Rechtsfragen beim Einsatz von generativer KI in gemeinnützigen Organisationen

von Dr. Till Kreutzer, Rechtsanwalt bei iRights.Law

DER PARITÄTISCHE
Gesamtverband

DER PARITÄTISCHE GESAMTVERBAND

Impressum

Herausgeber:
Deutscher Paritätischer Wohlfahrtsverband Gesamtverband e.V.
Oranienburger Str. 13-14
10178 Berlin
http://www.paritaet.org

Inhaltlich Verantwortlicher im Sinne des Presserechts:
Dr. Joachim Rock

Autor:
Dr. Till Kreutzer, Rechtsanwalt bei iRights.Law in Berlin

Redaktion:
Erika Koglin, Abteilungsleitung Recht, Der Paritätische Gesamtverband
Kay Schulze, Projekt #GleichImNetz, Der Paritätische Gesamtverband

Gestaltung:
Christine Maier, Der Paritätische Gesamtverband

Titelbild:
© PRASANNAPIX – AdobeStock

1. Auflage, Januar 2025

#gleich im NETZ

Gefördert vom:

Bundesministerium
für Familie, Senioren, Frauen
und Jugend

Inhalt

Vorwort

Dank der bemerkenswerten Leistungen moderner Systeme maschinellen Lernens befinden wir uns derzeit in einem technologischen Umbruch mit weitreichenden gesellschaftlichen und organisatorischen Folgen. Die neue Technik, die wir auch in dieser Veröffentlichung fortan mit dem Begriff „Künstliche Intelligenz“ (KI) bezeichnen wollen, wird zahlreiche digitale Prozesse sowie die Nutzung technischer Geräte insgesamt revolutionieren.

So verwundert es kaum, dass auch etliche Organisationen der Freien Wohlfahrt hochinteressiert sind an den Möglichkeiten, die KI ihnen bieten kann. Viele Einrichtungen testen KI in Pilotprojekten, einige implementieren entsprechende Systeme bereits in ihren Geschäftsprozessen und manche entwickeln eigene KI-unterstützte Dienstleistungsangebote.

Dabei treten oft ähnliche Fragen auf, z.B.:

- Was muss ich beachten, wenn ich KI-Anwendungen wie ChatGPT in meinen Arbeitsalltag integriere?
- Was ist zu beachten, wenn meine Organisation KI-basierte Online-Dienste für externe Nutzer bereitstellt?

Die vorliegende Rechtshilfe bietet einen kompakten Überblick über die wichtigsten Rechtsgebiete, die von diesen Szenarien berührt werden. Sie ergänzt die separat erhältliche KI-Textsammlung, die Ende 2024 umfassend überarbeitet wurde. Diese Textsammlung enthält u.a. einen Leitfaden für die agile Einführung von KI in Organisationen sowie eine Orientierungshilfe zur Erstellung interner KI-Leitlinien.

Die vorliegende Broschüre ist im Rahmen des Projekts #Gleichim Netz entstanden. Wir danken dem Autor Dr. Till Kreutzer für den gelungenen Einstieg in dieses neue Rechtsgebiet.

Zur besseren Lesbarkeit sind englische Ausdrücke kursiv geschrieben.

Der Paritätische Gesamtverband

Stand: Januar 2025

1. Einleitung: KI in der sozialen Arbeit

KI beschreibt im weiteren Sinn Technologien des maschinellen Lernens (*machine learning*). So bezeichnet man IT-Systeme, die menschliches Lernverhalten simulieren können. Sie verbessern ihre eigenen Fähigkeiten, indem sie bestimmte Aufgabenstellungen ständig wiederholen (sog. Trainieren). Der Unterschied zu herkömmlicher Software liegt darin, dass bei KI-Technologien Funktionen und Lösungsmechanismen nicht von vornherein vorgegeben sind. Sie sind selbst in der Lage, Lösungswege auszuprobieren, zu erlernen und zu optimieren. Hierfür brauchen sie viele Daten, an denen sie Zusammenhänge erkennen und lernen können. Je höher die Anzahl und Qualität von Trainingsdaten und Trainingsvorgängen, desto größer ist der Lerneffekt.

Definition von KI

Art. 3 Nr. 1 der 2024 verabschiedeten KI-Verordnung (KI-VO) der EU definiert ein KI-System als „ein maschinengestütztes System, das für einen in unterschiedlichem Grade autonomen Betrieb ausgelegt ist und das nach seiner Betriebsaufnahme anpassungsfähig sein kann und das aus den erhaltenen Eingaben für explizite oder implizite Ziele ableitet, wie Ausgaben wie etwa Vorhersagen, Inhalte, Empfehlungen oder Entscheidungen erstellt werden, die physische oder virtuelle Umgebungen beeinflussen können."

Mittlerweile gibt es eine Unmenge unterschiedlicher KI-Systeme. Die bekanntesten Beispiele gehören zur Kategorie der KI-Sprachassistenten (*Chatbots*) und der Bildgeneratoren. Bei solchen handelt es sich um Systeme der „generativen KI" (GenAI). Generative KI ist darauf spezialisiert, Inhalte zu erstellen. Das können Texte, Bilder, Musik, Videos oder sogar komplexe Designs sein. Die Bandbreite der Systeme ist groß. Manche sind für spezielle Aufgaben konzipiert, andere sind multifunktional und können die verschiedensten Aufgaben erfüllen.

Beispiele für generative KI-Systeme

Produkt	**Genre** (generierbare Inhalte)
ChatGPT	Multimodaler Chatbot/KI-Assistent
Gemini	Multimodaler Chatbot/KI-Assistent
Claude	Multimodaler Chatbot/KI-Assistent (Schwerpunkt auf ethische und sichere KI)
Copilot	Texterstellung, Code-Generierung, Datenanalysen
Copy.ai	Textgenerierung (speziell für Marketing, Verkaufstexte, Produktbeschreibungen)
Midjourney	Bildgenerierung
DALL-E	Bildgenerierung
Stable Diffusion	Bildgenerierung
Leonardo.AI	Erstellung von Game-Inhalten (Grafiken, Texturen usw.)
Jasper	Textgenerierung (Werbetexte, Blogs, Social-Media-Inhalte)
Runway Gen-2	Video-Generierung (speziell kurze Clips, Animationen)
Synthesia	Video-Generierung (speziell KI-Avatare und Präsentationen)
Invideo.AI	Videogenerierung (v. a. für Marketing, Social Media)
Adobe Firefly	Bild- und Videoanpassungen
Soundraw	Musikgenerierung
Descript	Audiobearbeitung und Sprachsynthese (Voiceovers, Podcasts)
Canva	Präsentationen
MagicSlides	Präsentationen (Plug-In)

Generative KI-Systeme basieren auf KI-Modellen (Basismodelle oder „*Foundation Models*"), wie beispielsweise große Sprachmodelle (*Large Language Models* - LLMs).

Merke: Ein KI-System ist eine voll funktionsfähige und regelmäßig mit einer Benutzeroberfläche ausgestattete KI-Anwendung, die auf einem KI-Modell basiert.

Die Entwicklung von Basismodellen ist sehr kosten- und ressourcenaufwändig. Sie werden dementsprechend in aller Regel nur von sehr großen IT-Unternehmen entwickelt. Beispiele sind GPT von OpenAI, Gemini von Google oder LLaMa von Meta. Auf den nachgelagerten Stufen der KI-Wertschöpfungskette werden die großen Modelle dann auch von kleineren Anbietern in ihre KI-Systeme integriert, um auf deren Basis eigene Produkte zu entwickeln. So ist ein ausdifferenzierter Markt entstanden, auf dem eine Vielzahl von KI-Lösungen verfügbar ist. Generative KI-Systeme werden den Nutzern in der Regel als vorkonfigurierte *„off-the-rack"* Standard-Lösungen angeboten und können als Service über Browser und Apps genutzt werden (*AI-as-a-Service*, AIaaS). Basisversionen werden meist kostenfrei angeboten, zumindest für einen gewissen (Test-) Zeitraum. Zudem bieten annähernd alle Anbieter kostenpflichtige Lizenzen für Vollversionen mit größerem Funktionsumfang an.

Neben Standard-Lösungen setzen Unternehmen und Organisationen zunehmend individuelle KI-Lösungen ein. Solche werden entweder an eigenen Daten trainiert oder als Datenbasis werden die Modelle großer Anbieter verwendet. Diese werden dann in der Regel für die eigenen Anforderungen angepasst und nachtrainiert (*Finetuning*) und in eigene Anwendungen implementiert. Die großen Basismodelle stehen hierzu unter unterschiedlichen Bedingungen zur Verfügung. Manche sind quelloffen und frei nutzbar (wie LLaMa), andere werden proprietär vermarktet und müssen für *Finetuning* und Integration in eigene Lösungen kostenpflichtig lizenziert werden (wie GPT).

Weil KI-Technologien nicht nur Potenziale, sondern auch Risiken bergen, sind sie komplexen Regularien unterworfen. Hierzu zählen unter anderem die KI-Verordnung aus dem Jahr 2024, die DS-GVO und das Urheberrecht. Welche Anforderungen und Pflichten sich aus diesen Regelungen ergeben, hängt von der jeweiligen Konstellation ab. Entscheidend hierfür sind Fragen wie:

- Um wen geht es, welche Rolle hat die Person, Einrichtung oder das Unternehmen im konkreten KI-Kontext (Entwickler, Anbieter, Betreiber, Nutzer*in)?
- Werden personenbezogene Daten verarbeitet?
- Um was für ein KI-Modell oder System handelt es sich? Welches Risiko geht hiervon aus?
- Wird das System rein intern verwendet oder zur Nutzung auch Dritten angeboten?
- Woher kommen die Daten und wo/bei wem sind sie gespeichert?

Die vorliegende Rechtshilfe soll einen Überblick über die wichtigsten Rechtsgebiete geben, die im Kontext des Einsatzes von KI in Organisationen der sozialen Arbeit relevant sind. Eingegangen wird auf die europäische KI-VO, das Datenschutz- und das Urheberrecht.

Die Rechtshilfe versteht sich als informativer Einstieg in die Thematik. Sie kann und soll Rechtsberatung im Einzelfall nicht ersetzen. **Viele Details des Rechtsrahmens sind noch nicht eindeutig geklärt. Wie die Technologie steht auch die rechtliche Entwicklung noch am Anfang und ist in ständiger Bewegung. Zudem handelt es sich bei den auf KI bezogenen Rechtsgebieten in der Regel um komplexe Materien, deren konkrete Einschätzung und Umsetzung im konkreten Fall Expert*innen überlassen werden sollte.**

2. KI-Verordnung

a. Hintergrund und Regelungsansatz

Die KI-Verordnung der Europäischen Union (Verordnung (EU) 2022/1689 vom 13. Juni 2024 zur Festlegung harmonisierter Vorschriften für künstliche Intelligenz – KI-VO) ist ein zentraler Bestandteil der europäischen Digitalstrategie. Sie hat das Ziel, einen Rechtsrahmen für den Einsatz von KI in der Europäischen Union zu schaffen. Um einen gerechten Ausgleich zwischen Innovationsinteressen, Schutz vor Risiken und Gewähr von Grundrechten zu erzielen, verfolgt die Verordnung einen abgestuften und risikobasierten Regulierungsansatz.

Zweck der KI-Verordnung (Art. 1 Abs. 1 KI-VO)

„Zweck dieser Verordnung ist es, das Funktionieren des Binnenmarkts zu verbessern und die Einführung einer auf den Menschen ausgerichteten und vertrauenswürdigen künstlichen Intelligenz (KI) zu fördern und gleichzeitig ein hohes Schutzniveau in Bezug auf Gesundheit, Sicherheit und die in der Charta der Grundrechte verankerten Grundrechte, einschließlich Demokratie, Rechtsstaatlichkeit und Umweltschutz, vor schädlichen Auswirkungen von Systemen der künstlichen Intelligenz (KI-Systeme) in der Union zu gewährleisten und die Innovation zu unterstützen."

Die Verordnung ist ein europäischer Rechtsakt der unmittelbar in allen Mitgliedstaaten Anwendung findet. Sie richtet sich an Anbieter von KI-Systemen, die in der EU entwickelt oder genutzt werden, sowie an Unternehmen und Organisationen, die solche Systeme einsetzen oder verbreiten. Auch Anbieter außerhalb der EU, deren KI-Systeme innerhalb Europas verwendet werden, fallen in den Anwendungsbereich.

