bitkom

Künstliche Intelligenz & Datenschutz

Praxisleitfaden

Inhalt

Geleitwort

Der vorliegende Leitfaden wurde federführend von den Mitgliedern des Arbeitskreises Datenschutz des Bitkom erstellt. Besonderer Dank gilt den folgenden Autorinnen und Autoren, die sich mit viel Mühe und Hingabe der Erstellung des Leitfadens gewidmet haben:

- Stefanie Bauer, ePrivacy GmbH
- Arnd Böken, GvW Graf von Westphalen Rechtsanwälte Steuerberater Partnerschaft mbB
- Dr. Nadja Christe, Bayer AG
- Jonas von Dall'Armi, Giesecke+Devrient GmbH
- Nils Freymuth, MSD Sharp & Dohme GmbH
- Dr. Alexander Fritz, OmegaLambdaTec GmbH
- Markus Frowein, RWE AG
- Ralf Herter, BASF SE
- Dr. Inka Knappertsbusch, LL.M., CMS Hasche Sigle Partnerschaft von Rechtsanwälten und Steuerberatern mbB
- Tobias Messerschmidt, DHL Group
- Petra Möritz, DATEV eG
- Paul Pink, DHL Group
- Dirk Refflinghaus, Finanz Informatik GmbH & Co. KG
- Janine Richter, BREDEX GmbH
- Lys Riemenschneider, Holisticon AG
- René Schneider, IPAI Aleph Alpha Research GmbH
- Jens Schreiber, medatixx GmbH & Co. KG
- Dr. iur. Dr. rer. pol. Hans Steege, CARIAD SE
- Florian Thoma, Accenture
- Jörn Wittmann, Volkswagen AG
- Susanna Wolf, DATEV eG

Ziel dieses Projektes ist es, das Spannungsverhältnis zwischen Künstlicher Intelligenz (KI) und Datenschutz sowohl in einer verständlichen als auch in einer fachlich anspruchsvollen Art und Weise darzustellen und konkrete Handlungsempfehlungen an die Hand zu reichen. Es ist unser Bestreben, dieses Dokument kontinuierlich zu überarbeiten und zu aktualisieren, um den neusten Entwicklungen auf dem Gebiet der KI Rechnung zu tragen. Wir laden daher alle Interessierten herzlich ein, sich aktiv an der Weiterentwicklung des Leitfadens zu beteiligen.

1 Einführung

Der vorliegende Praxisleitfaden »KI & Datenschutz« dient als umfassendes Nachschlagewerk, um Unternehmen und Organisationen bei der datenschutzkonformen Nutzung und Implementierung von KI-Technologien zu unterstützen. Ziel ist es, praxisnahe Anleitungen und rechtliche Grundlagen zu bieten, um sicherzustellen, dass die Verarbeitung personenbezogener Daten unter Einhaltung der DS-GVO (Datenschutz-Grundverordnung) und anderer relevanter Vorschriften erfolgt.

Der Praxisleitfaden zielt darauf ab, Klarheit und Sicherheit im Umgang mit bestehenden datenschutzrechtlichen Herausforderungen zu schaffen und richtet sich an Datenschutzbeauftragte, IT- und Compliance-Verantwortliche, Entwickler und Anwender von KI-Systemen sowie Entscheidungsträger in Unternehmen.
Mit Definitionen, konkreten Handlungsanweisungen und praktischen Checklisten bietet dieser Leitfaden einen direkten Mehrwert für Ihre tägliche Arbeit.

Während sich unser Leitfaden »Generative KI im Unternehmen« mit den allgemeinen Aspekten und Rahmenbedingungen der Nutzung von generativer KI beschäftigt, konzentriert sich dieser Praxisleitfaden speziell auf die datenschutzrechtlichen Anforderungen und ethischen Überlegungen bei der Nutzung von KI-Technologien.
Er bietet detaillierte, praxisnahe Anleitungen und Beispiele, die speziell auf die datenschutzkonforme Implementierung von KI-Anwendungen abzielen.

Eingangs beleuchtet der Leitfaden grundsätzliche Aspekte, einschließlich einer kurzen Einführung darüber, wann überhaupt von KI gesprochen werden kann, welche ethischen Aspekte bei der Nutzung von KI eine Rolle spielen und welche Gesetze bei der Nutzung von KI relevant werden können (1.). Im nächsten Teil werden die wesentlichen Vorschriften der DS-GVO und ihre jeweilige Bedeutung im Hinblick auf die Nutzung von KI erläutert. Daneben sollen die Anwenderinnen und Anwender von KI für die Erstellung von Berichtigungs- und Löschungskonzepten sowie von internen Richtlinien beim Einsatz von KI sensibilisiert werden (2.). Danach wird anhand der Nutzung von KI in der Automobilindustrie ein praktisches Anwendungsbeispiel für ein zielführendes Einsetzen von KI dargestellt (3.). Letztlich ist dem Leitfaden eine Checkliste angefügt, die eine Übersicht der wesentlichen Vorkehrungen zusammenfasst, die für eine datenschutzkonforme Nutzung von KI getroffen werden sollen (4.).

Wann sprechen wir überhaupt von Künstlicher Intelligenz?

Künstliche Intelligenz entstammt ursprünglich einem Forschungsgebiet der Informatik. Künstliche Intelligenz lässt sich als Simulation menschenähnlicher kognitiver Prozesse durch Computerprogramme bezeichnen. Anders als herkömmliche algorithmenbasierte IT-Systeme, die für sehr spezifische Aufgaben programmiert werden und nur auf festgelegten Regeln basieren, kann KI aus Daten lernen und ihre Leistung im Laufe der Zeit verbessern. Zum Beispiel kann ein herkömmliches algorithmenbasiertes System für die Lagerverwaltung zwar Bestandslisten aktualisieren, aber es kann nicht vorhersagen, welche Produkte in der nächsten Saison gefragt sein werden. Ein KI-System hingegen könnte durch die Analyse von Verkaufsdaten, Wetterbedingungen

und anderen Faktoren ziemlich genaue Vorhersagen treffen. Die 2024 verabschiedete KI-Verordnung der EU definiert ein KI-System als maschinengestütztes System, das für einen in wechselndem Maße autonomen Betrieb ausgelegt ist, das nach seiner Einführung anpassungsfähig sein kann und das aus den erhaltenen Eingaben für explizite oder implizite Ziele ableitet, wie Ergebnisse, etwa Vorhersagen, Inhalte, Empfehlungen oder Entscheidungen, hervorgebracht werden, die physischen oder virtuellen Umgebungen beeinflussen können.

Praxisrelevante Anwendungsfälle von KI-Modellen und/oder Systemen sind u. a.

- Automatisierte Entscheidungsfindung
- Biometrische Erkennung
- Generative KI (z. B. Bild- und Spracherkennung)
- Assistenzsysteme (z. B. Chatbots)
- Predicative Maintenance
- Überwachungsfunktionen von Maschinen (Videoüberwachung)
- Empfehlungssysteme basierend auf Nutzungsverhalten
- IT/Cybersecurity
- Autonome Fahrzeuge oder andere Produkte (Maschinen, Roboter etc.) mit integrierter KI oder Steuerung durch KI (»embedded AI«)

Der AI Act umfasst daneben auch KI-Modelle und Systeme mit allgemeinem Verwendungszweck (sog. Generative AI (GenAI)).

Abgrenzung zu »normaler« Software:

Von herkömmlicher Software unterscheiden sich KI-Systeme vor allem durch ihre Fähigkeit zur selbstständigen Problemlösung, Lern- und Analysefähigkeiten, Anpassung an neue Situationen und die Fähigkeit zur Lösung von komplexeren, von den Entwicklern nicht im Detail vorgegebenen Aufgaben.

Klassische Software ist darauf beschränkt, vordefinierte, statische Anweisungen in Form von Algorithmen auszuführen, die von einer Maschine gelesen und verstanden werden können, ohne die Fähigkeit zur Anpassung an neue Informationen/Aufgaben. Klassische Software ist damit regelbasiert. Dagegen verwenden KI-Systeme agilere Anweisungen und eine enorme Masse von Trainingsdaten, die in Kombination dazu führen, dass sich der Algorithmus durch seine Lern- und Analysefähigkeiten entsprechend selbst adaptiert, um die Aufgabe bestmöglich zu lösen. Dies wird auch in der Definition von KI in der KI-Verordnung deutlich.

Die KI-Verordnung sieht zwei Kategorien von Hochrisiko-KI-Systemen vor:

1. **KI-Systeme nach Anhang II der KI-Verordnung.**
   Dies sind KI-Systeme, die selbst Produkte oder Sicherheitskomponenten von Produkten sind, die gemäß der in Anhang II aufgeführten EU-Richtlinien eine Konformitätsbewertung für Gesundheit und Sicherheit durch Dritte durchlaufen müssen; wie z. B. zwei- oder dreirädrige Fahrzeuge, Autos, medizinische Geräte,

Flugzeuge, Eisenbahnsysteme, Aufzüge oder Spielzeug.

Als KI-Modelle fallen darunter z. B. virtuelle Assistenten, personalisierte Empfehlungen oder Steuerungssysteme/Überwachungssysteme als Teile solcher Produkte.

2. **KI-Systeme nach Anhang III der KI-Verordnung.**
   Dies sind KI-Systeme, die ein bedeutendes Risiko (significant risk) für Gesundheit, Sicherheit oder Grundrechte natürlicher Personen oder Umweltbeeinträchtigungen darstellen; z. B. KI-Systeme, die in folgenden Bereichen oder Zwecken eingesetzt werden.

   - Biometrische Identifizierung und Kategorisierung von natürlichen Personen
   - Verwaltung und Betrieb von kritischen Infrastrukturen
   - Bildung und Berufsausbildung
   - Beschäftigung, Verwaltung der Arbeitnehmer und Zugang zur Selbstständigkeit
   - Zugang zu und Inanspruchnahme von wesentlichen privaten und öffentlichen Diensten und Leistungen
   - Rechtsdurchsetzung
   - Verwaltung von Migration, Asyl und Grenzkontrollen
   - Rechtspflege und demokratische Prozesse (Wahlen)

Vorrangiges Ziel in der Entwicklung, Nutzung und Betrieb der Anbieter, Betreiber und Quasibetreiber muss sein, mögliche KI-Risiken prädikativ zu minimieren, um die Vorteile von KI für die Entwicklung und Nutzung von Produkten sowie Dienstleistungen nutzen zu können. Hierfür soll die sogenannte KI-Verordnung der EU als Gesetz über künstliche Intelligenz zusätzlich, neben der generellen Produktsicherheit von IT-Risikosystemen, gewährleisten, dass die auf dem EU-Markt in Verkehr gebrachten und in der Union verwendeten KI-Systeme sicher sind, wobei die bestehenden Grundrechte und die Werte der Union gewahrt bleiben sollen.

In einer aktuellen Bitkom-Umfrage zu KI in Cybersecurity von November 2023 sehen 57 Prozent der Unternehmen in Deutschland Gefahren durch KI, 35 Prozent erwarten eine Verbesserung der Cybersicherheit, aber nur jedes siebte Unternehmen hat sich mit dem KI-Einsatz für Cybersicherheit bereits beschäftigt. Die Ergebnisse basieren auf einer Befragung von 1.002 Unternehmen ab 10 Beschäftigten im Auftrag des Digitalverbands Bitkom. Hier ist klarer Handlungsbedarf und das vorliegende Dokument soll den Einsatz von KI im Zusammenhang mit unterschiedlichen datenschutzrelevanten Themen beleuchten.

Ethischer Rahmen: Vertrauenswürdige KI-Gestaltung strategisch verankern und umsetzen

Der mit dem zunehmenden Einsatz von KI verbundene Fortschritt bringt auch Herausforderungen mit sich, insbesondere im Hinblick auf einen vertrauenswürdigen KI-Einsatz. Inwieweit hierbei ethische Kriterien mit Blick auf die bestehende Regulatorik zum Tragen kommen und welche Rolle die Verantwortung von Unternehmen in diesem Zusammenhang spielt, behandelt der folgende Abschnitt.

Bei der ethischen Bewertung von KI-Produkten ist eine wertorientierte Technologiegestaltung entscheidend. Die Werte, die hierbei berücksichtigt werden sollten, hängen vom Einsatzkontext ab. Grundsätzliche Werte können in Codes of Conduct oder vergleichbaren Kodizes von Unternehmen und Institutionen festgehalten sein.

Ein mögliches Vorgehen zur Umsetzung des Wertes der Partnerschaftlichkeit ist die Pilotierung des KI-Produkts mit einer repräsentativen Gruppe von Stakeholdern. Hierbei sollte neben Usability-Fragen auch der Dialog mit den Stakeholdern zu ihrer Perspektive auf den Technologieeinsatz und den Chancen und Herausforderungen im Vordergrund stehen.

Es existieren verschiedene Ansätze für die ethische Bewertung von KI, wie beispielsweise die Ethik-Leitlinien der EU-Kommission oder die Stellungnahme des Deutschen Ethikrats.

Ein wichtiger Wert bei der Entwicklung und dem Betrieb von KI ist Transparenz. Hierbei sollten die Funktionen und Verarbeitungsmethoden der KI-Systeme mindestens für die relevanten Zielgruppen angemessen offen und verständlich sein, um das sogenannte Blackbox-Phänomen zu vermeiden.

Bei der Fairness von KI-Anwendungen sollte bereits bei der Entwicklung auf nicht intendierten Bias geachtet werden, um Verzerrungen in den Ergebnissen und Diskriminierung zu vermeiden. Stattdessen sollten KI-Anwendungen Vielfalt und Chancengleichheit fördern.

Stakeholder stufen KI-Anwendungen als vertrauenswürdig ein, wenn Unternehmen ethische Leitlinien und Selbstverpflichtungen haben. Um digitale Verantwortung nachhaltig zu integrieren und zu operationalisieren, können Unternehmen interne Strukturen im Bereich Corporate Digital Responsibility (CDR) aufbauen.

Insgesamt sollten ethische Kriterien in der KI-Entwicklung und im KI-Betrieb zu einem ganzheitlichen Risikomanagement beitragen, zentrale demokratische Werte wahren und deren Operationalisierung nachvollziehbar gestalten. Die ganzeinheitliche Übernahme von digitaler Verantwortung ist für den nachhaltigen Erfolg von KI-Anwendungen entscheidend und kann das Vertrauen von Stakeholdern und Mitarbeitenden sowie den Recruiting-Erfolg positiv beeinflussen.

Übergeordnete Überlegungen zum Rechtsrahmen (DS-GVO, KI-Verordnung, BDSG & Beschäftigtendatenschutz)

Die regulatorischen Ziele der Europäischen Union (EU) im Bereich Datenwirtschaft und künstliche Intelligenz sind von einer klaren Vision geprägt, die den Schutz der Bürgerrechte, die Förderung von Innovation und die Schaffung eines fairen und transparenten digitalen Binnenmarktes umfasst. Die EU versucht durch viele unterschiedliche, aber zusammenwirkende neue Vorschriften und ethische Leitlinien ein Gleichgewicht zwischen Innovation, individuellen Rechten und kommerziellen Interessen zu schaffen. Das Ziel ist auch, ungenutzte Daten, Informationen und das Potenzial der künstlichen Intelligenz kommerziell besser nutzbar zu machen, um die EU in der globalen digitalen Landschaft zu einem Vorreiter machen.

Ein zentraler Bestandteil bleibt die DS-GVO, die die Verarbeitung personenbezogener Daten reguliert und Standards für den Schutz der Privatsphäre setzt.

Der am 11.01.2024 in Kraft getretene Data Act (DA, Verordnung über harmonisierte Vorschriften für einen fairen Datenzugang und eine faire Datennutzung) zielt darauf ab, einen größeren und fairen Datenzugang für B2B, B2G, G2B und G2G zu ermöglichen und die Datennutzung in allen Wirtschaftszweigen zu fördern.

Er schafft neue Möglichkeiten für die kommerzielle Nutzung von nicht personenbezogenen Daten aus vernetzten Produkten und verbundenen Diensten und verringert und beseitigt Hindernisse für die Entwicklung der europäischen Datenwirtschaft. Die vielfältigen technischen, operativen, rechtlichen und kommerziellen Vorgaben des Datengesetzes sollen eine gerechte Verteilung der Wertschöpfung aus Daten der Datenwirtschaft gewährleisten.

In bestimmten Fällen kann die Nutzung von KI auch zur direkten Anwendung des DA führen, z. B. bei

- Nutzung von generierten Daten durch IoT-Geräte (Internet of Things-Geräte) zum Training für KI-Systeme und Modelle.
- Nutzung von KI-Systemen und Modellen zur Entwicklung und Betrieb von IoT-Geräten (z. B. KI-basierter Datenanalyse, Automatisierung und intelligente Entscheidungsfindung)
- Nutzung von KI-Systemen z. B. zur automatisierten Entscheidungsfindung beim Zugang, Abschluss, Unterbrechung oder Beendigung von »intelligenten Verträgen« (»Smart Contracts«) nach Art. 2 Nr. 39, Art. 36 DA (Verträge bei denen eine automatisierte Leistung/Ausführung unter Verwendung eines Computerprogramms erfolgt) die den Zweck haben, »Daten zur Verfügung zu stellen«. Smart Contracts mit anderen Zwecken fallen nicht unter den Anwendungsbereich des DA.
- Festlegung grundlegender Anforderungen der Interoperabilität, mit einem festgelegten Mechanismus, der öffentliche Stellen erlaubt in außergewöhnlichen Situationen (z. B. Notsituationen der Öffentlichkeit), Daten von Unternehmen anzufordern.

