungezielte Auslesen von Gesichtsbildern aus dem Internet oder von Überwachungsaufnahmen erstellen oder erweitern. Die Europäische Kommission hat **Leitlinien** zu diesem Thema veröffentlicht, in denen eine weitere Konkretisierung erfolgt und Beispiele genannt werden.

6. Welche Kriterien sind maßgeblich, damit ein KI-System als Hochrisiko-KI-System gilt?

Art. 6 KI-VO definiert, welche KI-Systeme als Hochrisiko-KI-Systeme im Sinne der KI-VO gelten. Das sind KI-Systeme, die dem Produktsicherheitsrecht unterliegen (Art. 6 Abs. 1 KI-VO) und KI-Systeme, die im Anhang III der KI-VO genannt werden, beispielsweise biometrische Fernidentifizierungssysteme oder KI-Systeme, die bestimmungsgemäß für die Kreditwürdigkeitsprüfung und Bonitätsbewertung natürlicher Personen verwendet werden sollen. Für die Entwicklung und den Betrieb von Hochrisiko-

KI-Systemen gelten die in Abschnitt 3 KI-VO (Art. 6 bis 49 KI-VO) geregelten Anforderungen und Pflichten.

7. Die KI-Verordnung enthält Pflichten für „Anbieter“ und „Betreiber“. Was ist der Unterschied?

Gemäß Art. 3 Nr. 3 KI-VO ist ein „Anbieter“ eine natürliche oder juristische Person, Behörde, Einrichtung oder sonstige Stelle, die ein KI-System oder ein KI-Modell mit allgemeinem Verwendungszweck entwickelt oder entwickeln lässt und es unter ihrem eigenen Namen oder ihrer Handelsmarke in Verkehr bringt oder das KI-System unter ihrem eigenen Namen oder ihrer Handelsmarke in Betrieb nimmt, sei es entgeltlich oder unentgeltlich. Anbieter von KI-Systemen oder KI-Modellen sind Adressaten umfangreicher in der KI-VO enthaltener Pflichten.

Gemäß Art. 3 Nr. 4 KI-VO ist ein „Betreiber“ eine natürliche oder juristische Person, Behörde, Einrichtung oder sonstige Stelle, die ein KI-System in eigener Verantwortung verwendet, es sei denn, das KI-System wird im Rahmen einer persönlichen und nicht beruflichen Tätigkeit verwendet. Falls ein Unternehmen beispielsweise ein KI-System ~~mit einem~~in der Form eines Chatbot für den Kundendialog einsetzt, ist es Betreiber eines KI-Systems. Ob das Unternehmen das KI-System selbst entwickelt hat, ist für die Betreibereigenschaft unbeachtlich.

8. Welche Anforderungen stellt die KI-VO an Betreiber von Hochrisiko-KI-Systemen hinsichtlich einer „Grundrechte-Folgenabschätzung“?

Gemäß Art. 27 KI-VO sind bestimmte Betreiber von Hochrisiko-KI-Systemen zur Durchführung einer Grundrechte-Folgenabschätzung verpflichtet. Deren Zielsetzung ist es gemäß Erwägungsgrund 96 KI-VO, dass der Betreiber die spezifischen Risiken für die Rechte von Einzelpersonen oder Gruppen von Einzelpersonen, die wahrscheinlich betroffen sein werden, ermittelt. Weiterhin sollen Maßnahmen ermittelt werden, die im Falle eines Eintretens dieser Risiken zu ergreifen sind.

Art. 27 Abs. 4 KI-VO regelt das Verhältnis zur Datenschutz-Folgenabschätzung gemäß Art. 35 DSGVO und stellt klar, dass die Grundrechte-Folgenabschätzung die Datenschutz-Folgenabschätzung

ergänzt. Für die Praxis folgt daraus erstens, dass die Grundrechte-Folgenabschätzung zusätzlich zur Datenschutz-Folgenabschätzung durchzuführen ist, soweit die erforderlichen Prüfungen nicht bereits im Rahmen der Datenschutz-Folgenabschätzung durchgeführt wurden. Zweitens können die von der Datenschutz-Folgenabschätzung bekannten Methoden und Schritten weitgehend auf die Grundrechte-Folgenabschätzung übertragen werden.

