

Erste Hilfe zur KI-Verordnung

KI-Kompetenz – Rechte – Pflichten

2	Was ist KI? _____	45
1.	Autonome Systeme _____	46
2.	Risiken des KI-Einsatzes _____	47
3	Wie kann KI eingesetzt werden? _____	48
3.	Einsatz zu nützlichen Zwecken _____	49
2.	Einsatz zu schädlichen Zwecken _____	50
4	Welche Pflichten hat der Betreiber nach der KI-VO? _____	51
1.	Verbote Zwecke _____	52
2.	Hochrisiko-Zwecke _____	52
3.	Betreiberpflichten für Hochrisiko-KI-Systeme _____	54
4.	Grenzenpflichten _____	55
	Übersicht der Pflichten: Vom Betreiber zum Anbieter per Zweckförderung _____	56
	Rechtsschutz für „alte“ KI _____	57
	Check: Geltung der Betreiberpflichten im Einzelfall _____	58



Inhaltsverzeichnis

I KI-Quickstart

1	KI-Quickstart	5
1.	Was ist KI?	6
2.	Was ist nicht KI?	6
3.	Worauf basiert KI und was hat es mit KI-Modellen und KI-Systemen auf sich?	6
4.	Was ist KI mit allgemeinem Verwendungszweck (GPAI)?	6
5.	Wie funktioniert KI?	7
6.	Warum gibt es besonderes KI-Recht (KI-VO)?	7
7.	Für wen gilt das KI-Recht?	7
8.	Was bedeutet es, ein KI-System zu betreiben?	8
9.	Muss jeder KI-Kompetenz nachweisen?	8
10.	Was bedeutet KI-Kompetenz konkret?	8
11.	Was ist der rechtliche Ansatz der KI-VO?	9
12.	Ist KI gefährlich?	9
13.	Wie funktioniert der risikobasierte Ansatz der KI-VO?	9
14.	Wann ist KI nicht riskant?	9
15.	Wann ist KI hochriskant?	9
16.	Gibt es Ausnahmen von der Einordnung eines KI-Systems als hochriskant?	10
17.	Wann darf man hochriskante KI konkret verwenden?	11
18.	Woran erkennt man hochriskante Anwendungen?	11
19.	Wann ist KI verboten?	11
20.	Wer haftet, wenn etwas bei der Verwendung von KI schiefgeht?	12
21.	Was müssen Arbeitnehmer und Arbeitgeber bei der Verwendung von KI beachten?	12
22.	Ab wann gilt das KI-Recht?	13
23.	Welches Recht muss man neben dem KI-Recht beachten?	14
2	Was ist KI?	15
1.	Autonome Systeme	16
2.	Risiken des KI-Einsatzes	17
3	Wie kann KI eingesetzt werden?	18
1.	Einsatz zu nützlichen Zwecken	19
2.	Einsatz zu schädlichen Zwecken	20
4	Welche Pflichten hat der Betreiber nach der KI-VO?	21
1.	Verbotene Zwecke	22
2.	Hochriskante Zwecke	22
3.	Betreiberpflichten für Hochrisiko-KI-Systeme	24
4.	Transparenzpflichten	25
5.	Wechsel der Pflichten: Vom Betreiber zum Anbieter per Zweckänderung	26
6.	KI-Kompetenz	28
7.	Bestandsschutz für „alte“ KI	28
8.	Schnellcheck: Geltung der Betreiberpflichten im Einzelfall	29

5 KI und Datenschutz

1. Anwendbarkeit des Datenschutzrechts	31
2. Ausnahme für private Nutzung	32
3. Rechtsgrundlagen	32
4. Dokumentationspflichten	34
5. Allgemeine Informationspflicht	34
6. Drittstaatentransfer	35

6 KI und Arbeitsrecht

1. KI in der Arbeitswelt	37
2. KI-Verbote: Was Arbeitgeber nicht dürfen	37
3. Hochrisikante KI am Arbeitsplatz	38
4. „Normales“ Arbeitsrecht	38

7 KI und Verbraucherschutzrecht

1. Verbraucherrechte	41
2. Recht auf Beschwerde	41
3. Recht auf Erläuterung	41
4. Whistleblowing	41
5. Transparenz	41

8 KI und Urheberrecht

1. Schutz der Eingabe	43
2. Schutz der Ausgabe	43
3. Wie schütze ich meine Werke vor einer Verwertung im KI-System?	45

Glossar

KI-Checkliste für Betreiber

Checkliste Prompts	48
--------------------	----

1 KI-Quickstart

Am 1. August 2024 ist die KI-Verordnung (KI-VO) in Kraft getreten. Sie setzt einen für die EU einheitlichen Rechtsrahmen für die Verwendung von KI. Wer die neue Technik verwendet, muss zentrale Fragen beantworten können und wissen, was das neue Recht voraussetzt, um die neue Rechtspflicht der KI-Kompetenz zu erfüllen. Dieses Kapitel dient dazu, Fragen aufzuwerfen und kurz zu beantworten. Damit soll ein Bewusstsein für die Probleme und deren Lösungen entstehen. Im Rahmen der folgenden Kapitel werden die Probleme aufgegriffen und genauer beantwortet.

1. Was ist KI?

Nach der Definition der KI-VO (Verordnung über Künstliche Intelligenz) ist KI eine besondere, insbesondere autonome und deshalb unbeherrschbare Technik, die sich ohne menschliches Zutun verändern kann. Man kann eine KI mit einem Tier vergleichen, dessen Wesen man nicht beherrschen kann. Allerdings kann man seine Verwendung verantworten, wenn man es verantwortlich einsetzt.

Vor allem muss man sich darüber klar sein, dass eine KI niemals Verantwortung tragen kann und im Rechtssinn auch keine Fehler begehen kann. Schuld ist immer der Mensch (siehe 2.1 Autonome Systeme).

2. Was ist nicht KI?

In vielen Fällen ist die KI-VO gar nicht einschlägig, weil es nicht um KI im Sinne der KI-Verordnung geht. Sogenannte Expertensysteme gleichen nur Muster ab, ohne für autonomen Betrieb angelegt und anpassungsfähig zu sein. Was sich nicht selbst verändern kann, unterfällt dem KI-Recht erst gar nicht (siehe 2.1 Autonome Systeme).

BEISPIEL

Setzt man solche Systeme im Unternehmen zur Auswertung von Verträgen ein, dann kann man ihnen Aufgaben nach festen Vorgaben stellen. Welcher Vertrag ist von Verjährung betroffen? Finden sich in Firmenunterlagen Hinweise darauf, ob eine Umstellung von Währungen in Verträgen bei der Zahlung Wechselkursgewinne ermöglicht? Wie entwickelt sich Kaufverhalten, so dass man beim Warenbestand besser disponieren kann? Ist eine Darmunebenheit oder eine Hautveränderung gutartig und wo muss man als Arzt genau hinsehen? Hier gelten etwa die Vorgaben der Datenschutz-Grundverordnung (DSGVO) für den maßgeblichen menschlichen Beitrag bei automatisierten Einzelentscheidungen, aber die KI-VO interessiert sich dafür nicht.

3. Worauf basiert KI und was hat es mit KI-Modellen und KI-Systemen auf sich?

Die KI-VO gilt für KI-Modelle und KI-Systeme.

KI-Modelle sind große Datenpools. Aus diesen Pools leiten KI-Systeme Inhalte ab, seien es Texte, Töne oder Bilder. Die KI-Modelle sind wie Wassertanks, die aus sehr vielen Quellen gespeist werden, über die die Anbieter der Modelle bestimmen. Modellanbieter erzeugen (trainieren) Datenpools mit enorm vielen Informationen, deren Herkunft und Auswahl so vielfältig wie bedeutsam und manipulierbar ist. Daten aus China erzeugen andere Werte als solche aus Europa oder den USA.

KI-Systeme sind wie Leitungen, die man an den Tank anschließt, etwa um eine Wasseraufbereitungsanlage herzustellen. KI-Systeme bauen damit für einen spezifischen Verwendungszweck auf dem Modell auf. Der Entwickler und Hersteller des Systems, in der KI-VO-Sprache der „Anbieter“ – legt damit den Verwendungszweck eines KI-Systems fest. Um im Bild zu bleiben: Der Anbieter hat eine Anlage zur Trinkwasseraufbereitung hergestellt und installiert. Der Betreiber kann nun bestimmen, wann und wie viel Wasser aufbereitet wird, aber er kann die Anlage nicht eigenständig in ein System zur industriellen Kühlwasseraufbereitung umfunktionieren. Der Betreiber – sei es ein Unternehmen, ein Angestellter, eine Lehrkraft oder ein Schüler – nutzt das System nur innerhalb dieses vordefinierten Rahmens. Ändert er diesen Rahmen und setzt er das KI-System zu einem neuen Zweck ein – eben der Kühlwasseraufbereitung –, wird er nach der KI-VO als Anbieter behandelt. Im Normalfall trägt aber der Anbieter die Verantwortung für die Sicherheit des KI-Systems während der Betreiber für den korrekten Einsatz im Rahmen der vorgegebenen Nutzung verantwortlich ist.

4. Was ist KI mit allgemeinem Verwendungszweck (GPAI)?

KI-Systeme mit allgemeinem Verwendungszweck (GPAI-Systeme) haben im Gegensatz zu regulären KI-Systemen keinen spezifischen Verwendungszweck. Man kann

sie – wie *ChatGPT* – für beliebige Zwecke nutzen. Der Bot kann Liebes- und Hassgedichte schreiben lassen. Was macht eine Software wie *ChatGPT* technisch so besonders? Obwohl KI Software ist, behandelt man sie rechtlich wie ein Produkt.

Man kann sich GPAI als Knetmasse vorstellen, aus der man je nachdem, was man möchte, unterschiedlich gefährliche Produkte formen kann. Aus derselben „digitalen Knetmasse“, kann man sowohl eine Wasserflasche als auch eine echte Flasche formen. Vielleicht hilft auch das Bild des 3D-Druckers. Man kann damit eine Tasse oder eine einsatzfähige Flasche oder ein Messer drucken. Spannend wird es bei KI-Systemen wie auch bei Produkten bei der Entscheidung über die konkrete Verwendung.

Auch die KI-VO macht die Verantwortung für die Verwendung der Knetmasse vom konkreten Verwendungszweck abhängig und weist die Verantwortung faktisch jedem zu, der KI einsetzt.