Mit der im März 2024 verabschiedeten KI-Verordnung erließ die EU das weltweit erste umfassende rechtliche Rahmenwerk für künstliche Intelligenz. Die KI-Verordnung fokussiert sich auf identifizierbare Risiken von KI-Systemen und fördert so verantwortungsbewusste Innovation in Europa. Durch die Gewährleistung der Sicherheit und der Grundrechte von Menschen und Unternehmen soll die Entwicklung, Bereitstellung und Nutzung von vertrauenswürdiger KI unterstützt werden. Die KI-Verordnung soll einen wesentlichen Beitrag zur Entwicklung globaler Regeln und Prinzipien für eine auf den Menschen ausgerichtete künstliche Intelligenz leisten.

Der kürzlich vorgestellte Entwurf einer neuen Produkthaftungsrichtlinie wird die Vorschriften über außervertragliche zivilrechtliche Haftung an künstliche Intelligenz anpassen. Die Richtlinie kategorisiert KI-Systeme nach ihrem Risikograd und sieht vor, dass Hersteller von KI-Systemen Schadensersatz leisten müssen, wenn sie Schwachstellen im Bereich der Cybersicherheit nicht schließen.

Bei der Entwicklung, Implementierung, Nutzung und Vermarktung von KI-Systemen müssen künftig eine Reihe von Gesetzen und Regularien eingehalten werden, um die Sicherheit, Transparenz, Nichtdiskriminierung und Rechtskonformität zu gewährleisten; dies sind insbesondere die DS-GVO, Rechte zum Schutz geistigen Eigentums (UrhG), KI-Verordnung, Data Act, die neue Produkthaftungsrichtlinie und die General Product Safety Regulation (GPSR).

Insofern normiert die KI-Verordnung auch direkt eine Pflicht zur datenschutzkonformen Gestaltung bzw. Einhaltung diverser datenschutzrechtlicher Normen. Beispielsweise müssen GPAI-Anbieter (GPAI: Global Partnership on Artifical Intelligence) eine Erklärung über die zum Teil personenbezogenen Daten erstellen und veröffentlichen, die zum Trainieren des KI-Modells oder Systems für allgemeine Zwecke verwendet werden.

Zudem enthält die KI-Verordnung auch eine direkte Pflicht zur Einhaltung der Urheberrechtsvorschriften und eine Pflicht sicherzustellen, dass Werke mit einem Nutzungsvorbehalt (Opt-Out) nach § 44b UrhG für Text- und Data-Mining (einschließlich Web-Scraping) nicht verwendet werden. In Bezug auf den Data Act stehen beispielhaft in Zukunft die Pflicht zur Zugänglichmachung von Produktdaten und verbundenen Dienstdaten aus Art. 3, 4, 8 DA und die Informationspflichten nach Art. 13 (1) f) DS-GVO – mit ihren jeweiligen Empfängerdefinitionen – selbstständig nebeneinander. Während der DA gerade zur Datenzugänglichmachung und Weitergabe von Daten für andere Nutzungszwecke verpflichtet, schränkt die DS-GVO solche Möglichkeiten gerade ein.

Aufgrund der klaren, gesetzlichen Vorrangregel der DS-GVO in Art. 1 (5) S.3 wird man in zukünftigen Zweifels- bzw. Konfliktfällen zwischen den DA-Vorgaben und der DS-GVO eine datenschutzkonforme Lösung wählen müssen.

2 DS-GVO-Anforderungen

Einführung: Personenbezug und Anonymität

Der Anwendungsbereich der DS-GVO erstreckt sich auf personenbezogene Daten. Nicht umfasst sind anonyme Daten, d.h. Angaben, die sich nicht auf eine identifizierte oder identifizierbare natürliche Person beziehen.

Gemäß Art. 1 DS-GVO ist eine Person identifizierbar, die direkt oder indirekt, z. B. mittels Zuordnung zu einem Namen, einer Kennnummer, Standortdaten oder zu einem oder mehreren besonderen Merkmalen, identifiziert werden kann. Gemäß Erwägungsgrund (ErwGr) 26 DS-GVO sind hierbei alle Mittel zu berücksichtigen, die von der oder dem Verantwortlichen oder einer anderen Person nach allgemeinem Ermessen wahrscheinlich genutzt werden, um die natürliche Person zu identifizieren. Bei der Bewertung, ob Mittel zur Identifizierung wahrscheinlich genutzt werden, sind alle objektiven Faktoren heranzuziehen, insbesondere der für die Identifizierung erforderliche Aufwand an Zeit, Kosten und Fähigkeiten sowie die zum Zeitpunkt der Verarbeitung verfügbare Technologie.

Die Identifizierung kann sich aus einer einzelnen Information selbst ergeben, zum Beispiel dem Namen oder einem anderen eindeutigen Kennzeichen. Identifizierbarkeit liegt darüber hinaus vor, wenn die Information (z. B. eine technische Kennung, ein Zitat oder eine Angabe zum Gesundheitszustand) in Verbindung mit anderen zu dieser Person verfügbaren Angaben der Person zugeordnet werden kann. Beispiel: Ein inhaltlich »unpersönlicher« Chat ist bei Kenntnis der IP-Adressen der Teilnehmer oder anderer technischer Informationen ggf. bestimmten Personen zuordenbar.

Ferner kommt es für die Definition des Personenbezugs nach der DS-GVO nicht darauf an, ob die Identifizierung durch einen bestimmten Verantwortlichen möglich ist. Es reicht aus, dass die Angabe für Dritte – die über entsprechendes Zusatzwissen verfügen – personenbeziehbar ist. Entscheidend ist dabei, ob Re-Identifizierungsmittel vom Verantwortlichen oder auch von Dritten nach allgemeinem Ermessen wahrscheinlich genutzt werden. Hieraus folgt, dass eine Re-Identifizierung, die – nach aktuellem Stand der Technik – praktisch nicht durchführbar oder auch unverhältnismäßig aufwendig und teuer wäre, außer Betracht bleiben kann. Nach der Rechtsprechung des Europäischen Gerichtshofs (EuGH) bleiben illegale Mittel ebenfalls außer Betracht.

Umgekehrt formuliert: Für eine Anonymisierung ist es ausreichend, wenn der Personenbezug derart aufgehoben wird, dass eine Re-Identifizierung entweder nur mit illegalen Mitteln oder praktisch nicht durchführbar ist, weil der Personenbezug nur mit einem unverhältnismäßig großen Aufwand an Zeit, Kosten und Arbeitskraft wiederhergestellt werden könnte. Diese Bewertung muss regelmäßig überprüft werden, da zum Beispiel neue technologische Entwicklungen eine Re-Identifikation zu einem späteren Zeitpunkt mit geringerem Aufwand ermöglichen könnten.

Unklarheit besteht allerdings noch in der Frage, ob es bei der Bewertung zulässig ist, bei entsprechend strikten Prozessgestaltungen in Bezug auf spezielle Verarbeitungen von Daten, spezifisch auf die Perspektive der jeweils verarbeitenden Stelle abzustellen. So ist zum Beispiel die Konstellation denkbar, dass ein Verantwortlicher, der Angaben

pseudonymisiert hat, diese Daten an einen Dritten weitergibt, jedoch ohne den Dritten die zur »Auflösung« der Pseudonyme erforderlichen Informationen mitzuliefern. Dürfen die Daten dann (nur) im Zusammenhang mit einer Nutzung oder Weiterverarbeitung durch diesen Dritten als »anonym« angesehen werden, wenn der Dritte auch ansonsten keine rechtmäßigen und vernünftigerweise nutzbaren Mittel hat, um die Pseudonymisierung aufzuheben? Mit dem Wortlaut sowie dem Sinn und Zweck der DS-GVO scheint diese für die Praxis potenziell sehr hilfreiche Möglichkeit vereinbar zu sein. Zuletzt fand diese Ansicht auch Wiedergabe in Rechtsprechung auf europäischer Ebene, allerdings bisher nur erstinstanzlich. Eine klare Positionierung des EuGH hierzu steht noch aus.

Für KI besteht im Kontext der Anonymisierung eine große Herausforderung darin, dass KI über ein großes »Hintergrundwissen« verfügt, aus dem sie Rückschlüsse ziehen könnte, die zu einer Identifizierbarkeit führen. Ein weiteres Problem liegt im sogenannten »Prompt Engineering«, über das die KI, insbesondere in Anwendungsfällen mit Eingabe in Form von Freitext, durch Anwenderinnen und Anwender gezielt dazu angeleitet werden könnte, eine eigentlich nicht vorgesehene De-Anonymisierung vorzunehmen oder Nutzerinnen und Nutzer die Möglichkeit haben, zusätzliches zu einer Identifizierbarkeit führendes Hintergrundwissen einzugeben.

Artikel 5 DS-GVO: Einhaltung der Datenschutzgrundsätze

Rechtmäßigkeit, Art. 5 (1) a) Alt. 1 DS-GVO

Die Rechtmäßigkeit der Verarbeitungen im Rahmen der Entwicklung, des Trainings und der Verwendung eines KI-Systems und Modells umfasst sowohl das »Ob« (Rechtsgrundlage) als auch das »Wie« der Verarbeitung (Einhaltung sonstiger Vorgaben aus den datenschutzrechtlichen Vorgaben). Auf der Grundlage von Art. 8 (II) der Charta der Grundrechte der Europäischen Union setzen u. a. Art. 5 (1) a) und Art. 6 DS-GVO diesen Grundsatz um. Zusätzlich sind auch etwaige nationale Vorgaben bei der Beurteilung der Rechtmäßigkeit heranzuziehen, soweit sie sich im Rechtsetzungsspielraum der DS-GVO halten. Grundsätzlich gilt – sowohl aus rechtlichen als auch aus praktischen Gesichtspunkten – entsprechend Art. 25 DS-GVO – alle datenschutzrechtlichen Anforderungen an KI-Modelle und Systeme direkt vom Beginn der Entwicklung und im Rahmen des Designs mitzuberücksichtigen. Dies erleichtert auch die spätere Umsetzung der Anforderungen.

Im Hinblick auf die verschiedenen Verarbeitungsphasen eines KI-Systems oder Modells (Erhebung, Training, Bereitstellung, Nutzung, Weiterentwicklung) können unterschiedliche Rechtsgrundlagen einschlägig sein. Grundsätzlich kommen hierfür alle Rechtsgrundlagen der DS-GVO, insb. Art. 6 und 9 DS-GVO, aber auch nationale Gesetze in Betracht, da die DS-GVO für KI-Modelle oder Systeme keine speziellen Rechtsgrundlagen vorsieht.

Schon im Rahmen der Erhebung/Veredlung von Trainingsdaten und der entsprechenden Nutzung dieser Daten für Zwecke des Trainings von KI-Systemen und Modellen wird regelmäßig die Frage nach der Zweckbindung bzw. Zweckänderung zu

stellen sein. Oftmals wird der Wunsch bestehen, vorhandene Datenbestände zum Training nutzbar zu machen. Für Zwecke der Innovationsförderung wird Art. 54 der KI-Verordnung in Zukunft unter engen Voraussetzungen eine Weiterverarbeitung für ursprünglich zu anderen Zwecken erhobene Daten innerhalb von Sandboxes zulassen (Art. 6 (4) DS-GVO i.V.m. Art. 54 Europäische Verordnung über künstliche Intelligenz (KI-Verordnung)).

Regelmäßig wird man Art. 6 (1) b) DS-GVO (Erfüllung eines Vertrages oder Durchführung vertraglicher Verpflichtungen) als Rechtsgrundlage heranziehen können, wenn nach objektiver Betrachtung die Verarbeitung des KI-Modells oder Systems wesentlicher Bestandteil der Hauptleistungspflichten eines Vertrags mit einem Datensubjekt bzw. einer betroffenen Person sind, so z. B. bei Nutzung von generativen KI-Modellen und Systemen zur Erzeugung von Texten oder Bildern. Schwieriger ist die Beurteilung, wenn in Zukunft KI-Modelle oder Systeme unterstützend bei der Erbringung vertraglicher Leistungen in unterschiedlichem Intensitäts- oder Wirkungsgrad eingesetzt werden, z. B. beim Einsatz von zum Teil personalisierten Chatbots im Rahmen des Kundenservices für eine Dienstleistung/ein Produkt. Hier ist eine Beurteilung im Einzelfall erforderlich, inwieweit die Verarbeitung noch objektiv erforderlich ist.

Soweit die Einwilligung als Rechtsgrundlage herangezogen werden soll, wird sich ein Schwerpunkt in der Bewertung der ausreichenden Verständlichkeit (»in informierter Weise«, s.a. Art. 4 (11) DS-GVO) und entsprechender Formulierung ergeben. Im Zusammenspiel mit den Informationspflichten (Art. 12 ff. DS-GVO) sollte auf eine konsistente Darstellung geachtet werden. Art. 12 (1) DS-GVO fordert zusätzlich eine »präzise« Information, welches der Transparenz in gewisser Weise entgegenstehen kann, soweit sehr komplexe technische Verarbeitungsvorgänge betroffen sind.

Problematisch sind auch Fälle des Widerrufs der Einwilligung, soweit im Modell noch personenbezogene Daten verarbeitet werden bzw. aus diesem extrahiert werden können. Die Sicherstellung dieser Anforderung muss unter dem Gesichtspunkt des Privacy-by-Design von Anfang an bedacht werden.

In diesem Zusammenhang sei letztlich auf die sorgfältige Prüfung der Freiwilligkeit (ErwGr 42 S. 5 DS-GVO) der Abgabe von Einwilligungserklärungen im Beschäftigtenverhältnis hingewiesen.

Treu und Glauben, Art. 5 (1) a) Alt. 2 DS-GVO

Der Grundsatz, wonach die Verarbeitung »nach Treu und Glauben« zu erfolgen hat, findet eher weniger praktische Relevanz. Bekannt ist er als unbestimmter Rechtsbegriff des Generaltatbestands nach § 242 BGB. Die deutsche Gesetzgebung definiert den Begriff als redlich, aufrichtiges Sozialverhalten. Als Norm für den Privatrechtsverkehr ist diese Definition jedoch nicht einfach auf die DS-GVO zu übertragen. Es liegt daher näher, auf die englische Fassung mit der Bezeichnung »Fairness« abzustellen, auch wenn diese ähnlich schwammig ist.

Unter unfairen Datenverarbeitungen werden beispielsweise verborgene, unerwartete oder unverhältnismäßige Verarbeitungen subsumiert. Dabei liegen Überschneidungen mit den anderen Grundsätzen auf der Hand und erklären abermals die untergeordnete Rolle dieses Grundsatzes.

In Bezug auf KI-Systeme und Modelle, bei denen wir vorwiegend von Lernsystemen sprechen, wären zu berücksichtigende Aspekte im Rahmen der unverhältnismäßigen Verarbeitungen, die Nutzung von Big Data als Grundlage für das Machine Learning und eben die Frage nach der Verhältnismäßigkeit und Erforderlichkeit dieses enormen Datenumfangs.

Hinsichtlich des gewählten Modells, der Lizenz, der Einbettung und der Konfiguration stellt sich die Frage, mit welchen Daten/aus welchen Datenquellen das Modell gelernt hat. Weiterhin stellt sich die Frage, ob auch anhand eigener Dateninputs gelernt wird, inwieweit das Recht auf Vergessenwerden beeinträchtigt wird und ob die eventuelle Verwendung und Speicherung unerwartet und unfair sein kann.

Selbst im Falle einer eingeholten Einwilligung zur Datenverarbeitung wäre die Umsetzung des »Rechts auf Widerruf« nur schwer durchsetzbar und damit nicht im Rahmen der geforderten »Fairness«. Noch »unfairer« wäre wohl, wenn die Nutzung einer KI für den Betroffenen erst gar nicht ersichtlich wäre (»verborgen«).

Darüber hinaus ist ein weiterer Aspekt zu berücksichtigen: Sofern die Trainingsquellen und herangezogenen Daten unbekannt sind, bleibt auch verborgen, ob der Ergebnis-Output aufgrund einseitiger, stereotypischer Daten erfolgt, und damit nicht repräsentativ wäre. Es entstünde eine »algorithmische Diskriminierung«.

Vor dem Einsatz eines KI-Systems oder Modells sollte daher innerhalb einer Risikoprüfung eruiert werden, ob eine Diskriminierung vorliegen könnte und somit Rechte und Freiheiten von betroffenen Personen gefährdet wären. Im Unternehmenskontext wäre ein zu berücksichtigender Punkt die Chancengleichheit am Arbeitsplatz.

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

Sobald ein Unternehmen personenbezogene Daten verarbeitet, ist es verpflichtet, dies transparent zu tun, indem es den Betroffenen verständlich, in einer einfachen Sprache informiert. Diese Pflicht bezieht sich sowohl auf eine Erhebung der Daten direkt bei der oder dem Betroffenen als auch bei einer Erhebung durch Dritte (Art. 13, 14 DS-GVO). Unternehmen haben daher beim Einsatz von KI die erste Hürde bereits dadurch zu überwinden, dass sie technisch anspruchsvolle KI-Lösungen in eine einfache und verständliche Sprache übersetzen müssen. Von der Datenverarbeitung betroffene Personen müssen über diese Datenverarbeitung dann hinreichend informiert werden.

Um dem Grundsatz einer hinreichenden Transparenz nachzukommen, ist Verantwortlichen zu empfehlen, ihre Datenschutzerklärungen und -richtlinien sowie Datenschutzhinweise im Arbeitsverhältnis dahingehend zu aktualisieren, dass der Einsatz sowie der Zweck verwendeter KI beschrieben wird. Auch die Logik hinter KI-gestützten automatisierten Entscheidungen sowie mögliche Risiken sollten verständlich dargelegt werden.