Ein zentrales Merkmal der Verordnung ist die Klassifizierung von KI-Technologien nach ihrem Risikopotenzial. Praktiken, die als untragbares Risiko gelten, wie etwa manipulative KI-Systeme, soziale Bewertungssysteme oder die biometrische Fernüberwachung in Echtzeit, werden ausdrücklich verboten. Hochrisiko-KI-Systeme, die beispielsweise in Bereichen wie kritischer Infrastruktur, Bildung, Personalwesen, Strafverfolgung oder Justiz eingesetzt werden, unterliegen strengen Vorgaben. Anbieter solcher Systeme müssen umfassende Risikobewertungen durchführen, strenge Transparenzanforderungen und hohe Standards an Sicherheit und Qualität erfüllen und dies muss auch nachweisbar sein. Auch Betreiber (Anwender) solcher Systeme sind erheblichen Auflagen unterworfen.

Systeme mit geringem oder minimalem Risiko, etwa KI-gestützte Chatbots und andere generative KI-Systeme, sind weniger stark reguliert. Hier gelten gewisse Informations- und Transparenzpflichten. Alle Auflagen und Pflichten sind an Rollen gebunden: Unterschiedliche Akteure verantworten unterschiedliche Risiken und unterliegen daher auch unterschiedlichen Regularien.

Transparenz ist ein zentraler Faktor der Regulierung: Je nach Risikoeinstufung stellt die Verordnung an die Transparenz und Nachvollziehbarkeit von KI-Systemen erhebliche Anforderungen. Nutzer*innen müssen darüber informiert werden, wenn sie mit KI interagieren. Zudem sollen Systeme erklärbar sein, damit die menschlichen Anwender*innen (*Operators*) KI-generierte Inhalte und Entscheidungen nachvollziehen können.

Hohe Auflagen gelten auch hinsichtlich der Datenqualität: KI-Systeme müssen mit hochwertigen, nicht-diskriminierenden (*bias*freien) Daten trainiert werden, um diskriminierende Praktiken zu verhindern. Bei Zuwiderhandlungen drohen drastische Strafen: Verstöße gegen die Verordnung können mit Geldstrafen bis zu 35 Millionen Euro oder sechs Prozent des weltweiten Jahresumsatzes eines Unternehmens geahndet werden. Auch Strafen und Bußgelder folgen einem abgestuften Ansatz. Sie orientieren sich an der Schwere des Verstoßes und der Rolle bzw. Leistungsfähigkeit des Akteurs.

Da die KI-VO technologieneutral und entwicklungsoffen angewendet werden soll, enthält sie eine Vielzahl von unbestimmten Rechtsbegriffen. Diese sind abstrakt-generell und bedürfen daher einer juristischen Auslegung im Einzelfall. Viele Rechtsfragen sind aufgrund der Neuheit der Regelungsmaterie weit von einer abschließenden Klärung entfernt. Gesetzgeber,

Regulierungsbehörden, Verbände und andere Stakeholder arbeiten mit Hochdruck daran, Einzelfragen der KI-VO weiter zu konkretisieren, um Rechtssicherheit zu schaffen. Die KI-VO ist das erste Gesetz seiner Art weltweit. Dass bei einer solch neuartigen und grundlegenden Regelungsmaterie für eine Übergangszeit allerhand Rechtsunsicherheit herrscht, liegt in der Natur der Sache. Hierauf soll dadurch Rücksicht genommen werden, dass die Verordnung nicht sofort, sondern erst innerhalb abgestufter Umsetzungsfristen wirksam wird.

b. Risikobasierter Ansatz, In-Kraft-Treten und Umsetzungsfristen

Die KI-VO wurde am 21. Mai 2024 verabschiedet, am 12. Juli 2024 im EU-Amtsblatt veröffentlicht und trat am 1. August 2024 in Kraft. Ihre Regelungen werden schrittweise für die Adressaten (also die Akteur*innen, die sie einhalten müssen) verbindlich. Die Umsetzungsfristen sind abhängig von der Risikoeinstufung (s. hierzu am Ende dieses Abschnitts in der „Übersicht über die verschiedenen Risikostufen und das abgestufte Inkrafttreten der jeweiligen Regelungen"). Die meisten Auflagen und Pflichten werden danach 24 Monate nach dem Inkrafttreten der Verordnung (also ab August 2026) wirksam.

Wie gesagt basiert die KI-VO auf einem risikobasierten Ansatz: KI-Systeme werden in Risikoklassen eingeteilt, in denen unterschiedliche rechtliche Vorgaben gelten. Auflagen und Pflichten werden den verschiedenen Akteursgruppen differenziert zugeordnet.

c. Verbotene Praktiken

Manipulative, ausbeuterische oder soziale Kontrollpraktiken sind vollständig verboten (Art. 5), da sie die Werte der EU verletzen. Hierzu zählen beispielsweise:

- Manipulation von Personen durch unterschwellige Techniken, um ihr Verhalten zu beeinträchtigen. Beispiel: Ein KI-System, das unbemerkt Kaufentscheidungen beeinflusst.

- Ausnutzung von Verwundbarkeiten bestimmter Gruppen (z. B. Kinder, ältere Menschen, Personen mit Behinderungen) mit schädlichen Folgen. Beispiel: KI-basierte Spielzeuge, die Kinder zum übermäßigen Geldausgeben verleiten.

- Sozialbewertung (*social scoring*) durch Behörden, die zu ungerechtfertigten Vor- oder Nachteilen führt. Beispiel: Ein staatliches Punktesystem, das Bürger*innen nach ihrem Verhalten einstuft und ihnen so den Zugang zu Leistungen verwehrt.

- Echtzeit-Fernidentifizierung im öffentlichen Raum durch Strafverfolgungsbehörden, außer in engen, gesetzlich geregelten Ausnahmen. Beispiel: Permanente Gesichtserkennung auf öffentlichen Plätzen.

d. Hochrisiko-KI-Systeme

Die KI-VO definiert bestimmte Anwendungsfälle, in denen KI-Systeme als hochriskant gelten, da sie sich negativ auf Gesundheit, Sicherheit oder Grundrechte auswirken können. Diese Einstufung hängt davon ab, wozu das KI-System dient (also wie es bestimmungsgemäß genutzt werden soll) und welche Risiken und Fehlanwendungen vorhersehbar sind. Um sicherzustellen, dass hochriskante KI-Systeme nur vertrauenswürdig und sicher auf dem EU-Markt angeboten, betrieben oder genutzt werden, fordert die KI-VO, dass sie eine Vielzahl z. T. komplexer und aufwändiger Compliance-Anforderungen erfüllen müssen (s. im Einzelnen hierzu unten: „Anforderungen und Rechtspflichten nach der KI-VO").

Zu unterscheiden ist zwischen:

- **den Eigenschaften**, die Hochrisiko-KI-Systeme aufweisen müssen (siehe Kapitel III, Abschnitt 2 (Art. 8 ff.) der KI-VO);

- **den Rechtspflichten**, die die einzelnen Akteure beim Umgang mit Hochrisiko-KI-Systemen erfüllen müssen (Pflichten der Anbieter: Kapitel III, Abschnitt 3 (Art. 16 ff.), Pflichten der Betreiber: Art. 26 ff., Pflichten der Bevollmächtigten von Anbietern: Art. 22, Pflichten der Importeure: Art. 23, Pflichten der Händler: Art. 24).

Darüber hinaus unterscheidet die Verordnung zwei Arten von Hochrisiko-KI-Systemen: solche, die unter die Vorschriften des Anhangs I fallen (spezielle durch andere EU-Regularien erfasste Produkte wie Fahrzeuge, Spielzeug oder sicherheitsrelevante Erzeugnisse) und solche, die in Anhang III aufgeführt sind. Anhang III ist nach Einsatzbereichen aufgeteilt und stellt auf die Funktion der KI-Lösungen ab. Aufgeführt werden hier Technologien, deren Einsatzgebiete für das Leben von Menschen von besonderer Bedeutung sind, so dass sich mögliche Diskriminierungen und Fehlentscheidungen besonders gravierend auswirken würden.

Übersicht: Beispiele für hochriskante Risikosysteme nach Art. 6 Abs. 2, Anhang III KI-VO

Bereich	Beispiele	Regelungsgrundlage in der KI-VO
Allgemeine und berufliche Bildung	KI-Systeme, die • bewerten, ob Personen zu Bildungseinrichtungen zugelassen werden sollten • Prüfungsergebnisse bewerten • Prüfungen überwachen	Art. 6 Abs. 2, Anhang III Nr. 3
Beschäftigung, Personalmanagement und Zugang zur Selbstständigkeit	KI-Systeme, die • im Recruiting eingesetzt werden, z. B. um Bewerbungen bewerten oder Stellenanzeigen zu schalten • Entscheidungen über Arbeitsverhältnisse beeinflussen, z. B. in Bezug auf Kündigungen, Aufgabenzuweisungen, Leistungsbewertungen	Art. 6 Abs. 2, Anhang III Nr. 4
Staatliche Unterstützungsleistungen	KI-Systeme, die • bei der Bewertung von Personen beim Abschluss von Kranken- oder Lebensversicherungen eingesetzt werden • bei der Bewertung des Zugangs zu Sozialversicherungssystemen eingesetzt werden • zur Bewertung und Klassifizierung von Notrufen eingesetzt werden	Art. 6 Abs. 2, Anhang III Nr. 5

e. Systeme mit geringem oder minimalem Risiko

KI-Systeme, die weder verboten noch hochriskant sind, werden als Systeme mit geringem oder minimalem Risiko bezeichnet. Die Begriffe werden in der KI-VO nicht verwendet, sie dienen daher nur zur Abgrenzung von den Hochrisiko-Systemen.

KI-Systeme mit geringem Risiko müssen (anders als Hochrisiko-Systeme) nicht generell Vorgaben erfüllen. Reguliert sind in den Art. 50 ff. vielmehr nur bestimmte Arten von KI-Systemen. Hierzu zählen beispielsweise Systeme, die dafür bestimmt sind, mit Menschen zu interagieren (etwa *Chatbots*), generative KI-Systeme (wie Bildgeneratoren), Emotionserkennungssysteme oder solche, die *Deepfakes* erzeugen. Die mit derartigen Systemen verbundenen Anforderungen an die Ausgestaltung und die Pflichten beim Betrieb unterscheiden wiederum nach Anbietern und Betreibern. Hierzu zählen beispielsweise Anforderungen an den Kompetenzaufbau und Transparenzpflichten gegenüber den Nutzer*innen. Beispielsweise müssen Anbieter von Systemen, die dafür bestimmt sind, mit Menschen zu interagieren (z. B. *Chatbots*) dafür sorgen, dass das System die Nutzer*innen darüber informiert, dass sie mit einem KI-System interagieren.

KI-Systeme, die nicht verboten und weder hoch- noch gering riskant sind, werden im Fachjargon als KI-Systeme mit minimalem Risiko bezeichnet. Hierzu können z. B. weitverbreitete administrative oder interne Anwendungen wie Spamfilter, KI-gestützte Videospiele oder Bestandsverwaltungssysteme zählen. Sie unterliegen keinen besonderen Anforderungen aus der KI-VO. Datenschutzrecht und andere Regularien sind natürlich auch hierbei zu beachten.

f. Besondere Anforderungen an die Anbieter von GPAI-Modellen

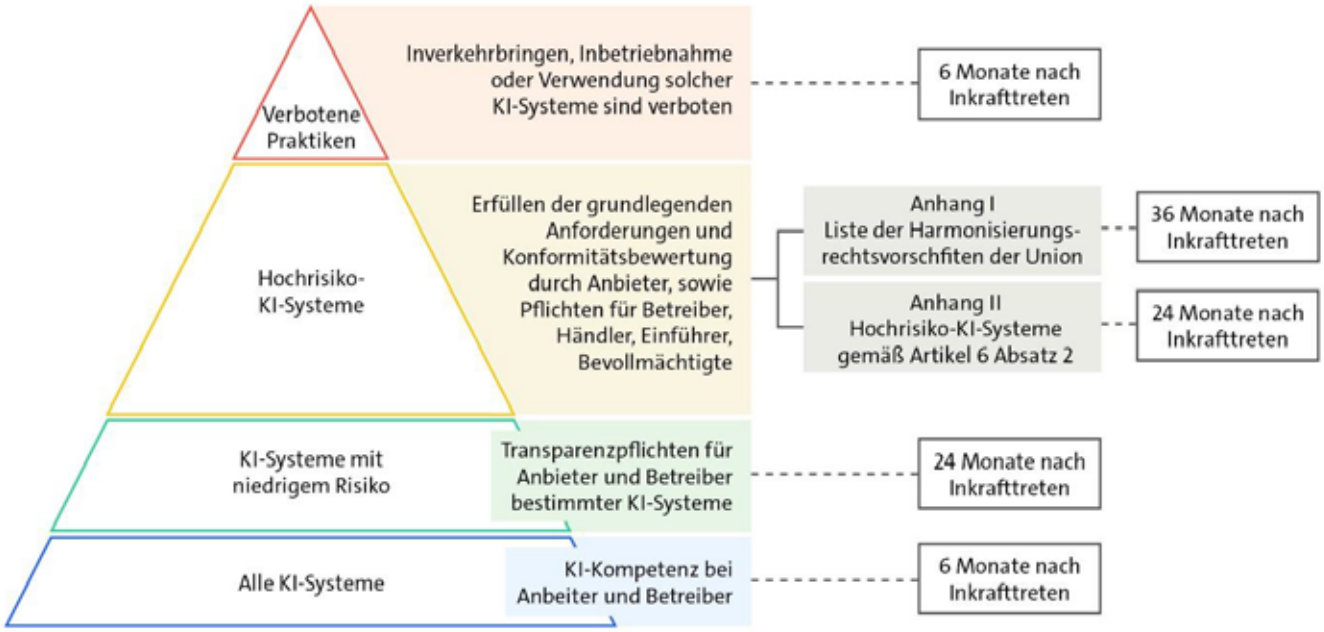
Anbieter von KI-Modellen für allgemeine Zwecke (GPAI-Modelle) unterliegen einem speziellen Regelungsregime. GPAI-Modelle, wie große Sprachmodelle (*Large Language Models* - LLMs) bilden die Grundlage für eine Vielzahl von möglichen Verwendungsformen in KI-Anwendungen. Beispielsweise ist das LLM „GPT" von OpenAI ein GPAI-Modell im Sinne der KI-VO. ChatGPT ist ein auf diesem Modell basierendes GPAI-System. Da es sich bei den Modellen nicht um nutzbare Anwendungen, sondern um deren technische Grundlagen handelt, weisen sie spezifische Risiken auf, die gesondert geregelt sind. Fehler auf der Ebene der Datengrundlage wirken sich potenziell in vielen unterschiedlichen Fällen aus, daher bedarf es hier einer besonderen regulativen Risikovorsorge.

