bitkom

Generative KI im Unternehmen

Rechtliche Fragen zum Einsatz generativer Künstlicher Intelligenz im Unternehmen

Herausgeber
Bitkom e. V.
Albrechtstraße 10
10117 Berlin
T 030 27576-0
bitkom@bitkom.org
www.bitkom.org

Ansprechpartner
Dr. iur. Pablo Schumacher | Referent für Digital Content & Licensing
T 030 27576-289 | p.schumacher@bitkom.org

Verantwortliches Bitkom-Gremium
AK Recht im Unternehmen & Compliance

Layout
Katrin Krause | Bitkom

Titelbild
© Nicolas Arnold – unsplash.com

Inhalt

Geleitwort

Der vorliegende Leitfaden entstand in einer Taskforce, bestehend aus Expertinnen und Experten aus folgenden Arbeitskreisen des Bitkom:

- ↗Personal & Arbeitsrecht
- ↗Intellectual Property
- ↗Artificial Intelligence
- ↗Legal Tech
- ↗Sicherheitspolitik
- ↗Arbeit 4.0
- ↗Datenpolitik & Datenräume
- ↗Datenschutz
- ↗Vertrags- & Rechtsgestaltung
- ↗Recht im Unternehmen & Compliance
- ↗Arbeitssicherheit

Besonderer Dank gilt folgenden Personen, die mit ihrer Expertise und wertvollen praktischen Erfahrung die Publikation erstellt haben:

- Martin Agethen, Deutsche Telekom AG
- Prof. Dr. Christoph Bauer, ePrivacy GmbH
- Thomas Bauer, Atruvia AG
- Stefanie Bauer, ePrivacy GmbH
- Dr. Malte Baumann, Nordemann Czychowski & Partner Rechtsanwältinnen und Rechtsanwälte mbB
- Dr. Nadja Christe, Bayer AG
- Dr. Christoph Cordes, YPOG Partnerschaft von Rechtsanwälten und Steuerberatern mbB Schnittker + Partner
- Jonas von Dall'Armi, Giesecke+Devrient GmbH
- Dr. Benedikt Flöter, YPOG Partnerschaft von Rechtsanwälten und Steuerberatern mbB Schnittker + Partner
- Dr. Anne Förster, Taylor Wessing Partnerschaftsgesellschaft von Rechtsanwälten und Steuerberatern mbB
- Mareike Gehrmann, Taylor Wessing Partnerschaftsgesellschaft von Rechtsanwälten und Steuerberatern mbB
- Dr. Inka Knappertsbusch, CMS Hasche Sigle Partnerschaft von Rechtsanwälten und Steuerberatern mbB
- Dr. Benedikt Kohn, Taylor Wessing Partnerschaftsgesellschaft von Rechtsanwälten und Steuerberatern mbB
- Tobias Kutscheidt, Deutsche Telekom AG
- Dilan Mienert, GÖRG Partnerschaft von Rechtsanwälten mbB
- Dr. Dominik Rabe, Rewe Zentralfinanz e.G.
- Alexandra Schmidt, Deutsche Telekom AG
- René Schneider, DataCo GmbH (DataGuard)
- Maike Scholz, Deutsche Telekom AG
- Marie Slowioczek, ELEMENT Insurance AG
- Dr. Gerald Spiegel, EnBW Energie Baden-Württemberg AG
- Ilka-Maria Sühling, Deutsche Telekom AG
- Dr. Maximilian Vonthien, LLM, CMS Hasche Sigle Partnerschaft von Rechtsanwälten und Steuerberatern mbB
- Dr. Nicolai Wiegand, Taylor Wessing Partnerschaftsgesellschaft von Rechtsanwälten und Steuerberatern mbB
- Dr. Marko Wolf, Robert Bosch GmbH

1 Einleitung

1.1 Ziel des Leitfadens

Der vorliegende Praxisleitfaden »Generative KI im Unternehmen – rechtliche Fragen zum Einsatz generativer Künstlicher Intelligenz im Unternehmen«, setzt sich mit den rechtlichen Herausforderungen und Aspekten bei der geschäftlichen Nutzung von generativer Künstlicher Intelligenz (KI) auseinander. Bevor wir tiefer in das Thema eintauchen, möchten wir nachfolgend klären, was genau unter »generativer KI« zu verstehen ist.