Detailliertere Ausführungen zur Transparenz finden sich unter dem Punkt »Transparenz und Informationspflichten«.

Neben den Anforderungen aus der Datenschutzgrundverordnung werden auch mit Inkrafttreten der KI-Verordnung weitere Transparenzpflichten zu erfüllen sein. Der Umfang wird abhängig von der Risikoklassifizierung variieren, als Mindestanforderung jedoch eine KI-Kennzeichnung sowie Transparenzerklärung beinhalten.

Zweckbindung, Art. 5 (1) b) DS-GVO

Fraglich ist weiterhin, ob bereits einmal zu einem bestimmten Zweck erhobene Daten durch eine weitere Verarbeitung in einem KI-System oder Modell und einem neu entstandenen Kontext eine nicht erlaubte Zweckänderung darstellen. Es bedarf jeweils einer Einzelfallprüfung, ob diese Weiterverarbeitung für

- Im öffentlichen Interesse liegende Archivzwecke,
- Wissenschaftliche oder
- Historische Forschungszwecke oder
- Statistische Zwecke

vorgenommen wurde, damit sie nicht als unvereinbar mit den ursprünglichen Zwecken und dem Kompatibilitätstest gilt.

Im Einklang mit dieser strengen Auslegung ist jedoch die Nutzung von Daten, die bereits keinen Personenzug mehr aufweisen, also anonymisiert wurden oder solche, die aus öffentlich zugänglichen Quellen eingesetzt werden.

Bei der Nutzung derartiger Daten würden die weiteren Tatbestandsmerkmale des Art. 5 DS-GVO, nämlich die Speicherbegrenzung sowie Integrität und Vertraulichkeit der personenbezogenen Daten nicht mehr tangiert.

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

Der Grundsatz der Datenminimierung sieht vor, dass Unternehmen personenbezogene Daten nur für bestimmte, erforderliche Zwecke verarbeiten und speichern. Es handelt sich um eine Ausprägung des Verhältnismäßigkeitsgrundsatzes, der eine Verhältnismäßigkeitsprüfung erfordert. Die Verarbeitung von personenbezogenen Daten ist daher grundsätzlich auf das notwendige Maß zu beschränken.

Auch hier besteht die Schwierigkeit in der Abwägung, da zur Durchführung des Trainings der KI i.d.R. große Datenmengen (»Big Data«) herangezogen werden. Sollten Unternehmen (bspw. zum Finetuning) ihr KI-Modell oder System trainieren, ist vorab eine Verhältnismäßigkeitsprüfung durchzuführen und die »Zweck-Mittel-Relation« zwischen dem Sammeln der Daten und Effizienz des Trainings zu bestimmen. Insbesondere das Mittel der (irreversiblen) Anonymisierung der Trainingssätze stellt eine Möglichkeit dar, datenschutzkonform vorzugehen.

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

Aufgrund der möglichen Konsequenzen von Falschinformationen für Betroffene verlangt die DS-GVO grundsätzlich, dass nur sachlich richtige personenbezogene Daten

verarbeitet werden. Unrichtige personenbezogene Daten sind unverzüglich zu löschen oder zu berichtigen.

Es kommt jedoch regelmäßig vor, dass ein Large Language Model Halluzinationen erzeugt. Es handelt sich hierbei um unrichtige Informationen, die zunächst plausibel erscheinen können. Halluzinationen kollidieren mit dem Grundsatz der Richtigkeit. Betroffenen Personen steht zudem nach Art. 16 DS-GVO ein Recht auf Berichtigung zu.

Ergebnisse der KI sind daher kritisch zu hinterfragen und zu verifizieren, auch im Falle von Plausibilität.

Rechenschaftspflicht, Art. 5 (2) DS-GVO

Die Rechenschaftspflicht stellt ein zentrales Prinzip der DS-GVO dar. Sie besagt, dass Verantwortliche nicht nur sicherstellen müssen, dass sie die Vorschriften der DS-GVO einhalten, sondern auch nachweisen können müssen, dass sie dies tun. Verantwortliche müssen angemessene technische und organisatorische Maßnahmen implementieren, um die Sicherheit und den Schutz personenbezogener Daten zu gewährleisten.

Noch immer konnten Unklarheiten bzgl. des Umfangs und der Form der Rechenschaft nicht in Gänze ausgeräumt werden, sodass einem restriktiven Verständnis, dass diesen Anforderungen nur mittels eines Datenschutzmanagementsystems begegnet werden könne, die Kritik am »One size fits all«-Ansatz und der Angemessenheit für kleinere Unternehmen entgegengesetzt wird.

Unabhängig davon, ob interne Mechanismen und Kontrollsysteme (Plan-Do-Check-Act, PDCA) eingerichtet, oder Prozesse mit Personenbezug aktenmäßig dokumentiert werden, ist zu gewährleisten, dass den Aufsichtsbehörden Nachweise vorgelegt werden können.

Zu den Maßnahmen zählen **insbesondere**

- Das Führen eines Verzeichnisses von Verarbeitungstätigkeiten
- Interne Datenschutzrichtlinien
- Die Dokumentation von Datenschutzverletzungen
- Prozessdokumentationen
- Der Abschluss von Auftragsverarbeitungsverträgen (ggf. Standardvertragsklauseln)
- Die Durchführung von Datenschutzfolgenabschätzungen
- Datenschutz durch Technikgestaltung und datenschutzfreundliche Voreinstellungen (Privacy by Design und Privacy by Default)
- Die Benennung eines Datenschutzbeauftragten

Diese Nachweise beziehen sich auch auf die Gewährleistung der datenschutzrechtlichen Grundsätze beim Einsatz von KI und deren Verschriftlichung.

Bereits das Global Privacy Assembly 2020 forderte im Falle einer Entwicklung oder Nutzung von KI u. a. zu folgenden Rechenschaftsmaßnahmen auf:

- Bewertung und Offenlegung der potenziellen Auswirkungen auf die Menschenrechte (einschließlich der Rechte auf den Schutz der Daten und der Privatsphäre) vor Nutzung der KI
- Führen von Verzeichnissen über die Folgenabschätzung, die Konzeption, die Entwicklung, die Prüfung und die Verwendung von KI
- Die Gewährleistung der Transparenz und Offenheit durch Offenlegung der Nutzung von KI, der verwendeten Daten und der Logik der KI

Um der Rechenschaftspflicht nachzukommen, ist insbesondere eine Dokumentation obiger Maßnahmen in einem angemessenen Rahmen umzusetzen.

Rechtsgrundlage bei der Nutzung von personenbezogenen Daten zu Trainingszwecken

Das Vorliegen einer Rechtsgrundlage ist ein wesentlicher Schritt für eine rechtskonforme Verarbeitung personenbezogener Daten für KI-Training. Als Rechtsgrundlage kommen grundsätzlich die Einwilligung (Art. 6 (1) a) DS-GVO, Art. 9 (2) a) DS-GVO), die Verarbeitung im Rahmen der Vertragserfüllung (Art. 6 (1) b) DS-GVO) sowie die Verarbeitung zur Wahrung berechtigter Interessen (Art. 6 (1) f) DS-GVO) für Unternehmen im nicht öffentlichen Bereich in Betracht.

Einwilligung

Die Einwilligung kann eine geeignete Rechtsgrundlage sein, wenn ein **direktes Verhältnis** zur betroffenen Person besteht. Es muss sichergestellt sein, dass die Einwilligung informiert erfolgt. Dies setzt voraus, dass die betroffene Person über die Verwendung ihrer Daten aufgeklärt wurde, aktiv und freiwillig eine Einwilligung abgegeben hat und die Einwilligung jederzeit und ohne Angabe von Gründen widerrufen werden kann. Kernvoraussetzung für die Wirksamkeit ist, dass die betroffene Person eine echte Wahlmöglichkeit hinsichtlich der Verwendung ihrer Daten hat. Kurzum muss die betroffene Person die Möglichkeit haben, ihre Einwilligung für jeden Zweck separat zu erteilen. Dies ist insbesondere dann erforderlich, wenn das Unternehmen mehrere unterschiedliche Zwecke verfolgt (z. B. Verwendung von Daten für Marketingzwecke sowie für den Aufbau eines Datensatzes zum KI-Training).

Im Falle eines »Machtungleichgewichts« zwischen der betroffenen Person und dem für die Verarbeitung Verantwortlichen (z. B. im Rahmen eines Arbeitsverhältnisses) ist die Einholung einer Einwilligung zu Trainingszwecken nahezu ausgeschlossen, da es in der Praxis schwer sein wird, zu beweisen, dass die Einwilligung ohne jeglichen Druck abgegeben wurde sowie im Arbeitsverhältnis die Mitarbeitenden einen klaren rechtlichen oder wirtschaftlichen Vorteil erlangen oder gleichgelagerte Interessen vorliegen müssten.

Eine Einwilligung scheidet in aller Regel dann aus, wenn der Verantwortliche **keine direkte Beziehung** zur betroffenen Person hat. In diesem Fall kann keine wirksame

Einwilligung nach Art. 6 (1) a) DS-GVO eingeholt werden (z. B. Datenerhebung durch Web-Scraping). Wenn eine Person ihre Einwilligung zur Verwendung ihrer personenbezogenen Daten widerruft, ist es in den meisten Fällen schwierig, die bereits verwendeten Trainingsdaten zu entfernen. Dies kann dazu führen, dass in solchen Fällen mangels einer »Extrahierungsmöglichkeit« die KI-Anwendung vollständig neu trainiert werden müsste.

Vertragserfüllung

Eine Datenverarbeitung kann nur dann auf die Erfüllung eines Vertrages gestützt werden, wenn ein konkretes vertragliches oder vorvertragliches Verhältnis zwischen der betroffenen Person sowie der oder dem Verantwortlichen besteht. Dabei ist zu beachten, dass die Verarbeitung ein objektiv erforderlicher Vertragsbestandteil sein muss. In Hinblick auf die Verarbeitung von personenbezogenen Daten zu KI-Trainingszwecken scheidet ein vertragliches Verhältnis in aller Regel aus, da die Verarbeitung objektiv nicht zur Vertragserfüllung erforderlich ist und eine entsprechende Regelung einer AGB-Kontrolle nicht standhalten würde. Anders sieht es nur dann aus, wenn das KI-Training Vertragsbestandteil ist (z. B. Erstellung eines KI-Sprachgenerators, der mit der Stimme der betroffenen Person trainiert wird)[1]. Zu beachten ist, dass keine vertragliche Vereinbarung geschlossen werden kann, die die Verarbeitung personenbezogener Daten von Dritten zu KI-Trainingszwecken legitimiert.

Berechtigte Interessen

Eine für die Praxis sehr relevante Rechtsgrundlage ist Art. 6 (1) f) DS-GVO, die Verarbeitung aufgrund berechtigter Interessen.

Danach ist die Verarbeitung personenbezogener Daten rechtmäßig, wenn sie zur Wahrung der berechtigten Interessen des Verantwortlichen oder eines Dritten erforderlich ist, sofern nicht die Interessen, Grundrechte oder Grundfreiheiten der betroffenen Person, die den Schutz personenbezogener Daten erfordern, überwiegen.

Datenschutzexpertinnen und Datenschutzexperten und einzelne Aufsichtsbehörden (wie der Landesbeauftragte für den Datenschutz und die Informationsfreiheit in Baden-Württemberg (LfDI BW)) messen Art. 6 (1) f) DS-GVO eine besondere Bedeutung als Rechtsgrundlage für die Verarbeitung personenbezogener Daten in KI-Systemen und Modellen zu. Das LfDI BW weist aber auch auf die damit derzeit noch verbundenen Rechtsunsicherheiten hin.[2]

Das berechtigte Interesse kommt als Rechtsgrundlage in Betracht, wenn die nachfolgenden drei Voraussetzungen kumulativ erfüllt sind:

1. **Nachweis des berechtigten Interesses (der »Zweck-Test«)**
   Ein ausreichendes berechtigtes Interesse liegt regelmäßig vor, sofern dieses legitim und rechtmäßig ist, wie z. B. Entwicklung/Verbesserung der Leistungsfähigkeit,

[1] Vgl. Landesbeauftragter für den Datenschutz und die Informationsfreiheit Baden-Württemberg, Rechtsgrundlagen im Datenschutz beim Einsatz von Künstlicher Intelligenz, 2024, S. 13.
[2] Vgl. Landesbeauftragter für den Datenschutz und die Informationsfreiheit Baden-Württemberg, Rechtsgrundlagen im Datenschutz beim Einsatz von Künstlicher Intelligenz, 2024, S. 16.

Genauigkeit, Nutzbarkeit, Vermarktbarkeit, Sicherheit von KI-Systemen und Modellen oder deren Nutzung für effizientere Geschäftstätigkeiten.

2. **Verarbeitung muss zur Verwirklichung dieses Interesses erforderlich sein (Prüfung der »Erforderlichkeit«)**
   Die Erforderlichkeit liegt vor, wenn die Verarbeitung der personenbezogenen Daten aus objektiver Sicht notwendig oder sachgerecht ist, das avisierte Interesse zu erreichen. Entscheidend ist dabei, ob es weniger invasive Wege gibt, um das Ziel zu erreichen, ohne personenbezogene Daten zu verarbeiten. Der Verantwortliche besitzt hier einen – ggf. zu begründenden – Ermessensspielraum und muss sich nicht auf deutlich schlechtere, teurere, aufwendigere, weniger Erfolg versprechende oder weniger effektive Alternativen verweisen lassen. Die Entwicklung und/oder Nutzung von KI-Systemen und Modellen ganz ohne Verarbeitung personenbezogener Daten wird aus vielerlei Gründen regelmäßig nicht möglich, sachgerecht oder zielführend sein. Solange der Grundsatz der Datenminimierung berücksichtigt ist, dürfte die Erforderlichkeit in der Regel jedenfalls begründbar sein.

3. **Abwägung mit den Interessen, Rechten und Freiheiten der betroffenen Personen (»Abwägungsprüfung«)**
   Entscheidend für die Einschlägigkeit von Art. 6 (1) f) DS-GVO als Rechtsgrundlage ist daher die Abwägung der Interessen des Verantwortlichen (bzw. eines Dritten) an der Datenverarbeitung mit den Interessen der betroffenen Person(en), dass eine Datenverarbeitung unterbleibt. Die bestehenden Interessen müssen benannt und gewichtet werden. Danach können die beiderseitigen Interessen gegeneinander abgewogen werden. Sofern die Interessen der betroffenen Person(en) den Interessen des Verantwortlichen (oder eines Dritten) **nicht** überwiegen, ist die Datenverarbeitung grundsätzlich zulässig.

Zu berücksichtigende Aspekte sind dabei **insbesondere**

- Die Art der Daten
- Die Menge der Daten und der betroffenen Personen
- Die Quelle der Daten
- Die Anzahl der Verarbeiter, die in die Verarbeitung involviert sind
- Die Dauer der Datenverarbeitung
- Die Sicherheit der Verarbeitung und
- Die vernünftigen Erwartungen der betroffenen Person(en), also ob diese zum Zeitpunkt der Erhebung der personenbezogenen Daten und angesichts der Umstände, unter denen diese Erhebung erfolgt, vernünftigerweise absehen konnte, dass möglicherweise eine Verarbeitung für den angestrebten Zweck erfolgen wird (ErwGr 47 S. 1 bis 3)
- Die breite öffentliche Verfügbarkeit solcher Daten im Internet, die klar darauf hinweisenden Nutzungs- und Datenschutzbedingungen vieler Systeme, Modelle oder Plattformen sowie die stetige und sehr breite öffentliche Berichterstattung über die Nutzung personenbezogener Daten für Entwicklung und Nutzung von Software, Plattformen und KI.

Unter anderem aus diesen Gründen dürfte oft gut begründbar sein, dass die Gruppe der betroffenen Personen oder die betroffene Einzelperson vernünftigerweise annehmen musste, dass ihre personenbezogenen Daten für KI genutzt werden.

Um sicherzustellen, dass die Verarbeitung verhältnismäßig ist, sollte, wenn die Datenminimierung nicht ausreicht und/oder keine Verwässerung der »individuellen Betroffenheit« (EuG 557/20) vorliegt, zusätzlich eine Anonymisierung oder Pseudonymisierung der Daten in Betracht gezogen werden.

Es ist ferner sicherzustellen, dass keine besonders schutzwürdigen Daten (Art. 9 DS-GVO) verarbeitet werden und darüber hinaus Auswahlkriterien festlegt sind, um die Erhebung personenbezogener Daten auf die für die Verarbeitung relevanten und erforderlichen Daten zu beschränken. Bei der Nutzung von KI-Systemen und Modellen für Trainingszwecke müssen ergänzend weitere Kriterien berücksichtigt und im Rahmen der Abwägung bewertet werden. Hierzu zählen insbesondere die Detailgenauigkeit, der Umfang der Trainingsdaten sowie die Auswirkungen auf die betroffene Person (»Eingriffsintensität«). Die Intensität des Eingriffs variiert je nach Art der Datenverarbeitung, d.h., dass das Trainieren von umfangreichen Sprachmodellen stärker in die Rechte betroffener Personen eingreifen kann als das Trainieren statistischer Modelle. Wichtig ist, dass im Einzelfall eine »substanzielle Auseinandersetzung« der gegenläufigen Interessen im Rahmen einer Abwägung erfolgt. Angemerkt sei, dass die Wahrung berechtigter Interessen keine ausreichende Rechtsgrundlage für die Verarbeitung personenbezogener Daten aus besonderen Kategorien nach Art. 9 DS-GVO (z. B. biometrische Daten) darstellt, es sei denn, Art. 9 (2) e) DS-GVO ist anwendbar, was bei KI-Trainingsdaten, die im Internet auf öffentlich zugänglichen Webseiten nach § 44b UrhG generiert werden, regelmäßig der Fall sein kann.