Unterscheidung von KI-Modellen und KI-Systemen mit allgemeinem Verwendungszweck
(General Purpose AI - GPAI)

Definition des GPAI-Models in Art. 3 Nr. 63 der KI-VO:
„KI-Modell mit allgemeinem Verwendungszweck" ein KI-Modell – einschließlich der Fälle, in denen ein solches KI-Modell mit einer großen Datenmenge unter umfassender Selbstüberwachung trainiert wird –, 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, ausgenommen KI-Modelle, die vor ihrem Inverkehrbringen für Forschungs- und Entwicklungstätigkeiten oder die Konzipierung von Prototypen eingesetzt werden;

Definition des GPAI-Systems in Art. 3 Nr. 66 der KI-VO:
„KI-System mit allgemeinem Verwendungszweck" ein KI-System, das auf einem KI-Modell mit allgemeinem Verwendungszweck beruht und in der Lage ist, einer Vielzahl von Zwecken sowohl für die direkte Verwendung als auch für die Integration in andere KI-Systeme zu dienen".

Übersicht über die verschiedenen Risikostufen und das abgestufte Inkrafttreten der jeweiligen Regelungen
Grafik: BITKOM (Hrsg.), Umsetzungsleitfaden zur KI-Verordnung, 2024, S. 14.



g. Ausnahmen

Die KI-VO sieht für spezifische KI-Systeme und Akteursgruppen Ausnahmen vor. Solche finden sich u.a. in Art. 2. Ausnahmen von der Anwendung der KI-VO sind vor allem für militärische Zwecke, die Rechtspflege und Strafverfolgung, Forschung und Entwicklung, freie quelloffene KI-Systeme und die Verwendung im Privatbereich vorgesehen.

Merke: **Eine allgemeine Ausnahme für die Verwendung von KI in gemeinnützigen Organisationen oder generell in nicht kommerziellen Kontexten sieht die KI-VO nicht vor.** Der regulative Ansatz der KI-VO zielt auf einen denkbar weiten Anwendungsbereich ab, um möglichst alle Risiken zu erfassen. Die Austarierung von Pflichten und Freiheiten und damit die Feinabstimmung der regulativen Belastung auf die Leistungsfähigkeit der Akteure erfolgt über die Abstufung der Rechtsfolgen (Auflagen, Pflichten, Sanktionen). So soll erreicht werden, dass einerseits die Risiken möglichst effektiv minimiert, andererseits die involvierten Akteure aber nicht über Gebühr mit überbordenden Anforderungen belastet werden. Ob dieser Ausgleich gelungen ist, wird sich im Laufe der nächsten Jahre herausstellen.

h. Die Akteure: Anbieter und Betreiber

In die Wertschöpfungskette von KI-Technologien sind, von der Entwicklung über den Vertrieb bis zum Einsatz, verschiedene Akteure involviert. Sie alle haben unterschiedlichen Einfluss auf die Ausgestaltung und damit die Potenziale und Risiken von KI. Mit anderen Worten: Sie agieren in verschiedenen Risiko- und Verantwortungssphären. Ein Beispiel: Die Firma OpenAI hat als Entwickler von ChatGPT vollen Einfluss auf alle Aspekte dieser KI-Lösung, einschließlich des Basismodells und der konkreten Anwendung ChatGPT. Sie hat die Entscheidungshoheit über Auswahl und Bewertung der Trainingsdaten, die für Dritte angebotenen Schnittstellen, die Features der Anwendungen, die Vermarktung usw..

Entsprechend trägt OpenAI als „Anbieter" die Verantwortung dafür, Risiken bei der Entwicklung und Vermarktung zu minimieren und technische und rechtliche Standards bei der Ausgestaltung einzuhalten. Eine Organisation, die ihren Mitarbeitenden ChatGPT zur Nutzung zur Verfügung stellt, hat dagegen weder auf dessen Entwicklung noch auf die Ausgestaltung Einfluss. Sie ist daher allenfalls Betreiber einer ChatGPT-Instanz und als solcher z. B. verpflichtet, Mitarbeiter*innen zu schulen, auf einen ordnungsgemäßen Umgang mit der KI-Lösung zu achten oder für Transparenz gegenüber Dritten zu sorgen.

Um den unterschiedlichen Einflussmöglichkeiten, Risiko- und Verantwortungssphären Rechnung zu tragen, weist die KI-VO den verschiedenen Akteuren differenzierte Anforderungen und Pflichten zu. Durch diese abgestufte Aufgaben- und Verantwortungsverteilung sollen Risiken und Auflagen möglichst effizient verteilt werden. Damit soll auch der Umsetzungsaufwand für Hersteller, Anbieter und Betreiber und andere Beteiligte in einem angemessenen Rahmen gehalten werden.

Die unterschiedlichen Regelungsadressaten nach Art. 2 Abs. 1 KI-VO

„Diese Verordnung gilt für

a. **Anbieter**, die in der Union KI-Systeme in Verkehr bringen oder in Betrieb nehmen oder KI-Modelle mit allgemeinem Verwendungszweck in Verkehr bringen, unabhängig davon, ob diese Anbieter in der Union oder in einem Drittland niedergelassen sind;

b. **Betreiber** von KI-Systemen, die ihren Sitz in der Union haben oder sich in der Union befinden;

c. **Anbieter und Betreiber** von KI-Systemen, die ihren Sitz in einem Drittland haben oder sich in einem Drittland befinden, wenn das vom System hervorgebrachte Ergebnis in der Union verwendet wird;

d. **Einführer und Händler** von KI-Systemen;

e. **Produkthersteller**, die KI-Systeme zusammen mit ihrem Produkt unter ihrem eigenen Namen oder ihrer Handelsmarke in Verkehr bringen oder in Betrieb nehmen;

f. **Bevollmächtigte von Anbietern**, die nicht in der Union niedergelassen sind;

g. **betroffene Personen**, die sich in der Union befinden."

Da sich diese Rechtshilfe auf deutsche Einrichtungen der sozialen Arbeit fokussiert, stehen wirtschaftliche und ausländische Akteure nicht im Fokus. Entsprechend werden nachstehend lediglich die Rollen erläutert, die in diesem Kontext relevant erscheinen. Dies sind „Anbieter" und „Betreiber".

Anbieter

Die meisten Anforderungen und Rechtspflichten aus der KI-VO betreffen primär den Anbieter. Der Anbieter-Begriff wird in Art. 3 Nr. 3 KI-VO weit definiert. Hierin heißt es:

„Anbieter" [ist] 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;"

Für die Anbietereigenschaft müssen also zwei Faktoren zusammenkommen:

1. Anbieter ist, wer ein KI-System oder ein GPAI-Modell entweder selbst entwickelt oder entwickeln lässt **und**

2. die KI-Technologie entweder in Verkehr bringt, also vermarktet **oder**

3. sie bestimmungsgemäß in Betrieb nimmt.

Der Anbieter-Begriff entspricht damit im Wesentlichen dem Begriff des „Herstellers", wie er beispielsweise in der Produktregulierung gebräuchlich ist. Über das Verständnis im allgemeinen Sprachgebrauch geht er insofern hinaus, als es einerseits nicht darauf ankommt, ob der Vertrieb oder die Inbetriebnahme kommerziellen Zwecken dient und andererseits auch der Eigengebrauch einer KI-Technologie die Anbietereigenschaft begründen kann. Zwar begründet der Eigengebrauch einer eingekauften KI-Lösung (z. B. eines Standard-*ChatBot*) keine Anbieterstellung. **Wenn aber individuelle KI-Systeme selbst oder im Auftrag für die eigene Nutzung entwickelt oder Standardlösungen von Dritten nicht nur verwendet, sondern maßgeblich verändert oder angepasst werden, kann aus einem reinen Betreiber ein Anbieter werden.**

Beispielsweise wird häufig die Konstellation auftreten, dass eine Organisation ein großes Basismodell eines Dritten lizenziert, um hierauf eine individuelle KI-Anwendung aufzusetzen, die den eigenen speziellen Anforderungen entspricht. GPAI-Basismodelle sind – wie gesagt – gerade dafür gedacht, in verschiedenen Kontexten und Applikationen eingesetzt zu werden. In einem solchen Szenario sind zahlreiche Varianten denkbar. Beispielsweise könnte die Organisation dem Basismodell eigene Daten hinzufügen, an denen die KI weiter trainiert wird (*Finetuning*). Oder es könnten eigene Datenbanken per Schnittstelle (API) mit dem Modell verknüpft werden, um der KI spezielle Informationen zur Verfügung zu stellen (so genannte *„Retrieval Augmented Generation"* oder *„Pre-Prompt Engineering"*). Je nach konkreter Ausgestaltung, können sich verschiedene Fragen stellen, die einer schwierigen Einschätzung im Einzelfall bedürfen. Solche können z. B. lauten:

- Wird die Organisation durch die Ergänzung weiterer Datenquellen zur Anbieterin des GPAI-Basismodells?

- Wird die Organisation durch das *Finetuning/die Weiterentwicklung des Modells zur Anbieterin des GPAI-Basismodells?*

- Welche Maßnahmen fallen noch unter eine bestimmungsgemäße Benutzung des ursprünglichen Modells und bei welchen Änderungen ist die Schwelle zu einer so substanziellen Weiterentwicklung überschritten, dass von einer „wesentlichen Veränderung" (Art. 3 Nr. 23 KI-VO) auszugehen ist, die eine neue Anbietereigenschaft begründet? Gemäß Art. 25, 16 ff. KI-VO gelten auch für Betreiber die Anbieterpflichten, wenn sie: das Hochrisiko-System eines Dritten mit ihrem Namen oder ihrer Handelsmarke versehen; das Hochrisiko-System eines Dritten wesentlich verändern; ein bereits in den Verkehr gebrachtes KI-System mit geringem oder minimalen Risiko so verändern, dass hieraus ein Hochrisiko-System wird.

- Ist die Organisation in Bezug auf ihr eigenes KI-System als Anbieterin anzusehen?

- Handelt es sich hierbei um ein KI-System mit niedrigem oder hohem Risiko? Hiervon hängt die Intensität und Reichweite der Auflagen und Pflichten beim Betrieb ab.

Viele dieser praxisrelevanten Fragen sind derzeit offen und in Diskussion. Für weitere Informationen hierzu siehe BITKOM (Hrsg.), Umsetzungsleitfaden zur KI-Verordnung, 2024, S. 29 ff.

Betreiber

Betreiber haben auf die Ausgestaltung von KI-Technologien keinen Einfluss. Sie müssen daher nicht die Anforderungen an die Ausgestaltung von KI-Modellen und -systemen erfüllen. Auch sie unterliegen zwar gewissen Rechtspflichten, mit denen Risiken aus ihrer eigenen Risikosphäre begegnet werden soll. Diese sind jedoch (abhängig von der Risikoklasse) deutlich geringer als die der Anbieter.