Weil GPAIS alles kann und der Unterschied zwischen Gut und Böse oder Gleich- und Ungleichbehandlung am Ende nicht rechtskonform programmiert werden kann, trägt – von Ausnahmen abgesehen – die alleinige Verantwortung für die Zwecke der Nutzung deren Verwender (Betreiber) (siehe 4.5 Wechsel der Pflichten: Vom Betreiber zum Anbieter per Zweckänderung).

5. Wie funktioniert KI?

KI basiert auf Computeranwendungen, die große Datenseiten verarbeiten, um für Menschen gut verständliche Ergebnisse zu erzeugen. Im Fokus stehen derzeit Anwendungen, die wie *ChatGPT* Inhalte in menschlicher Sprache oder wie Dall-E Bilder hervorbringen, die von menschlich erzeugten Inhalten gar nicht oder nur schwer zu unterscheiden sind. Anders als Menschen „sehen“ Maschinen aber nicht mit Augen und sind nicht „kreativ“ wie Menschen. Stattdessen nutzen sie fortschrittliche statistische Methoden, um Muster in ihren Trainingsdaten zu erkennen und die wahrscheinlichsten Wort- oder Bildfolgen zu generieren. Was oft als KI-„Kreativität“ bezeichnet wird, ist in Wahrheit eine hochentwickelte Form der Mustererkennung und -reproduktion. KI-Systeme „verstehen“ nicht wirklich die Inhalte, die

sie produzieren; ein KI-System weiß nicht, was es sagt oder zeichnet. Stattdessen wählt es aufgrund bekannter Muster die Wörter oder Bildteile aus, die am ehesten zusammenpassen könnten. Es ist, als würde man ein Puzzle zusammensetzen, weil die Teile zusammenpassen, ohne die Bedeutung des daraus entstehenden Bildes zu verstehen. Obwohl die Ergebnisse oft erstaunlich gut sind, hat die KI kein echtes Verständnis oder eigene Gedanken wie ein Mensch. Sie kann nur das wiedergeben und neu mischen, was sie von Menschen gelernt hat. Dennoch können KI-Systeme Zusammenhänge in Datensätzen erkennen, die ein einzelner Mensch nicht sieht, weil sie mehr Muster gelernt haben (siehe 2.2 Risiken des KI-Einsatzes).

6. Warum gibt es besonderes KI-Recht (KI-VO)?

In Unternehmen und Behörden haben sich KI-Systeme, die Texte und andere Inhalte generieren, teilweise bereits durchgesetzt. Experten sehen sowohl Chancen für bedeutende Fortschritte als auch Risiken für (unbeabsichtigte) schädliche Folgen für die Gesundheit, Sicherheit und Grundrechte von Bürgern. Die Verordnung über künstliche Intelligenz soll Fortschritt im Rahmen des rechtlich Zulässigen ermöglichen, nicht verhindern. Das ist wichtig, denn die Menschen in Europa sind ebenso wie die Wirtschaft auf Fortschritt angewiesen. Deshalb hat man in der Europäischen Union die KI-Verordnung verabschiedet.

7. Für wen gilt das KI-Recht?

Die KI-VO gilt faktisch für jedermann. Sie regelt nämlich von der Entwicklung bis hin zum Betrieb, sprich die Verwendung, von KI-Systemen wie *ChatGPT* die gesamte Wertschöpfungskette. Auch für Betreiber legt sie dabei Pflichten fest.

BEISPIEL
Wer also als Handwerker seine Mitarbeiter oder Kunden per <i>ChatGPT</i> anspricht oder sich einen Werbeflyer von einer Bild-KI erzeugen lässt, ist Betreiber,

denn er verwendet ein KI-System in eigener Verantwortung für die berufliche und nicht private Tätigkeit.

Ebenso sind Lehrer, die ihren Schülern die Verwendung von *ChatGPT* zur Unterstützung bei den Hausaufgaben zur Verfügung stellen, Betreiber von KI-Systemen. Schließlich sind Hausaufgaben keine Privatsache, sondern werden im Rahmen der staatlichen Schulpflicht gemacht.

Insgesamt sind **Schulen**, die den Einsatz von KI für Schüler und Lehrer gestatten, oder **Unternehmer**, die ihren Mitarbeitern die Verwendung von KI gestatten, Betreiber. Auch Beschäftigte, die KI ohne Wissen oder Erlaubnis des Unternehmens einsetzen (zum Beispiel Übersetzungssoftware), verwenden KI und unterfallen als Betreiber dem neuen Recht, denn die Verwendung erfolgt ja zu beruflichen Zwecken (siehe 4. Welche Pflichten hat der Betreiber nach der KI-VO?).

8. Was bedeutet es, ein KI-System zu betreiben?

Wer als Bäcker einen Ofen verwendet, der betreibt ihn im Rechtssinne. Wer ein KI-System wie *ChatGPT* verwendet, der betreibt es dementsprechend. KI-Systeme sind keine Spielzeuge, sondern mächtige technische Instrumente. Da der Gesetzgeber KI-Systeme je nach deren Verwendungszweck für gefährlicher hält als Rasenmäher, verlangt er von jedermann Kompetenznachweise für die Verwendung. Auch wenn es, was unrealistisch wäre, keine Führerscheinpflicht für den Einsatz von KI gibt, muss man sich dennoch um Kompetenz bemühen. Das ist erforderlich, damit man Erlaubtes von Verbotenem unterscheiden kann und es möglichst nicht zu Rechtsverstößen kommt. Die KI-VO knüpft dementsprechend die Verwendung von KI außerhalb des privaten Bereichs an die Rechtspflicht zur Vermittlung von KI-Kompetenz (siehe 4. Welche Pflichten hat der Betreiber nach der KI-VO?).

9. Muss jeder KI-Kompetenz nachweisen?

Die KI-VO schreibt dazu jedem, der ein KI-System wie *ChatGPT* außerhalb des privaten Bereichs verwendet, also betreibt, die Vermittlung von KI-Kompetenz (Artikel 4 KI-VO) vor. Da diese Pflicht für Betreiber, also Verwender von KI-Systemen gilt, sind nicht nur Unternehmen und Behörden verpflichtet, sondern auch jede natürliche Person, die ein KI-System wie *ChatGPT* nicht zu persönlichen Zwecken nutzt. Diese Pflicht muss im **Februar 2025 umgesetzt** sein. Wer diese Broschüre gelesen und verstanden hat, ist einen wichtigen und rechtlich nötigen Schritt gegangen.

Im Rahmen einer allgemeinen Pflicht muss jeder Anbieter und Betreiber von KI-Systemen, also faktisch jeder der Systeme wie *ChatGPT* in seinem Geschäftskreis einsetzt, sicherstellen, dass seinem Personal und anderen Personen, die in seinem Auftrag KI-Systeme nutzen und betreiben etwa durch Schulungen KI-Kompetenz vermittelt wird. Das muss nach besten Kräften erfolgen, darf also nicht nachlässig geschehen (siehe 4.6 KI-Kompetenz).

10. Was bedeutet KI-Kompetenz konkret?

Konkrete Fragen die jeder beim Umgang mit KI-Systemen wie *ChatGPT* beantworten können muss, lauten etwa wie folgt (siehe auch 4.6 KI-Kompetenz):

- Was ist ein KI-System, was ist ein KI-Modell, was ist der Unterschied?
- Was bedeutet Autonomie von KI?
- Warum kann KI nicht denken und trotzdem mit mir sprechen?
- Welche Nutzung von KI-Systemen ist gefahrlos möglich? Wo muss ich aufpassen?
- Was bedeutet „prompten“ und wie geht das?
- Wie setze ich mich mit KI-Ergebnissen auseinander?
- Wie behalte ich als Mensch die Kontrolle über das Werkzeug KI?
- Was bedeutet der Einsatz von KI im beruflichen Alltag? Wo kann mir die Technik helfen, wo nicht?

11. Was ist der rechtliche Ansatz der KI-VO?

Die KI-VO wählt einen rechtlichen Ansatz, der aus zwei Kernelementen besteht. Sie steckt zunächst einen gesetzlichen Rahmen für die Entwicklung und den Betrieb künstlicher Intelligenz ab und ordnet die Nutzung der Technik in Risikoklassen ein. Sodann löst die KI-VO das Problem der Sicherung der menschlichen Verantwortung bei maschineller Hilfe, indem sie den Menschen in die Pflicht nimmt, die autonome Technik selbstbestimmt zu stoppen, wenn es sein muss. Jenseits der Grenzen dieses Rechtsrahmens zum Schutz der Menschen und ihrer Rechte herrscht Freiheit zum Einsatz von KI, soweit nicht das von der KI-VO unberührte und unabhängig davon geltende sonstige Recht – etwa das Datenschutz- oder Urheberrecht – ohnehin Grenzen setzt (siehe 4. Welche Pflichten hat der Betreiber nach der KI-VO?).

12. Ist KI gefährlich?

Das kommt es auf den konkreten Verwendungszweck an. Von diesem macht die KI-VO auch die rechtlichen Grenzen des Einsatzes abhängig und den bestimmt derjenige, der die KI verwendet (siehe 4. Welche Pflichten hat der Betreiber nach der KI-VO?).

BEISPIEL

Stellen Sie sich ein KI-Übersetzungsprogramm vor, das wie ein digitaler Dolmetscher funktioniert. Wenn Sie es im Urlaub benutzen, um ein Restaurant-Menü zu übersetzen, ist das nicht riskant und deshalb ohne rechtliche Konsequenzen erlaubt. Nutzt ein Gericht es für die Vernehmung mit fremdsprachigen Zeugen, muss sichergestellt werden, dass es sehr genau arbeitet und von Menschen überprüft wird. Würde aber jemand dieses Programm so umbauen, dass es bei der Übersetzung von Mietverträgen für fremdsprachige Mieter wichtige Informationen auslässt oder verändert, um Klauseln zum Mieterschutz zu unterschlagen, wäre dies verboten.

13. Wie funktioniert der risikobasierte Ansatz der KI-VO?

Die KI-VO stuft das Risiko in Kategorien ein. Sie lauten wie in einer Pyramide erstens risikolos und erlaubt, zweitens hochriskant und nur unter strengen Voraussetzungen zulässig und drittens verboten. Ist der Einsatzzweck hochriskant, gelten sehr strenge und spezifische Pflichten für den Betrieb eines KI-Systems (siehe 4. Welche Pflichten hat der Betreiber nach der KI-VO?).