Rechtliche Verpflichtung; Wahrnehmung einer Aufgabe von öffentlichem Interesse; Betriebsvereinbarung

Die Anforderungen, welche für Art. 6 (1) b) DS-GVO »Erfüllung eines Vertrages« gelten, sind ebenso gültig für Datenverarbeitungen, welche auf die Art. 6 (1) c) bis e) DS-GVO gestützt werden. Jedoch gibt es bisher keine rechtliche Verpflichtung für nicht öffentliche Stellen zum Einsatz von KI zu Trainingszwecken.

Der Einsatz von KI kann grundsätzlich auch in einer Betriebsvereinbarung geregelt werden. Allerdings darf das Schutzniveau der DS-GVO in diesem Fall nicht unterschritten werden. Zudem muss das anwendbare Tarifrecht sowie Betriebsverfassungsrecht gewahrt werden. Aufgrund der eben genannten Hürden, der wegen des Vorlagebeschlusses des EuGH bestehenden Rechtsunsicherheiten[3] sowie dem Aspekt, dass der Einsatz von KI zu Trainingszwecken mit hoher Wahrscheinlichkeit die Interessen der betroffenen Personen nicht überwiegt, eignet sich eine Betriebsvereinbarung als Legitimationsgrundlage nicht.

Die aktuell öffentlich zugängliche Fassung der KI-Verordnung sieht die Möglichkeit der Weiterverarbeitung zuvor rechtmäßig erhobener personenbezogener Daten zur Entwicklung bestimmter KI-Systeme und Modelle im öffentlichen Interesse im KI-Real-

[3] AG Vorlagebeschluss an EuGH vom 22.09.2022 – 8 AZR 209/21.

labor grundsätzlich vor. Aufgrund des engen Anwendungsbereichs und der sehr hohen Anforderungen wird diese Regelung jedoch nur in Ausnahmefällen zur Anwendung kommen können.

Zweckänderung

Eine Zweckänderung ist nach Art. 6 (4) DS-GVO nur dann zulässig, wenn sie mit dem ursprünglichen Zweck vereinbar ist und eine rechtliche Grundlage hat. Verantwortliche müssen die Konformität ihrer eigenen Verarbeitungsoperationen vollständig analysieren. Insbesondere muss der Verantwortliche die Transparenzanforderungen erfüllen. Es wird daher nur schwer möglich sein, vorhandene, für einen ursprünglichen Zweck vorgesehen Daten unter dem Gesichtspunkt der Zweckänderung für Trainingszwecke zu verwenden.

Verwendung anonymisierter oder aggregierter Daten

Als mögliche Lösung dieser vielfältigen Probleme kommt die Verwendung anonymisierter oder aggregierter Daten in Betracht (hinsichtlich der Begriffe »relative und absolute Anonymität« siehe Kapitel: Einführung Begriffe »relative und absolute Anonymität« auf Seite 12). Da es sich bei diesen Daten nicht mehr um personenbezogene Daten handelt, ist der Anwendungsbereich der DS-GVO nicht eröffnet.

Für das Training könnten sich daher aggregierte Daten eignen, da bei der Aggregation von Daten Fallgruppen im aktiven Datensatz zu einzelnen Fällen kombiniert, die dann als separate aggregierte Datei abgespeichert werden. In diesem Fall kann der Personenbezug entfallen, wenn sich die aggregierten Daten nicht auf eine bestimmte natürliche Person, sondern auf eine Personengruppe beziehen. Dies bedarf einer Prüfung des jeweiligen Einzelfalls.

Sowohl die Anonymisierung von Daten als auch die Erstellung aggregierter Datensätze stellt nach vorherrschender Meinung eine Verarbeitung personenbezogener Daten dar, sodass eine Rechtsgrundlage hierfür notwendig ist. In Betracht kommt das berechtigte Interesse – wobei in der Feststellung dessen die Abwägung eine entscheidende Rolle spielt.

Problematisch ist dabei jedoch, dass die betroffenen Personen ein Widerspruchsrecht haben. Wenn die Daten bereits anonymisiert sind, besteht aufgrund der Anonymisierung keine Möglichkeit mehr, diesem Recht effektiv Rechnung zu tragen. Dieser scheinbare Widerspruch kann jedoch damit aufgelöst werden, da die DS-GVO nach der Anonymisierung der Daten keine Anwendung mehr findet.

Im Allgemeinen ist auch der Begriff der Anonymität nicht einheitlich und klar gesetzlich definiert. Der Begriff war bereits Streitpunkt vieler gerichtlicher Entscheidungen, sowohl auf nationaler als auch auf internationaler Ebene. Es lohnt sich daher, in die Absicherung der Anonymisierung von Daten einen hohen Aufwand zu investieren. Dabei können die Ausführungen in der Checkliste im letzten Kapitel dieses Papiers helfen.

Artikel 9 DS-GVO: Verarbeitung besonderer Kategorien personenbezogener Daten

Art. 9 (1) DS-GVO listet die folgenden besonderen Kategorien personenbezogener Daten auf: Daten zur rassischen und ethnischen Herkunft, politische Meinung, religiöse oder weltanschauliche Überzeugungen, Gewerkschaftszugehörigkeit, genetische und biometrische Daten, Gesundheitsdaten sowie Daten zum Sexualleben oder der sexuellen Orientierung einer natürlichen Person.

Da diese besonderen Kategorien erhöhte Risiken für die Grundrechte und Freiheiten der betroffenen Personen bergen, ist ihre Verarbeitung grundsätzlich untersagt (Art. 9 (1) DS-GVO sowie ErwGr 51). Es ist zu beachten, dass hier analog zur Regelung der Rechtsgrundlagen zur Verarbeitung »normaler« Datenkategorien in Artikel 6 ein eigenständiges Verarbeitungsverbot mit eigenen Erlaubnisvorbehalten gilt. Art. 9 (2) a) bis j) regelt Ausnahmen vom Verbot der Verarbeitung besonderer Datenkategorien. Zu diesen Ausnahmen gehören die ausdrückliche Einwilligung der betroffenen Person, die Notwendigkeit für die Geltendmachung, Ausübung oder Verteidigung von Rechtsansprüchen oder spezifische Anforderungen im Bereich des Arbeitsrechts und des Sozialschutzes. Die Verarbeitung besonderer Datenkategorien ist auch zulässig für Daten, die die betroffene Person selbst öffentlich gemacht hat, beispielsweise in öffentlich zugänglichen Profilen in sozialen Netzwerken.

Der Einsatz von KI-Systemen und Modellen zur Verarbeitung besonderer Kategorien personenbezogener Daten wirft verschiedene rechtliche Herausforderungen auf. Ein Kernproblem ist die Einholung einer wirksamen Einwilligung, die sowohl spezifisch als auch informiert sein muss. KI-Systeme und Modelle oder einzelne KI-Komponenten können jedoch undurchsichtig sein, was die Erfüllung dieser Anforderungen erschwert. Auch die Möglichkeit des Widerrufs nach Art. 7 (3) kann für die Nutzung von KI ein Problem darstellen.

Beispiele für die Verarbeitung besonderer Kategorien personenbezogener Daten mittels KI sind:

- die KI-gestützte Befundung in der bildgebenden Diagnostik, Applikationen zur Unterstützung von Patientinnen und Patienten (z. B. in Form von Chatbots);
- die Verarbeitung von Porträtfotos und Videos (z. B. zur Erstellung von Avataren). Porträtfotos stellen besondere Kategorien personenbezogener Daten dar, da sie Information zu ethnischer Herkunft und gegebenenfalls auch zu religiöser Zugehörigkeit oder Erkrankungen liefern.

Auch im Rahmen von Profiling können besondere Kategorien personenbezogener Daten verarbeitet werden. Hierbei ist zu beachten, dass besondere Kategorien personenbezogener Daten auch während der Verarbeitung entstehen können, z. B. wenn auf Basis von Kauf- und Surfverhalten die Wahrscheinlichkeit einer Schwangerschaft oder einer bestimmten Erkrankung für Werbe- oder Forschungszwecke kalkuliert wird.

Wie bereits dargestellt, stellen Einwilligungen eine gängige Rechtsgrundlage zur Verarbeitung besonderer Kategorien personenbezogener Daten dar. Bei den oben beschrieben Profiling-Beispielen stellt dies jedoch eine Herausforderung dar, da die

betroffenen Personen über die Nutzung von Daten aus verschiedenen Quellen informiert werden müssen. Surfverhalten wird üblicherweise mittels Cookies oder vergleichbaren Technologien ermittelt. Da hier grundsätzlich eine Einwilligung eingeholt werden muss, kann man in diesen Fällen die Einwilligung für Profiling ebenfalls einholen. Dabei ist zu beachten, dass dies, inklusive bei Kombination mit Daten aus anderen Quellen (Kategorien benennen!), sowie das mögliche Generieren besonderer Kategorien personenbezogener Daten, bereits im ersten Level des Cookie-Banners erwähnt wird. Mit der entsprechenden Cookie-Kategorie können die Verarbeitung bzw. Partner genauer erläutert werden.

Werden besondere Kategorien basierend auf den Ausnahmetatbeständen nach Art. 9 (2) b) - f) DS-GVO verarbeitet, so ist dies sehr genau zu prüfen und zu dokumentieren.

Transparenz und Informationspflichten

Wie bereits im Kapitel »Einhaltung der Datenschutzgrundsätze des Art. 5 (1) DS-GVO« erläutert, müssen bei der personenbezogenen Datenverarbeitung gewisse Transparenzgrundsätze eingehalten werden. Eine Analyse der Informationen vom LfDI Rheinland-Pfalz hat gezeigt, dass erheblicher Bedarf für Nachfragen besteht, um bewerten zu können, ob die Regelungen der DS-GVO bei der Datenverarbeitung durch ChatGPT eingehalten werden. Die Forderung nach mehr Transparenz für die Nutzerinnen und Nutzer ist dabei eine wesentliche Komponente. Bei der Entwicklung und dem Einsatz von KI sollte darauf stets besonderes Augenmerk gelegt werden. Einerseits ist Transparenz wichtig, um Akzeptanz und Vertrauen in Anwendungen zu schaffen. Des Weiteren sind gesetzliche Vorgaben zur Transparenz einzuhalten. Neben Transparenzanforderungen, die sich u. a. aus der KI-Verordnung ergeben werden, bleiben die Anforderungen aus der DS-GVO anwendbar. Zu nennen sind insbesondere:

- Das Grundprinzip der Transparenz aus Art. 5 (1) lit. a DS-GVO
- Die Anforderungen aus Kapitel III Abschnitte 1 und 2 (Art. 12-14) DS-GVO
- Sofern Einwilligungen eingeholt werden, die Sicherstellung einer informierten Einwilligung (Art. 6 (1) lit. a, Art. 7 DS-GVO)

Art. 5 (1) a) DS-GVO macht eine generelle Vorgabe, die Verarbeitung personenbezogener Daten transparent zu gestalten (wofür der Verantwortliche nach Art. 5 (2) rechenschaftspflichtig ist). Als umfassendes Prinzip ausgestaltet, lässt die Bestimmung aber offen, wie das im Einzelnen zu erreichen ist. Auch Art. 12 DS-GVO fordert zunächst nur, dass angemessene Maßnahmen zu ergreifen sind, um Informationen in einer transparenten Weise bereitzustellen, verweist aber bereits auf Art. 13 und 14 DS-GVO.

Diese beiden Artikel beinhalten detaillierte Anforderungskataloge, wobei Art. 13 DS-GVO anzuwenden ist, wenn Informationen beim Betroffenen erhoben werden; Art. 14 DS-GVO findet demgegenüber Anwendung, wenn die Informationen nicht beim Betroffenen, sondern vielmehr ohne seine Mitwirkung bei Dritten erhoben werden. Die Anforderungen lassen sich wie folgt zusammenfassen (im Einzelfall sollte der Wortlaut der Vorschriften herangezogen werden):

- Informationen zum Verantwortlichen und zum Datenschutzbeauftragten

- Zwecke der Verarbeitung und Rechtsgrundlage(n), ggf. auch eine Darlegung der verfolgten legitimen Interessen und der Verarbeitung für weitere Zwecke
- Empfänger von Daten und Drittlandberührungen
- Dauer der Verarbeitung
- Rechte und Beschwerdemöglichkeiten des Betroffenen
- Notwendigkeit einer Bereitstellung der Daten (gesetzlich oder vorvertraglich, beziehungsweise vertraglich)
- Vorliegen einer automatisierten Entscheidung, ggf. Profiling

Art. 14 DS-GVO hat naturgemäß die weitere Anforderung, die betroffene Person über die Herkunft der Daten zu informieren, um ihr so eine Herkunftskontrolle zu ermöglichen.

Im Rahmen von KI sind zwei Aspekte von besonderer Bedeutung: die Frage, wie die Transparenz konkret hergestellt wird und der Punkt zu automatisierten Einzelentscheidungen und Profiling.

a. Es wird oben deutlich, dass eine Reihe von Einzelangaben zu machen sind. Je nach der Art der Anwendung (z. B. Einbindung in umfassendere Systeme, begrenzter Bildschirminhalt, sprach- oder gestengesteuerte Systeme, Verarbeitung im Hintergrund) stellen sich die gleichen Herausforderungen wie bei vielen klassischen KI-freien Anwendungen. Die Komplexität steigt jedoch, wenn zur Tatsache des KI-Einsatzes und zu relevanten Umständen ebenfalls umfassende Angaben zu machen sind. Dieser Aspekt wird sich durch die KI-spezifischen Anforderungen der EU KI-Verordnung weiter verschärfen. Hier kämen im Einzelfall – nicht immer mit der Garantie völliger Rechtssicherheit – die Verwendung von Links, QR-Codes, Symbolen und Piktogrammen oder auch die Bereithaltung klassischer Papierdokumente in Betracht. Hier kommt es ggf. zu besonderen Herausforderungen, wenn Informationen nur auf Umwegen bereitgestellt werden – z. B. wenn in einem Laden oder an öffentlichen Plätzen lediglich Links bzw. QR-Codes publiziert werden, über die weitere Informationen abrufbar sind, während nicht vorausgesetzt werden kann, dass alle Betroffenen diese Informationen zumutbar zeitgleich abrufen können.

b. Wortgleich ist nach Art. 13 und 14 DS-GVO über »das Bestehen einer automatisierten Entscheidungsfindung einschließlich Profiling gemäß Artikel 22 (1,4) DS-GVO« zu informieren und es sind darüber hinaus »zumindest in diesen Fällen – aussagekräftige Informationen über die involvierte Logik sowie die Tragweite und die angestrebten Auswirkungen einer derartigen Verarbeitung für die betroffene Person« bereitzustellen.

    a. Art. 22 (1) DS-GVO bezieht sich dabei auf das Recht des Betroffenen, nicht einer ausschließlich automatisierten Entscheidung mit rechtlicher Wirkung oder vergleichbaren Auswirkungen unterworfen zu werden.

    b. Art. 22 (4) DS-GVO bestimmt enge Grenzen für die Verwendung besondere Arten personenbezogener Daten, wenn in Abweichung von Abs. 1 eine automatisierte Entscheidung ausnahmsweise nach Abs. 2 zulässig ist.

Aufgrund der denkbaren Eingriffstiefe wird häufig intensiverer Kontakt mit einem Betroffenen (z. B. als Kunde, Bewerber) bestehen, sodass diese Informationen außerhalb der eigentlichen KI-Anwendung gegeben werden können. Beabsichtigt eine Versicherung etwa, kleinere Kraftfahrschäden mit geringer Schadenshöhe durch KI-Einsatz zu automatisieren, sind entsprechende Angaben erforderlich. Das bringt dann ggf. die Gefahr mit sich, dass durch Angaben zur Logik das System oder Modell unsachgemäß beeinflusst und ausgenutzt werden könnte.

Oft wird auch eine Datenschutz-Folgenabschätzung nach Art. 35 (1,3) a) DS-GVO erforderlich sein. Diese wird im Folgenden auf Seite 31 detailliert erläutert.

Schließlich ist es auch empfehlenswert, signifikante Entscheidungen nicht ausschließlich auf eine automatisierte Entscheidung zu stützen (»human in the loop«, d. h. menschliche Entscheidungsfindung, die ggf. durch KI vorbereitet und unterstützt wird – dieser Schritt sollte aber nicht auf eine rein formelle Prüfung, gleichsam ein Abnicken des KI-Entscheidungsvorschlages, reduziert werden).

Artikel 12 ff. DS-GVO: Umsetzung von Betroffenenrechten

Aus Art. 12 ff. DS-GVO ergeben sich Betroffenenrechte, die relevant werden, wenn ein Unternehmen mittels KI personenbezogene Daten verarbeitet.

Das Unternehmen hat die Betroffenen umfassend über die Verarbeitung ihrer Daten zu informieren. Nicht mehr benötigte Daten sind zu löschen und unrichtige Daten zu berichtigen. Bei einer Änderung der Zweckbestimmung der Daten, wie z. B. der Weiterverwendung für das Training einer KI, sind die betroffenen Personen grundsätzlich zu informieren.