Betreiber ist nach Art. 3 Nr. 4 KI-VO:

„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"

Der Betreiber-Begriff ist damit weit gefasst. **Auch reine „Verwender" sind Betreiber und damit Regelungsadressaten der KI-VO, unabhängig davon, ob es sich um eine natürliche Person, eine Firma oder gemeinnützige Einrichtung handelt oder ob mit der Verwendung betriebliche oder kommerzielle Zwecke verfolgt werden.** Ausgeschlossen vom Anwendungsbereich der KI-VO sind lediglich rein persönliche, nicht berufliche Verwendungen (sog. Haushaltsausnahme). Nach der Rechtsprechung des EuGH zur Haushaltsausnahme im Datenschutzrecht (Art. 2 Abs. 2.c DS-GVO) ist diese eng zu verstehen. **Nicht einmal Handlungen zu rein ehrenamtlichen Zwecken sollen hierunter fallen, wenn sie öffentlichkeitsrelevant sind (denn hier sei der höchstpersönliche Bereich jedenfalls verlassen).**

i. Anforderungen und Rechtspflichten nach der KI-VO

Die KI-VO sieht eine Vielzahl von Anforderungen an die Ausgestaltung (vor allem von Hochrisiko-KI-Systemen) sowie allerhand Rechtspflichten vor. Gemäß dem risikobasierten, abgestuften Regulierungskonzept hängt ihre Anwendbarkeit von der jeweiligen Konstellation ab. **Welche Pflichten einzuhalten sind, ist danach zu entscheiden, um was für eine Art KI-Technologie und Akteur es sich handelt.**

Der Katalog der Anforderungen und Rechtspflichten ist äußerst komplex und umfangreich. Eine umfassende Darstellung würde den Rahmen dieser Rechtshilfe weit sprengen. Die nachstehende Übersicht enthält daher nur eine Auswahl und eine kurze Beschreibung einiger Anforderungen an Anbieter und Betreiber aus der KI-VO (s. zu weiteren Anforderungen aus der DS-GVO unten, Teil 4). Sie dient lediglich dazu, sich einen ersten Überblick zu verschaffen. Wer eine konkrete Prüfung der Pflichten und Anforderungen im Einzelfall vornehmen will, dem sei der „Umsetzungsleitfaden zur KI-Verordnung" des BITKOM empfohlen, der sowohl als PDF (220 Seiten) als auch als *Click-Through-Tool* verfügbar ist.

Anforderung, Rechtspflicht	Adressaten (wer?)	Kategorie des KI-Systems (was?)	Wo steht das in der KI-VO?
Ordnungsgemäße Benutzung: Es ist sicherzustellen, dass das System gemäß der Gebrauchsanweisung genutzt wird.	Betreiber	Hochrisiko-KI-System	Art. 26 Abs. 1
Menschliche Aufsicht: Kompetente und geschulte Personen haben die Benutzung zu beaufsichtigen und die Nutzer zu unterstützen.	Betreiber	Hochrisiko-KI-System	Art. 26 Abs. 2
Technische Überwachung: Der Betrieb des Systems ist zu überwachen.	Betreiber	Hochrisiko-KI-System	Art. 26 Abs. 5
Grundrechte-Folgenabschätzung: Vor Inbetriebnahme ist eine Abschätzung vorzunehmen, welche Personen in welchem Maß in ihren Grundrechten betroffen sein könnten.	Betreiber, bei denen es sich um Einrichtungen des öffentlichen Rechts oder private Einrichtungen, die öffentliche Dienste erbringen, handelt Betreiber, von Hochrisiko-Systemen, die bei der Kreditwürdigkeitsprüfung oder der Risikobewertung und Preisbildung von Lebens- und Krankenversicherungen verwendet werden	Hochrisiko-Systeme (ggf. spezifische)	Art. 27 Abs. 2 Anhang III Nr. 5 lit. b und c
KI-Kompetenz: Es sind Maßnahmen zu ergreifen, um sicherzustellen, dass Personal, das mit Betrieb und Nutzung von KI-Systemen befasst ist, ausreichend kompetent ist.	Betreiber Anbieter	Alle KI-Systeme, soweit nicht aus dem Anwendungsbereich ausgenommen oder verboten	Art. 4
Transparenz bei interaktiver KI: Interagiert ein KI-System mit Menschen, muss es so konzipiert und ausgestaltet werden, dass die Nutzer*inen darüber informiert werden, dass sie mit einer KI interagieren.	Anbieter	Alle KI-Systeme, die für die direkte Interaktion mit Menschen bestimmt sind	Art. 50 Abs. 1
Transparenz bei KI-Systemen, die synthetische Audio-, Text-, Bild- und Videoinhalte erzeugen: Der Output generativer KI-Systeme ist erkennbar als künstlich erzeugt oder manipuliert zu kennzeichnen.	Anbieter	Alle KI-Systeme, die synthetische Audio-, Text-, Bild- und Videoinhalte erzeugen	Art. 50 Abs. 2
Risikomanagement: Es ist ein Risikomanagement-System aufzusetzen und kontinuierlich zu aktualisieren/weiterzuentwickeln, mit dem potenzielle Risiken des KI-Systems ermittelt, bewertet und gesteuert werden.	Anbieter	Hochrisiko-Systeme	Art. 9

Anforderung, Rechtspflicht	Adressaten (wer?)	Kategorie des KI-Systems (was?)	Wo steht das in der KI-VO?
Daten-Governance: Sofern Hochrisiko-KI-Systeme mit Daten trainiert werden, müssen Trainings-, Validierungs- und Testdatensätze verwendet werden, die bestimmten Qualitätskriterien entsprechen.	Anbieter	Hochrisiko-Systeme	Art. 10
Technische Dokumentation: Bevor ein Hochrisiko-KI-System in Verkehr gebracht oder in Betrieb genommen wird, ist eine technische Dokumentation anzufertigen. Sie hat bestimmte Spezifika zu erfüllen und ist laufend zu aktualisieren.	Anbieter	Hochrisiko-Systeme	Art. 11
Aufzeichnung und Protokollierung: Hochrisiko-KI-Systeme müssen technisch so ausgestattet sein, dass sie in der Lage sind, bestimmte „Ereignisse" automatisch dauerhaft aufzuzeichnen. Dies ist die technische Grundlage, damit die Betreiber ihrer Beobachtungspflicht nachkommen können.	Anbieter	Hochrisiko-Systeme	Art. 12
Transparenz über Funktionsweise: Hochriskante KI-Systeme müssen mit Bedienungsanleitungen und Dokumentationen versehen werden. Dies ist die Voraussetzung, damit die Betreiber ihren Pflichten zur ordnungsgemäßen Benutzung nachkommen können. Diese Pflicht soll dem Black-Box-Phänomen entgegenwirken.	Anbieter	Hochrisiko-Systeme	Art. 13
Technische Ermöglichung menschlicher Aufsicht: Hochriskante KI-Systeme sind so zu konzipieren, dass sie vom Betreiber mit angemessenen Mitteln wirksam menschlich überwacht werden können. Das setzt u. a. voraus, dass ihre Funktionsweise, Fähigkeiten und Grenzen verständlich sind. Auch diese Pflicht soll dem Black-Box-Phänomen entgegenwirken.	Anbieter	Hochrisiko-Systeme	Art. 14
Gewähr von Genauigkeit, Robustheit und Cybersicherheit: Hochsicherheits-KI-Systeme müssen ein angemessenes Maß an Genauigkeit, Robustheit und Cybersicherheit gewährleisten und über die gesamte Lebenszeit aufrechterhalten.	Anbieter	Hochrisiko-Systeme	Art. 15

Anforderung, Rechtspflicht	Adressaten (wer?)	Kategorie des KI-Systems (was?)	Wo steht das in der KI-VO?
Qualitätsmanagementsystem: Anbieter von Hochrisiko-KI-Systemen müssen ein Qualitätsmanagementsystem einrichten, das u. a. Konzepte zur Einhaltung von Regulierungsvorschriften, Maßnahmen für Entwurfs- und Qualitätskontrolle, Test- und Validierungsverfahren umfassen muss.	Anbieter	Hochrisiko-Systeme	Art. 17
Korrekturmaßnahmen und Informationspflichten: Sofern der Verdacht besteht, dass ein KI-System nicht mit der KI-VO vereinbar ist oder Risiken für die Gesundheit oder Sicherheit oder Grundrechte von Personen verursacht, ist unverzüglich gegenzusteuern. Es entstehen Informationspflichten und Pflichten zur Kooperation mit den zuständigen Behörden.	Anbieter	Hochrisiko-Systeme	Art. 20 Art. 21
Konformitätserklärungen und CE-Kennzeichnungen: Bei Hochrisiko-KI-Systemen müssen Konformitätsprüfungen vorgenommen und Konformität muss entsprechend erklärt werden. Hochrisiko-KI-Systeme müssen zudem mit einer CE-Kennzeichnung versehen sein, aus der ihre Konformität mit der KI-Verordnung hervorgeht.	Anbieter	Hochrisiko-Systeme	Art. 42 - 47 (Konformität) Art. 48 (CE-Kennzeichnung)

j. Sanktionen bei Verstößen

Auch bei Sanktionen und Bußgeldern folgt die KI-VO einem abgestuften Konzept. Je nach Art und Schwere des Verstoßes bzw. Akteur sieht die Verordnung Bußgelder von bis zu 35 Millionen Euro vor. Bei Unternehmen kann sich die Höhe nach dem Umsatz richten. Die Sanktion richtet sich an den jeweiligen Regelungsadressaten, der seine Pflichten verletzt hat. Sanktionen drohen bei Verstößen also allen von der KI-VO erfassten Akteursgruppen.

Neben den unmittelbar in der KI-VO geregelten Sanktionen (Art. 99) können und werden die Mitgliedstaaten ergänzende eigene Rechtsfolgen regeln, die von Land zu Land variieren.

k. Weiterführende Literatur, Checklisten und Praxistipps zur KI-Verordnung

Im Literaturverzeichnis am Ende dieser Rechtshilfe findet sich eine Übersicht mit Literaturtipps, Checklisten und anderen hilfreichen Materialien zur KI-VO.

3. KI und Datenschutzrecht

Der Einsatz von KI-Systemen führt oft zur Verarbeitung von personenbezogenen Daten. Hierfür gilt das Datenschutzrecht, das – zusätzlich zur KI-VO – zahlreiche Anforderungen aufstellt. Wenn personenbezogene Daten verarbeitet werden sollen, sind diese Verarbeitungsvorgänge bereits bei der Planung für KI-Lösungen einzubeziehen und im weiteren Betrieb zu gewährleisten. Die hierfür v. a. einschlägige Datenschutz-Grundverordnung (DS-GVO) und das Bundesdatenschutzgesetz (BDSG) stellen nicht nur Anforderungen auf, unter welchen Umständen personenbezogene Daten überhaupt verarbeitet werden dürfen, sondern sie machen auch Vorgaben, wie eine solche Verarbeitung in IT-Systemen zu organisieren und abzusichern ist. Hiermit kann und wird in der Regel ganz erheblicher Aufwand einhergehen. Auch wird Fachexpertise benötigt, die – wenn nicht intern vorhanden – externe Beratung erfordert.

Merke: Planung ist alles! Sowohl die KI-VO als auch die DS-GVO stellen eine Vielzahl regulativer Anforderungen auf. Ob und in welchem Ausmaß sie auf das eigene Vorhaben Anwendung finden, ist im Einzelfall zu prüfen. Das kostet Zeit, Ressourcen und im Zweifel auch Geld. Regulative Aspekte sollten daher in jedem Fall möglichst frühzeitig in die Planung einbezogen werden.

a. Datenverarbeitung in verschiedenen Phasen der Entwicklung und des Einsatzes von KI

Der Begriff der Datenverarbeitung (Art. 4 Nr. 4 DS-GVO) ist sehr weit gefasst. Prinzipiell wird hiermit jeder Vorgang im Zusammenhang mit personenbezogenen Daten erfasst.

Datenschutzrecht, Personenbezug und Anonymisierung

Datenschutzrecht ist nur zu beachten, wenn personenbezogene Daten verarbeitet werden. Der „Personenbezug" der mit einer KI-Technologie verarbeiteten Daten ist daher ein entscheidender Faktor für die Frage, welche *Compliance*-Anforderungen erfüllt werden müssen. Der datenschutzrechtliche Begriff des Personenbezugs wird sehr weit verstanden. Passender wäre eigentlich „Personenbeziehbarkeit". Gem. Art. 4 Nr. 1 DS-GVO sind Daten personenbezogen, die sich auf identifizierte oder identifizierbare natürliche Personen beziehen und eine direkte oder indirekte Zuordnung zu einem Individuum ermöglichen, entweder allein oder in Kombination mit weiteren Daten. Beispiele sind: Name, Adresse, Geburtsdatum, Geschlecht, Fingerabdruck, Steuer-ID, Kontonummer, Standortdaten der Person, KFZ-Kennzeichen, Inhalte in *Social-Media*-Profilen, IP-Adressen.