14. Wann ist KI nicht riskant?

Ob eine KI riskant ist oder nicht, hängt von ihrem Einsatzzweck ab – egal ob es sich um ein spezialisiertes System oder einen vielseitigen digitalen Assistenten (GPAI) handelt. Für alltägliche Aufgaben wie Fotosortierer oder Einkaufslisten-Ersteller gelten keine besonderen Regeln. Wird die KI jedoch in sensiblen Bereichen wie Bildung oder Personalwesen eingesetzt, stuft man sie als hochriskant ein und reguliert sie strenger. Als unbedenklich gelten KI-Systeme, die lediglich Routineaufgaben erledigen, menschliche Arbeit unterstützen oder Muster erkennen, ohne eigenständig wichtige Entscheidungen zu treffen. Sobald die KI aber persönliche Profile erstellt, gilt sie automatisch als hochriskant.

BEISPIEL

Ein Bäcker nutzt eine KI-App „Bäcker-Berater“ für Geschäftsideen. Die App analysiert Verkaufszahlen und Trends, verarbeitet die Vorschläge des Bäckers und empfiehlt zum Beispiel eine glutenfreie Produktlinie oder Mittagssnacks. Das KI-System unterstützt nur bei der Ideenfindung, trifft aber keine Entscheidungen.

15. Wann ist KI hochriskant?

Wann KI hochriskant ist, bestimmt das Recht selbst und benennt dafür konkrete Bereiche. So ist KI, die die Bedingungen von **Arbeitsverhältnissen** beeinflussen

kann oder die für die Bewertung von Lernergebnissen im **Bildungsbereich** also bei Schülern, Auszubildenden oder Studierenden verwendet wird, hochriskant. Andere Bereiche sind **Gesundheit** und **Justiz**.

Da man GPAI für allgemeine und beliebige Zwecke verwenden kann, verlangt der Einsatz dieser KI jedermann, der sie verwendet, eine schwierige Entscheidung ab. Er muss bei der Verwendung bewerten, ob sie im konkreten Fall hochriskant ist. Das hängt allein vom Zweck der Verwendung ab (siehe 4.2 Hochriskante Zwecke).

⇒ BEISPIEL

Wie kann man GPAI im Beschäftigtenkontext einsetzen? Indem man sie a) für die Erstellung des Speiseplanes für das persönliche Mittagessen im Büro nutzt. Das ist ein privater Zweck. Man kann den Bot b) die Kolorierung einer Präsentation des Speiseplans für gesundes Kantinenessen erstellen lassen. Das ist dienstlich, hat aber keinen Einfluss auf die Bedingungen von Arbeitsverhältnissen. Wenn man die KI aber c) den Speiseplan des kommenden Monats für gesundes Essen für alle Mitarbeiter – mit der Maßgabe entweder vegetarisch oder mit Fleisch – erstellen lässt, dann hat das Bedeutung für die Bedingungen der Arbeitsverhältnisse. Je nach Entscheidung der KI könnte dies weitreichende Auswirkungen auf die Arbeitsbedingungen haben: Eine einseitige Bevorzugung vegetarischer oder fleischhaltiger Gerichte könnte bestimmte Mitarbeitergruppen benachteiligen, kulturelle oder religiöse Sensibilitäten verletzen oder gesundheitliche Aspekte vernachlässigen. Lässt man die KI – ebenfalls c) eine (Vor-)auswahl für Beförderungen oder Kündigungen treffen, dann hat das auch Bedeutung für den Job. Man kann die KI auch d) dazu nutzen, die Qualität des Kantinenessens für Lowperformer oder unliebsame Mitarbeiter zu verringern. In der Risiko-Pyramide der KI-VO kommt a) als private Verwendung nicht vor, b) ist risikolos und ohne Vorgaben erlaubt, c) ist hochriskant und unter sehr strengen Voraussetzungen erlaubt und d) ist verboten. In der Praxis ist die Grenze der Verwendungen zwischen b) und c) interessant. Ob man als Beschäftigter beim Einsatz von KI ohne Vorgaben erlaubt oder streng

reguliert agiert, richtet sich nach einer Entscheidung des Gesetzgebers im Anhang zur KI-VO. Alles, was Auswirkungen auf die Bedingungen von Arbeitsverhältnissen hat, ist nach dem Gesetz hochriskant.

16. Gibt es Ausnahmen von der Einordnung eines KI-Systems als hochriskant?

Das Gesetz enthält aber **Ausnahmen** von dieser Einordnung als hochriskant. Das ist dann der Fall, wenn die KI nur unmaßgebliche Hilfsaufgaben übernimmt und ein zuvor gefundenes menschliches Ergebnis optimiert, aber nicht wesentlich beeinflusst.

Die Fragen, wo die Grenzen des Hilfseinsatzes der KI liegen, hängen vom Zweck der Verwendung ab (siehe 4.2 Hochriskante Zwecke):

- **Gesundheit:** Wo verlaufen bei der Diagnose bis hin zur Entscheidung über Leben und Tod (Triage) die Grenzen für den faktisch autonom agierenden KIArzt?
- **Unternehmen:** Wie weit darf der Rat des Kollegen Chatbot gehen, wenn es um Personalentscheidungen im Betrieb geht?
- **Gericht:** Darf ein Bot am Ende dem Richter helfen und gar Tipps für faire Gerichtsurteile geben?
- **Schule:** Was darf der „KI-Lehrer“ bei der Benotung von Schülern?

⇒ BEISPIEL

Wenn ein Lehrer sich bei der Bewertung einer Klassenarbeit von ChatGPT helfen lässt, dann ist dieser Zweck als hochriskant eingestuft. Das könnte in bestimmten Fällen aber zu streng sein. Deshalb benennt die KI-VO abschließend vier Fälle, in denen die Verwendung der KI im Kontext der Bewertung von Schülern nicht hochriskant sein soll. Erstens: Lässt der Lehrer im Rahmen der Bewertung nur eng gefasste Verwaltungsaufgaben erledigen, etwa die Schüler in alphabetischer Reihenfolge oder nach zuvor vergebener Note sortieren, greift die

Ausnahme. Zweitens: Hat der Lehrer bereits eine Note vergeben und begründet, dann kann er per KI Impulse für das Überdenken seiner Bewertung einholen. Drittens: Hat der Lehrer die Noten vergeben und begründet und möchte er danach wissen, ob es Muster oder Abweichungen von Mustern bei der Notenvergabe gab, kann er diese per KI ermitteln lassen. Die Note darf das aber nicht beeinflussen. Viertens: Der Lehrer will nur eine „vorbereitende Aufgabe“ für eine Benotung vornehmen lassen. Da damit aber Vorbereitungsmaßnahmen wie Indexierung, Suche sowie Text- und Sprachverarbeitung gemeint sind, lässt diese Ausnahme keine Vorbewertungen oder Bewertungsentwürfe von Schularbeiten zu. Bei näherem Hinsehen zeigt sich, dass die Ausnahmen keine Zweckbestimmungen legitimieren, mit denen KI-Systemen maßgebliche Aufgaben übertragen werden können. Insofern ist ihr Nutzen fragwürdig.

„Hinweise zur Beförderung oder Kündigung“) ein, dann verändert er durch die Zweckänderung die Risikoklasse von harmlos in hochriskant. Deshalb muss sich der Beschäftigte anstelle des Anbieters des KI-Systems nach den sehr strengen Regeln für den Anbieter verantworten. So will es die KI-VO.

Hat der Arbeitgeber den Einsatz eines GPAIS zu dem hochriskanten Zweck gestattet, dann trifft die Verantwortung diesen. Er wird insoweit anstelle des Anbieters zum Verantwortlichen. In diesem Fall treffen die rechtlichen Folgen, sei es wegen eines Verstoßes nach KI-Recht oder gegen sonstiges Recht, den Betreiber.

17. Wann darf man hochriskante KI konkret verwenden?

Wenn man KI verwendet, die als hochriskant eingestuft ist, dann muss man die strengen **Pflichten** einhalten, die die KI-VO daran knüpft. Diese bestehen etwa darin, passende Eingabedaten auszuwählen, den Betrieb des KI-Systems zu überwachen, von dem System erzeugte Protokolle aufzubewahren und von der Verwendung des Systems betroffene Arbeitnehmer zu informieren. Zudem muss eine menschliche Aufsicht installiert werden. Behörden müssen sich schließlich mit der Frage auseinandersetzen, wie der Einsatz des KI-Systems die Grundrechte der betroffenen Personen beeinflusst (siehe 4.3 Betreiberpflichten für Hochrisiko-KI-Systeme).

BEISPIEL

Was bedeutet das? Setzt ein Arbeitnehmer, dem die Nutzung von ChatGPT am Arbeitsplatz nicht gestattet ist, dieses KI-System für die oben geschilderten hochriskanten Zwecke der Rubrik c) („Kantinenessen vegetarisch oder mit Fleisch“) oder

18. Woran erkennt man hochriskante Anwendungen?

Abstrakt klingt das einfach. Zum Schwur kommt es in konkreten Situationen. Oft kann man hochriskantes und nicht hochriskantes nur schwer auseinanderhalten (siehe 4.2 Hochriskante Zwecke).

⇒ FAUSTFORMEL

Immer dann, wenn der Einsatz der KI einen Menschen in Rechten betreffen kann, also bei der Bewertung in Beruf oder Schule oder bei der Erbringung öffentlicher Leistungen sollte man zurückhaltend sein. Sich von der KI eine Geschichte erzählen oder einen Reisetipp geben zu lassen, ist demgegenüber unproblematisch.

19. Wann ist KI verboten?

Die verbotenen Zwecke legt die KI-VO ebenso fest, wie die erlaubten und hochriskanten Zwecke. Dazu zählt unter anderem sogenanntes Social Scoring, bei dem etwa der Staat seine Bürger per KI manipuliert und klassifiziert und von dieser Klassifizierung deren staatliche Behandlung abhängig macht (siehe 4.1 Verbotene Zwecke).

⇒ BEISPIEL

Eine Recycling-KI bewertet Bürger anhand ihrer Mülltrennung. Wer schlechter trennt, muss länger auf Termine im Bürgeramt warten und zahlt höhere Gebühren für städtische Dienstleistungen.