Darüber hinaus regeln Art. 13 (2) f) und Art. 14 (2) g) DS-GVO besondere Informationspflichten bei automatisierten Entscheidungen, insbesondere Profiling. In diesen Fällen hat das Unternehmen die betroffene Person über das Bestehen einer automatisierten Entscheidungsfindung einschließlich Profiling zu informieren und zudem aussagekräftige Informationen über die involvierte Logik sowie die Tragweite und die angestrebten Auswirkungen einer derartigen Verarbeitung für die betroffene Person bereitzustellen. Dies kann Unternehmen speziell dann vor Schwierigkeiten stellen, wenn diese KI verwenden, die von anderen Anbietern entwickelt wurde und sie daher mitunter keine Kenntnis über die involvierte Logik haben.

Neben dem Informationsrecht sind auch im Rahmen des Auskunftsrechts der betroffenen Person gemäß Art. 15 (1) h) DS-GVO zusätzliche Pflichten geregelt, wenn automatisierte Entscheidungsfindungen, einschließlich Profiling, durchgeführt werden. Betroffene können verlangen, detaillierte Informationen über die Logik, Reichweite und die beabsichtigten Auswirkungen solcher Verarbeitungen zu erhalten. Bei der Anwendung von KI kann das Bereitstellen präziser Informationen herausfordernd sein, besonders bei komplexen Datenverarbeitungsprozessen. Deshalb müssen Unternehmen gewährleisten, dass sie auch bei anspruchsvollen KI-gestützten Vorgängen in der Lage sind, klare und nachvollziehbare Informationen bezüglich der Datenverarbeitung zu vermitteln.

Insbesondere: Sicherstellung der Betroffenenrechte im »Data Lake«

Auch wenn zum Training der KI große Datenmengen betroffener Personen verarbeitet werden und in einem »Data Lake« zusammengefasst werden, stehen den Betroffenen die Rechte zu, die oben im Abschnitt 2 beschrieben wurden. Beim »Data Lake« stellen sich in der Praxis vor allem zwei Herausforderungen: 1) Information der Betroffenen und 2) Auskunftsrechte.

Werden Trainingsdaten erhoben, so sind die betroffenen Personen zu informieren. Allerdings ist keine Information notwendig (Art. 14 (5) DS-GVO), wenn sie mit einem unverhältnismäßigen Aufwand verbunden wäre, beispielsweise wegen der großen Zahl der Personen und eventueller Schwierigkeiten bei der Ermittlung. Dies bedeutet, dass der Verantwortliche vor der Entscheidung über die Information den entstehenden Aufwand mit den Informationsinteressen der betroffenen Personen abwägen muss. Waren Daten bereits vorher öffentlich zugänglich, so fällt dies bei der Abwägung entscheidend zugunsten des Verantwortlichen ins Gewicht, wenn die Einbeziehung solcher Daten in das KI-Training die Interessen des Betroffenen nicht gravierend beeinträchtigt. Hier wird die Informationspflicht häufig entfallen. Gleiches dürfte wohl gelten, wenn eine »Verwässerung der individuellen Betroffenheit« aufgrund einer massenhaften Verarbeitung personenbezogener Daten beim Training/Finetuning von KI-Modellen und Systemen eintritt, sodass die nach DS-GVO erforderliche Betroffenheit unterschritten wird (im Lichte des EuG-Urteils vom 26.4.2023 (Az: T-557/20).

Ein ähnliches Problem stellt sich, wenn ein Betroffener Auskunft über seine Daten verlangt, Art. 15 DS-GVO. In vielen Fällen wird der Verantwortliche gar nicht in der Lage sein, einem Betroffenen bestimmte Daten zuzuordnen. Dabei ist im Zusammenhang mit dem Auskunftsverlangen vor allem die Regelung des Art. 11 DS-GVO wichtig. Wenn der Verantwortliche den Betroffenen nicht identifizieren kann, entfällt die Auskunftsverpflichtung. Der Verantwortliche braucht auch keine Informationen aufzubewahren oder zu erheben, um die Identifizierung zu ermöglichen. Im Sinne der Datensparsamkeit ist es sinnvoll, solche Daten zur Identifizierung so früh wie möglich zu löschen.

Artikel 24 ff. DS-GVO: Datenschutzrechtliche Verantwortlichkeit

Werden im Rahmen der Entwicklung, des Trainierens oder der Verwendung von KI-Modellen/KI-Systemen personenbezogene Daten verarbeitet, muss zunächst die Verantwortlichkeit für die Datenverarbeitung geklärt werden. An diese Feststellung knüpfen dann Pflichten und Rechtsfolgen aus den datenschutzrechtlichen Vorschriften. Es können unterschiedliche Konstellationen von Verantwortlichkeiten vorliegen, die im Einzelfall zu prüfen sind:

1. **Alleinige Verantwortlichkeit (Independent Controller)**
   Eine eigenständige Verantwortlichkeit liegt vor, wenn der Verantwortliche unabhängig von anderen über die Zwecke und Mittel der Verarbeitung von personenbezogenen Daten entscheidet.

Beispiel:
Allein verantwortlich für die Datenverarbeitung ist in der Regel ein Unternehmen, das ein KI-basiertes System selbst entwickelt, bereitstellt und nutzt und somit über die Zwecke und Mittel der Datenverarbeitung allein entscheidet.[4]

2. **Gemeinsame Verantwortlichkeit (Joint Controller)**
Entscheiden zwei oder mehr Parteien gemeinsam über die Zwecke und Mittel der Verarbeitung personenbezogener Daten, sind diese gemeinsame Verantwortliche nach Art. 26 (1) S. 1 DS-GVO. Wäre die Verarbeitung ohne die Beteiligung beider Parteien nicht möglich und ist die Verarbeitung der Parteien untrennbar miteinander verbunden, liegt ein weiterer Hinweis auf eine gemeinsame Verantwortlichkeit vor.

Bei einer gemeinsamen Verantwortlichkeit müssen beide oder mehrere Verantwortliche eine Vereinbarung über die gemeinsame Verantwortlichkeit nach Art. 26. (1) S. 2 DS-GVO abschließen (sog. Joint Controller Agreement, JCA). In dieser Vereinbarung werden insbesondere die Verpflichtungen und Verantwortlichkeiten zu den Betroffenenrechten und Informationspflichten nach Art. 12, 13 und 14 DS-GVO festgelegt.

Beispiel:
Eine gemeinsame Verantwortlichkeit wäre denkbar, wenn das KI-System oder Modell Datensätze verschiedener Unternehmen für das Trainieren einer gemeinsamen KI verwendet.[5] Eine solche kann auch vorliegen, wenn verschiedene Unternehmen mit ihren Kompetenzen gezielt zusammenwirken, z. B. durch Zurverfügungstellung der Verarbeitungstechnologie durch ein Unternehmen und

das Training des KI-Systems oder des KI-Modells mit Daten eines anderen Unternehmens zur Erreichung eines gemeinsamen Zwecks. Bei reiner Lizenzierung von Daten für Trainingszwecke wird im Zweifel eher eine alleinige Verantwortung des Verarbeiters bestehen.

3. **Auftragsverarbeitung (Data Processor)**
Verarbeitet ein Unternehmen nur im Auftrag und auf Weisung eines Verantwortlichen personenbezogene Daten, gilt dieser als Auftragsverarbeiter. Liegt ein Auftragsverarbeitungsverhältnis vor, muss zwischen dem Verantwortlichen und dem Auftragsverarbeiter ein Auftragsverarbeitungsvertrag nach Art. 28 (3) DS-GVO abgeschlossen werden.

Beispiel:
Eine Auftragsverarbeitung liegt bspw. vor, wenn ein Entwickler oder Anbieter eines KI-Systems bzw. Modells personenbezogene Daten unter Weisung eines

[4] Diskussionspapier »Rechtsgrundlagen im Datenschutz beim Einsatz von Künstlicher Intelligenz« des LfDi BW: https://www.baden-wuerttemberg.datenschutz.de/rechtsgrundlagen-datenschutz-ki/
[5] Diskussionspapier »Rechtsgrundlagen im Datenschutz beim Einsatz von Künstlicher Intelligenz« des LfDi BW: https://www.baden-wuerttemberg.datenschutz.de/rechtsgrundlagen-datenschutz-ki/

Verantwortlichen verarbeitet und ausschließlich zu Zwecken des Verantwortlichen mit diesen Daten trainiert.[6]

Ein weiteres Beispiel für eine Auftragsverarbeitung kann die Nutzung eines bereits bestehenden, Cloud-basierten, online zur Verfügung stehenden KI-Systems eines KI-Anbieters sein.

Achtung: Es ist eine genaue Prüfung erforderlich, um festzustellen, ob eine Auftragsverarbeitung vorliegt und ob der KI-Anbieter die entsprechenden Vereinbarungen einhalten kann. Dies ist von besonderer Bedeutung, da der KI-Anbieter in einigen Fällen nicht in der Lage ist, Aspekte des Modells zu beeinflussen.

Viele KI-Anbieter verarbeiten die erhobenen Daten zur Weiterentwicklung ihrer Modelle, was zu einer gemeinsamen Verantwortlichkeit führen kann. Es ist unbedingt zu berücksichtigen, dass die Entwicklung und das Training von KI-Systemen und Modellen dynamische Prozesse darstellen, die stetig neue Datenschutzherausforderungen mit sich bringen. Modelle, Datensätze und Ziele können sich täglich ändern, was die Bewertungskriterien für die Verantwortlichkeit stetig beeinflusst. Auch die Zusammenarbeit zwischen externen Dienstleistern und dem KI-Anbieter muss berücksichtigt werden. Eine frühzeitige und konstante Einbindung des Datenschutzteams sowie eine kontinuierlich aktualisierte Dokumentation sind essenziell, um den Transparenzpflichten der DS-GVO und der KI-Verordnung hinreichend nachzukommen. Was bei der Entwicklung nicht ausreichend und richtig dokumentiert wird, ist später in der Wertschöpfungskette praktisch nicht mehr korrigierbar, sodass eine Nutzung des KI-Systems oder Modells dann kaum noch rechtskonform möglich ist.

Die Entwicklungs-, Trainings-, interne Verwendungsphase und externe Bereitstellung sowie die Verantwortlichkeiten je nach Verarbeitungszweck müssen separat betrachtet werden. Diese Verantwortlichkeiten können sich innerhalb dieser Phasen im Laufe der Zeit ändern. Auch Art. 6 c) der Produkthaftungsrichtlinie (PLD) erkennt an, dass ein KI-System schon aufgrund von Kenntnissen, die nach der Implementierung erworben/gelernt wurden, fehlerhaft werden kann, wodurch die Haftung des Verantwortlichen auf solche Zeiträume und Weiterentwicklungen ausgedehnt wird.

Bei der Nutzung sind ebenfalls diverse Konstellationen im Detail zu überprüfen. Gerade in Konzernen mit vielen Gesellschaften bei denen z. B. die Muttergesellschaft das KI-System oder Modell einkauft und betreibt, die Tochtergesellschaften diese aber nutzen ist besondere Vorsicht geboten. Hier können sich die Verantwortlichkeiten für die jeweils konkrete Eingabe, Speicherung, Übertragung, Weitergabe oder sonstige Verarbeitung von personenbezogenen Daten durch das KI-System oder das Modell überlagern bzw. an einem bestimmten Punkt ändern.

Aus den Vorschriften der KI-Verordnung zum »Quasi Provider« (Art. 25 KI-Verordnung) ergeben sich ggf. ebenfalls Auswirkungen auf die datenschutzrechtliche Verantwortlichkeit. Art. 25 (1) c) KI-Verordnung regelt, dass Betreiber die den beabsichtigten Nutzungszweck des KI-Systems oder des Modells ändern oder substanzielle Änderungen daran vornehmen, wie ein Anbieter behandelt wird. Es ist schwer

[6] Diskussionspapier »Rechtsgrundlagen im Datenschutz beim Einsatz von Künstlicher Intelligenz« des LfDi BW: https://www.baden-wuerttemberg.datenschutz.de/rechtsgrundlagen-datenschutz-ki/

vorstellbar, dass eine Änderung des beabsichtigten Nutzungszwecks der KI nicht auch zu einer (mindestens gemeinsamen) datenschutzrechtlichen Verantwortlichkeit führt.

Artikel 25 ff. DS-GVO: Privacy by Design/Privacy by Default und Einsatz von geeigneten technischen und organisatorischen Maßnahmen

Im Zusammenhang mit KI erlangen die Konzepte »Privacy by Design« und »Privacy by Default« zunehmend an Bedeutung. Diese Ansätze sind nicht nur integraler Bestandteil der DS-GVO, sondern stellen auch einen zentralen Aspekt im Umgang mit personenbezogenen Daten innerhalb von KI-Systemen und Modellen dar. Der folgende Abschnitt befasst sich mit der Anwendung dieser Prinzipien im Kontext von KI und beleuchtet die Rolle geeigneter technischer und organisatorischer Maßnahmen (TOM) zur Gewährleistung des Datenschutzes.

»Privacy by Design« bezeichnet einen Ansatz, bei dem Datenschutz bereits in der Entwicklungsphase von Produkten und Systemen berücksichtigt wird. »Privacy by Default« hingegen sichert durch die Umsetzung datenschutzfreundlicher Voreinstellungen, dass standardmäßig nur die für den jeweiligen Zweck notwendigen personenbezogenen Daten verarbeitet werden. Beide Konzepte sind in Artikel 25 DS-GVO verankert und verpflichten Entwickler und Anbieter von KI-Systemen, Datenschutz von Anfang an in ihre Systeme und Prozesse zu integrieren.

Die Implementierung geeigneter technischer und organisatorischer Maßnahmen ist essenziell, um den Anforderungen des Datenschutzes in KI-Systemen und Modellen gerecht zu werden. Bei der Auswahl der Maßnahmen sind Art, Umfang, Umstände und Zwecke der Verarbeitung sowie die aus der Verarbeitung hervorgehenden Risiken für die betroffenen Personen und der Stand der Technik zu berücksichtigen. Umgesetzte TOM sollten kontinuierlich überprüft und ggf. angepasst werden, um aktuellen Risiken und technologischen Entwicklungen Rechnung zu tragen.

Beispielhaft bieten sich als Maßnahmen zur Datenminimierung unter anderem vorgeschaltete Prompt-Filter oder vergleichbare Technologien zur Einschränkung/Standardisierung des Inputs an, um die Eingabe personenbezogener Daten so weit wie möglich zu verhindern. Weiterhin sollten Beschäftigte mittels Schulungen und Informationsmaterial dabei unterstützt werden, ihre Eingaben in die KI-Anwendung möglichst datensparsam gestalten zu können. Hinweise und Regelungen zur (Nicht-)Eingabe personenbezogener Daten in KI-Anwendungen können auch in Form von Nutzungsbedingungen, Richtlinien oder einem Verhaltenskodex geregelt werden. Die Erstellung pseudonymisierter Benutzeraccounts und die Anonymisierung der in das KI-System oder dem Modell eingehenden Daten, dient weiterhin dem Ziel der Datenminimierung.

Um die Einhaltung der Zweckbindung sicherzustellen, sollte mit dem Anbieter vertraglich ausgeschlossen werden, dass die in die KI-Anwendung eingegebenen Daten zum weiteren Training der KI durch den Anbieter weiterverarbeitet werden. In manchen Anwendungen kann dies auch über entsprechende Konfigurationsmöglichkeiten technisch umgesetzt werden, genauso wie ein Opt-Out aus dem anwendungsseitigen Anlegen eines Nutzungsverlaufs. Falls solche Einstellungen nur auf Nutzerebene

vorgenommen werden können, sollten die Nutzerinnen und Nutzer entsprechend instruiert werden. Weiterhin sollte sichergestellt sein, dass die mit der KI erzeugten Inhalte, sofern sie personenbezogene Daten enthalten, nur nach dem Need-to-know-Prinzip verarbeitet werden.

Für den Einsatz von KI-Anwendungen liegt bei der Umsetzung von TOM eine Herausforderung darin, dass bei den unterschiedlichen auf dem Markt verfügbaren Einsatzformen von KI (z. B. »KI as a Service«, öffentlich verfügbare Systeme oder individuell erstellte/angepasste Systeme) unterschiedlicher Spielraum für das Ergreifen eigener Maßnahmen seitens des Verantwortlichen besteht. So wird in der Regel bei selbst oder spezifisch im Auftrag erstellten KI-Anwendungen auch umfangreiches Customizing möglich sein, während bei der Nutzung eines allgemein z. B. per Browser aufrufbaren Standardprodukts kaum Einfluss auf die technische Gestaltung genommen werden kann. Entsprechend sind je nach Anwendung unterschiedliche technische Maßnahmen aufseiten des Verantwortlichen umsetzbar oder ggf. nicht umsetzbar. Organisatorische Maßnahmen sind von dieser Variabilität weniger betroffen, wodurch der Verantwortliche sie weitgehend eigenständig und unabhängig von der konkreten Anwendung implementieren kann.

Bei Einsatzformen von KI-Systemen und Modellen, die es einsetzenden Unternehmen nicht ermöglichen, technische Maßnahmen in der Anwendung selbst umzusetzen, sollten - insbesondere mit Blick auf die Transparenz - zumindest entsprechende Hinweise implementiert werden. Dies kann z. B. bei der Nutzung allgemein verfügbarer online KI-Anwendungen auf dienstlichen Endgeräten durch eine in den Browser integrierte Warnung erfolgen, in der bei Besuch einer entsprechenden Website auf bestehende Acceptable Use-Richtlinien, Betriebsvereinbarungen o.Ä. hingewiesen wird. Über die Risiken der Nutzung solcher KI-Anwendungen, mit deren Anbietern in der Regel gerade keine gesonderten vertraglichen Vereinbarungen bzgl. Vertraulichkeit, Zweckbindung etc. bestehen, sollten die Nutzer unabhängig davon auch mittels Schulungen und Informationsmaterial sensibilisiert werden.