9. Wie müssen Anbieter und Betreiber sicherstellen, dass ihr Personal über ausreichende KI-Kompetenz verfügt?

Gemäß Art. 3 Nr. 56 KI-VO sind unter „KI-Kompetenz" die Fähigkeiten zu verstehen, „die Kenntnisse und das Verständnis, die es Anbietern, Betreibern und Betroffenen unter Berücksichtigung ihrer jeweiligen Rechte und Pflichten im Rahmen dieser Verordnung ermöglichen, KI-Systeme sachkundig einzusetzen sowie sich der Chancen und Risiken von KI und möglicher Schäden, die sie verursachen kann, bewusst zu werden".

Art. 4 KI-VO verpflichtet Anbieter und Betreiber von KI-Systemen, „sicherzustellen, dass ihr Personal und von ihnen beauftragte Personen, die mit dem Betrieb und der Nutzung von KI-Systemen befasst sind, über ausreichende KI-Kompetenzen verfügen müssen". In der Praxis sollte bei der Einstellung von Beschäftigten auf deren technische und juristische Qualifikation geachtet werden. Zudem sollten geeignete Schulungen und Fortbildungen der Beschäftigten erfolgen. Die Bundesnetzagentur hat umfangreiches **Informationsmaterial** zur KI-Kompetenz veröffentlicht.

Wenn in KI-Systemen personenbezogene Daten verarbeitet werden, müssen Anwenderinnen und Anwender im Rahmen der KI-Kompetenz auch zu spezifischen datenschutzrechtlichen Risiken geschult werden. ~~Hierzu gehört unter anderem das Risiko der Ausgabe von falschen personenbezogenen Daten sowie über das Risiko von Bias und **Halluzinationen** in Ausgabedaten und das Verbot einer automatisierten Entscheidung.~~Hierzu gehören unter anderem die Gefahr

- der Ausgabe von falschen personenbezogenen Daten,
- eines Bias im Einhebungsprozess der Trainingsdaten, der sich in den Ausgabedaten widerspiegelt,
- von **Halluzinationen** in Ausgabedaten

sowie das Verbot einer automatisierten Entscheidung.

10. Was ist vor dem Einsatz von Large-Language-Modellen im Büroalltag zu beachten?

Vor dem Einsatz sind folgende ~~rechtliche~~technische und organisatorische Maßnahmen (~~TOM~~toM) sicherzustellen:

1. Risikoeinstufung durchführen im Sinne der KI-VO und Umsetzung der daraus resultierenden rechtlichen Anforderungen

2. Klare Definition der zulässigen Anwendungsbereiche, Zwecke und konkreten Einsatzmodalitäten des KI-Systems

3. Datenschutzmaßnahmen bei Verarbeitung personenbezogener Daten:

- Sicherstellen, dass das eingesetzte KI-Modell nicht rechtswidrig trainiert wurde
- Prüfen, ob eine geeignete Rechtsgrundlage für die Verarbeitung personenbezogener Daten vorhanden ist
- Geeignete ~~TOM~~technische und organisatorische Maßnahmen (TOM) identifizieren und implementieren
- Einhaltung der Datenschutz-Grundsätze gewährleisten (z.B. Datenminimierung, Zweckbindung)
- Sicherstellen, dass Betroffenenrechte gewährt werden können (z.B. Auskunft, Löschung)
- Schwellwertprüfung für Datenschutz-Folgenabschätzung (DSFA), und falls erforderlich, eine DSFA durchführen

4. Gegebenenfalls Grundrechte-Folgenabschätzung durchführen

5. KI-Kompetenz bei den Beschäftigten sicherstellen (z.B. durch Schulungen)

6. Erstellung einer Dienstanweisung oder Betriebsanweisung, die die Bedingungen und Vorgaben für den Einsatz des KI-Systems regelt.