20. Wer haftet, wenn etwas bei der Verwendung von KI schiefgeht?

Ein Verstoß gegen Pflichten der KI-VO können mit enormen Bußgeldern in der Spurze in Höhe von vielen Millionen Euro belegt werden. Da zahlreiche Normen der KI-VO dem Schutz der Menschen dienen, die von den Anwendungen betroffen sind, tritt zudem die Haftung der Betreiber nach dem Schadensersatzrecht des Bürgerlichen Gesetzbuches ein, wenn sie gegen diese Normen der KI-VO verstossen. Da dieselben Handlungen mit Datenverarbeitungen verbunden sind, dürfte zusätzlich auch an die Haftung auf Schadensersatz nach der DSGVO zu denken sein. Wie gesagt: Wenn Anbieter, also Hersteller von KI-Modellen und KI-Systemen wie Open AI bei *ChatGPT* diese für beliebige und allgemeine Zwecke anbieten, dann sind sie nicht ohne Weiteres für konkrete Anwendungen durch Betreiber verantwortlich. Zunächst kommt eine Haftung des Anwenders für Rechtsverletzungen in Betracht. Veröffentlicht der KI-Anwender etwa KI-generierte Inhalte, die urheberrechtlich geschützte Werke enthalten, drohen Schadensersatzforderungen des Rechteinhabers. Gleicher gilt, wenn ein KI-System persönlichkeitsrechtsverletzende Inhalte generiert, die der Anwender veröffentlicht. Insbesondere im Urheberrecht wird von den Gerichten ein strenger Maßstab für den Vorwurf der Fahrlässigkeit angelegt. Betreiber sollten deshalb sicherstellen, dass ihre Mitarbeitenden darüber aufgeklärt sind, dass auch KI-generierte Inhalte wie Bilder oder Texte eine Vervielfältigung eines geschützten Werkes im Sinne des Urheberrechts sein können. Daneben kommt eine Haftung in Betracht, wenn Betreiber zum Beispiel ihre Transparenzpflichten nach Art. 50 KI-VO nicht erfüllen. Daneben drohen bei Verstößen gegen die KI-VO und gegen die DS-GVO Bußgelder. Die beiden Gesetze sehen

umfangreiche Bußgeldkataloge mit beträchtlichen Höchststrafen vor. Im Anwendungsbereich der DS-GVO haben die Datenschutzaufsichtsbehörden Bußgelder bisher meist angemessen berechnet. Es darf gehofft werden, dass auch im Anwendungsbereich der KI-VO die Höchstbeträge für Ausnahmefälle reserviert bleiben.

⇒ BEISPIEL

Wer also als Arbeitgeber ein GPAIS mit Informationen über seine Beschäftigten füttelt, und sich von der KI Empfehlungen für deren Zusammenarbeit und dafür, wen man vielleicht kündigen soll geben lässt, der nutzt eine KI, die nur für allgemeine Zwecke angeboten wird, eigenmächtig für hoch-riskante Zwecke im Beschäftigtenkontext. Dafür kann der Anbieter des KI-Systems nichts. Auch der Hersteller von Kühlwasser kann nichts dafür, wenn Menschen es in Trinkwasserflaschen füllen und es dann auch trinken. Ebenso kann der Anbieter eines KI-Systems nichts dafür, wenn man es zu Zwecken verwendet, für die es nicht gedacht ist. Deshalb haftet er auch nicht für Ergebnisse, die unter Verstoß gegen das Urheber-, Marken- oder Datenschutzrecht erzeugt werden.

→ HINWEIS

Wer dieses Risiko als Betreiber beherrschen will, der kann dafür sorgen, dass ein KI-System nur auf seinen Datenpool mit Verträgen, Warenbestand etc. zugreift. Solche Angebote sind am Markt verfügbar, aber teuer, weil der Systemanbieter das Risiko übernimmt und es sich bezahlen lässt.

21. Was müssen Arbeitnehmer und Arbeitgeber bei der Verwendung von KI beachten?

Die KI-VO gilt unter anderem für jede natürliche oder juristische Person, die ein KI-System in eigener Verantwortung verwendet. Persönliche und nicht berufliche Tätigkeiten sind ausgenommen. Viele Unternehmen lassen die Nutzung von KI zu. Sie müssen dafür nach der KI-VO als Betreiber und nach der DS-GVO als

2 Was ist KI?

Verantwortliche geradestehen. Andere Arbeitgeber gestatten nicht, dass ihre Beschäftigten Denk- und Prüfaufgaben an Computer übertragen, die autonom und insofern für Arbeitgeber und Mitarbeiter letztlich unbekannter agieren. Es muss ja nicht jedem geheuer sein, bei der Arbeit Maschinen zu benutzen, die dem Menschen nicht wie Suchmaschinen nach kalkulierbaren technischen Vorgaben beim Finden von Ergebnissen helfen, sondern die Lösungen nach nicht beherrschbaren Regeln autonom erfinden.

Sich als Arbeitgeber vor dem unautorisierten Einsatz von KI zu schützen, ist schwierig. Beschäftigte können sich leicht über ein betriebliches Verbot hinwegsetzen. Dazu können sie ein privat erworbene GPAI-System nutzen. Das geschieht dann allerdings in eigener Verantwortung nach der KI-VO auf der einen und sonstigem Recht, etwa der DS-GVO auf der anderen Seite.

Beschäftigte können sich in diesem Fall nicht aus der Verantwortung stehlen, weil die private KI im Unternehmen „im Rahmen einer persönlichen und nicht beruflichen Tätigkeit verwendet“ wird. Der Einsatz dient beruflichen Zwecken des Arbeitgebers. Genauso handelt ein Beschäftigter nicht privat, wenn er alle dienstlichen Dokumente mit eigener Software auf einem privaten Rechner erstellt und sie mit einem privaten Stift unterschreibt (siehe 4.5 Wechsel der Pflichten: Vom Betreiber zum Anbieter per Zweckänderung).

22. Ab wann gilt das KI-Recht?

Die KI-VO ist am 1. August 2024 in Kraft getreten. Damit sich Staat, Gesellschaft und Wirtschaft an den neuen Rechtsrahmen gewöhnen können, gelten die Regelungen stufenweise. Die Berechnung aller Geltungsfristen beginnt am 2. August 2024. Die wesentlichen Geltungsschritte für Unternehmen sind:

Bereits ab dem **2. Februar 2025** müssen Anbieter und Betreiber von KI-Systemen – letzteres ist jeder, der KI im beruflichen Kontext in eigener Verantwortung nutzt – sicherstellen, dass ihr Personal über ausreichende **KI-Kompetenz** verfügt. Die KI-VO verlangt damit ein grundlegendes Verständnis für die Systeme sowie alle Fähigkeiten und Kenntnisse, die ihren sachkundigen

Einsatz ermöglichen. Wer KI-Systeme verwendet, soll sich der Chancen von KI, aber auch ihrer Risiken und möglicher Schäden bei ihrem Einsatz bewusst sein. Unternehmen und Behörden, die KI-Systeme in ihre Prozesse integrieren wollen oder sie bereits integriert haben, sollten deshalb schon jetzt ein Konzept zur Weiterbildung ihrer Mitarbeiter entwickeln.

Ebenfalls ab dem **2. Februar 2025** gelten die **Verbote für bestimmte KI-Praktiken**. Die Verwendung von KI-Systemen zur unterschweligen Beeinflussung, zur sozialen Bewertung oder etwa zur Ableitung von Emotionen ist ab diesem Tag untersagt. Der Gesetzgeber begründet die vorgezogene Geltung mit dem unannehbaren Risiko, das von diesen Praktiken ausgeht. Wer KI-Systeme einsetzt, sollte sich so früh wie möglich mit der Frage auseinandersetzen, ob die konkrete Verwendung einer der verbotenen Praktiken unterfällt. So ist etwa von einer verbotenen Ableitung von Emotionen auszugehen, wenn die Prüfungsangst von Schülern oder die Zufriedenheit von Arbeitnehmern durch einen KI-basierten Chatbot ermittelt wird. Eine falsche Einschätzung führt zwar zunächst nicht zu staatlichen Sanktionen, da die Sanktionsvorschriften erst zu einem späteren Zeitpunkt Geltung beanspruchen. Als gesetzliches Verbot könnten aber Verträge zur Erstellung, Nutzung oder Vertrieb einer solchen KI-Anwendung nichtig sein. Sie kann aber bereits Schadensersatzansprüche auslösen.

Am **2. August 2025** wird sodann der **infrastrukturelle Grundstein** für die umfängliche Geltung der KI-VO gelegt: Ab diesem Tag gelten die Vorschriften der KI-VO, die die Durchsetzung des Rechtsakts sicherstellen sollen. Der Staat muss bis zum 2. August 2025 deshalb eine Leitungsstruktur aufbauen und Verfahren etablieren, sodass er die Einhaltung der KI-VO überwachen kann. Dazu muss er insbesondere die zuständige Marktüberwachungsbehörde benennen. In Deutschland deutet derzeit einiges auf die Bundesnetzagentur hin, die Datenschutzaufsichtsbehörden haben allerdings ebenfalls ein Interesse an der Übernahme der Aufsicht angemeldet.

Ein Jahr später, am **2. August 2026**, beginnt die **Geltung des größten Teils des Rechtsakts**. Ab diesem Tag müssen **Hochrisiko-KI-Systeme** die besonderen Anforderungen an Transparenz, Datenqualität, Genauigkeit,

Robustheit und vieles mehr erfüllen. Für bestimmte KI-Systeme gelten ab diesem Tag zudem besondere **Transparenzpflichten**. Insbesondere der Einsatz von KI-Systemen zur Generierung von Inhalten wie Texten, Bildern, Videos oder Musik muss dann grundsätzlich kenntlich gemacht werden. Unternehmen und Behörden, die KI-Systeme einsetzen, sollten diesen Tag rot im Kalender markieren. Denn eine Vielzahl der dann geltenden Vorschriften betrifft Hochrisiko-KI-Systeme. Darunter fallen etwa Systeme, die bestimmungsgemäß im **Bildungssektor** oder im **Beschäftigungskontext** eingesetzt werden. Werden KI-Systeme zu einem der genannten Zwecke verwendet, müssen die verantwortlichen Betreiber, das heißt die Unternehmen und Behörden, unter anderem einen Einsatz im Einklang mit der Betriebsanleitung sicherstellen und eine **menschliche Aufsicht** installieren, die den Betrieb überwacht. Für Behörden, die KI-Systeme in hochriskanten Anwendungsbereichen einsetzen, tritt die Pflicht hinzu, sich über die Auswirkungen des KI-Einsatzes auf die Grundrechte der betroffenen Personen nachweislich zu vergewissern (**Grundrechte-Folgenabschätzung**). Die Umsetzung dieser Pflichten dürfte einige Organisation beanspruchen. Vorbereitende Maßnahmen sollten deshalb schon jetzt ergripen werden.