Ein Personenbezug liegt auch dann vor, wenn die Identifizierung der Personen nur mit einer Kombination von Daten möglich ist. Das gilt selbst dann, wenn die erforderlichen Daten in verschiedenen Händen liegen und gar nicht ohne weiteres beim Datenverarbeiter zusammengeführt werden können. Beispielsweise ist die in einem *Server-Logfile* gespeicherte IP-Adresse ein dem Datenschutz unterliegendes personenbezogenes Datum. Daran ändert nichts, dass nur der Server-Betreiber oder Zugangs-Provider den Bezug zwischen der IP-Adresse und einem konkreten Anschlussinhaber*in/Nutzer*in herstellen kann und diese bei der Speicherung und Herausgabe

dieser Informationen strengen Regeln unterliegen. IP-Adressen dürfen daher auch von Dritten – beispielsweise von Strafverfolgungsbehörden oder Urheberrechtsinhabern zur Rechtsverfolgung – nur unter Einhaltung des Datenschutzrechts verarbeitet werden.

Nicht personenbezogen sind Daten, die entweder keinerlei Personenbezug aufweisen (wie z. B. Wetter- oder Geodaten) oder die vollständig anonymisiert wurden. Eine Pseudonymisierung führt nicht zum Entfall des Personenbezugs, da hier eine Identifizierung (durch den Inhaber des Zuordnungsschlüssels) noch möglich ist. Welche Maßnahmen zu einer vollständigen Anonymisierung von personenbezogenen Daten erforderlich sind, ist eine umstrittene und komplexe Frage, die noch immer nicht vollständig geklärt ist. Wie gesagt, genügt im Grundsatz eine Re-Identifizierbarkeit durch einen Dritten, um (noch) von einem Personenbezug auszugehen. Allerdings soll nach der Rechtsprechung des Europäischen Gerichtshofs (EuGH) eine rein theoretische Re-Identifizierungsgefahr nicht genügen. Ist die Re-Identifizierung nach aktuellem Stand der Technik praktisch nicht durchführbar oder nur mit immensem Aufwand machbar, ist auszuschließen, dass sie vorgenommen wird. Solche rein theoretischen Möglichkeiten führen, wie auch illegale Re-Identifizierungsmethoden, nicht zu einem Personenbezug.

Ebenfalls weit verstanden wird der Begriff der (Daten-)Verarbeitung. Datenschutzrelevante Verarbeitungen sind daher auf allen Stufen der Entwicklung, des Trainierens und der Verwendung von KI-Technologien denkbar. Für jede Verarbeitung, also das Erheben, Speichern, Verändern usw. bedarf es einer eigenen Rechtsgrundlage für den jeweils Verantwortlichen.

Unterschieden werden können die folgenden Phasen von Datenverarbeitung (nach: Landesbeauftragter für den Datenschutz und die Informationsfreiheit in Baden-Württemberg (LfDI BW), Diskussionspapier „Rechtsgrundlagen im Datenschutz beim Einsatz von Künstlicher Intelligenz"):

Erhebung von Trainingsdaten

Im Rahmen der Konzeption von KI-Modellen werden in der Regel Trainings-, Test- und Anwendungsdaten gesammelt, generiert, strukturiert und kategorisiert, ggf. auch anonymisiert. Je nach System werden in diesem Zuge öffentlich zugängliche Daten aus dem Internet oder anderen öffentlichen Quellen *gescrapt* (z. B. bei LLMs) oder eigene Daten erhoben (beispielsweise Daten von Verkehrsüberwachungskameras oder Nutzungsdaten von Internet-Diensten).

Verarbeitung von Trainingsdaten

Im Rahmen des Trainings eines KI-Systems werden die erhobenen Daten weiterverarbeitet. Zu datenschutzrelevanten Vorgängen kann es sowohl beim initialen Trainieren des KI-Modells als auch beim Nachtrainieren, Ergänzen und *Finetunen* kommen.

Anbieten von KI-Systemen

Die Verarbeitung personenbezogener Daten im Rahmen der Erstellung des Basismodells ist zu unterscheiden von Datenverarbeitungen im Rahmen der Entwicklung und des Angebots von KI-Systemen (zu den Unterschieden zwischen KI-Modellen und KI-Systemen siehe oben Abschnitt 2.f). Die Verantwortlichkeit kann hier einheitlich sein, wenn die Daten von demselben Unternehmen für das Basismodell erhoben und dann bestimmungsgemäß in dessen eigenem KI-System weiterverarbeitet werden (Beispiel: OpenAI, das sowohl das *Large Language Model* GPT entwickelt und trainiert als auch das KI-System ChatGPT anbietet und vertreibt).

Schwieriger ist der Fall, in dem der Entwickler des Basismodells und der Anbieter des KI-Systems auseinanderfallen (wie in dem oben in Abschnitt 2.h genannten Beispielsfall: Eine Organisation setzt eine eigene KI-Anwendung auf einem großen GPAI-Basismodell wie GPT auf). Ob hier die Bereitstellung des KI-Systems zu einer (weiteren) Verarbeitung von im Basismodell evtl. enthaltenen personenbezogenen Daten führt, wird derzeit lebhaft diskutiert.

Der Hamburgische Beauftragte für Datenschutz und Informationsfreiheit (HmbBfDI) vertritt in einem Diskussionspapier, dass jedenfalls der Datenbestand von LLMs bei technischer Betrachtung keine personenbezogenen Daten enthalten soll. Der Landesbeauftragte für den Datenschutz und die Informationsfreiheit in Baden-Württemberg (LfDI BW) vertritt eine differenzierte Auffassung. Ebenso der Europäische Datenschutzausschuss (EDSA) in einer aktuellen Stellungnahme.

Nutzung von KI-Systemen (Eingaben)

Werden bei Eingaben (*Prompts*) in ein KI-System personenbezogene Daten verwendet, bedarf es hierfür einer datenschutzrechtlichen Grundlage. Bei fremd-*gehosteten* Systemen ist zu berücksichtigen, dass diese Daten an den Anbieter übermittelt werden, der sich häufig in einem Drittstaat außerhalb der EU befinden wird.

Verwendung von Ergebnissen der Künstlichen Intelligenz (*Output)*

KI-generierte Texte oder Bilder können reale personenbezogene Daten oder Abbildungen enthalten. Bei deren Weiterverarbeitung (insbesondere, aber nicht nur bei Publikationen) ist auf eine datenschutzrechtliche Rechtsgrundlage zu achten. Außerdem sollte bedacht werden, dass generative KI zu Halluzinationen neigen. Aufgrund des Grundsatzes der Richtigkeit (s. u.) sollten von solchen Systemen produzierte personenbezogene Angaben wie Namen, Anschriften, Geburtsdaten usw. auf ihre Richtigkeit überprüft werden.

b. Zuweisung von Verantwortlichkeiten

Ein Datenschutzkonzept für KI-Lösungen bedarf der Festlegung von datenschutzrechtlichen Verantwortlichkeiten. Adressat des Datenschutzrechts ist der „Verantwortliche". An ihn richten sich die datenschutzrechtlichen Grundsätze aus Art. 5 Abs. 1 DS-GVO: Rechtmäßigkeit, Verarbeitung nach Treu und Glauben, Transparenz, Zweckbindung, Datenminimierung, Richtigkeit, Speicherbegrenzung, Integrität und Vertraulichkeit der Verarbeitung personenbezogener Daten (s. hierzu unten). Verantwortlich ist, wer allein oder gemeinsam mit anderen über die Zwecke und Mittel (das „Wie" und „Warum") der Verarbeitung von personenbezogenen Daten entscheidet (Art. 4 Abs. 1 Nr. 7 DS-GVO). Dies kann eine Person, eine Behörde, ein Unternehmen oder eine sonstige Stelle sein.

Das Datenschutzrecht kennt drei Formen der Verantwortungsverteilung: Alleinige Verantwortlichkeit, gemeinsame Verantwortlichkeit und Auftragsverarbeitung.

Verantwortlichkeitsformen nach dem Datenschutzrecht

Art der Verantwortlichkeit	Beispiel	Regelungsgrundlage in der DS-GVO
Alleinige Verantwortlichkeit: Der Verantwortliche entscheidet unabhängig von anderen über die Zwecke und Mittel der Verarbeitung von personenbezogenen Daten	Eine Einrichtung oder ein Unternehmen entwickelt ein KI-System selbst, stellt es bereit und nutzt es und entscheidet allein über die Zwecke und Mittel der Datenverarbeitung	Art. 24
Gemeinsame Verantwortlichkeit: Zwei oder mehr Parteien entscheiden gemeinsam über die Zwecke und Mittel der Verarbeitung personenbezogener Daten. Sie müssen eine Vereinbarung über die gemeinsame Verantwortlichkeit (sog. *Joint Controller Agreement*, JCA) schließen, in der insbesondere die Verpflichtungen und Verantwortlichkeiten zu den Betroffenenrechten und Informationspflichten festgelegt werden.	Ein KI-System oder -Modell wird mit Datensätzen verschiedener Einrichtungen trainiert. Sie treffen gemeinsam die wesentlichen Entscheidungen darüber, welche personenbezogenen Datenbestände zu welchen Zwecken erhoben und verarbeitet werden.	Art. 26
Auftragsverarbeitung: Ein Unternehmen verarbeitet nur im Auftrag und auf Weisung einer Organisation (Verantwortlicher) personenbezogene Daten. Zwischen dem Verantwortlichen und dem Auftragsverarbeiter muss eine Auftragsdatenverarbeitungsvereinbarung geschlossen werden.	Ein Anbieter eines KI-Systems bzw. Modells verarbeitet personenbezogene Daten unter Weisung und ausschließlich zu Zwecken eines Verantwortlichen und trainiert das System mit dessen Daten.	Art. 28

Wie die KI-VO, unterscheidet die DS-GVO nach Verantwortlichkeits- und Risikosphären. Wer über die Zwecke und Mittel der Verarbeitung personenbezogener Daten entscheidet, ist für die Einhaltung des Datenschutzrechts für die unter seiner Kontrolle erfolgenden Verarbeitungsvorgänge verantwortlich. Das bedeutet beispielsweise, dass der Entwickler/Anbieter eines KI-Basismodells für die Erhebung und weitere Verarbeitung von personenbezogenen Daten bei der Erstellung und Verbreitung des Modells verantwortlich ist. Der Anwender/Betreiber eines auf diesem Modell basierenden KI-Systems hat auf diese Sphäre keinen Einfluss. Sein Einflussbereich – und damit seine Verantwortlichkeit – bezieht sich dagegen beispielsweise auf:

- Die Entwicklung eigener KI-Systeme auf Basis von GPAI-Modellen, bspw. indem neue Datenquellen einbezogen werden oder das KI-Modell mit eigenen Daten nachtrainiert wird (wie bei „*Retrieval Augmented Generation" oder „Fine-tuning*", siehe den oben unter Abschnitt 2.h genannten Beispielsfall);

- Eingaben (*Prompts*) in das System: Werden personenbezogene Daten in ein KI-System eingegeben, ist vom Anwender auf das Datenschutzrecht zu achten. Personenbezogene oder geheime (Unternehmens-)Daten sollten daher nicht in solche Systeme eingegeben werden. Das gilt v. a. bei fremdgehosteten Systemen, zumal wenn sie von außereuropäischen Anbietern stammen;

- Die Verwendung von KI-Ergebnissen (*Output*): Werden KI-Erzeugnisse veröffentlicht oder anderweitig weiterverarbeitet, ist das Datenschutzrecht in Bezug auf hierin enthaltene personenbezogene Daten vom Anwender zu beachten.

a. Grundsätze des Datenschutzrechts beim Einsatz von KI

Bei der Entwicklung und dem Einsatz von KI-Lösungen sind die allgemeinen, in Art. 5 DS-GVO genannten „Grundsätze für die Verarbeitung personenbezogener Daten" zu beachten. Deren wesentliche Aspekte werden im Anschluss kurz erklärt.