Letztendlich bemisst sich der Umfang und die Auswahl der zu ergreifenden TOM an den mit dem konkreten Use Case verbundenen Risiken und Umständen, muss also stets in einer Einzelfallbetrachtung ermittelt werden. Neben den hier beispielhaft genannten eher KI-spezifischen Maßnahmen sind dabei auch zur grundsätzlichen Gewährleistung der Datensicherheit in IT-Systemen eingesetzte TOM, z. B. Verschlüsselung und Zugriffskontrollen, zu berücksichtigen. Wo dem Verantwortlichen selbst die Implementierung insbesondere technischer Maßnahmen vor dem oben dargestellten Hintergrund nicht ausreichend möglich ist, muss auf andere Weise (z. B. vertraglich mit dem Anbieter) sichergestellt werden, dass trotzdem ein angemessenes Schutzniveau sichergestellt ist.

Die Integration von »Privacy by Design« und »Privacy by Default« in KI-Systeme oder Modelle sowie die Implementierung geeigneter TOM sind nicht nur rechtliche Erfordernisse, sondern tragen auch zur Vertrauensbildung bei Nutzenden und zur Förderung ethischer Standards in der Technologie bei. Ein einheitliches Ordnungsschema zur Identifikation und Umsetzung relevanter technischer und organisatorischer Maßnahmen hinsichtlich Privacy by Design und Sicherheit der Verarbeitung bietet das

vom Bitkom erstellte Datenschutz-Reifegradmodell zur Abbildung von technisch-organisatorischen Maßnahmen bei der Auftragsverarbeitung.[7]

Artikel 30 DS-GVO: Aufnahme der Verarbeitung in das Verzeichnis von Verarbeitungstätigkeiten

Sollten personenbezogene Daten mithilfe einer KI verarbeitet werden, so sind diese Verarbeitungstätigkeiten gemäß Art. 30 DS-GVO in einem Verzeichnis von Verarbeitungstätigkeiten zu hinterlegen.

Details zur Definition eines solchen Verarbeitungsverzeichnisses, zu Grundlagen sowie zum Prozess zur Erstellung einer solchen Dokumentation finden Sie im **Bitkom-Leitfaden »Das Verarbeitungsverzeichnis«**[8].

Hier besteht die Gefahr von intransparenter Dokumentation. Diese Intransparenz tritt aufgrund von ungenauen Angaben in Bezug auf die Verarbeitung der Daten der einzelnen KI-Anwendungen auf. Oft ist unklar, wie, wozu und wie lange die verwendeten Daten auf Servern von KI-Anbietern verarbeitet werden. Hierbei handelt es sich um eine »Black Box«, die sich aufgrund von selbstlernenden und komplexen Strukturen bildet. [1]

Um dieser Intransparenz entgegenzuwirken, sollten Kontrollmaßnahmen festgehalten werden, die detaillierte Informationen über die Verarbeitung der Daten sicherstellen, damit eine präzise Dokumentation des Verzeichnisses von Verarbeitungstätigkeiten gewährleistet werden kann.

Holthausen in Recht der Arbeit, 2021 Heft 1: Big Data, People Analytics, KI und Gestaltung von Betriebsvereinbarungen – Grund-, arbeits- und datenschutzrechtliche An- und Herausforderungen

Artikel 33, 34 DS-GVO: Prozess Datenschutzvorfall

Der Datenschutzvorfall-Prozess beim Einsatz von KI unterscheidet sich grundsätzlich nicht vom allgemeinen Vorgehen bei einer DSV (Datenschutzverletzung). Besonderes Augenmerk ist auf die Analysephase und den Faktor Zeit zu legen, Details s. u. *Herausforderungen*.

Allgemeine Informationen hinsichtlich der Verletzung des Schutzes personenbezogener Daten im Sinne der DS-GVO, einer ggf. notwendigen Meldung und Benachrichtigung betroffener Personen, sowie zur Auslegung der Art. 33 und 34 DS-GVO sind im **Bitkom-Leitfaden Datenschutzverletzung und Meldung im Kontext des**

[7] Datenschutz-Reifegradmodell zur Abbildung von technischorganisatorischen Maßnahmen bei der Auftragsverarbeitung, Betaversion | Leitfaden 2022 | Bitkom e. V.
[8] Das Verarbeitungsverzeichnis | Leitfaden 2017 | Bitkom e. V.

»Hafnium Hacks«[9] dargestellt. Dort finden sich auch Ausführungen zu Auswirkungen auf eine Auftragsverarbeitung.

Anwendungsfälle:

Beim Einsatz von KI sind insbesondere folgende **Use Cases** im Kontext von Datenschutzvorfällen vorstellbar:

- **Anwendungsfall 1:** Weiterverbreitung oder Veröffentlichung von personenbezogenen Daten durch KI-Systeme und Modelle ohne Wissen und ohne Rechtsgrundlage (Verstoß gegen Grundsatz der Transparenz, Rechtmäßigkeit)
- **Anwendungsfall 2**: KI funktioniert nicht ordnungsgemäß – Softwarefehler in KI-Lösung verursacht eine unbefugte Offenbarung personenbezogener Daten von Nutzerinnen und Nutzer
  - Beispiel: im Beschäftigtenkontext werden Daten wegen fehlerhafter KI mit Kolleginnen und Kollegen geteilt, die keine Zugriffsberechtigung besitzen
- **Anwendungsfall 3:** KI-gesteuerter Hackerangriff auf ein Unternehmen, bei dem personenbezogene Daten z. B. von Kundinnen und Kunden oder Beschäftigten kompromittiert werden
  - Beispiel: Ein KI-gesteuertes Chatbot-System könnte von einem Angreifer gehackt werden, der dann gefälschte Unterhaltungen führt, um persönliche Informationen von Nutzern zu stehlen, indem er sich als legitimer Service ausgibt.
- **Anwendungsfall 4:** Kompromittierung von personenbezogenen Daten bei Vorgängen ohne menschliche Kontrolle bzw. Interaktion mit direkten rechtlichen Auswirkungen auf betroffene Personen z. B. im Rahmen von automatisierten Entscheidungsfindungen oder KI-gestütztem Scoring
  - Beispiel 1: Anwendung für Bonitätsrating/Kreditvergabe
  - Beispiel 2: Ein Bewerbungs-Tracking-System, das auf KI basiert, könnte aufgrund von Voreingenommenheit in den Trainingsdaten Bewerberinnen und Bewerber bestimmter ethnischer Gruppen benachteiligen, indem es sie fälschlicherweise ausschließt.

Hinweis: Kommt der Verantwortliche bei der Analyse zu dem Ergebnis, dass die DSV voraussichtlich ein hohes Risiko für Rechte und Freiheiten der Betroffenen zur Folge hat, muss er diese gemäß Art. 34 (1) DS-GVO unverzüglich von der Verletzung benachrichtigen.[10]

[9] Datenschutzverletzung und Meldung im Kontext des »Hafnium Hacks« | Leitfaden 2021 | Bitkom e. V.
[10] Zu Modalitäten und Inhalt einer Benachrichtigung s. Art. 34 Abs. 1 und 2 DS-GVODS-GVO, zu Ausnahmen s. Art. 34 (3) DS-GVODS-GVO. Details s. Bitkom-LF aaO, vgl. Fn. 15

Herausforderungen:

Bei der Beurteilung von Datenschutzvorfällen in Verbindung mit KI gibt es spezifische Herausforderungen, welche den nachfolgenden Aspekten zugeordnet werden können:

- **Feststellung einer Datenschutzverletzung**: Zu Transparenzanforderungen in Verbindung mit KI (Siehe die Ausführungen im Abschnitt Transparenz- und Informationspflichten auf Seite 26.) Aufgrund der technischen Komplexität von KI leidet die Nachvollziehbarkeit ihrer Funktionsweise. Verantwortlichen fehlt häufig das nötige technische Verständnis. Dieses ist jedoch wichtig, um beurteilen zu können, ob KI ordnungsgemäß funktioniert. Die hohe Komplexität von KI-Systemen oder Modellen erschwert zudem die Risikoeinschätzung im Einzelfall. Die Feststellung, ob überhaupt eine Verletzung der Sicherheit und damit eine meldepflichtige Datenschutzverletzung vorliegt, kann somit in der Praxis herausfordernd sein. Oft sehen sich Verantwortliche mit intransparenten und wenig verständlichen Informationen der KI-Hersteller konfrontiert. Wichtig dabei ist, im Einzelfall zu untersuchen, welche personenbezogenen Daten konkret z. B. für ein Training von Systemen verwendet werden. Diese Analyse sollte bestenfalls schon im Rahmen der Dokumentation für das Verzeichnis von Verarbeitungstätigkeiten i.S.v. Art. 30 DS-GVO geschehen – bevor es zu einem Vorfall kommt.

- **Verantwortlichkeit:** Der Aspekt der datenschutzrechtlichen Verantwortlichkeit kann bei einer DSV vor dem Hintergrund der gesetzlichen 72h-Stunden-Frist besonders herausfordernd werden. Es wird daher empfohlen, diese Frage frühzeitig zu klären, bevor es zu einer Datenschutzverletzung kommt. An dieser Stelle wird auf die Verpflichtung zur *unverzüglichen* Meldung von Auftragsverarbeitern gegenüber dem Verantwortlichen gemäß Art. 33 (2) DS-GVO hingewiesen[11].

- **Faktor Zeit**: Bei Datenschutzverletzungen ist vor dem Hintergrund der strengen gesetzlichen Anforderungen (72-Stunden-Meldefrist für Verantwortliche, »unverzüglich« für Auftragsverarbeiter) besonders wichtig, nach Bekanntwerden unverzüglich mit einer Analyse zu beginnen. Denn im Zusammenhang mit KI sind mögliche Folgen für die Datensicherheit aufgrund der Komplexität der KI (z. B. wo kommen die Daten der KI her) (s. o.) schwerer abschätzbar als in Szenarien ohne KI. Hilfreich ist die Möglichkeit nach Art. 33 (4) DS-GVO bzgl. eines schrittweisen Vorgehens, wenn Informationen nicht zur gleichen Zeit bereitgestellt werden können. Dann können diese schrittweise der zuständigen Aufsichtsbehörde zur Verfügung gestellt werden.

- **Fehlende Praxisfälle und unterschiedliche Positionierung der Datenschutz-Aufsichtsbehörden:** Angesichts der rasanten Entwicklung der KI-Thematik stehen Verantwortliche (Unternehmen) mit ihren Erfahrungen zu Datenschutzverletzungen noch am Anfang. Herausfordernd ist auch die unterschiedliche Definition und Positionierung von Datenschutz-Aufsichtsbehörden von bzw. zu KI per se. Mit zunehmender praktischer Erfahrung seitens der Verantwortlichen und spezifischer Hilfestellungen seitens der Datenschutz-Aufsichtsbehörden konkret

[11] Details s. Bitkom-Leitfaden, aaO, s. Fn. 15.

zum Thema Datenschutzverletzung bei KI-Nutzung werden Verantwortliche und Auftragsverarbeiter mehr Orientierung und Rechtssicherheit erhalten.

- Praktische Hilfestellung im Vorfeld bzw. bei der Analyse von Datenschutzvorfällen bieten folgende **Fragestellungen**:
- Integration von KI: Auf welche Art wurde die KI integriert?
- Beispiel: sog. »model serving« oder »model training«?
- Reichweite: Ist die KI-Anwendung nach »außen« exponiert, i.S.v. öffentlich zugänglich?
- Technische und organisatorische Maßnahmen (TOM):
  - Welche technischen und organisatorischen Maßnahmen müssen ergriffen werden, um sicherzustellen, dass personenbezogene Daten im Rahmen der Eingabe oder dem Abruf von Ergebnissen vor unbefugter Offenlegung, Veränderung oder Verlust der Verfügbarkeit geschützt sind?
  - Beispiele für TOM: KI-Governance, Privacy-by-design/default, Pseudonymisierung, Anonymisierung, Verschlüsselung, sichere Speicherung, usw.
  - Greifen die Schutzmaßnahmen wie geplant?
- Vertragliche Aspekte:
  - Bei Beschaffung von Generativer KI wird empfohlen, die Vertragsbedingungen sorgfältig zu prüfen, unter denen ein KI-System oder Modell erworben bzw. lizenziert wird. Ist der Umgang mit Datenschutzverletzungen geregelt und insbesondere welche Vertragspartei trägt welche Pflichten?

Artikel 35 DS-GVO: Durchführung einer Datenschutzfolgenabschätzung/ Folgenabschätzung

Nach der KI-Verordnung ist vor Einführung oder Nutzung von Hoch-Risiko-Systemen und General-Purpose-KI-Modellen und Systemen mit systemischen Risiken ein »fundamental rights impact assessment« (»FRIA«) (Grundrechtsfolgenabschätzung) durchzuführen. Diese Folgenabschätzung ist mit einer Datenschutz-Folgen-abschätzung nach DS-GVO nicht identisch, sondern eine **zusätzliche** Anforderung der KI-Verordnung.

Beim Einsatz von KI spielt das Thema Datenschutz-Folgenabschätzung eine große Rolle, da die Verwendung von KI-Systemen und Modellen angesichts einer potenziellen Diskriminierungsgefahr sowie fehlender Kontrollmöglichkeiten mit hohen Risi-ken für die Rechte und Freiheiten der betroffenen Personen verbunden sein kann.

Allgemeine Ausführungen zum Thema Datenschutz-Folgenabschätzung (DSFA) sowie zum Vorgehen bei der Prüfung der Pflicht zur Durchführung einer DSFA enthält der **Bitkom-Leitfaden »Risk Assessment & Datenschutz-Folgenabschätzung«**[12].

Erforderlichkeit einer DSFA/FRIA

Ob eine DSFA durchzuführen ist, ergibt sich aus einer Abschätzung der Risiken der Verarbeitungsvorgänge **(»Schwellwertanalyse«**). Ergibt diese ein voraussichtlich hohes Risiko, dann ist eine DSFA durchzuführen. Wird festgestellt, dass der Verarbeitungsvorgang kein hohes Risiko aufweist, dann ist eine DSFA nicht zwingend erforderlich. In jedem Fall ist die Entscheidung über die Durchführung oder Nichtdurchführung der DSFA mit Angabe der maßgeblichen Gründe für den konkreten Verarbeitungsvorgang schriftlich zu dokumentieren.

Bei der Bewertung des konkreten Risikos muss zunächst die Auswirkung der gesetzlichen Kategorisierung von KI-Systemen nach der KI-Verordnung (Art. 6, Annex II) berücksichtigt werden. Wenn KI-Systeme bereits nach der KI-Verordnung als Hochrisikosystem gelistet/angesehen werden, ist recht unwahrscheinlich, dass die gleichen Systeme rein datenschutzrechtlich zu keinem hohen Risiko führen.

Im Rahmen der FRIA werden u. a. auch datenschutzrelevante Aspekte geprüft, wie

- Kategorien von natürlichen Personen und Gruppen, die von der Nutzung des Systems betroffen sein könnten
- Vereinbarkeit der Nutzung des Systems mit den einschlägigen Rechtsvorschriften der Union und der Mitgliedstaaten über die Grundrechte (= DS-GVO)
- die nach vernünftigem Ermessen vorhersehbaren Auswirkungen des Einsatzes des Hochrisiko-KI-Systems auf die Grundrechte
- spezifische Schadensrisiken, die sich auf marginalisierte Personen oder schutzbedürftige Gruppen auswirken können

Daher ist ebenfalls unwahrscheinlich, dass FRIA und DFSA bei den Risiken zu erheblich abweichenden Ergebnissen kommen.

Art. 35 (3) DS-GVO benennt – nicht abschließend vgl. *»insbesondere«* - einige Faktoren, die wahrscheinlich zu einem hohen Risiko i.S.d. Art. 35 (1) DS-GVO und damit zu einer entsprechenden Pflicht zur Durchführung einer DSFA führen:

a) systematische und umfassende Bewertung persönlicher Aspekte natürlicher Personen, die sich auf automatisierte Verarbeitung einschließlich Profiling gründet und die ihrerseits als Grundlage für Entscheidungen dient, die Rechtswirkung gegenüber natürlichen Personen entfalten oder diese in ähnlich erheblicher Weise beeinträchtigen;

b) umfangreiche Verarbeitung besonderer Kategorien von personenbezogenen Daten gemäß Artikel 9 (1) DS-GVO oder von personenbezogenen Daten über strafrechtliche Verurteilungen und Straftaten gemäß Artikel 10 DS-GVO oder

[12] Bitkom-Leitfaden Risk Assessment & Datenschutz-Folgenabschätzung, s. Risk Assessment & Datenschutz-Folgenabschätzung | Leitfaden 2017 | Bitkom e. V.

c) systematische, umfangreiche Überwachung öffentlich zugänglicher Bereiche.