Ab dem **2. August 2027** finden die Vorschriften für Hochrisiko-KI-Systeme auch Anwendung auf KI-Systeme, die als Sicherheitsbauteile spezifisch regulierter Produkte dienen oder selbst entsprechende Produkte sind.

Für bestimmte Systeme, die vor dem 2. August 2026 bzw. vor dem 2. August 2027 in Verkehr gebracht oder in Betrieb genommen wurden, gelten besondere Ausnahmen mit Blick auf die Einhaltung der Vorschriften der KI-VO. Für diese läuft eine letzte Frist am **31. Dezember 2030** ab.

23. Welches Recht muss man neben dem KI-Recht beachten?

Mit der KI-VO ist es nicht getan. Wer KI-Systeme verwendet, der muss nicht nur die Regeln der KI-VO einhalten. Da bei der Verwendung von KI-Systemen Texte und Bilder entstehen, deren Grundlagen geschützt sind und auf Inhalte zugegriffen wird, die geschützt sind, muss man zusätzlich das Urheberrecht und das Markenrecht beachten. Das gilt bei der Verwendung von KI ohnehin und ebenso wie das Datenschutzrecht, da Daten verarbeitet werden. Das Verbraucherschutzrecht, das Arbeitsrecht und das Jugendschutzrecht gelten ebenso (siehe 5. KI und Datenschutz, 6. KI und Arbeitsrecht, 7. KI und Verbraucherschutzrecht und 8. KI und Urheberrecht).

BEISPIEL

Beim Einsatz von KI-Systemen muss man unabhängig von der KI-VO das sonstige Recht beachten. Das ist immer so. Wer zum Beispiel etwas stiehlt, macht sich strafbar und muss sich nach dem Strafgesetzbuch verantworten. Zugleich muss der Dieb die Sache nach dem Bürgerlichen Gesetzbuch dem Eigentümer zurückgeben.

2

Was ist KI?

Diese Broschüre soll eine erste Hilfe zum rechtssicheren Einsatz künstlicher Intelligenz (KI) in Ihrem Unternehmen bieten. Das wirft die Frage auf, was eigentlich unter KI zu verstehen ist. Unter „künstlich“ können wir uns noch etwas vorstellen, aber wie lässt sich „Intelligenz“ definieren? Seit in einem Forschungsantrag im Jahr 1956 erstmals der Begriff der künstlichen Intelligenz erwähnt wurde, hat sich sein Verständnis stetig gewandelt. Erscheint eine neue Technologie, verliert die KI oft die ihr zugeschriebene Aura des Außergewöhnlichen; es entsteht der Eindruck, KI sei lediglich das, was maschinelle Systeme aktuell noch nicht leisten können. Für die rechtliche Auseinandersetzung mit KI bedarf es einer präziseren Definition. Entscheidend ist die Frage, ob es Technologien gibt, die eine besondere rechtliche Betrachtung erfordern.

1. Autonome Systeme

Immer mehr Computerprogramme verfügen heute über **Autonomie**. In diesem Kontext meint Autonomie die Fähigkeit eines Programms, eine Handlung **ohne menschlichen Eingriff** auszuführen. Die relevante Handlung liegt allerdings nicht in der Erfüllung einer vom Nutzer gestellten Aufgabe, sondern in der **Veränderung der Handlungsanweisungen**, die der menschliche Programmierer zur Aufgabenerfüllung vorgegeben hat. Das mag im ersten Zugriff kontraintuitiv erscheinen, trifft aber im Kern den Grund einer besonderen Regulierung dieser Technologien.

⇒ BEISPIEL: AMPELANLAGE

Denken Sie an eine Ampelanlage, welche den Verkehr ohne menschlichen Eingriff regelt. Die Ampel wird ihre Aufgabe immer auf die Weise ausführen, die ihr vorgegeben wurde. Für manche Kreuzungen und bestimmte Uhrzeiten erzielt sie damit gute Ergebnisse, ändern sich die Bedingungen, frustriert sie Verkehrsteilnehmer mit kurzen Grünphasen zu Stoßzeiten oder Stillstand trotz fehlendem Verkehr. Das mag für manche Verkehrsteilnehmer ärgerlich sein, die hinter der Ampelanlage stehende Technologie ist aber nicht gefährlich – sie führt von Menschen bestimmte Regeln aus. Eine autonome Ampelanlage ist dagegen befähigt, ihre Schaltung ohne menschlichen Eingriff zu verändern und auf das konkrete Verkehrsaufkommen maßzuschneidern. Sie könnte etwa erkennen, dass die Überquerung für Fußgänger mit Kindern länger dauert oder der Bremsweg eines 16-Tonners ein früheres Rotzeichen benötigt, als ein Moped. Der erhoffte technologische Vorteil dieser Systeme ist die maschinell durchgeführte Personalisierung von Diensten.

Die Lösungen für manche Probleme kann der Mensch nicht so in Programmiersprache übersetzen, dass eine Maschine die Aufgabe nach seinen Vorstellungen übernehmen kann. Das kann daran liegen, dass er die Lösung selbst nicht kennt oder aber daran, dass er sie nicht ausdrücken kann. Für diese Fälle wurden technologische Verfahren entwickelt, in denen die Maschine ihre eigene Programmierung übernimmt, indem sie die

Handlungsanweisungen des Menschen **interpretiert und optimiert**. Dieses Verfahren wird in der Informationstechnik als **maschinelles Lernen** bezeichnet. Zuvor beteiligte Verkehrsteilnehmer könnten sich über eine derart autonome Ampelanlage freuen. Optimierung ist aber stets eine Frage der Perspektive: Was für die Maschine eine Verbesserung der Umstände bedeutet, kann dem Menschen schaden. Das wäre kein Problem, wenn der Mensch nachverfolgen könnte, wie das System die vorgegebenen Handlungsanweisungen verändert. Bei einer Fehlentwicklung könnte er nachbessern, im schlimmsten Fall sprichwörtlich den Stecker ziehen. Eine Überprüfung des Systems ist allerdings mit erheblichen Schwierigkeiten verbunden: Sobald das System einmal begonnen hat, die Handlungsanweisungen zu verändern, ist für den menschlichen Programmierer **nicht mehr nachvollziehbar**, was im Maschinenraum vor sich geht. Deshalb werden autonomen Systemen auch als **Blackbox** bezeichnet.

⇒ BEISPIEL: AUTONOMIE

Der Jäger trainiert seinen Hund monatlang darauf, eine geschossene Ente zu apportieren. Auf der Jagd gehorcht der Hund zunächst, einige Jahre lang schafft er brav die Enten heran. Eines Tages allerdings bringt er statt der geschossenen Ente eine Nachbarskatze, die er selbst erlegt hat. Die Handlungsanweisungen des Jägers waren klar: „Bring mir die geschossene Ente.“ Irgendwann hat der Jagdhund diese Anweisungen allerdings verallgemeinert: „Bring mir die Beute.“ Dass der Jäger weder Interesse an einer erlegten Katze noch an einem Nachbarschaftsstreit hatte, konnte der Hund nicht wissen. Dem Jäger wurden die Grenzen der Beherrschbarkeit seines Tieres aufgezeigt. Gleichwohl trägt er die Konsequenzen. Die Autonomie der Tiere zeigt sich nicht nur im Bild des eigenständig jagenden Hundes. Der Schoßhund macht nicht immer Sitz, wenn er soll, und der Umgang mit domestizierten Wildtieren bleibt stets gefährlich.

Geht es um eine Ampelanlage wird das Risiko dieser Technologie deutlich: Geht ein nicht nachvollziehbares, autonomes Ampelsystem fehl, kann das lebensgefährlich sein.

Wie kann KI eingesetzt

Der europäische Gesetzgeber hat das Gefahrenpotential dieser Technologie erkannt und die Autonomie als **entscheidendes Merkmal regulierungsbedürftiger KI-Systeme** festgelegt. Eine der zentralen Pflichten, die Unternehmen künftig im Umgang mit KI-Systemen erfüllen müssen, ist die Vermittlung von **KI-Kompetenz**: Mitarbeiter, die mit dem Betrieb und der Nutzung von KI-Systemen befasst sind, sollen sich deren Fähigkeiten, Unfähigkeiten, damit verbundenen Chancen und Risiken, sowie möglicher Schäden, die sie verursachen können bewusst werden.

→ HINWEIS

KI-Systeme sind autonom. Autonomie bezeichnet die Fähigkeit eines Systems, die Handlungsanweisungen des menschlichen Programmierers aufgrund neuer Daten eigenständig zu verändern.

2. Risiken des KI-Einsatzes

Trotz der Autonomie sind KI-generierte Inhalte kein Ergebnis eines kreativen Schaffensprozesses, sondern einer komplexen **Wahrscheinlichkeitsrechnung**. Autonomie bedeutet, dass das System die Werte in dieser Wahrscheinlichkeitsrechnung ohne menschlichen Eingriff verändern kann. Inhalte eines Chatbots wie *ChatGPT* sind aus Sicht des Nutzers sinnvoll zusammenhängende Texte, aus Sicht des Systems hingegen **statistische Notwendigkeit**. Die Texte geben daher mehr oder weniger wahrscheinliche Wortketten.

→ BEISPIEL:

Auf die Wortkette „Ich wünsche dir einen Guten ...“ folgt wahrscheinlich als nächstes Wort „Morgen“ oder „Abend“. Weniger wahrscheinlich ist hingegen das Wort „Mittag“, auch wenn der Satz „Ich wünsche dir einen guten Mittag“ inhaltlich und grammatisch korrekt ist. Das führt bisweilen zu unerwünschten Ergebnissen. Die Systeme verknüpfen

sachlich falsche Informationen (**Halluzination**), folgen Tendenzen, die Nutzereingaben zu entnehmen sind und geben aufgrund einer entsprechenden Repräsentation in historischen Datensätzen Antworten aus, die aus heutiger Sicht anstößig oder gar verboten sind (**Bias**).