11. Welche datenschutzrechtlichen Vorgaben gibt es beim Einsatz von KI-Systemen?

Beim Einsatz von KI-Systemen sind folgende Vorgaben zu beachten:

1. Alle Vorschriften der **Datenschutz-Grundverordnung** (DSGVO)

2. Je nach Anwendungsbereich sind weitere spezifische Datenschutzgesetze zu berücksichtigen (z. B. das Bundesdatenschutzgesetz und das Telekommunikation-Digitale-Dienste-Datenschutz-Gesetz)

- **Insbesondere Wahrung der Datenschutz-Grundsätze (Art. 5 DSGVO), unter anderem** Rechtmäßigkeit, Verarbeitung nach Treu und Glauben, Transparenz, Zweckbindung, Datenminimierung und Richtigkeit

3. **Gewährleistung der Betroffenenrechte (z.** B. Auskunft, Berichtigung, Löschung, Widerspruch)

12. Weil sie KI-Chatbots nicht auf ihrem Arbeitsrechner benutzen dürfen, weichen viele Beschäftigte auf ihre Privatgeräte aus, um berufliche Aufgaben zu erledigen. Was sind die Risiken?

Diese unautorisierte Nutzung von KI-Systemen im beruflichen Kontext wird auch als „Schatten-KI" verstanden. Dies ist der Fall, wenn Mitarbeiterinnen oder Mitarbeiter im Web verfügbare KI-Systeme wie ChatGPT, Gemini oder Claude nutzen, ohne dass diese in die IT-Infrastruktur der Behörde oder des Unternehmens integriert sind und die Nutzung genehmigt wurde.

Dies birgt vor allem ~~das Risiko~~die Gefahr, gegen die gesetzlichen Vorgaben sowohl der KI-Verordnung als auch der Datenschutz-Grundverordnung zu verstoßen. Je nachdem, in welche Risikostufen die Schatten-KI einzuordnen ist, unterscheidet sich das Risiko von Gesetzesverstößen. Ein Verstoß gegen die Pflicht zur Gewährleistung von KI-Kompetenz steht immer im Raum. Datenschutzrechtlich besteht vor allem ~~das Risiko~~die Gefahr, dass personenbezogene Daten unkontrolliert ohne Rechtsgrundlage verarbeitet werden. Um solche Verstöße gegen die DSGVO zu vermeiden, sollte der Einsatz von Schatten-KI von den Verantwortlichen nicht toleriert werden. So kann die Nutzung von Schatten-KI zum Beispiel in einer Dienstanweisung adressiert und untersagt werden. Ergänzend sollten Behörden und Unternehmen bewusst entscheiden, welche KI-Systeme in der Organisation unter welchen Bedingungen eingesetzt werden sollen.

13. Was ist zu tun, damit der Einsatz von KI-Systemen nicht gegen das Verbot automatisierter Entscheidungen verstößt?

Gemäß Artikel 22 DSGVO haben Betroffene das Recht, nicht einer ausschließlich auf einer automatisierten Verarbeitung beruhenden Entscheidung unterworfen zu werden, die ihr gegenüber rechtlicher Wirkung entfaltet oder sie in ähnlicher Weise erheblich beeinträchtigt. Dies sind beispielsweise Entscheidungen über den Erlass eines Verwaltungsakts, Entscheidungen über Einstellungen und Beförderungen von Mitarbeiterinnen und ~~Mitarbeiter~~Mitarbeitern oder Entscheidungen über die Kreditwürdigkeit. In solchen Bereichen darf ein KI-System zwar Vorschläge oder Empfehlungen machen, die Entscheidung muss aber stets bei einer natürlichen Person liegen. Es muss in jedem Fall eine substanzielle Prüfung erfolgen; ein bloßes Abnicken der Prüfungsvorschläge eines KI-Systems ist nicht ausreichend.

14. Wie kann ich als Betreiber einer Webseite verhindern, dass die dort veröffentlichten Daten für das Training von KI-Modellen verwendet werden, wie dies insbesondere bei der Entwicklung der LLMs erfolgt ist?

Betreiber von Webseiten sollten sich bewusst sein, dass auf ihren Webseiten veröffentlichte personenbezogene Daten mittels Web-Scraping verarbeitet und insbesondere für ein Training von KI-Modellen genutzt werden. Im Rahmen ihrer datenschutzrechtlichen Verantwortlichkeit sollten sie technische und organisatorische Maßnahmen ergreifen, um dies erforderlichenfalls zu verhindern.

Es sollte z.B. darauf verzichtet werden, personenbezogene Daten in öffentlichen Bereichen zur Verfügung zu stellen, wenn stattdessen die Einrichtung geschlossener Bereiche für einen bekannten Benutzerkreis möglich ist, auf deren Quelltext Dritten keinen Zugriff haben. Außerdem kann die robots.txt-Datei dazu genutzt werden, Programmen zum Web-Scraping mitzuteilen, dass auf der Webseite kein Web-Scraping erfolgen soll.