Generative KI ist ein Unterbereich der Künstlichen Intelligenz, der darauf ausgerichtet ist, neue Daten oder Inhalte zu generieren. Diese Technologie findet Anwendung in einer Vielzahl von Bereichen, einschließlich der Generierung von Texten, der Erstellung künstlicher Bilder oder Videos, automatischer Übersetzungen, und sogar in der Musikkomposition. Prominente Beispiele hierfür sind GANs (Generative Adversarial Networks) zur Erzeugung realistischer Bilder oder Textmodelle wie GPT-4 für die automatische Textgenerierung.

Bitte beachten Sie, dass sich dieser Leitfaden grundsätzlich auf generative KI konzentriert und nicht die allgemeinen rechtlichen Aspekte der Künstlichen Intelligenz behandelt. Die mit der Nutzung generativer KI verbundenen spezifischen Herausforderungen und Risiken sind komplex und verdienen eine eigenständige Betrachtung. Dennoch sind die Grenzen zu anderen KI-Technologien stellenweise fließend und insbesondere machen rechtliche Fragestellungen nicht an den Grenzen der Technologie halt. Deshalb lässt es sich nicht vermeiden, rechtliche Fragen aufzuwerfen und zu beantworten, die nicht (nur) primär generative KI, sondern KI generell betreffen.
Die Anwendung von generativer KI bringt eine Reihe von rechtlichen Fragestellungen mit sich, die für Unternehmen und Nutzende von Bedeutung sind. Dies umfasst sowohl die Daten, die in die KI-Modelle eingegeben werden, als auch die Verwendung der von der KI generierten Ergebnisse. Relevante Themenbereiche sind dabei unter anderem geistiges Eigentum, insbesondere Urheberrecht, Datenschutz, Haftungsfragen und Arbeitsrecht.

Der vorliegende Praxisleitfaden hat das Ziel, Ihnen eine klare Übersicht über die rechtlichen Aspekte im Kontext generativer KI zu bieten. Zudem werden wir praktische Handlungsempfehlungen darlegen, wie mit diesen rechtlichen Herausforderungen umzugehen ist. Ein Exkurs zu ethischen Aspekten bei der Beschaffung und dem Einsatz von generativer KI rundet den Leitfaden ab.

Da der Bereich der generativen KI äußerst dynamisch ist, ist es unser Anspruch, diesen Leitfaden regelmäßig zu aktualisieren und an neue Erkenntnisse sowie Entwicklungen anzupassen. Der Stand des KI-Leitfadens ist Februar 2024. Insbesondere in Bezug auf die EU-KI-Verordnung (AI Act) lag zu diesem Zeitpunkt lediglich ein vom Rat der Europäischen Union beschlossener Text, jedoch noch kein finaler und amtlich veröffentlichter Text der Verordnung vor. Wir behalten uns diesbezügliche Änderungen ausdrücklich vor.

Der Leitfaden richtet sich insbesondere an Unternehmen als Anwender von generativer KI.
Der Fokus liegt also auf einer Anwenderperspektive in der Einsatzphase von generativer KI.