Das ungewünschte Ergebnis eines KI-Systems muss nicht zwangsläufig zu Konsequenzen in der analogen Welt führen. Wäre der Mensch stets in der Lage, diese Phänomene zu erkennen und darauf zu reagieren, könnte er entsprechende Ergebnisse aussortieren oder überarbeiten. Er neigt aber dazu, der Technologie zu vertrauen. In der Psychologie wird das als **Automatisierungsbias** bezeichnet. Diesem unterliegend übernehmen Autoren die falschen Vorschläge einer automatisierten Rechtschreibprüfung oder folgen Autofahrer ihrem Navigationssystem in den nächsten See. Besonders bei Chatbots wird diese Gefahr virulent, da die Programme ihr mangelndes Wissen durch überzeugend klingende Texte kaschieren.

→ BEISPIEL:

Ein Anwalt in den USA fragte die KI nach Präzedenzfällen für ein Klageverfahren. Die KI nannte dem Anwalt mehrere vergleichbare Fälle mit Aktenzeichen. Der zuständige Richter stellte im Laufe der Verhandlung fest, dass die zitierten Aktenzeichen von der KI erfunden waren.

→ HINWEIS

Erkenntnis für den rechtssicheren Einsatz: KI ist in manchen Kontexten verlässlich, sie ist allerdings nie beherrschbar. Trotz ihrer Autonomie kann die KI keine Verantwortung tragen. Vielmehr knüpft die Rechtsordnung an die Verantwortlichkeit des Individuums an. Zentraler Faktor des KI-Einsatzes im Unternehmen muss deshalb der Mensch sein.

3

Wie kann KI eingesetzt werden?

Ende 2022 ist ein autonomes KI-System für die breite Öffentlichkeit veröffentlicht worden: ChatGPT. Es ist ein KI-System zur Verarbeitung natürlicher Sprache (Natural Language Processing). Das heißt, dass aufgrund menschlicher Eingaben passende, weil wahrscheinliche Ausgaben produziert werden, die unserer Sprache nachgebildet sind. Doch die Abbildung menschlicher Sprache ermöglicht nicht nur die Übersetzung oder das Korrigieren menschlicher Texte: Wenn ein KI-System wahrscheinliche Sprachtexte generieren kann, beginnt es auch dahinterstehende sprachliche Konzepte wie Logik, Emotionen oder Rhetorik nachzuahmen.

Was können diese KI-Systeme wirklich? Sogenannte generative KI-Systeme wie ChatGPT werden als die ersten KI-Systeme bezeichnet, die einen allgemeinen Verwendungszweck haben. Ein KI-System mit allgemeinem Verwendungszweck kann zu nützlichen Zwecken eingesetzt werden, aber auch zu schädlichen. Einige Beispiele werden in diesem Kapitel aufgeführt.

4

Welche Pflichten hat der KI-Sprachsystem? (Vorlesung)

1. Einsatz zu nützlichen Zwecken

Die Verarbeitung natürlicher Sprache durch Computer hat verschiedene Funktionalitäten. Jeder dieser Funktionalitäten können **zahlreiche Anwendungsfälle** zugeordnet werden.

a) Generieren von Texten

Die wohl prominenteste Funktion der neuen Generation von KI-Sprachsystemen ist das Generieren von Texten. Dieser Anwendungsfall kann etwa vom **Betreiber eines Online-Shops** genutzt werden:

- **Kundenservice:** KI-Chatbots können eine erste Hilfe bei mehr oder weniger typischen Fragestellungen bieten oder Bestellungen abwickeln.
- **Öffentliche Kommunikation:** Das Generieren von Texten kann darüber hinaus beim Verfassen einer Pressemitteilung oder der Konzeption einer Werbekampagne helfen. Das System kann etwa einprägsame Werbeslogans vorschlagen.
- **Website-Gestaltung:** können mit KI automatisch Produktbeschreibungen oder SEO-optimierte Texte erstellen lassen, um Zeit zu sparen und die Sichtbarkeit in Suchmaschinen zu verbessern.

b) Bearbeiten von Texten

Die KI-Sprachsysteme können auch für die Bearbeitung eines fertiggestellten Textes eingesetzt werden. **Anwendungsfälle sind beispielsweise:**

- die redaktionelle Überarbeitung;
- die Kürzung;
- die Übersetzung in andere Sprachen.

In diesen Fällen ist von einem Menschen zu prüfen, ob das KI-System tatsächlich eine rein formelle Überarbeitung vorgenommen hat oder ob die sprachlichen Änderungen mit einer Veränderung auf inhaltlicher Ebene einhergehen.

⇒ BEISPIEL:

Die KI soll einen Text über eine „Bank“ korrigieren. Der menschliche Nutzer sollte das Ergebnis daraufhin überprüfen, ob die KI den richtigen Kontext erfasst hat. Es könnte sein, dass die KI den Begriff „Bank“ in Zusammenhang mit einem Finanzinstitut verstanden hat. Der Nutzer wollte hingegen einen Text über Sitzgelegenheiten schreiben.

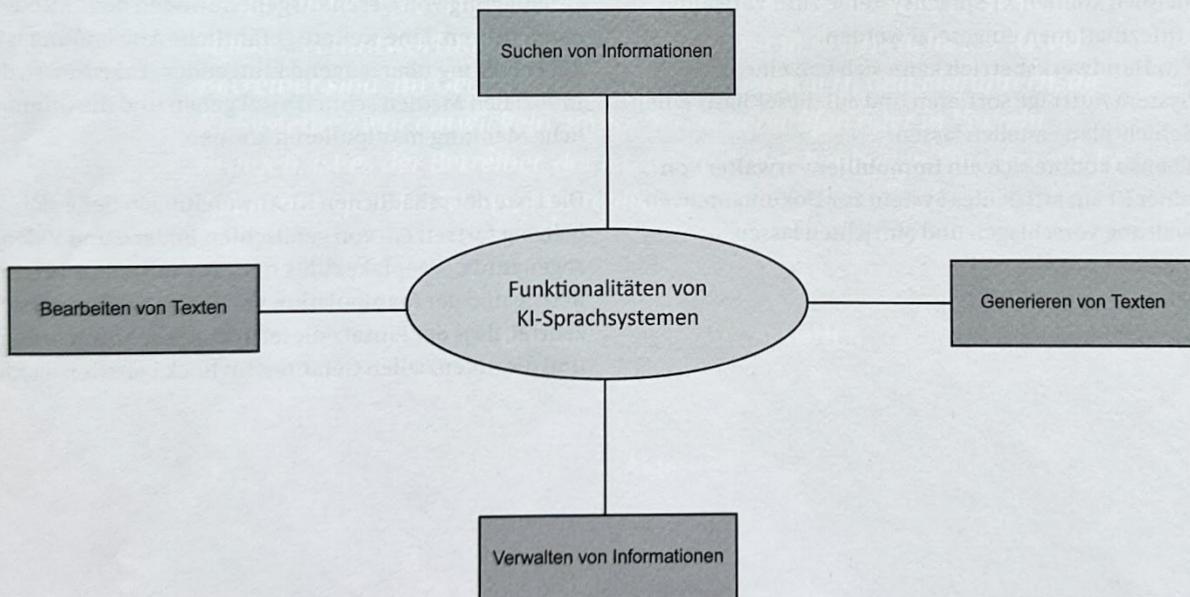


Abbildung: Funktionalitäten von KI-Sprachsystemen

c) Suchen von Informationen

KI-Sprachsysteme können zum Suchen von Informationen in großen Dokumentenmengen eingesetzt werden. Sie eignen sich daher etwa zum Wissensmanagement in einem Unternehmen.

Hier liegen große Schwächen beim Einsatz von KI-Sprachsystemen: Weil sie nur wahrscheinliche Ergebnisse generieren und nicht zwingend richtige Ergebnisse, müssen die gefundenen Informationen überprüft werden. Insbesondere bei der Suche nach schwer auffindbaren Informationen laufen KI-Systeme Gefahr, überzeugend, weil wahrscheinlich passende – aber tatsächlich unrichtige Informationen zu liefern.



BEISPIEL:

Wird die KI gefragt, wann Olaf Scholz geboren ist, wird das richtige Ergebnis „14. Juni 1958“ genannt. Bei weniger prominenten Personen, deren Geburtsdatum nicht auf Wikipedia und anderen bekannten Quellen veröffentlicht ist, kann es sein, dass die KI ein falsches Geburtsdatum „erfindet“.

d) Verwalten von Informationen

Schließlich können KI-Sprachsysteme zum Verwalten von Informationen eingesetzt werden.

- Ein **Handwerksbetrieb** kann sich von einem KI-System Aufträge sortieren und auf dieser Basis einen Schichtplan erstellen lassen.
- Ebenso könnte sich ein **Immobilienverwalter** von einer KI ein effizientes System zur Dokumentenverwaltung vorschlagen und einrichten lassen.

2. Einsatz zu schädlichen Zwecken

Diese nützlichen Zwecke zeigen, dass KI-Systeme ein enormes Potenzial bieten. Allerdings auch für schädliche und illegale Zwecke: Hacker nutzen KI-gestützte Werkzeuge, um Sicherheitslücken zu finden, Passwörter zu knacken oder Schadsoftware zu entwickeln. Auch das Versenden von personalisierten Spam-Nachrichten wird durch KI erleichtert. Beispielsweise enthalten Phishing-Mails eine persönliche Anrede, einen grammatisch einwandfreien Text und täuschend echt aussehende Links und Anhänge. Besonders gefährlich ist die Möglichkeit, dass KI-Systeme detaillierte Anleitungen zu illegalen Aktivitäten wie dem Bau von Waffen oder Bomben erstellen. Selbst wenn der Zugang zu solchen Inhalten eingeschränkt wird und die Ausgabe der Informationen durch Filter verhindert wird, besteht die Gefahr, dass sie in falsche Hände geraten und zu Gewalttaten führen. Zudem hat der enorme Energieverbrauch beim Training und Einsatz großer KI-Modelle negative Auswirkungen auf die Umwelt.