Glossar

Künstliche Intelligenz (KI)

Es gibt keine allgemeingültige Definition des Begriffs „künstlicher Intelligenz“ (KI). Grundsätzlich ist KI ein Teilgebiet der Informatik und bezeichnet die Fähigkeit einer Maschine, menschliche Fähigkeiten, wie logisches Denken, Lernen, Planen und Kreativität zu imitieren (siehe dazu auch Europäisches Parlament).

Die KI-Verordnung enthält in Art. 3 Nr. 1 KI-VO eine rechtliche Definition des Begriffs KI-System, der in Erwägungsgrund 12 KI-VO erläutert und in den **Leitlinien der Kommission zur Definition eines Systems der künstlichen Intelligenz** weiter präzisiert wird. Danach beinhaltet die Definition die folgenden sieben Hauptelemente:

1. ein maschinengestütztes System,
2. das für einen in unterschiedlichem Grade autonomen Betrieb ausgelegt ist,
3. das nach seiner Betriebsaufnahme anpassungsfähig sein kann,
4. das für explizite oder implizite Ziele
5. aus den erhaltenen Eingaben ableitet,
6. wie Ausgaben wie etwa Vorhersagen, Inhalte, Empfehlungen oder Entscheidungen erstellt werden, die
7. physische oder virtuelle Umgebungen beeinflussen können.

Zusammengefasst ist ein KI-System:

1. maschinengestützt,
2. autonom und
3. in der Lage, aus Input Output abzuleiten, der Auswirkungen auf die physische oder virtuelle Umgebung hat.

Ein KI-System kann nach der Betriebsaufnahme anpassungsfähig sein.

KI-Systeme

Art. 3 Nr. 1 KI-VO definiert ein KI-System „als ein maschinengestütztes System, das für einen in unterschiedlichem Grad 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".

Beispiele für KI-Systeme sind KI-Chatbots wie ChatGPT oder Gemini, KI-Bildgeneratoren wie Stable Diffusion oder Dell-E, KI-basierte Übersetzungsplattformen wie DeepL oder Google Translator. „Herzstück" eines KI-Systems ist ein oder sind mehrere KI-Modelle. Daneben gibt es weitere Bestandteile wie beispielsweise die Benutzeroberfläche oder Schnittstellen zu Datenbanken.

KI-Modelle

Die KI-Verordnung enthält keine Legaldefinition des Begriffes „KI-Modell", sondern setzt diesen beim Begriff „KI-Modell mit allgemeinem Verwendungszweck" voraus. Ein KI-Modell ist das „Herzstück" eines ~~generativen~~ KI-Systems, in dem als Reaktion auf einen Prompt oder einen sonstigen Input autonom ~~Ergebnisse~~Output, beispielsweise ~~neue~~ Texte oder Bilder, erstellt werden ~~können~~kann. Bekannte KI-Modelle für Textgenerierung sind beispielsweise Claude Opus 4.5, Gemini 3.1 Pro, GPT 5.5, DeepSeek R1 und Grok 4.1 Fast. Midjourney V6.1 und ~~Aurora~~Stable Diffusion 3.5 Large sind KI-Modelle, die für Bildgenerierung genutzt werden können.

KI-Modelle mit allgemeinem Verwendungszweck

Art. 3 Nr. 63 KI-VO definiert ein KI-Modell mit allgemeinem Verwendungszweck als ein KI-Modell, das „eine erhebliche allgemeine Verwendbarkeit aufweist und in der Lage ist, unabhängig von der Art und Weise seines Inverkehrbringens ein breites Spektrum unterschiedlicher Aufgaben kompetent zu erfüllen, und das in eine Vielzahl nachgelagerter Systeme oder Anwendungen integriert werden kann". Umfasst werden auch Fälle, in denen ein solches KI-Modell mit einer großen Datenmenge unter umfassender Selbstüberwachung trainiert wird.

Ausgenommen sind KI-Modelle, die vor ihrem Inverkehrbringen für Forschungs- und Entwicklungstätigkeiten oder die Konzipierung von Prototypen eingesetzt werden. Erwägungsgrund 99 KI-VO nennt als Beispiel für KI-Modelle mit allgemeinem Verwendungszweck große generative KI-Modelle, da sie eine flexible Erzeugung von Inhalten ermöglichen, etwa in Form von Text-, Audio-, Bild- oder Videoinhalten, die leicht ein breites Spektrum unterschiedlicher Aufgaben umfassen können.