Die DSK hat eine **Muss-Liste** der Verarbeitungsvorgänge i.S.v. Art. 35 (4) S. 1 DS-GVO erstellt, für die eine Datenschutz-Folgenabschätzung durchzuführen ist (sogenannte »Blacklist«)[13]. Diese nimmt in Ziff. 11 explizit Bezug auf den Einsatz von KI:

Liste der Verarbeitungstätigkeiten, für die eine DSFA durchzuführen ist

Nr.	Maßgebliche Beschreibung der Verarbeitungstätigkeit	Typische Einsatzfelder	Beispiele
11	Einsatz von künstlicher Intelligenz zur Verarbeitung personenbezogener Daten zur Steuerung der Interaktion mit den Betroffenen oder zur Bewertung persönlicher Aspekte der betroffenen Person	Kundensupport mittels künstlicher Intelligenz	Ein Callcenter wertet automatisiert die Stimmungslage der Anrufer aus. Ein Unternehmen setzt ein System ein, welches mit Kunden durch Konversation interagiert und für deren Beratung personenbezogene Daten durch eine künstliche Intelligenz verarbeitet werden.

Die Liste ist nicht abschließend, sondern ergänzt die in den Absätzen 1 und 3 des Artikels 35 DS-GVO enthaltenen allgemeinen Regelungen. Die Liste orientiert sich wiederum an der allgemeinen Vorgehensweise wie beschrieben in Arbeitspapier 248 Rev. 1 der früheren Artikel 29-Gruppe[14] und ergänzt und konkretisiert diese.

Von der Möglichkeit, entsprechende **Braucht-Nicht-Listen** ohne DSFA-Pflicht gemäß Art. 35 (5) DS-GVO zu erstellen (»Whitelists«), haben die deutschen Aufsichtsbehörden bislang keinen Gebrauch gemacht. **Vorherige Konsultation der Aufsichtsbehörde gemäß Art. 36 DS-GVO**

Kommt der Verantwortliche bei der Durchführung einer DSFA zu dem Ergebnis, dass seine geplante Verarbeitung im Rahmen der Nutzung von KI ein hohes Risiko zur Folge hätte, muss er vor der Verarbeitung die Aufsichtsbehörde konsultieren, soweit er keine Maßnahmen zur Eindämmung des Risikos trifft.

Pflicht zur DSB-Benennung gemäß § 38 BDSG

Bei positiver Feststellung einer DSFA-Pflicht beim Einsatz von KI ist die nach § 38 I 2 BDSG resultierende Verpflichtung zur Benennung eines bzw. einer Datenschutz-beauftragen (DSB) zu beachten – und zwar unabhängig von der Anzahl der mit der

[13] Liste der Verarbeitungstätigkeiten, für die eine DSFA durchzuführen ist, s. Ziff. 11; Link s. Offizielles Kurzpapier der DSK (bayern.de)
[14] Datenschutzgruppe nach Artikel 29 (ersetzt durch EDSA seit 25.05.2018); vgl. Arbeitspapier 248 Rev. 1 *Leitlinien zur Datenschutz-Folgenabschätzung (DSFA) und Beantwortung der Frage, ob eine Verarbeitung im Sinne der Verordnung 2016/679 »wahrscheinlich ein hohes Risiko mit sich bringt«, s. wp248 rev.01_de (datenschutzkonferenz-online.de).*

Verarbeitung beschäftigten Personen[15]. Die Benennungspflicht richtet sich sowohl an Verantwortliche als auch an Auftragsverarbeiter und kann für diese eine Herausforderung darstellen.

Dies kann in der Praxis beispielsweise in folgenden Szenarien relevant sein:

Beispiel: Der Einsatz von KI-basierten Lösungen durch Verantwortliche, die selbst nicht Anbieter der Lösung sind, jedoch Produkte nutzen, in denen KI verbaut ist. Allein die Nutzung kann eine DSFA-Pflicht auslösen und damit zwangsweise die Bestellung eines DSB.

Beispiel: Die Nutzung von Online-Office-Suiten durch KMU.

Die vorgenannte, nationale Regelung steht in der Kritik. So hat sich u. a. Bitkom im Rahmen des Konsultationsprozesses zum geplanten BDSG-Änderungsgesetz des BMI im Jahr 2023 für eine Streichung von § 38 I 2 1. Fall BDSG ausgesprochen, um das Datenschutzrecht mit dem Recht auf unternehmerische Freiheit in eine angemessene Balance zu bringen, eine Gleichbehandlung von Verantwortlichen in der EU und Wettbewerbsgleichheit sicherzustellen sowie innovative Geschäftsmodelle im Zuge der digitalen Transformation zu unterstützen.[16] Bitkom wird das Gesetzgebungsverfahren zur BDSG-Änderung weiter beobachten.

Ergänzend wird auf die Notwendigkeit zusätzlicher Prüfungen beim Einsatz von Künstlicher Intelligenz gemäß der KI-Verordnung hingewiesen.

Fazit DSFA:

Werden mithilfe von KI rein automatisierte Entscheidungen getroffen bzw. vorbereitet oder erfolgt eine systematische und umfassende Bewertung persönlicher Aspekte natürlicher Personen, führt die Schwellwertanalyse in der Regel zu dem Ergebnis, dass eine Datenschutz-Folgenabschätzung gem. Art. 35 DS-GVO durchgeführt werden muss.

Besonders die Beurteilung der Risikolage kann in der Praxis herausfordernd sein, da Verantwortliche oftmals keine bzw. nicht hinreichend Transparenz zu Verarbeitungsdetails, verwendeten Algorithmen und involvierter Logik haben. Diese Herausforderung besteht ebenfalls, wenn KI-Anwendungen im Rahmen einer Auftragsverarbeitung genutzt werden. Hier sind Verantwortliche auf die entsprechenden Informationen der Hersteller angewiesen, um ihren datenschutzrechtlichen Verpflichtungen nachkommen zu können. Details zum Aspekt Transparenz und Informationspflichten s.o.

In der unternehmerischen Praxis zeigt sich, dass gerade bei KI-Projekten ein großer Zeitdruck besteht, da sich Unternehmen mit einer schnellen Realisierung Wettbewerbsvorteile sichern möchten. Verantwortlichen wird daher empfohlen, sich frühzeitig mit der Frage bzgl. Durchführung einer DSFA auseinanderzusetzen, zumal diese vor dem Start der Verarbeitung personenbezogener Daten erfolgen muss.
Da die Durchführung einer DSFA inklusive Erstellung eines DSFA-Berichts einer Vorbe-

[15] Vgl. § 38 Abs. 1 Satz 1 BDSG *»…in der Regel mindestens 20 Personen ständig mit der automatisierten Verarbeitung personenbezogener Daten beschäftigen«.*
[16] Details s. Bitkom Stellungnahme zum BDSG-Änderungsgesetz s. Bundesdatenschutzgesetz 2023 | Stellungnahme 2023 | Bitkom e. V.

reitung bedarf und zeitlich aufwendig ist, wird geraten, entsprechend Vorlaufzeit einzuplanen. Erfahrungswerte aus der unternehmerischen Praxis liegen im Bereich von ca. 3 bis 6 Monaten Dauer je Vorhaben und abhängig von der Komplexität im Einzelfall. Eine ggf. nötige Konsultation der Aufsichtsbehörde verlängert den Prozess.

Weiterführende Informationen zu DSFA in Verbindung mit KI sowie zu DSFA allgemein:

- Positionspapier der DSK zu empfohlenen technischen und organisatorischen Maßnahmen bei der Entwicklung und dem Betrieb von KI[17]
- Hambacher Erklärung der DSK zur Künstlichen Intelligenz[18]
- DSK-Kurzpapier Nr. 5 Datenschutz-Folgenabschätzung nach Art. 35 DS-GVO[19] mit allgemeinen Ausführungen zur DSFA-Thematik.
- Bitkom Leitfaden Risk Assessment & Datenschutz-Folgenabschätzung[20]

Berechtigungskonzept

Ein Berechtigungskonzept für KI ist ein wichtiger Baustein für den verantwortungsvollen und rechtskonformen Einsatz von KI-Anwendungen. Es sollte die Anforderungen und Besonderheiten von KI berücksichtigen und die Rollen, Rechte und Pflichten der beteiligten Akteure klar regeln.

Insbesondere folgende Fragen bieten u. a. eine Hilfestellung:

- Muss der Datenzugriff auf einen bestimmten Personenkreis, z. B. Administratoren, beschränkt werden?
- Welche Vertraulichkeit haben die im KI-System oder Modell verarbeiteten personenbezogenen Daten?
- Welche Rolle spielt eine Zugangskontrolle (z. B. Authentifizierungssystem)?

Löschkonzept

Die Verarbeitung personenbezogener Daten durch KI-Systeme und Modelle wirft spezifische Fragen auf, die im Rahmen eines Löschkonzepts berücksichtigt werden müssen. Zunächst muss geprüft werden, ob innerhalb einer bzw. eines LLM gelöscht werden kann.

KI-Systeme und Modelle nutzen oft komplexe und verteilte Speichersysteme, einschließlich Cloud-basierter Dienste. Diese Systeme können die Lokalisierung und Löschung von Daten erschweren, insbesondere wenn Daten über verschiedene

[17] Positionspapier der Konferenz der unabhängigen Datenschutzaufsichtsbehörden des Bundes und der Länder v. 06.11.2019 s.20191106_positionspapier_kuenstliche_intelligenz.pdf (datenschutzkonferenz-online.de).
[18] Entschließung der 97. Konferenz der unabhängigen Datenschutzaufsichtsbehörden des Bundes und der Länder Hambacher Schloss 3. April 2019, Hambacher Erklärung zur Künstlichen Intelligenz, Sieben datenschutzrechtliche Anforderungen, vgl. Ziff. 6; Link s. 20190405_hambacher_erklaerung.pdf (datenschutzkonferenz-online.de).
[19] s. DSK_KPNr_5_Datenschutz-Folgenabschätzung_Lizenzvermerk (datenschutzkonferenz-online.de).
[20] s. o. Fn. 1.

Standorte und Jurisdiktionen hinweg gespeichert werden. Ein effektives Löschkonzept muss daher die spezifischen Speicherstrukturen und Zugriffsmechanismen berücksichtigen, die in KI-Systemen und Modellen verwendet werden.

Das Recht auf Vergessenwerden (Art. 17 DS-GVO) ermöglicht es Einzelpersonen, die Löschung ihrer personenbezogenen Daten unter bestimmten Umständen zu fordern. Eine KI muss in der Lage sein, solche Anforderungen effizient und vollständig umzusetzen.

Automatisierte Löschmechanismen können eingerichtet werden, um Daten nach Ablauf ihrer Relevanz oder auf Anfrage automatisch zu löschen.

Es ist wichtig, den Prozess der Datenlöschung zu dokumentieren, um die Einhaltung von Datenschutzbestimmungen nachzuweisen. Dies ist wichtig, um bei Anfragen von Datenschutzbehörden oder betroffenen Personen Rechenschaft ablegen zu können.

Interne Richtlinien zur Nutzung von KI

Unternehmen, die sich mit der Nutzung von KI befassen, müssen zunächst vielfältige technische, kommerzielle und rechtliche Bewertungen und Festlegungen der internen und externen Anforderungen an die KI durchführen.[21] Dazu gehört auch die Erstellung einer unternehmensinternen Richtlinie zur Nutzung von (generativer) KI[22]. Datenschutzrechtliche Aspekte müssen bei der Entwicklung, Implementierung und Nutzung von KI-Systemen bzw. Modellen eine zentrale Rolle spielen. Die Richtlinien sollten sicherstellen, dass personenbezogene Daten rechtmäßig und zweckgebunden erhoben, verarbeitet und bei Dateneingaben und -ausgaben auf das für die Zwecke der Verarbeitung notwendige Maß beschränkt werden. Es sollten – soweit nötig – Maßnahmen ergriffen werden, um die Datensicherheit zu gewährleisten, wie beispielsweise die Anonymisierung oder Pseudonymisierung von Daten. Die Richtlinie sollte einen Prozess zur Erfüllung der gegebenenfalls bestehenden Rechte der betroffenen Personen, wie das Recht auf Information, Berichtigung und Löschung beinhalten. Zudem sollten die Richtlinien sicherstellen, dass die KI-Systeme und Modelle auch aus Sicht der Verarbeitung personenbezogener Daten transparent und erklärbar sind, um den Schutz der Privatsphäre und die Wahrung der Rechte der betroffenen Personen zu gewährleisten.

Wichtig ist es auch, von Anfang an eine Dokumentation von KI-Modellen und Systemen, Tools und Anwendungsfällen sicherzustellen, damit ausreichend Klarheit besteht und nicht zu einem späteren Zeitpunkt erst mühsam begonnen werden muss, den KI-Bestand im Unternehmen zu ermitteln und die jeweiligen Risiken im Einzelnen zu bewerten.

Zu berücksichtigen ist auch die Vielfalt der Technologien und Anwendungsfelder, die Schnelligkeit der Entwicklung und der notwendige Detaillierungsgrad. Es wird sich daher oft empfehlen, eine generelle und ggf. abstrakte Richtlinie (Policy) durch

[21] siehe insbesondere Bitkom Leitfaden Generative KI im Unternehmen, Kapitel 2.2 (https://www.bitkom.org/Bitkom/Publikationen/Generative-KI-im-Unternehmen)
[22] Bitkom Leitfaden Generative KI im Unternehmen, Kapitel 3.5.5 (https://www.bitkom.org/Bitkom/Publikationen/Generative-KI-im-Unternehmen)

detailliertere Vorgaben zu ergänzen, die einzelne Aspekte eingehender regeln. Die weiteren Regelungen (z. B. »Standards«) bieten dann die Flexibilität, einzelne Aspekte nicht nur spezifisch zu regeln, sondern auch die Regeln bei Bedarf zu aktualisieren, ohne stets die grundlegende Richtlinie neu fassen und publizieren zu müssen. Der Zuschnitt dieser detaillierten Regelungen kann nicht allgemeingültig empfohlen werden, sondern ergibt sich aus Art und Umfang der KI-Nutzung im Unternehmen. Werden z. B. eigene Modelle oder Anwendungen entwickelt, sind mehr Regelungen erforderlich, als wenn ein Unternehmen lediglich am Markt verfügbare Anwendungen einsetzt.

1. **Die Richtlinie sollte insbesondere Regelungen zu folgenden Punkten enthalten:**
   - Anwendungsbereich, Definitionen
   - KI-Prinzipien oder Leitbild des Unternehmens bzgl. der Entwicklung und Nutzung von KI
   - Grundsätze der Nutzung
   - Dokumentations- und Transparenzanforderungen
   - Zuständigkeiten, Genehmigungserfordernisse, Aufklärungs- und Sanktionsmechanismen bei Missachtung der Richtlinien
   - Bezüge zu anderen internen Richtlinien und zu externen regulatorischen Vorgaben
   - Verweis auf weitere Regelungen

2. **Die weiteren Regelungen (z. B. »Standards«) können beliebige weitere Felder abdecken. Beispiele sind:**
   - Details der Dokumentation von KI-Systemen und Modellen
   - Verfahren zur Sicherstellung von Fairness, Transparenz und Interpretierbarkeit, Zuverlässigkeit, Vermeidung von Voreingenommenheit (»bias«) Compliance und Regulierung: Sicherstellung der Einhaltung geltender Gesetze wie der KI-Verordnung und weiterer Regulierungen, Terms of Use und branchenspezifischer Standards im Kontext der Entwicklung/Nutzung von KI-Systemen und Modellen
   - Vorgaben für die Beschaffung/den Einkauf von Anwendungen mit KI-Komponenten
   - Handlungsanforderungen/-empfehlungen für die Angestellten
   - Einzelheiten des Risikomanagementsystems und des Qualitätsmanagementsystems IP Recht (u. a. urheberrechtliche Aspekte bei Nutzung von Daten zu Training/Entwicklung/Adaptierung von KI-Modellen und Systemen)
   - Überwachung von Modellen im Einsatz, die mit hohem Risiko behaftet sind
   - Nutzung von personenbezogenen und nicht personenbezogenen Daten für die Entwicklung, das Testen und den Einsatz von Modellen
   - Sicherheit und Informationssicherheit von KI-Anwendungen

- Regelungen für einzelne Anwendungsfelder (»Use Cases«) – etwa KI-Einsatz im Bewerbungsverfahren, in der Personaladministration
- Verhältnis zu benachbarten Themenfeldern, z. B. Datenschutz
- evtl. ergänzende oder abweichende Regeln für unterschiedliche Länder.

Dabei sollte auch berücksichtigt werden, dass die Mitarbeiter in geeigneter Weise geschult werden, um die jeweils relevanten Regelungen kennen und anwenden zu können.

3 Nutzung von KI

Ziel dieses Kapitels ist es, anhand eines praktischen Beispiels zu zeigen, wie KI in einem Unternehmen eingesetzt werden kann. Im Zuge des Reviewprozesses des Leitfadens soll eine Umfrage unter den Mitgliedsunternehmen durchgeführt werden, um die Nutzung und Adaption von KI und verschiedenen KI-Modellen und Systemen zu evaluieren. Ein weiterer Schwerpunkt der Umfrage soll die Identifikation der größten Herausforderungen bei der KI-Nutzung sein. Schließlich sollen auch die relevantesten Anwendungsbereiche der KI-Nutzung zusammengefasst werden. Die Ergebnisse dieser Umfrage sollen Unternehmen helfen, ihre KI-Strategien zu verbessern und einen Überblick über den aktuellen Stand der KI-Nutzung zu geben.

Während der von der DSK veröffentlichte KI-Datenschutz-Leitfaden anhand einfacher Beispiele weitere mögliche KI-Anwendungen und Einsatzkonzeptionen beleuchtet[23], konzentriert sich dieser Leitfaden im Folgenden auf eine detaillierte, datenschutzkonforme Darstellung des Einsatzes von KI in der Automobilindustrie.