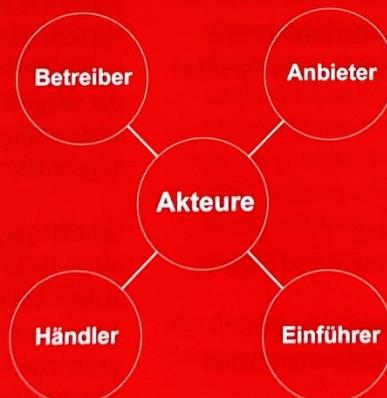
Unternehmen sehen sich ebenfalls bedroht, da Angreifer durch bestimmte Angriffstechniken vertrauliche Daten aus KI-Modellen extrahieren können, die während des Trainings verwendet wurden. Dies kann zur Offenlegung von Geschäftsgeheimnissen oder Kundendaten führen. Eine weitere gefährliche Anwendung ist die Erstellung überzeugender „Fake News“, die in sozialen Medien schnell viral gehen und die öffentliche Meinung manipulieren können.

Die Liste der schädlichen KI-Anwendungen ließe sich beliebig fortsetzen: von gefälschten Bildern und Videos, sogenannte „Deepfakes“, bis hin zum automatisierten Betrug und der Manipulation von Wahlen. Es ist daher zentral, dass der Einsatz dieser Technik achtsam erfolgt und die potenziellen Gefahren im Blick behalten werden.

4

Welche Pflichten hat der Betreiber nach der KI-VO?

KI birgt erhebliche Potenziale, aber auch große Risiken für Bürger und Gesellschaft. Diese Risiken versucht der europäische Gesetzgeber mit seinem Gesetz über künstliche Intelligenz, die sogenannte KI-Verordnung (KI-VO) zu verringern. Die KI-VO folgt einem risikobasierten Ansatz: je risikoreicher der Zweck, für den ein KI-System eingesetzt wird, desto strenger wird geregelt, ob und unter welchen Voraussetzungen es auf den Markt kommen darf. Einige KI-Systeme sind nach der KI-VO folglich verboten, für hochriskante Systeme gelten strenge Vorgaben und einfache Systeme unterliegen keiner besonderen Regulierung. Der risikoorientierte Ansatz der KI-VO leuchtet ein: Warum sollten für das Schreiben einer Geburtstagskarte mit ChatGPT die gleichen Voraussetzungen gelten, wie für das Abfassen eines Urteils oder die Bewertung einer Bachelorarbeit? In diesem Kapitel wird beschrieben, welche Pflichten Unternehmen und Behörden konkret erfüllen müssen, wenn sie eingekaufte KI-Systeme in eigener Verantwortung verwenden und damit in der Sprache der KI-VO **Betreiber** sind. Im Gegensatz zu Anbietern, Händlern und Einführern, die KI-Systeme zwar entwickeln, importieren, verkaufen oder in Verkehr bringen, ist es der Betreiber, der letztendlich den konkreten Einsatz und die Auswirkungen des KI-Systems realisiert.



1. Verbotene Zwecke

Einige Praktiken im KI-Bereich sind nach Einschätzung des europäischen Gesetzgebers so gefährlich, dass sie grundsätzlich verboten wurden. Dazu zählen etwa die unterschwellige Beeinflussung oder die Bewertung natürlicher Personen anhand ihres sozialen Verhaltens (so genanntes **Social Scoring**) durch autonome KI-Systeme. Während die meisten dieser Verbote in erster Linie im staatlichen Kontext relevant werden und den typischen Betreiber eines KI-Systems kaum betreffen, ist einer der verbotenen Einsatzzwecke von besonderer Bedeutung: Verboten sind die Inbetriebnahme und die Verwendung eines KI-Systems zur Ableitung von Emotionen einer natürlichen Person am Arbeitsplatz und in Bildungseinrichtungen. KI-gestützte Maßnahmen zur Feststellung des Betriebsklimas oder von Prüfungsangst sind damit unzulässig. Bei Verstößen drohen hohe Bußgelder. Eine Ausnahme gilt nur für den Einsatz des Systems aus medizinischen oder Sicherheitsgründen. So könnte beispielsweise ein KI-System, das zur Vermeidung des Sekundenschlafs bei LKW-Fahrern, deren Gesichtszüge während der Fahrt analysiert nicht zwingend verboten sein.

Wo steht es?

Art. 5 KI-VO

- Strafverfolgung;
- Migration, Asyl und Grenzkontrolle;
- Rechtspflege und demokratische Prozesse;

Wo steht es?

Art. 6 KI-VO in Verbindung mit Anhang III

a) KI-Einsatz im hochriskanten Bereich

Um diesen abstrakten Begriffen etwas Leben einzuhauen, seien hier einige Anwendungsfälle beschrieben:

In einer Studie der Universität Cambridge wurde festgestellt, dass das einem KI-System auf Basis des Sprachmodells *GPT-4*, auf dem auch *ChatGPT* beruht, bei der Beurteilung von Augenproblemen und der Beratung von Patienten mehr zuzutrauen sei als Ärzten. Genau genommen schnitt das KI-System nur im Vergleich zu unerfahrenen Assistenzärzten besser ab, die Leistung ist aber dennoch beachtlich. Sie verleitete die Forscher in einer Pressemitteilung zu der Aussage, dass KI realistischerweise beim Triagieren von Patienten mit Augenproblemen eingesetzt werden könne, um zu entscheiden, wann ein Notfall Priorität hat. Hier wird es komplex. Wie beschrieben sind KI-Systeme autonom und damit nicht beherrschbar. Bei der Triage geht es darum, Überlebenswahrscheinlichkeiten in Notfällen zu berücksichtigen. Die komplexen Wertungen hinter der Triage können schwerlich allein per KI vorgenommen werden, die in ihrer Unbeherrschbarkeit aus Trainingsdatensätzen möglicherweise diskriminierende Folgerungen abgeleitet hat. Deshalb wertet die KI-VO die Triage berechtigterweise als hochriskanten Zweck des KI-Einsatzes. Bereits der Entwickler des Systems muss deshalb sicherstellen, die Qualität des KI-Systems sowie das vorzeitige Erkennen und Abstellen möglicher Risiken durch technische und organisatorische Maßnahmen sicherzustellen. Der Anwender darf es nur für die vorgeschriebenen Zwecke nutzen und dabei nicht übermäßig in die von dem System hervorgebrachten Ergebnisse vertrauen.

2. Hochriskante Zwecke

Ist ein System nicht verboten, aber hochriskant, werden strenge Anforderungen an die Entwicklung und den Betrieb gestellt. Was genau aber ist unter hochriskanten KI-Systemen zu verstehen?

Die meisten Fälle, in denen ein KI-System als hochriskant eingestuft wird, sind in der Verordnung aufgeführt.

Sie umfassen folgende Bereiche:

- Biometrie;
- kritische Infrastruktur;
- allgemeine und berufliche Bildung;
- Beschäftigung, Personalmanagement und Zugang zur Selbstständigkeit;
- Zugänglichkeit und Inanspruchnahme grundlegender privater und grundlegender öffentlicher Dienste und Leistungen;

Erhebliche Auswirkungen auf betroffene Personen kann auch die Prüfungsbewertung mittels eines KI-Systems haben. Im Vereinigten Königreich wurden wegen der Corona-Pandemie Abschlussprüfungen für Schulabgänger gestrichen. An deren Stelle berechnete ein KI-System die Abschlussnote. In der Folge erhielten tausende Schüler,

insbesondere aus sozial schwachen Verhältnissen, schlechtere Noten, als ihre Lehrer antizipiert hatten. Sicher geglaubte Studienplätze gingen dadurch zunächst verloren. Die Regierung sah sich daraufhin gezwungen, die berechneten Noten zurückzurufen und die Beurteilung den jeweiligen Lehrkräften zu überlassen. Das Beispiel zeigt, dass auch bei der Leistungsbewertung strenge Anforderungen an den KI-Einsatz zu stellen sind – so sieht es auch die KI-VO vor.

Die KI-VO qualifiziert allerdings nicht alle gefährlichen Anwendungsmöglichkeiten künstlicher Intelligenz als hochriskant. KI-Systeme können erheblichen Einfluss auf unsere Meinungsbildung haben. Indem generative Systeme wie *ChatGPT* Texte und Bilder erzeugen, vermitteln sie den Eindruck eines faktenbasierten Austauschs von Informationen. Weil dahinter aber nur die wahrscheinlichste Wortfolge oder Pixelzusammensetzung auf Basis vergangener Daten ausgegeben wird, spiegeln diese Systeme nicht unsere Welt, sondern ihr Training wider – ein vom Internet geprägter Ausschnitt der Realität. Lässt man *ChatGPT* aktuelle gesellschaftliche Fragen beantworten, ergibt sich also eine politische Tendenz auf Basis der Trainingsdaten aus dem Internet, mit der Bürger bei der Nutzung des Systems konfrontiert werden. Trotzdem gilt der KI-Einsatz im Bereich der Meinungsbildung nicht als hochriskant. Im Bereich demokratischer Prozesse sind die Anforderungen an hochriskante KI-Systeme zwar zu erfüllen. Konkret sind damit aber lediglich Systeme angesprochen, die dazu verwendet werden sollen, das Ergebnis einer Wahl oder das Wahlverhalten des Einzelnen zu beeinflussen. Vielfaltssichernde Regelungen für KI-Systeme sieht die KI-VO hingegen nicht vor.

Die Sicherung der Meinungsvielfalt im Internet – gegebenenfalls durch KI-Systeme – ist Aufgabe anderer europäischer Gesetze, etwa des Digital Services Acts.

BEISPIELE

- Biometrie:**
Leseunterstützung mit automatischen Übersetzungen und Begriffsklärungen aufgrund von Pupillenerweiterung

- Kritische Infrastruktur:**
Automatisierte Routenplanung für die Müllabfuhr in einer Kommune
- Allgemeine und berufliche Bildung:**
KI-gestützte Analyse von Schülerverhalten, um deren Eignung oder Lernfähigkeit zu bewerten
- Beschäftigung, Personalmanagement und Zugang zur Selbstständigkeit:**
KI-basierte Analyse von Soft Skills in Bewerbungsvideos für ein Praktikum
- Zugänglichkeit und Inanspruchnahme grundlegender privater und öffentlicher Dienste und Leistungen:**
Automatisierte Priorisierung von Anträgen in der Hotline eines Stromanbieters
- Strafverfolgung:**
Einsatz von KI zur Erkennung von verdächtigem Verhalten in Überwachungsvideos eines Einkaufszentrums
- Migration, Asyl und Grenzkontrolle:**
Automatische Prüfung der Vollständigkeit von Anträgen
- Rechtspflege und demokratische Prozesse:**
KI-unterstützte Erstellung von Sitzungsprotokollen in einem Gemeinderat

Ausnahmsweise Risikominderung

Die beschriebenen Einsatzzwecke führen aber nicht immer zu einer besonderen Regulierung durch die KI-VO. Wenn das System die menschliche Entscheidungsfindung nicht wesentlich beeinflusst, gilt der Einsatz auch in den genannten Bereichen nicht als hochriskant.