Generative KI

Der Begriff generative KI bezieht sich auf ~~computergestützte~~maschinengestützte Technologie, die in der Lage ist, ~~aus~~auf der Grundlage von Trainingsdaten in Kombination mit Prompts von Nutzern oder sonstigen Input scheinbar neue, aussagekräftigen Inhalt zu erzeugen, wie Texte, Bilder oder Audioinhalte.[1] ~~.~~ Generative KI zeichnet sich dadurch aus, dass als Reaktion auf Nutzereingaben ~~neue~~ Inhalte, beispielsweise Texte oder Bilder, erzeugt werden können. ~~Generative~~Große generative KI-Modelle sind KI-Modelle mit allgemeinem Verwendungszweck im Sinne von Art. 3 Nr. 63 KI-VO, für die gemäß Art. 53 KI-VO besondere Pflichten gelten.

Large-Language-Modelle

Large-Language-Modelle (LLMs) sind dem Bereich der generativen KI-Modelle zuzuordnen. LLMs ~~generieren~~können dafür genutzt werden, um Texte auf Basis stochastischer Korrelation zu generieren, die sie während der Entwicklung mit sehr großen Mengen Trainingsdaten gelernt haben.[2] Aufgrund von Wahrscheinlichkeiten wird die nächste Wortfolge vorhergesagt, um einen Text von hoher sprachlicher Qualität zu erzeugen. Beispiele für LLMs sind GPT-5, Mistral 7B oder Grok 4.1.

Halluzinieren von Large-Language-Modellen

Beim Einsatz von Large-Language-Modellen zeigt sich gelegentlich das Phänomen des sogenannten Halluzinierens. Darunter wird die Ausgabe erfundener oder falscher Informationen verstanden, die

[1] Feuerriegel, S., Hartmann, J., Janiesch, C., & Zschech, P. (2024). Generative AI. Business and Information Systems Engineering, 66(1), 111-126. https://doi.org/10.1007/s12599-023-00834-7

[2] BSI, Generative KI-Modelle- Chance und Risiken für Industrie und Behörden, S. 7.

oftmals auf den ersten Blick plausibel erscheinen, oder qualitativ hochwertig formuliert.[3] Die Ursache dieses Phänomens ist, dass die Ausgabedaten nicht auf in Datenbanken gespeicherten Fakten, sondern auf ~~statistische~~statistischen Wahrscheinlichkeiten beruhen. Sofern die Ausgabedaten beim Halluzinieren personenbezogene Daten enthalten, kann dadurch gegen den Grundsatz der Datenrichtigkeit (Art. 5 Abs. 1 lit. d DSGVO) verstoßen werden.

Prompt

Ein Prompt ist die Eingabeaufforderung des Benutzers an ein (generatives) KI-System, auf dessen Grundlage dieses eine Ausgabe erstellt. Wenn im Prompt personenbezogene Daten enthalten sind, stellt die Eingabe des Prompts eine Verarbeitung im Sinne von Art. 4 Nr. 2 DSGVO dar, für die eine Rechtsgrundlage erforderlich ist.

Web-Scraping

Unter Web-Scraping wird das automatisierte Extrahieren von Daten aus Webseiten mit Hilfe von spezieller Software verstanden. Diese Technik kann dazu verwendet werden, um Daten für das Training von KI-Modellen zu gewinnen. Sofern dabei personenbezogene Daten verarbeitet werden, sind die Vorgaben der DSGVO einzuhalten. Insbesondere muss der Verantwortliche diese Datenverarbeitung auf eine Rechtsgrundlage stützen können und er muss die Gewährleistung der Betroffenenrechte sicherstellen.

[3] https://www.iese.fraunhofer.de/blog/halluzinationen-generative-ki-llm/

Der Landesbeauftragte für den Datenschutz Niedersachsen

Adresse Prinzenstraße 5
30159 Hannover

Telefon 0511 120-4500

Fax 0511 120-4599

E-Mail poststelle@lfd.niedersachsen.de

Internet https://lfd.niedersachsen.de