Anwendungsbeispiel: KI in der Automobilindustrie

KI kommt in der Automobilindustrie etwa zum Einsatz, wenn es um das Training von hoch- und voll automatisierten (§§ 1a ff. StVG) sowie autonomen Fahrfunktionen (§§ 1d ff. StVG) geht. Für das Training sind vielfältige Daten aus Erprobungsfahrten erforderlich, mittels derer etwa die Objekterkennung bei sicherheitsrelevanten Verkehrssituationen verbessert wird. Dafür ist es notwendig, möglichst viele Verkehrssituationen und -szenarien sowie verschiedene Objekte und Straßenverkehrsteilnehmer sowie Passanten zu erfassen.

Bei der Datenerhebung und -verarbeitung sind die in Art. 5 DS-GVO enthaltenen Datenschutzgrundsätze zu beachten. Aufgrund des Grundsatzes der

[23] https://www.datenschutzkonferenz-online.de/media/oh/20240506_DSK_Orientierungshilfe_KI_und_Datenschutz.pdf

Datenminimierung sind die verarbeiteten Daten auf das benötigte Minimum zu reduzieren. Das benötigte Minimum an Daten kann dabei auch nicht verpixelte Rohbilder der Umgebung des Fahrzeuges umfassen, sofern diese für die Entwicklung und Weiterentwicklung von Technologie erforderlich ist. Dies ist insbesondere dann der Fall, wenn anhand der Daten Algorithmen trainiert werden sollen, die Verkehrssituationen möglichst genau entsprechend der Realität erkennen können sollen.

In Fällen, in denen die Datenverarbeitung auf Grundlage einer Interessenabwägung gem. Art. 6 (1) f) DS-GVO durchgeführt wird, ist zu beachten, dass gerade intelligente Fahrerassistenz- und generell autonome Fahrsysteme in aller Regel der Steigerung der allgemeinen Verkehrssicherheit und damit im allgemeinen öffentlichen Interesse sind. Vor diesem Hintergrund können zu diesen Zwecken allgemein umfassend, hauptsächlich aber auch Daten von datenschutzrechtlich besonders schützenswerten Bevölkerungsgruppen (z. B. Kinder oder in der Mobilität eingeschränkte Personen) verarbeitet werden, um die Erhöhung der Verkehrssicherheit in deren und dem allgemeinen Interesse zu ermöglichen.

Die Trainings-, Validierungs- und Testdatensätze müssen gem. Art. 10 (3)3 KI-VO so weit wie möglich »fehlerfrei und vollständig« sein. Erforderlich ist daher nach KI-VO, Daten zu erheben und zu erfassen, die eine sichere und nachvollziehbare Nutzung von KI-Systemen und Modellen absichern. Auch wenn die KI-VO die DS-GVO grundsätzlich unberührt lässt, ist in diesem Fall Art. 10 (3) KI-VO lex specialis und kann begründen, dass etwa die umfassende Verarbeitung von Validierungsdaten mit umfasst wird. Auch mit Nichtannahme einer datenschutzrechtlichen Spezialregelung, stellt die KI-VO aber klar, dass für einen grundrechtskonformen Einsatz von KI-Systemen und Modellen die umfassende Nutzung von Trainings-, Validierungs- und Testdatensätzen technisch erforderlich ist. Dies ist auch bei der datenschutzrechtlichen Bewertung zwingend zu berücksichtigen. Klarstellend kann zumindest durch Art. 10 (3) KI-VO mithin davon ausgegangen werden, dass die Verarbeitung derartiger Daten mit dem Grundsatz der Datenminimierung nach Art. 5 DS-GVO vereinbar ist.

Um eine zuverlässige Objektdetektierung und -klassifizierung zu gewährleisten, müssen nicht nur vielfältige Objekte erfasst werden, sondern auch verschiedene Personen, wobei etwa Größe, Gewicht, Alter, Hautfarbe etc. divergieren können. Diese Pflicht folgt aus Art. 10 (3) KI-VO. Je nachdem wird oftmals angenommen, es handele sich in bestimmten Situationen um besondere Kategorien von personenbezogenen Daten i.S. des Art. 9 (1) DS-GVO, etwa wenn Personen im Rollstuhl erfasst werden. Häufig wird das gerade bei Datenerfassungen und -verarbeitungen zum Zweck der Entwicklung und Weiterentwicklung von sicherheitsrelevanten Fahrerassistenz- oder autonomen Fahrsystemen nicht der Fall sein, da die Datenverarbeitung nicht auf Daten gem. Art. 9 DS-GVO ausgelegt ist.[24] Aber selbst wenn man zu der datenschutzrechtlichen Einschätzung kommt, dass es sich um Daten i.S.v. Art. 9 (1) DS-GVO handelt, so müssen diese nicht nur aufgrund von Art. 10 (3) KI-VO verarbeitet werden, um vollständige Trainings-, Validierungs- und Testdatensätze zu ermöglichen, sondern dürfen gem. Art. 10 (5) KI-VO i.V.m. ErwG 44c KI-VO verarbeitet werden, da dies ein besonderes öffentliches Interesse nach Art. 9 (2) g) DS-GVO darstellt.

[24] DSK, Positionspapier zur audiovisuellen Umgebungserfassung im Rahmen von Entwicklungsfahrten, 2023, S. 7; EDSA-Leitlinien 3/2019, Verarbeitung personenbezogener Daten durch Videogeräte, Rn. 62 ff.

Training eigener KI

Sofern man den Einsatz von KI-Techniken in seinem Unternehmen etablieren möchte, kann auch die Entwicklung eines eigenen KI-Systems oder Modells in Betracht gezogen werden. Damit ein solches System oder Modell jedoch effektiv arbeitet und letztlich positive Effekte erzielt werden können, muss es intensiv trainiert werden. Dieser Leitfaden verzichtet an dieser Stelle auf detaillierte Ausführungen, da die bisherigen Erfahrungen mit dem Training eigener KI noch nicht hinreichend ausgereift sind, aber auch die letztendliche Relevanz des Einsatzes eigener KI noch unbekannt ist. Ein Blick in die Praxis zeigt, dass bislang vor allem die Nutzung bereits etablierter KI-Modelle im Vordergrund steht und die Vielzahl unternehmenseigener KI-Systeme nicht über die Entwicklungsphase hinaus besteht.

Aufgrund dessen finden sich in Checkliste ab Seite 42 nur grundlegende Richtlinien zum Training der eigenen KI und kein ausformulierter Abschnitt. Eine detaillierte Auseinandersetzung zum Thema »Training eigener KI«, wird im Laufe des Reviewprozesses des Leitfadens neu evaluiert und ggf. in Form eines Ergänzungsleitfadens genauer eruiert.

4 Anlage Checkliste

Wie in diesem Leitfaden erörtert, bietet die Nutzung von KI im Unternehmen Chancen zur Prozessoptimierung und Erschließung neuer Geschäftsmodelle. Sie bringt jedoch auch rechtliche und ethische Herausforderungen mit sich, insbesondere im Datenschutz. Diese Checkliste unterstützt Entscheidungsträger, Projektmanager und technische Fachkräfte bei der Einführung und Nutzung von KI-Technologien. Sie gewährt eine strukturierte Übersicht der wesentlichen Schritte und Maßnahmen zur Einhaltung der DS-GVO und Minimierung von Risiken für die Rechte der betroffenen Personen.

Die Checkliste ist in zwei Bereiche unterteilt:

1. **Training eigener KI-Modelle:**
   Dieser Teil der Checkliste behandelt die Schritte und Maßnahmen, die bei der Entwicklung und dem Training von KI-Modellen notwendig sind. Dazu gehören beispielsweise die Auswahl und Dokumentation des Modells, die Klassifizierung und Verarbeitung von Trainingsdaten unter Berücksichtigung datenschutzrechtlicher Anforderungen.

2. **Nutzung von KI-Systemen:**
   Dieser Teil der Checkliste fokussiert sich auf die Nutzung bestehender KI-Systeme und die Einhaltung datenschutzrechtlicher Vorgaben im operativen Einsatz. Hier werden Aspekte wie die Risikobewertung, die Dokumentationspflichten und die Sensibilisierung der Mitarbeiter behandelt.

Hinweis:
Die Checkliste ist nicht abschließend und erhebt keinen Anspruch auf Vollständigkeit.

Je nach Anwendungsfall empfiehlt es sich, die Checkliste ggf. um Aspekte außerhalb des Datenschutzrechts sowie spezielle betriebliche Vorgaben zu ergänzen.

1. Training eigener KI-Modelle und Systeme

- ☐ Festlegung und Beschreibung der KI-Technologie und Anlage entsprechender Dokumentation
- ☐ Klassifizierung der Trainingsdaten, Beurteilung der Personenbeziehbarkeit und Bewertung, ob das Training unter die DS-GVO fällt
- ☐ Ggf. Einbindung des Datenschutzbeauftragten
- ☐ Berücksichtigung von Privacy by Design/Privacy by Default, insbesondere durch Beurteilung der Erforderlichkeit für die Verwendung personenbezogener Daten unter dem Aspekt, ob die Verwendung
  - ☐ anonymer oder anonymisierter Daten oder zumindest
  - ☐ pseudonymer Daten möglich ist; sowie
  - ☐ entsprechende Bereinigung der Trainingsdaten.
- ☐ Beurteilung der Rechtsgrundlage für die Verarbeitung personenbezogener Daten in Hinblick auf die verfolgten Verarbeitungszwecke (Art. 6 DS-GVO)
- ☐ Beurteilung der besonderen Voraussetzungen im Fall von besonderen Kategorien personenbezogener Daten (Art. 9 DS-GVO)
- ☐ Prüfung der Datenqualität (Datenrichtigkeit) auch in Hinblick auf Voreingenommenheit (Bias) und Diskriminierungspotential
- ☐ Beurteilung der Risiken für die Rechte und Freiheiten der Betroffenen/ Erstellung eines Risikomodells
- ☐ Beurteilung besonderer Anforderungen an den besonderen Schutzbedarf von Kindern und Jugendlichen
- ☐ Durchführung einer Datenschutzfolgenabschätzung (Art. 35 DS-GVO) (zumindest bei Bestehen eines hohen Betroffenenrisikos nach Durchführung einer Schwellwertanalyse oder entsprechender Vorgaben seitens Aufsichtsbehörden)
- ☐ Bei Drittstaatentransfer: Berücksichtigung und Dokumentation der Anforderungen aus Kapitel V der DS-GVO
- ☐ Festlegung und Sicherstellung angemessener technischer und organisatorischer Maßnahmen (einschließlich Zugriffs-/Berechtigungs- und Löschkonzept) (Art. 32 DS-GVO)

- ☐ Aufnahme in das Verzeichnis der Verarbeitungstätigkeiten und Dokumentation zur Erfüllung der Rechenschaftspflicht (Art. 5 (2) DS-GVO) unter Angabe insbesondere von
  - ☐ Datenquelle/Herkunft der Daten
    Bei flüchtigen Daten, insbesondere von Webseiten, Speicherung einer vollständigen Kopie der Dateninhalte/Webseite
  - ☐ Rechtsgrundlage (Art. 6 DS-GVO)
  - ☐ Erfüllung der Transparenzanforderungen (Art. 12 ff. DS-GVO)
- ☐ Bei Inanspruchnahme von (externen) Dienstleistern:
  - ☐ Sorgfältige Auswahl unter Berücksichtigung der besonderen Eignung des Dienstleisters
  - ☐ Festlegung des Verantwortungsbereichs des Dienstleisters einschließlich der Verarbeitungszwecke
  - ☐ Abschluss notwendiger Verträge unter Vorgabe entsprechender Leistungsanforderungen/Weisungen (insbesondere im Fall der Auftragsverarbeitung nach Art. 28 DS-GVO)
- ☐ Bewertung und Validierung des trainierten KI-Modells bzw. Systems in Hinblick auf ggf. vorhandene Personenbezüge
- ☐ Implementierung notwendiger Prozesse zur Erfüllung von Betroffenenrechten wie
  - ☐ Auskunftsersuchen (Art. 15 DS-GVO)
  - ☐ Berichtigungs-/Löschersuchen (Art. 16, 17 DS-GVO)
  - ☐ Anfragen zur Einschränkung der Verarbeitung (Art. 18 DS-GVO)
  - ☐ Anfragen zur Datenportabilität (Art. 20 DS-GVO)
  - ☐ Widersprüchen (Art. 21 DS-GVO)
  - ☐ Widerrufe von zuvor erteilten Einwilligungen insb. bei automatisierten Entscheidungen i.S.v. Art. 22 DS-GVO

2. Nutzung von KI-Systemen und Modellen

- ☐ Festlegung und Beschreibung der KI-Technologie und Anlage entsprechender Dokumentation unter Berücksichtigung
  - ☐ der Inhaberschaft und Verantwortlichkeit hinsichtlich des KI-Modells/Systems (eigenes vs. fremdes KI-Modell/System)
  - ☐ der verwendeten KI-Umgebung/KI-Anwendung
- ☐ Einbindung des Datenschutzbeauftragten
- ☐ Bei Inanspruchnahme eines (externen) Dienstleisters:
  - ☐ Sorgfältige Auswahl unter Berücksichtigung der besonderen Eignung des Dienstleisters

- ☐ Festlegung des Verantwortungsbereichs des Dienstleisters einschließlich der Verarbeitungszwecke
- ☐ Abschluss notwendiger Verträge unter Vorgabe entsprechender Leistungsanforderungen/Weisungen (insbesondere im Fall der Auftragsverarbeitung nach Art. 28 DS-GVO)

☐ Beurteilung der Risiken für die Rechte und Freiheiten der Betroffenen/ Erstellung eines Risikomodells

☐ Durchführung einer Datenschutzfolgenabschätzung (Art. 35 DS-GVO) (zumindest bei Bestehen eines hohen Betroffenenrisikos nach Durchführung einer Schwellwertanalyse oder entsprechender Vorgaben seitens Aufsichtsbehörden)

☐ Aufnahme in das Verzeichnis der Verarbeitungstätigkeiten und Dokumentation zur Erfüllung der Rechenschaftspflicht (Art. 5 (2) DS-GVO) unter Angabe insbesondere von

- ☐ Rechtsgrundlage (Art. 6 DS-GVO)
- ☐ Erfüllung der Transparenzanforderungen (Art. 12 ff. DS-GVO)

☐ Entwicklung und Implementierung verbindlicher Verhaltensregeln hinsichtlich der Eingabe personenbezogener Daten (Prompt), insbesondere unter Berücksichtigung besonderer Kategorien personenbezogener Daten (Art. 9 DS-GVO), und der Verwendung der Ergebnisse (Output)

☐ Angebote zur Sensibilisierung der Beschäftigten zum Umgang mit KI

☐ Berücksichtigung von Privacy by Design/Privacy by Default, insbesondere durch Beurteilung der Erforderlichkeit der Protokollierung der KI-Nutzung in Form von

- ☐ anonymisierter Protokolldaten oder zumindest
- ☐ pseudonymer Protokolldaten

☐ Implementierung notwendiger Prozesse zur Erfüllung von Betroffenenrechten wie

- ☐ Auskunftsersuchen (Art. 15 DS-GVO)
- ☐ Berichtigungs-/Löschersuchen (Art. 16, 17 DS-GVO)
- ☐ Anfragen zur Einschränkung der Verarbeitung (Art. 18 DS-GVO)
- ☐ Anfragen zur Datenportabilität (Art. 20 DS-GVO)
- ☐ Widersprüchen (Art. 21 DS-GVO)
- ☐ Widerrufe von zuvor erteilten Einwilligungen insb. bei automatisierten Entscheidungen i.S.v. Art. 22 DS-GVO

Bitkom vertritt mehr als 2.200 Mitgliedsunternehmen aus der digitalen Wirtschaft. Sie generieren in Deutschland gut 200 Milliarden Euro Umsatz mit digitalen Technologien und Lösungen und beschäftigen mehr als 2 Millionen Menschen. Zu den Mitgliedern zählen mehr als 1.000 Mittelständler, über 500 Startups und nahezu alle Global Player. Sie bieten Software, IT-Services, Telekommunikations- oder Internetdienste an, stellen Geräte und Bauteile her, sind im Bereich der digitalen Medien tätig, kreieren Content, bieten Plattformen an oder sind in anderer Weise Teil der digitalen Wirtschaft. 82 Prozent der im Bitkom engagierten Unternehmen haben ihren Hauptsitz in Deutschland, weitere 8 Prozent kommen aus dem restlichen Europa und 7 Prozent aus den USA. 3 Prozent stammen aus anderen Regionen der Welt. Bitkom fördert und treibt die digitale Transformation der deutschen Wirtschaft und setzt sich für eine breite gesellschaftliche Teilhabe an den digitalen Entwicklungen ein. Ziel ist es, Deutschland zu einem leistungsfähigen und souveränen Digitalstandort zu machen.

Herausgeber
Bitkom e.V.
Albrechtstr. 10 | 10117 Berlin

Ansprechpartner/in
Paul Breitbarth | Wissenschaftlicher Mitarbeiter
T 030 27576-259 | p.breitbarth@bitkom.org

Isabelle Stroot | Referentin Datenschutz
T 030 27576-228 | i.stroot@bitkom.org

Felix Kuhlenkamp | Referent für Sicherheitspolitik
T 030 27576-279 | f.kuhlenkamp@bitkom.org

Verantwortliches Bitkom-Gremium
AK Datenschutz

Titelbild
© Irina Vodneva – istockphoto.com.