Das ist der Fall, wenn das System dazu bestimmt ist:

- eine eng gefasste Verfahrensaufgabe durchzuführen;
- das Ergebnis einer zuvor abgeschlossenen menschlichen Tätigkeit zu verbessern;
- Entscheidungsmuster oder Abweichungen von früheren Entscheidungsmustern zu erkennen, oder
- eine vorbereitende Aufgabe für eine Bewertung durchzuführen, die für hochriskante Zwecke relevant ist.

Für die oben genannten Beispiele für Hochrisiko-Systeme wie die KI-gesteuerte Routenplanung für die

Müllabfuhr bedeutet das Folgendes: Wenn das System nur Empfehlungen gibt, die von Mitarbeitern überprüft und gegebenenfalls angepasst werden, ist der Einsatz nicht hochriskant. Auch bei der intelligenten Lernsoftware, die den Lernstil eines Schülers erkennt, wäre der Einsatz nicht hochriskant, solange diese Information nur dem Betroffenen mitgeteilt wird, um besser zu lernen und kein Lehrer hiervon erfährt, sodass eine Bewertung ausgeschlossen ist. Bei der Priorisierung von Anfragen in der Hotline eines Stromanbieters gilt Ähnliches: Wenn die KI nur unterstützend wirkt und die wesentliche Entscheidung letztlich durch einen Menschen getroffen wird, ist der Einsatz nicht hochriskant. Im Einzelfall ist die Grenze zwischen einer rein formalen Unterstützung und einer wesentlichen Beeinflussung der Entscheidungsfindung oft fließend, was die Abgrenzung erschwert. Es wird von der konkreten Ausgestaltung des Systems, der Art der Entscheidung und dem Kontext des Einsatzes abhängen. Wer sich auf einen der genannten Ausnahmetatbestände beruft, muss eine entsprechend eingeschränkte Verwendung des KI-Systems dokumentieren und auf Verlangen der Behörden nachweisen.

Wie praxistauglich die Ausnahmevorschrift ist, wird sich daher zeigen müssen. Die Anwendungsfälle, die dem europäischen Gesetzgeber bei der Schaffung der Norm vorschwebten, können oftmals ebenso von einfacher, nicht-autonomer Software übernommen werden. Unternehmen und Behörden sollten sich daher fragen, ob für den konkreten Anwendungsfall ein KI-System angeschafft werden soll, das nur unter der Bedingung einer umfangreichen Dokumentation nicht als hochriskant zu bewerten ist, oder ob sie die Aufgabe lieber einer klassisch programmierten Software übertragen, auf welche die KI-VO keine Anwendung findet.

Wo steht es?

Art. 6 Abs. 3 KI-VO

3. Betreiberpflichten für Hochrisiko-KI-Systeme

Betreiber von Hochrisiko-KI-Systemen müssen zwar nicht die umfassenden Anforderungen an die Entwicklung und das Inverkehrbringen der Systeme erfüllen.

Auch sie treffen aber besondere Pflichten, die einen sicheren Einsatz gewährleisten sollen.

a) Allgemeine Betreiberpflichten

Zunächst hat der Betreiber mindestens einem Mitarbeitenden die **Aufsicht** über das Hochrisiko-KI-System zu übertragen. Die menschliche Aufsicht muss über die erforderliche Kompetenz, Ausbildung und Befugnis verfügen: Sie muss also insbesondere in der Lage sein, die Gefahr eines Automatisierungsbias zeitnah zu erkennen und Maßnahmen zu ergreifen, die ein übermäßiges Vertrauen in die Ergebnisse des Systems verhindern. Der menschlichen Aufsicht muss auch die Befugnis übertragen werden, den Betrieb des Hochrisiko-KI-Systems zu stoppen, wenn sie unbekerrschbare Risiken erkennt. Eine praktikable Ausgestaltung dieser Befugnis erfordert auch eine gewisse Unabhängigkeit der menschlichen Aufsicht. Andernfalls besteht die Gefahr, dass den Betrieb beschränkende Maßnahmen aus Angst vor Repressalien nicht ergriffen werden.

Erster Schritt – Datenqualität: Betreiber müssen auch sicherstellen, dass **Daten**, die in das KI-System eingespeist oder von diesem erfasst werden, der Zweckbestimmung des Systems entsprechen und ausreichend repräsentativ sind.

Zweiter Schritt – Dokumentation: Protokolle, die das System erzeugt, sind von dem Betreiber grundsätzlich für mindestens sechs Monate aufzubewahren.

Dritter Schritt – Information und Transparenz: Sofern ein KI-System am Arbeitsplatz eingesetzt wird, etwa um Abläufe zu optimieren oder Schichtpläne effizient zu erzeugen, muss der Arbeitgeber die Arbeitnehmervertreter und die betroffenen Arbeitnehmer darüber **informieren**, dass sie der Verwendung eines Hochrisiko-KI-Systems unterliegen werden.

Gleiches gilt gegenüber anderen betroffenen Personen, die keine Mitarbeitenden des Betreibers sind. Wird also etwa ein KI-System eingesetzt, um Bewerbungen nach vorbestimmten Kriterien auszuschließen, sind Bewerber über den Einsatz des Systems zu informieren.

Vierter Schritt – Kontinuierliche Überwachung: Schließlich sind Betreiber dafür verantwortlich, den

ordnungsgemäßen Betrieb des Hochrisiko-KI-Systems zu überwachen. Stellen sie dabei fest, dass das eingesetzte KI-System ein Risiko für die Gesundheit, Sicherheit oder Grundrechte birgt, müssen sie den Anbieter oder Händler und die zuständige Behörde informieren und die Verwendung des Systems aussetzen. Gleches gilt, wenn ein schwerwiegender Vorfall festgestellt wird. Ein schwerwiegender Vorfall liegt vor, wenn der Betrieb eine besonders schwere Folge hat, zum Beispiel den Tod oder die schwere gesundheitliche Schädigung einer Person oder eine schwere und unumkehrbare Störung der Verwaltung und des Betriebs kritischer Infrastrukturen.

Konkret heißt das: Um sicherzustellen, dass der Betrieb gemäß diesen Anforderungen erfolgt, hat der Betreiber geeignete **technische und organisatorische Maßnahmen** zu treffen. Dabei hat er auch die Vorstellungen des Anbieters von einem ordnungsgemäßen Betrieb zu beachten. Für welchen Einsatz der Anbieter das jeweilige KI-System entwickelt hat, ergibt sich aus der **Betriebsanleitung**, die Anbieter den Betreibern zur Verfügung stellen müssen. Zu organisatorischen Maßnahmen zählen etwa Anweisungen an die Mitarbeitenden, in denen die von Arbeitgeberseite erlaubten Einsatzzwecke klar definiert sind, oder die angesprochene Unabhängigkeit der menschlichen Aufsicht.

b) Grundrechte-Folgenabschätzung durch öffentliche Stellen

Handelt es sich bei dem Betreiber um eine Einrichtung des öffentlichen Rechts oder um eine private Einrichtung, die öffentliche Dienste erbringt, ist darüber hinaus eine **Grundrechte-Folgenabschätzung** erforderlich. Diese setzt voraus, dass der vorgesehene Einsatz klar definiert wird. Dazu zählt eine Beschreibung der Zweckbestimmung und des Verfahrens sowie des Zeitraums und der Häufigkeit der Verwendung des Systems. Der Betreiber muss sich darüber im Klaren sein, welche Personengruppen von dem Einsatz betroffen und welche Schäden zu befürchten sind. Es genügt dazu nicht, abstrakte Überlegungen zu Schadensrisiken anzuführen. Vielmehr muss sich der Betreiber im Einzelfall darüber im Klaren sein, ob bei einschneidenden Lebensereignissen wie staatlichen Prüfungen, der Beantragung sozialer Leistungen zur Existenzsicherung oder der Frage über den Zugang zur Daseinsvorsorge die

Unbeherrschbarkeit des jeweiligen KI-Systems akzeptabel ist. Zu einer vollständigen Folgenabschätzung zählt schließlich auch eine Auseinandersetzung mit der Frage, wie auf Beschwerden reagiert werden kann.

Während die vorbenannten Pflichten das „Wie“ des Einsatzes betreffen, soll sich der öffentliche Sektor im Rahmen der Grundrechte-Folgenabschätzung darüber Gedanken machen, „ob“ der konkrete Einsatz **grundrechtskonform** ist.

→ HINWEIS

Das heißt auch, dass Behörden auf bestimmte Einsätze verzichten müssen, bis sie die Grundrechte-Folgenabschätzung getroffen haben. Wenn eine Schule zur Bewertung von Prüfungsleistungen ein hochrisikantes KI-System einsetzen möchte, ist das rechtlich erst erlaubt, wenn diese Grundrechte-Folgenabschätzung mit einem positiven Ergebnis abgeschlossen werden konnte.

Die Grundrechte-Folgenabschätzung hat damit das Ziel einer behördlichen Vergewisserung: Ist der konkret geplante Einsatz mit den Beeinträchtigungen für die Grundrechte betroffener Personen vereinbar? Wird die KI-basierte Benotung die Chancengleichheit berühren und wie lässt sich das verhindern? Möglicherweise geht das bei einem autonomen System nur dadurch, dass man es zu diesem Zweck nicht einsetzt. Die Folgenabschätzung würde dann negativ ausfallen und der Einsatz müsste ausgeschlossen sein.

4. Transparenzpflichten

Bei der Verwendung von Hochrisiko-KI-Systemen haben Betreiber die betroffenen Personen über den Einsatz zu informieren. Für zulässige Emotionserkennungssysteme und Systeme zur biometrischen Kategorisierung gilt zudem eine spezielle Informationspflicht. Daneben sieht die KI-VO Transparenzpflichten für bestimmte generative KI-Systeme vor. So müssen die Betreiber eines KI-Systems, das **Bild-, Ton- oder Videoinhalte** erzeugt oder manipuliert offenlegen, dass die Inhalte künstlich erzeugt oder manipuliert wurden. Zur Aufdeckung von Straftaten soll diese Pflicht nicht gelten, im Kontext offensichtlich künstlerischer, kreativer, satirischer oder