Rechtmäßigkeit der Verarbeitung (Art. 5 Abs. 1.a Alt. 1, Art. 6 DS-GVO)

Die Verarbeitung personenbezogener Daten bei KI-Modellen und -Systemen durch den Verantwortlichen ist rechtmäßig, wenn sie auf einer einschlägigen Rechtsgrundlage basiert und die sonstigen datenschutzrechtlichen Vorgaben eingehalten werden. Bereits in der Entwicklungs- und Designphase müssen gemäß Art. 25 DS-GVO Datenschutzanforderungen berücksichtigt werden („*Privacy by Design*"). Je nach Verarbeitungsphase – wie Datenerhebung, Training, Nutzung oder Weiterentwicklung – können unterschiedliche Rechtsgrundlagen relevant sein. In Betracht kommen insbesondere:

- die Einwilligung (Art. 6 Abs. 1.a DS-GVO, Art. 9 Abs. 2.a DS-GVO);

- die Verarbeitung im Rahmen der Vertragserfüllung (Art. 6 Abs. 1.b DS-GVO);

- die Verarbeitung zur Wahrung berechtigter Interessen (Art. 6 Abs. 1.f DS-GVO);

- sonstige/spezifische nationale Rechtsgrundlagen.

Datenschutzrechtliche Rechtsgrundlagen im Kontext von KI

Anders als in anderen Bereichen, sind Einwilligungen für Datenverarbeitungsvorgänge im KI-Kontext häufig keine sinnvolle Rechtsgrundlage und zwar aus verschiedenen Gründen:

Die Einwilligung kommt als Rechtsgrundlage für Konstellationen in Betracht, bei denen der Verantwortliche mit dem „Betroffenen" (die Person, deren personenbezogene Daten verarbeitet werden), ein direktes Verhältnis hat. Das ist bei der massenhaften Datenverarbeitung („*Big Data*"), insbesondere durch die Anbieter generativer KI-Modelle, in aller Regel nicht der Fall.

Einwilligungen können jederzeit widerrufen werden können, woraufhin die Daten zu löschen sind bzw. nicht mehr verwendet werden dürfen. Ob und wie dies insbesondere bei personenbezogenen Daten umgesetzt werden soll, die in ein KI-Datenmodell eingeflossen sind, ist unklar.

Einwilligungen müssen auf informierten Entscheidungen basieren. Den Betroffenen muss also z. B. verständlich vermittelt werden, welchen Datenverarbeitungsprozessen sie zustimmen, welche Daten erhoben und wie sie verarbeitet werden, wie die Technologie grundsätzlich funktioniert und welchen Zwecken die Datenverarbeitung dient. Diese Transparenz herzustellen stellt bei KI-Systemen aufgrund ihrer Komplexität eine besondere Herausforderung dar. Ist sie nicht gegeben, scheidet die Einwilligung als Rechtsgrundlage für die Verarbeitung aus.

Einwilligungen müssen freiwillig abgegeben werden. In Arbeits- und anderen Abhängigkeitsverhältnissen sind sie als Rechtsgrundlage für Datenverarbeitungen daher schon per se wenig geeignet.

Von vorrangiger Bedeutung dürfte als Rechtsgrundlage, jedenfalls für das Anlernen und Training großer KI-Modelle, das „berechtigte Interesse" sein. Hiernach ist es gestattet, personenbezogene Daten (ohne Einwilligung) zu verarbeiten, wenn dies zur Wahrung der berechtigten Interessen des Verantwortlichen oder eines Dritten erforderlich ist und dem nicht überwiegende Interessen oder Grundrechte und Grundfreiheiten der betroffenen Person entgegenstehen. Die Rechtsgrundlage basiert auf einer Interessenabwägung. Beispielsweise kann hiernach die Verarbeitung von Gesundheitsdaten zum Trainieren einer KI gerechtfertigt sein, wenn diese dazu dient und in der Lage ist, wichtige Erkenntnisse zu Krankheiten und Behandlungsmethoden zu ermitteln. Ob dies tatsächlich zutrifft, muss im Einzelfall geprüft werden. Der Europäische Datenschutzausschuss (EDSA) hat in einer Stellungnahme jüngst vertreten, dass sich die Entwicklung und der Einsatz von KI-Modellen grundsätzlich auf ein berechtigtes Interesse stützen können. Ob dies jeweils der Fall ist, soll von einer dreistufigen Prüfung abhängen, ob (a) der Anspruch auf Datenverarbeitung legitim ist (b) die Datenverarbeitung notwendig ist und sie (c) im Einklang mit den Grundrechten der Betroffenen steht.

Merke: Auch mit Blick auf das Datenschutzrecht ist eine gründliche Konzeptionierung und Planung von KI-Lösungen das A und O!

Transparenz (Art. 5 Abs. 1.a Alt. 3 DS-GVO)

Ein weiteres zentrales Element der DS-GVO ist die Verpflichtung zur Transparenz. Betroffene sind vom Verantwortlichen über die Details der Verarbeitung ihrer personenbezogenen Daten zu informieren. Sie sollen verstehen können, welche Daten zu welchen Zwecken erhoben und wie sie verarbeitet werden. Aus Art. 13 und 14 DS-GVO ergeben sich erhebliche Anforderungen an die den Betroffenen mitzuteilenden Pflichtangaben. Transparenz wird durch Datenschutzerklärungen, -richtlinien und -hinweisen geschaffen. Bei komplexen KI-Systemen stellt sich auch diesbezüglich die besondere Herausforderung, die Prozesse, Abläufe und Logiken verständlich zu erklären.

Zweckbindung (Art. 5 Abs. 1.b DS-GVO) und Speicherbegrenzung (Art. 5 Abs. 1.e DS-GVO)

Der Zweckbindungsgrundsatz stellt sicher, dass Daten nur für die Zwecke genutzt werden, für die sie erhoben wurden. Eine Verarbeitung zu einem anderen Zweck erfordert in der Regel eine eigene Rechtsgrundlage. Beispielsweise könnte sich die Frage stellen, ob Personaldaten nach ihrer Erhebung (die im Zweifel zur Erfüllung des Arbeitsvertrags nötig und damit auch ohne Einwilligung zulässig war) ohne Weiteres zum Training einer KI genutzt werden dürfen, die Personalentscheidungen erleichtern soll. Hierin könnte eine unzulässige Zweckänderung liegen, die einer eigenen Rechtsgrundlage bedürfte. Diese wiederum könnte in einem berechtigten Interesse des Arbeitgebers liegen, was angesichts der konkreten Umstände des jeweiligen Einzelfalls geprüft werden muss.

Merke: Bereits vorhandene personenbezogene Daten dürfen nicht ohne Weiteres zu anderen Zwecken weiterverwendet werden!

Mit dem Zweckbindungsgrundsatz eng zusammen hängt der Grundsatz der Speicherbegrenzung. Wann der Zweck einer Erhebung von personenbezogenen Daten für KI-Systeme erreicht ist, hängt von der Ausgestaltung und dem Verwendungszweck ab. Wenn eine Identifizierbarkeit der Betroffenen für den ursprünglichen Erhebungszweck nicht mehr erforderlich ist, ist der Personenbezug aufzuheben. Grundsätzlich bedeutet das, dass die Daten entweder gelöscht, anonymisiert oder zumindest pseudonymisiert werden müssen. Geschieht dies nicht, dürfen Weiterverarbeitungen nur vorgenommen werden, wenn hierfür eine eigene Rechtsgrundlage gegeben ist. Werden die Daten dagegen vollständig anonymisiert (wodurch der Personenbezug restlos entfällt), können die Daten beliebig aufbewahrt und weiterverwendet werden.

Datenminimierung (Art. 5 Abs. 1.c DS-GVO)

Personenbezogene Daten dürfen nach dem Grundsatz der Datenminimierung nur insofern verarbeitet werden, als es für den Verarbeitungszweck notwendig ist. Dieser Grundsatz ist eine konkrete Ausprägung des Verhältnismäßigkeitsprinzips, das eine sorgfältige Abwägung erfordert. Demnach muss die Verarbeitung personenbezogener Daten stets auf das erforderliche Minimum beschränkt werden.

Besonders herausfordernd ist diese Abwägung im Kontext des KI-Trainings, da hierfür in der Regel große Datenmengen (*Big Data*) benötigt werden. Wenn Organisationen beispielsweise ein KI-Modell mit eigenen Daten anreichern und/oder finetunen wollen, müssen sie im Vorfeld eine Verhältnismäßigkeitsprüfung durchführen. Der Zweck der Verarbeitung der Daten muss im Verhältnis zu den hiermit einhergehenden Eingriffen in die Interessen der Betroffenen stehen. Eine Anonymisierung der Daten ist – soweit möglich und zweckkonform – bei der Abwägung zu berücksichtigen.

Richtigkeit (Art. 5 Abs. 1.d DS-GVO)

Personenbezogene Daten dürfen nur verarbeitet werden, wenn sie sachlich richtig und aktuell sind. Unrichtige Daten müssen unverzüglich berichtigt oder gelöscht werden.

Anbieter von großen KI-Modellen könnte dieser Grundsatz vor große Herausforderungen stellen. Es dürfte nur unter erheblichen Schwierigkeiten, wenn überhaupt, möglich sein, aus einem *Large Language Model* einzelne Daten zu extrahieren, um sie zu berichtigen. Für Anwender solcher Systeme bedeutet der Grundsatz der Richtigkeit z. B., bei der Nutzung von im *Output* enthaltenen personenbezogenen Daten auf deren Richtigkeit zu achten. Hierbei ist zu bedenken, dass generative KI-Systeme zu Halluzinationen neigen. Sie können und sie werden häufig falsche (auch personenbezogene) Informationen generieren. Die Ergebnisse sollten daher kritisch überprüft werden, bevor sie veröffentlicht oder sonstwie weiterverwendet werden.

Grundsatz der Integrität und Vertraulichkeit (Art. 5 Abs. 1.f DS-GVO)

Dieser allgemeine Grundsatz zielt auf die Datensicherheit ab. Personenbezogene Daten sollen sicher verarbeitet werden. Es soll u. a. vermieden werden, dass Unbefugte hierauf zugreifen, sie verloren oder verfälscht werden können. Zu diesem Zweck sind entsprechende, dem jeweiligen Risiko angemessene, technische und organisatorische Maßnahmen (TOMs) vorzusehen.

Rechenschaftspflicht (Art. 5 Abs. 2 DS-GVO)

Ein zentrales Prinzip der DS-GVO ist die Rechenschaftspflicht. Verantwortliche müssen Datenschutzvorschriften nicht nur einhalten, sondern dies auch dokumentieren und gegenüber den Aufsichtsbehörden nachweisen können. Auch hierfür sind technische und organisatorische Maßnahmen vorzusehen (Art. 24 Abs. 1 DS-GVO). Sie umfassen beispielsweise die Erstellung von Verzeichnissen der Verarbeitungstätigkeiten, die Benennung eines Datenschutzbeauftragten, Datenschutz-Folgeabschätzungen, die Implementierung von internen Richtlinien oder die Ein- und Durchführung eines Datenschutzmanagementsystems. Welche Maßnahmen konkret zu treffen sind, hängt vom jeweiligen Nutzungskontext ab.

Privacy by Design, *Privacy by Default* und technische und organisatorische Maßnahmen (Art. 25 ff. DS-GVO)

Datenschutzmaßnahmen müssen bereits in der Entwicklungsphase von KI-Systemen berücksichtigt (*Privacy by Design*) und datenschutzfreundliche Voreinstellungen müssen implementiert werden (*Privacy by Default*). Bei Datenverarbeitungen, die hohe Risiken für die Betroffenen bergen, sind Datenschutz-Folgeabschätzungen vorzunehmen und ggf. Datenschutz-Managementsysteme zu implementieren.

Beispiele für technische und organisatorische Maßnahmen (TOMs) zur Gewährung des Datenschutzes beim Einsatz von generativen GPAI-Systemen in einer Organisation

- Schulungen und anderen datenschutzbezogene Sensibilisierungsmaßnahmen des mit der eingesetzten KI-Lösung befassten Personals
- Einsatz von *Prompt*-Filtern, um die Verwendung personenbezogener Daten beim *Input* zu minimieren
- Einführung interner Richtlinien, *Policies* oder Verhaltenskodizes über die Verwendung personenbezogener Daten bei Eingaben in das GPAI-System
- Vergabe von pseudonymen oder anonymen Benutzeraccounts für das KI-System
- Kurzfristige automatisierte Löschung oder Deaktivierung der Verlaufshistorie für Benutzereingaben
- Zustimmung zur Weiterverarbeitung von Eingaben durch den Anbieter nicht erteilen/deaktivieren
- Der Nutzung Hinweise auf die datenschutzfreundliche Verwendung des Systems vorschalten

Welche Maßnahmen sinnvoll sind und umgesetzt werden können, hängt von der jeweiligen Konstellation ab. Der Einfluss auf technische Maßnahmen zur Gewähr des Datenschutzes ist bei Standard-Produkten naturgemäß deutlich geringer als bei selbst entwickelten oder angepassten Lösungen. Auch im Datenschutzrecht gilt: Es wird nichts verlangt, das unmöglich ist! Bei gründlicher Analyse und Planung sind die Auflagen in aller Regel auch machbar.