1.2 Definition: Generative KI

Künstliche Intelligenz ist ein Forschungsgebiet der Informatik. In dem Sinne existiert auch (noch) keine abschließende Definition von KI. Künstliche Intelligenz lässt sich dennoch 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. Der AI Act lehnt sich an die OECD-Definition an und definiert ein KI-System (frei übersetzt aus dem englischen Text) als ein maschinengestütztes System, das so konzipiert ist, dass es mit unterschiedlichem Grad an Autonomie arbeiten und nach seiner Einführung eine Anpassungsfähigkeit aufweisen kann, und das für explizite oder implizite Ziele aus den Eingaben, die es erhält, ableitet, wie es Ergebnisse wie Vorhersagen, Inhalte, Empfehlungen oder Entscheidungen erzeugen kann, die physische oder virtuelle Umgebungen beeinflussen können.[1]

Generative KI ist ein Teilbereich der KI, der darauf ausgerichtet ist, neue Inhalte oder Daten zu schaffen. Diese Art von KI unterscheidet sich von anderen Formen wie der »diskriminativen KI«, die hauptsächlich darauf trainiert ist, Unterschiede zwischen verschiedenen Arten von Daten zu erkennen. Ein Einsatzbereich ist zum Beispiel, Spam-E-Mails von legitimen E-Mails zu unterscheiden. Während diskriminative KI für Klassifizierungsaufgaben eingesetzt wird, ist generative KI darauf spezialisiert, originelle Daten zu erzeugen, sei es Text, Bilder oder sogar Musik.

Generative KI kann als ein spezialisierter Zweig innerhalb der General Purpose AI (GPAI) betrachtet werden. Während GPAI auf vielfältige Weise eingesetzt werden kann und oft als »Allzweck-KI« bezeichnet wird, fokussiert sich generative KI auf die Schaffung neuer Inhalte. Basismodelle sind oft die Grundlage für beide Typen und können entweder für generative oder diskriminative Aufgaben spezialisiert werden, je nachdem, wie sie trainiert wurden.

Schließlich ist es wichtig, zwischen KI-Technologie und KI-Systemen zu unterscheiden. KI-Technologie bezieht sich auf die Algorithmen und Modelle (z. B. Large Language Models), die die Grundlage für KI-Anwendungen (z. B. ChatGPT) bilden. Ein KI-System hingegen ist eine vollständige Lösung, die KI-Technologie enthält, aber auch andere Elemente wie Benutzeroberflächen, Datenmanagement-Tools und Anwendungslogik, also zusätzlich noch eine Applikationskomponente umfasst. Ein KI-System kann beispielsweise eine generative KI-Technologie nutzen, um personalisierte Inhalte für eine Marketingkampagne zu erzeugen, aber es würde auch Tools für die Verwaltung dieser Kampagne und die Analyse der Ergebnisse enthalten.

1 Die offizielle englische Definition lautet: »An AI system is a machine-based system designed to operate with varying levels of autonomy and that may exhibit adaptiveness after deployment and that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments.«

1.3 Technische Grundlagen generativer KI

Die Grundlage für jedes KI-System ist die Phase des Anlernens mit Daten, die als »Training« bezeichnet wird. In dieser Phase wird ein KI-Modell mit einer großen Menge an Daten gefüttert, um bestimmte Muster oder Beziehungen in den Daten zu erkennen. Die Qualität dieser Datenbasis hat einen entscheidenden Einfluss auf die Qualität der Ergebnisse, die das KI-System später erzeugt. Obwohl sich dieser Leitfaden auf die Nutzungsphase konzentriert, ist es wichtig zu beachten, dass KI-Systeme auch in der Lage sein können, aus den Daten der Nutzungsphase zu lernen, abhängig davon, ob sie dynamisch oder statisch eingestellt sind. Im letzten Fall lernen sie nichts mehr durch die Nutzung dazu. Der Speicherort dieser Datenbasis sowie der Kreis der Zugriffsberechtigten können rechtliche Implikationen haben, die in späteren Abschnitten dieses Leitfadens behandelt werden.

In der Nutzungsphase benötigt ein generatives KI-System Eingabeinformationen, sogenannte Prompts, um spezifische Aufgaben zu erfüllen. Diese Prompts können in Form von Textanweisungen erscheinen, aber auch durch Bilder (dann »multimodal«), Texte oder Audiodaten ergänzt werden. Als Ergebnis generiert das System entsprechende Texte, Bilder oder Audiodaten.
KI-Systeme haben die Fähigkeit, in Sitzungen iterativ vorzugehen. Das bedeutet, dass sie in der Lage sind, auf der Grundlage von Rückmeldungen oder zusätzlichen Eingaben ihre Ergebnisse kontinuierlich zu verbessern und zu verfeinern.

Der Begriff der »Explainable AI« (XAI) beschreibt den Ansatz, KI-Entscheidungen nachvollziehbar und erklärbar zu gestalten. Dies ist besonders im Kontext des AI Acts für Hochrisikosysteme von Bedeutung. Allerdings stellt XAI für generative KI generell eine besondere Herausforderung dar, da generative KI nicht nach einem festgelegten Algorithmus arbeitet, sondern auf statistischen Methoden basiert. Oft ist es deshalb schwierig, genau zu erläutern, warum und wie ein bestimmtes Ergebnis aus den Eingabeinformationen entstanden ist. Die mehrmalige Eingabe desselben Prompts kann daher zu verschiedenen Ergebnissen, dem sog. Output, führen.

1.4 Anwendungsbeispiele für generative KI in Unternehmen

Generative KI bietet eine Vielzahl von Anwendungsmöglichkeiten, die von einfachen Anwendungsfällen bis hin zu komplexeren reicht, bei denen die Transparenz des Prozesses, die Erklärbarkeit des Outputs sowie Zukunftssicherheit eine Rolle spielen können. Beispiele sind hier nicht nur die Generierung von (Software-)Codes oder Algorithmen, sondern auch die Automatisierung von Prozessen oder die Generierung kreativer Inhalte.

1.4.1 Textgenerierung

- Content-Erstellung: Automatische Erstellung von Artikeln, Blogposts und Social-Media-Inhalten.
- Automatische Übersetzungen: Schnelle und präzise Übersetzungen von Texten in verschiedene Sprachen.
- E-Mail-Automatisierung: Generierung von personalisierten E-Mails für Marketingkampagnen.
- Chatbots: Kundenservice-Chatbots, die natürliche Sprache verstehen und generieren können.
- SEO-Optimierung: Automatische Erstellung von SEO-optimierten Texten.
- Finanzberichterstattung: Automatische Generierung von Finanzberichten- und Analysen.
- Vertragsanalyse: Automatische Zusammenfassungen und Analysen von rechtlichen Dokumenten.
- FAQs: Automatische Erstellung und Aktualisierung von häufig gestellten Fragen.
- Wissenschaftliche Forschung: Literaturrecherche und Generierung von Forschungsanträgen.
- E-Learning: Erstellung von Quizfragen und Lernmaterialien.
- Verständliche Formulierung von Texten, bspw. mit dem Prompt: »Formuliere den Text so um, dass ihn ein 12-Jähriger versteht.«
- Einheitliche Formulierung von Texten nach Unternehmensvorgaben, bspw. mit dem Prompt: »Erzeuge aus dem mitgelieferten Text einen Text, welcher unseren Unternehmensvorgaben entspricht.«
- Prüfung von Eingangsmails und Ergänzung von Lösungsvorschlägen für die nachgelagerte Weiterbearbeitung durch Mitarbeitende (die vollautomatische Beantwortung durch KI ist im Bankenumfeld kritisch und wahrscheinlich ihn den meisten Fällen nicht zulässig)
- Grundsätzliche Muster bei der Textgenerierung:
  - Frage (Prompt) an Modell senden. Modell antwortet auf Basis dessen, was es gelernt hat (Problem z. B. bei bestimmten Anwendungen hier: Quelle des Wissens kann aufgrund der Menge nicht einfach genannt werden.)