Keine automatisierte Letztentscheidung (Art. 22 DS-GVO)

Sofern die Betroffenen nicht ausdrücklich eingewilligt haben, dürfen Entscheidungen, die „Rechtswirkung" auf die Betroffenen haben, nicht ausschließlich auf einer vollautomatisierten Datenverarbeitung basieren. Die Letztentscheidung über Einstellungen, die Kreditwürdigkeit, Asyl- oder Bürgergeldanträge, Vertragsabschlüsse usw. muss also stets von einem Menschen getroffen werden. Zwar dürfen KI-Systeme (unter Einhaltung der Rechtsgrundlagen) eingesetzt werden, um solche Entscheidungen vorzubereiten. Der für die Letztentscheidung zuständigen Person müssen aber tatsächliche Entscheidungsfreiheit und -spielräume zustehen. Wird KI in solchen Kontexten eingesetzt, erfordert dies beispielsweise, dass die Entscheider deren Funktion in gewissem Maß verstehen, um Entscheidungsvorschläge der KI hinterfragen zu können. Diese Anforderung ist sowohl bei der Gestaltung betrieblicher Prozesse zu beachten als auch bei der Konzeption und Ausgestaltung der Systeme durch deren Anbieter (s. oben Abschnitt 2.i).

d. Weiterführende Literatur, Checklisten und Praxistipps zur Umsetzung datenschutzrechtlicher Maßnahmen bei KI

Viele Vorgaben, insbesondere die dargestellten Prinzipien der Datenverarbeitung, sind hochabstrakt und müssen auf operative Vorgaben heruntergebrochen werden. Welche TOMs sind konkret erforderlich? Welche datenschutzfreundlichen Techniken müssen eingesetzt werden? Welche Dokumente müssen produziert werden? Für die Klärung von solchen Detailfragen führen die Datenschutzbehörden mit Diskussionspapieren und Entwürfen einen Diskurs, dessen Ergebnis bislang nur als vorläufig gelten kann. Es ist daher unabdingbar, sich auf dem Laufenden zu halten und die Entwicklung im Blick zu behalten.

Im Literaturverzeichnis am Ende dieser Rechtshilfe findet sich eine Übersicht mit Literaturtipps, Checklisten und anderen hilfreichen Materialien zum Thema Datenschutzrecht und KI.

Merke: Offizielle Stellungnahmen von Datenschutzaufsichtsbehörden, der „Konferenz der unabhängigen Datenschutzaufsichtsbehörden des Bundes und der Länder" (DSK) oder des „Europäischen Datenschutzausschusses" (EDSA) geben einen guten Einblick in die Haltung und Denkweise der zuständigen Aufsichtsbehörden. Sie können daher dazu dienen, Risiken zu minimieren, sind aber in aller Regel nicht verbindlich und können sich mit der Zeit auch ändern.

4. Urheberrecht beim Einsatz von KI

Das Urheberrecht schützt kreative Werke wie Bilder, Texte, Software, Videos oder Musik. Besteht ein Urheberrechtsschutz, darf das Werk nur genutzt, also kopiert, verbreitet oder öffentlich wiedergegeben (beispielsweise online gestellt und geteilt) werden, wenn hierfür eine Rechtsgrundlage besteht. Eine solche kann in einer Erlaubnis des Rechteinhabers liegen, beispielsweise einem individuellen Nutzungsrechtsvertrag oder einer offenen Lizenz wie Creative Commons. Manche Nutzungen sind auch per Gesetz durch eine Nutzungsfreiheit („Schrankenbestimmung") gestattet.

Urheberrecht in der Paritätischen Arbeit

Weitere Details über das Urheberrecht im Kontext der Paritätischen Arbeit werden im Leitfaden „Urheberrecht und Persönlichkeitsrechte in der Paritätischen Arbeit" erklärt.

Im KI-Kontext kann das Urheberrecht in verschiedener Hinsicht von Bedeutung sein. Urheberrechtliche Fragen stellen sich z. B. in folgender Hinsicht:

- Training von KI-Modellen:
  - Darf man urheberrechtlich geschütztes Material zum Trainieren verwenden?
  - Kann ich als Inhaber*in von Urheberrechten verhindern, dass mein Material von Dritten zum Trainieren verwendet wird?
  - Müssen KI-Anbieter mir als Urheber eine Vergütung zahlen?
- *Prompting* und Verwendung von KI-Systemen:
  - Sind *Prompts urheberrechtlich geschützt oder darf sie jeder verwenden?*
  - *Kann man durch Prompts Urheberrechte verletzen?*
- Output (v. a. von generativen KI-Systemen):
  - Sind die Bilder oder Texte, die ich von der KI generieren lasse, urheberrechtlich geschützt?
  - Können sie von anderen nachgenutzt werden, ohne dass ich dem zustimme?

Wie auch andere Rechtsfragen sind viele Aspekte des Urheberrechts im Kontext mit KI heute noch nicht abschließend geklärt. Zahlreiche Rechtsstreitigkeiten sind vor Gerichten in verschiedenen Ländern anhängig, in denen unter anderem die oben genannten Fragen diskutiert werden. Die Klärung der wesentlichen Fragen dürfte noch einige Zeit in Anspruch nehmen.

a. Urheberrecht in der Anlernphase

Beim Training von KI-Modellen werden häufig urheberrechtlich geschützte Werke kopiert. Um ein Foundation Model einer Bildgenerierungs-KI wie Midjourney oder Stable Diffusion zu erstellen, werden beispielsweise Millionen von zumeist öffentlich verfügbaren Bilddateien genutzt, damit Abbildung und Bildunterschrift miteinander abgeglichen werden können. Auf diese Weise lernt die KI, was auf den Bildern zu sehen ist, um dadurch in die Lage versetzt zu werden, selbst Bilder entsprechend den Eingaben des Nutzers generieren zu können.

Im Zuge dieses Lernprozesses werden die Bilder im Zweifel zumindest vorübergehend gespeichert. Aus urheberrechtlicher Sicht sind dies „Vervielfältigungen" (§ 16 Urheberrechtsgesetz – UrhG). Sie sind nur gestattet, wenn es aufgrund eines (Lizenz-) Vertrags vom Rechteinhaber oder durch eine gesetzliche Schrankenbestimmung gestattet ist.

Gerade in dem genannten Beispielsfall – dem Training eines generativen GPAI-Modells – ist es keine realistische Option, für jedes verwendete Werk eine Erlaubnis einzuholen. Das würde erfordern, Millionen oder gar Milliarden von individuellen Vereinbarungen zu schließen, was faktisch unmöglich wäre. Amerikanische KI-Unternehmen wie OpenAI berufen sich daher auf gesetzliche Nutzungserlaubnisse wie die so genannte *fair use doctrine* (Sec. 107 US Copyright Act). Hierbei handelt es sich um eine Ausnahme vom US Copyright, nach der unter bestimmten Umständen Nutzungen von geschütztem Material ohne Zustimmung des Rechteinhabers zulässig sind. Ob *fair use* für das Training von KI-Modellen greift, ist rechtlich umstritten und derzeit Gegenstand zahlreicher Rechtsstreitigkeiten.

Nach europäischem Urheberrecht fallen Kopien, die in der Lernphase von KI-Modellen entstehen, unter Umständen unter Art. 3, 4 der DSM-Richtlinie (Richtlinie (EU) 2019/790 vom 17. April 2019 über das Urheberrecht und die verwandten Schutzrechte im digitalen Binnenmarkt). Die hierin enthaltenen Schrankenbestimmungen wurden in Deutschland in den Paragrafen 44b und 60d UrhG umgesetzt. Sie besagen, dass urheberrechtlich geschütztes Material ohne Zustimmung kopiert werden darf, wenn dies dem „Text und Data Mining" (TDM) dient.

> ***Text und Data Mining* gemäß § 44b Abs. 1 UrhG**
>
> *„Text und Data Mining* ist die automatisierte Analyse von einzelnen oder mehreren digitalen oder digitalisierten Werken, um daraus Informationen insbesondere über Muster, Trends und Korrelationen zu gewinnen."

Ob das Trainieren von KI als „*Text und Data Mining*" angesehen werden kann, ist umstritten. In der ersten in Europa zu dieser Frage ergangenen Gerichtsentscheidung hat sich das Landgericht Hamburg (Urteil vom 27.09.2024 – Az. 310 O 227/23) mit differenzierten Argumenten für die Anwendbarkeit ausgesprochen. Auch Art. 53 Abs. 1.c der KI-Verordnung spricht dafür. Hiernach sollen Anbieter von KI-Modellen für allgemeine Zwecke eine Politik zur Einhaltung von Art. 4 Abs. 3 der DSM-Richtlinie vorsehen. Diese Regel gibt Rechteinhabern unter gewissen Umständen das Recht, durch ein technisches „*opt-out*" zu erklären, dass ihre Veröffentlichungen nicht für Zwecke des Text und Data Minings genutzt werden dürfen. Die Details dieser Zusammenhänge sind komplex und umstritten. Dennoch erkennen die meisten juristischen Kommentator:innen (zurecht) an, dass dieser Verweis keinen Sinn ergäbe, wenn der EU-Gesetzgeber nicht davon ausginge, dass die TDM-Schranken auf das KI-Training anwendbar sind. Bewahrheitet sich dies in späteren Entscheidungen durch höhere Gerichte (insbesondere den Europäischen Gerichtshof, EuGH), wären Nutzungen im Zuge von KI-bezogenen Datenanalysen bzw. des KI-Trainings urheberrechtlich generell zulässig. Den betroffenen Rechteinhabern stünde es aber – technische Lösungen vorausgesetzt – frei, ihre Werke durch eine *Opt-out*-Erklärung von der Nutzung auszuschließen (s. § 44 Abs. 3 UrhG).

b. Urheberrecht und *Prompting*

„*Prompts*" sind die Befehle, die ein*e Anwender*in in eine (generative) KI eingibt, damit sie einen Output – wie ein Bild oder einen Text – erzeugen kann. Einen guten *Prompt* zu verfassen ist eine menschliche Leistung, die durchaus kreativ sein kann. In der Regel werden die *Prompts* dennoch kein Werk darstellen und damit urheberrechtlich auch nicht geschützt sein. Gibt jemand etwa in einen Bildgenerator ein: „Gestalte mir eine fotorealistische Darstellung von einem Eichhörnchen, das auf einem Rennrad im Sommer durch das Brandenburger Tor fährt, und dabei die Flagge Frankreichs schwenkt" mag das zwar eine originelle Idee sein. Die Idee ist jedoch urheberrechtlich nicht geschützt und bei den Worten handelt es sich nicht um ein literarisches Werk, sondern lediglich um Anweisungen, die in Alltagssprache formuliert sind. Simple Sätze, einzelne Worte, Parameter und Eckdaten sind keine urheberrechtlich geschützten Werke. Dies ändert sich auch nicht, wenn die Einzelangaben weiter verfeinert und (ggf. eine Vielzahl) weiterer Parameter hinzugefügt werden.

Letztlich ist zwar denkbar, dass *Prompts* in seltenen Einzelfällen Urheberrechtsschutz zukommt. Das wird jedoch eher eine Ausnahme als der Regelfall sein. Und selbst wenn ein *Prompt* ausnahmsweise urheberrechtlich geschützt ist, gilt das nicht automatisch auch für das durch die KI auf dieser Basis generierte Erzeugnis (siehe hierzu im nächsten Abschnitt). Dieses wird nicht vom Verwender gestaltet, sondern von der KI. Das zeigt sich an dem Umstand, dass die KI, selbst bei identischem *Prompt*, jedesmal ein anderes Bild erzeugen wird.

c. KI-*Output und Urheberrecht*

Urheberrechtsschutz von KI-generierten Gestaltungen

Eine viel diskutierte Frage im Kontext von KI und Urheberrecht ist, ob der *Output,* also die durch generative KI erzeugten Ergebnisse, urheberrechtlich geschützt sind. Die Frage ist praktisch äußerst relevant: Besteht kein Urheberrechtsschutz, kann der *Output* ohne Zustimmung beliebig weiterverwendet, also kopiert, geteilt, auf Werbeplakate gedruckt usw., werden. Zahllose technische Gestaltungen wären damit gemeinfrei und könnten frei nachgenutzt werden.

Das Grundprinzip zur Beantwortung dieser Frage ist einfach: **Das Urheberrecht ist ein personenbezogenes Schutzrecht. Es gilt ausschließlich für von Menschen gemachte kreative Gestaltungen. Bei reinen KI-Kreationen fehlt es am „Faktor Mensch", ein Urheberrecht kommt hierfür nicht in Betracht.**

Wird der KI-*Output* dagegen von Menschen überarbeitet, ist es möglich, dass die durch das Zusammenwirken von Mensch und Maschine hervorgebrachte „hybride Schöpfung" urheberrechtlich geschützt ist. Dies ist der Fall, wenn es sich bei dem menschlichen Beitrag um eine „persönliche geistige Schöpfung" (§ 2 Abs. 1 UrhG) handelt. Bearbeitet also die Anwender*in das im Beispiel genannte Eichhörnchen-Bild der KI mit einem Bildbearbeitungs-Programm oder druckt es aus und übermalt oder ergänzt es, kann und wird häufig ein urheberrechtlich geschütztes Werk entstehen.

Gleiches gilt, wenn KI-generierter *Output* mit menschlichen Gestaltungen kombiniert wird. In einem frühen Fall hat das US-Copyright-Office beispielsweise entschieden, dass an den von einer KI generierten Abbildungen eines Comics kein Copyright besteht, aber durchaus an dem von der Autorin des Comics geschriebenen Text. **Dies zeigt, dass bei Fragen des Urheberrechtsschutzes stets zwischen den Bestandteilen eines Erzeugnisses – sofern sie abgrenzbar sind – unterschieden werden muss.** So käme im Beispiel des bearbeiteten KI-Eichhörnchen-Bildes Urheberrechtsschutz nur für die bearbeitete Fassung in Betracht, während das ursprüngliche, rein von der KI generierte, Bild gemeinfrei ist.

Urheberrechtsverletzung durch Veröffentlichung von KI-*Output*

Grundsätzlich kann unbearbeiteter, von einer KI generierter *Output* ohne weiteres veröffentlicht und in jeder Form und zu jedem Zweck (auch beispielsweise kommerziell, in der Werbung) nachgenutzt werden. Urheberrechtliche Bedenken könnten jedoch entstehen, wenn ein KI-Erzeugnis einem bestimmten Trainingsinhalt gleicht oder sehr ähnelt. Denn die Veröffentlichung einer Reproduktion eines geschützten Werkes ist urheberrechtlich nur mit Zustimmung des Rechteinhabers zulässig. Wie die Reproduktion zustande gekommen ist, ob mittels einer KI, durch Abschreiben, nachzeichnen oder per Foto, ist belanglos.

Ob generative KI identische oder sehr ähnliche Abbilder von Trainingsdaten reproduzieren können, hängt von der Technik ab. Stilimitationen (im Sinne von: „Generiere mir ein Bild im Stil von Andy Warhol") werden generell möglich sein. Dies ist – auch wenn sie publiziert werden – urheberrechtlich jedoch unproblematisch. Der Stil ist als solcher nicht geschützt. Jeder darf Werke erstellen und veröffentlichen, die denen anderer Künstler*innen ähneln und sich hieran anlehnen. Die Grenze zur Urheberrechtsverletzung ist erst überschritten, wenn es zu erheblichen Gemeinsamkeiten mit einem konkreten Werk kommt.

Für (nahezu) identische Reproduktionen bestimmter Werke sind generative KI-Systeme generell nicht gedacht: Wie der Name schon sagt, sollen sie eigene Inhalte generieren und nicht vorbestehendes Material kopieren. Die Anbieter solcher Systeme haben ein eigenes Interesse, dass 1:1-Reproduktionen nicht möglich sind. Das würde die Debatte um die Urheberrechtswidrigkeit von KI-Erzeugnissen nur weiter verschärfen. Wenn mit generativer KI (nahezu) identische Kopien erzeugt werden können, dürfte das also eher auf einem Fehler als auf einer

gewünschten Funktionalität basieren. Mit allerhand technischem Geschick und Know-How mag es dennoch möglich sein, Reproduktionen zu erzeugen, die konkreten Werken sehr ähnlich sind. Amerikanischen Wissenschaftler*innen ist dies im Rahmen einer Studie offenbar gelungen. Ist dies der Fall, sollten sie nicht veröffentlicht oder geteilt werden, um Urheberrechtsverletzungen zu vermeiden. Eine private Nutzung ist gem. § 53 Abs. 1 UrhG in aller Regel erlaubt.

5. Schlussbemerkung

Rechtsfragen im Kontext von KI sind derzeit noch stark im Fluss. **Neue Regularien wie die KI-VO sind in vielen Details noch sehr unklar. Auch bei der Anwendung bestehender Regelungen wie dem Datenschutz- und Urheberrecht auf KI bleiben noch viele Fragen offen.** Daher gilt es, die Entwicklung im Blick zu behalten, wenn man KI in der Organisation einführen und/oder selbst entwickeln will. Rechtlichen Rat einzuholen, wird sich in vielen Fällen empfehlen. Solche Herausforderungen sollten jedoch nicht vom Einsatz von KI abschrecken. Die Technologie hat großes Potenzial, das nur realisiert werden kann, wenn auch gewisse Unsicherheiten und Risiken in Kauf genommen werden.

Literatur

Die nachstehende Übersicht enthält Hinweise auf Literatur, Checklisten und Praxistipps, die aus Sicht des Autors dieser Rechtshilfe besonders nützlich ist. Sie erhebt keinen Anspruch auf Vollständigkeit oder Aktualität. Da viele dieser Publikationen (v. a. solche von offiziellen Stellen wie den Datenschutzbehörden) aufgrund der Dynamik des Themas mehr oder weniger kurzfristig aktualisiert werden, sollte man darauf achten, die jeweils aktuelle Fassung zu verwenden.

KI-VO

- Bitkom (Hrsg.): **Umsetzungsleitfaden zur KI-Verordnung**, Version 1.0, 29.10.2024, https://www.bitkom.org/Bitkom/Publikationen/Umsetzungsleitfaden-zur-KI-Verordnung-EU-2024-1689
  Umfangreicher Leitfaden zur KI-VO mit Checklisten, der auch als Schritt-für-Schritt-Anleitung verwendet werden kann. Er ist sowohl als PDF (220 Seiten) als auch als Click-Through-Tool verfügbar.

- European AI Office (Hrsg.): **General-Purpose AI Code of Practice**, 1. Entwurf 14.11.2025, https://digital-strategy.ec.europa.eu/en/policies/ai-code-practice
  Entwurf für einen Kodex zur Umsetzung der KI-VO für Anbieter von KI-Modellen.

Diverse

- iRights.info (Hrsg.): **Texte zu Rechtsfragen und Künstlicher Intelligenz**, erscheint fortlaufend, https://irights.info/kategorie/kuenstliche-intelligenz
  *Eine Reihe allgemeinverständlicher Texte verschiedener Autor*innen zum Thema KI und Recht (u. a. Urheberrecht, Datenschutz, freie Lizenzen).*

- Der Paritätische Gesamtverband(Hrsg.): **Künstliche Intelligenz in der Sozialen Arbeit**, 2. Auflage, 12/2024, https://www.der-paritaetische.de/fileadmin/user_upload/Schwerpunkte/Digitalisierung/doc/ki/KI_Textsammlung_Update2024_final.pdf
  Textsammlung mit vielen nützlichen Praxistipps, einem „Leitfaden zur Einführung von generativer KI in Non-Profit-Organisationen" und vielem mehr (v. a. Praxistipps, Begriffserklärungen).

- Bitkom (Hrsg.): **Generative KI im Unternehmen Rechtliche Fragen zum Einsatz generativer Künstlicher Intelligenz im Unternehmen**, 2024, https://www.bitkom.org/Bitkom/Publikationen/Generative-KI-im-Unternehmen
  Umfangreicher Leitfaden zu verschiedenen Rechtsthemen im Zusammenhang mit KI (u. a. Datenschutz, Arbeitsrecht, Urheberrecht, Haftungsrecht).

- Evangelisches Werk für Diakonie und Entwicklung e. V. (Hrsg.): **Leitlinien zur Nutzung von Künstlicher Intelligenz**, 27.2.2024 https://www.diakonie.de/diakonie_de/user_upload/diakonie.de/PDFs/Publikationen/Leitlinien_zur_Nutzung_von_KI_im_EWDE__März_2024__extern.pdf
  Leitlinien für den KI-Einsatz, die u.a. ein Kapitel über Datenschutz enthalten. Sie haben Empfehlungscharakter und sollen eine Orientierungshilfe für alle Mitarbeitenden des EWDE bieten.

- Bertelsmann Stiftung, iRights.Law (Hrsg.): **KI in Unternehmen – Ein Praxisleitfaden für rechtliche Fragen**, 1. Auflage, 1.2.2021, https://www.bertelsmann-stiftung.de/de/publikationen/publikation/did/ki-in-unternehmen-all
  Umfangreiche Handreichung zu verschiedenen Rechtsthemen im Zusammenhang mit KI (u. a. Datenschutz, Urheberrecht, Haftungsrecht).

Datenschutzrecht

- Konferenz der unabhängigen Datenschutzaufsichtsbehörden des Bundes und der Länder (DSK) (Hrsg.): **Orientierungshilfe der Konferenz der unabhängigen Datenschutzaufsichtsbehörden des Bundes und der Länder vom 6. Mai 2024: Künstliche Intelligenz und Datenschutz.**, 1. Version vom 6.5.2024, https://www.datenschutzkonferenz-online.de/media/oh/20240506_DSK_Orientierungshilfe_KI_und_Datenschutz.pdf
  Anwendungsorientierte praktische Leitlinien zur Auswahl, Implementierung und Nutzung von KI-Anwendungen für Behörden und Unternehmen mit Schwerpunkt auf Large Language Models (LLMs). Sie richtet sich primär an datenschutzrechtlich Verantwortliche, mittelbar aber auch an Entwickler und Anbieter von KI-Systemen. Die Orientierungshilfe ist kein verbindlicher Beschluss der DSK, gibt aber einen übersichtlichen und gut strukturierten Einblick in die Haltung der Datenschutzaufsichtsbehörden.

- Europäischer Datenschutzausschuss (EDSA) (Hrsg.): **Stellungnahme zur Verwendung personenbezogener Daten für die Entwicklung und Einführung von KI-Modellen**, 18.12.2024, https://www.edpb.europa.eu/news/news/2024/edpb-opinion-ai-models-gdpr-principles-support-responsible-ai_de
  Einschätzung des Datenschutzausschusses zum Datenschutzrecht bei KI-Modellen.

- Hamburgischer Beauftragter für Datenschutz und Informationsfreiheit (HmbBfDI) (Hrsg.): **Checkliste zum Einsatz LLM-basierter Chatbots**, 13.11.2023, https://datenschutz-hamburg.de/fileadmin/user_upload/HmbBfDI/Datenschutz/Informationen/20231113_Checkliste_LLM_Chatbots_DE.pdf
  Speziell auf Anwender von Chatbots wie ChatGPT, Luminous oder Bard ausgerichtete kurze Checkliste.

- Landesbeauftragter für den Datenschutz und die Informationsfreiheit in Baden-Württemberg (LfDI BW) (Hrsg.): **Rechtsgrundlagen im Datenschutz beim Einsatz von Künstlicher Intelligenz – Wann und wie dürfen personenbezogene Daten für das Training und die Anwendung von Künstlicher Intelligenz verarbeitet werden?**, 2. Version vom 17.10.2024, https://www.baden-wuerttemberg.datenschutz.de/rechtsgrundlagen-datenschutz-ki/
  Diskussionspapier zur Auslegung des Datenschutzrechts beim Einsatz von KI. Enthält eine übersichtliche Checkliste zur Einhaltung des Datenschutzes beim Einsatz von KI.

- Bitkom(Hrsg.): **Praxisleitfaden Künstliche Intelligenz**, 2024, https://www.bitkom.org/sites/main/files/2024-07/202407bitkom-leitfaden-ki-datenschutz.pdf
  Umfangreicher spezieller Leitfaden zum Thema Datenschutz mit Checkliste.

Urheber- und Persönlichkeitsrechte

- Der Paritätische Gesamtverband (Hrsg.), Till Kreutzer (Autor): **Urheberrecht und Persönlichkeitsrechte in der Paritätischen Arbeit**, 2020, https://www.der-paritaetische.de/fileadmin/user_upload/Publikationen/doc/Handreichung_UrheberrechtundPersoenlichkeitsrechte_final.pdf
  Spezieller Leitfaden zu Urheberrecht und Persönlichkeitsrechte in der Paritätischen Arbeit (ohne konkreten Bezug zu KI-Themen).

Urheberrecht

- Heise.de (Hrsg.), Till Jaeger (Autor): **Künstliche Intelligenz: Kampf um das Urheberrecht**, 2023, https://www.heise.de/hintergrund/Kuenstliche-Intelligenz-Kampf-um-das-Urheberrecht-7518607.html
  Ausführlicher Text zu vielen Aspekten im Kontext Urheberrecht und KI.