- Frage (Prompt) an Modell senden und auszuwertendes Know-how bzw. Kontext mitliefern. Das Modell ist dabei für das Verstehen der natürlichsprachigen Anfrage und für die Formulierung einer Antwort in natürlicher Sprache zuständig. Vorteil: Die Wissensquelle ist bekannt. Nachteil: Die Abfragekosten steigen mit Größe des mitgelieferten Kontexts.

1.4.2 Bildgenerierung

- Design-Prototypen: Schnelle Erstellung von Design-Entwürfen für Websites oder Produkte.
- Bildbearbeitung: Automatische Retusche oder Anpassung von Bildern.
- Kunst: Erstellung künstlerischer Werke für digitale oder physische Medien.
- Architektur: Generierung von architektonischen Entwürfen und Modellen.
- Spieleentwicklung: Erstellung von Charakteren, Szenarien und Texturen.
- Retail: Automatische Generierung von Produktbildern in verschiedenen Umgebungen.
- Augmented Reality: Erstellung von AR-Inhalten.
- 3D-Druck: Entwurf von 3D-Modellen für den Druck.
- Landkarten und Geodaten: Generierung topografischer oder thematischer Karten.

1.4.3 Audiogenerierung (Sprache und Musik)

- Sprachassistenten: Generierung natürlicher Sprachausgabe für virtuelle Assistenten.
- Musikproduktion: Automatische Komposition von Musikstücken.
- Hörbucherstellung: Generierung einer menschenähnlichen Stimme für Hörbücher.
- Sounddesign: Erstellung von Soundeffekten für Filme oder Spiele.
- Audio-Ads: Automatische Erstellung von Audio-Werbung.
- Automatische Transkription: Umwandlung gesprochener Sprache in Text.
- Sprachanalyse: Automatische Erkennung und Analyse von Stimmungen oder Emotionen in Sprachaufnahmen.
- Telekommunikation: Verbesserung der Audioqualität in Kommunikationssystemen.
- Podcast-Erstellung: Automatische Schnitte, Zusammenfassungen und Highlights für Podcasts.
- Sprachbildung: Erstellung von Übungs- und Trainingsmaterial für Sprachkurse.

2 Rechtliche Aspekte bei der Beschaffung von KI

2.1 Verteilung der Verantwortlichkeiten entlang der Wertschöpfungskette nach AI Act

An der Konzeption, Entwicklung und Bereitstellung von KI-Systemen sind unterschiedliche Akteure beteiligt. Die Verantwortlichkeiten sind demnach unter den Akteuren der Lieferkette aufzuteilen.

Der AI Act unterscheidet in der Fassung des Rats zwischen KI-Basismodellen (im Entwurf des Europäischen Parlaments »foundation model«, jetzt »General Purpose AI model«) und KI-Systemen mit allgemeinem Verwendungszweck (»General Purpose AI system«):

Ein Basismodell ist ein KI-Systemmodell, das auf einer breiten Datenbasis in großem Umfang trainiert wurde, auf eine allgemeine Ausgabe ausgelegt ist und an ein breites Spektrum unterschiedlicher Aufgaben angepasst werden kann.

Ein KI-System mit allgemeinem Verwendungszweck ist ein KI-System, das in einem breiten Spektrum von Anwendungen eingesetzt und an diese angepasst werden kann, für die es nicht absichtlich und speziell entwickelt wurde. Ein solches GPAI-System kann eine Implementierung eines Basismodells sein.

Verpflichtungen für KI-Basismodelle

Basismodelle sollen nach dem Vorschlag des Europäischen Parlaments unabhängig vom Risiko reguliert werden, was damit der ursprünglichen risikobasierten Ausrichtung des AI Acts widerspricht. Basismodelle sollen vor dem Markteintritt in der EU eine Reihe von Transparenzanforderungen erfüllen müssen, zum Beispiel mit Bezug auf Verwendung, Architektur, Trainingsdaten und die weitere technische Dokumentation.

Verpflichtungen für Einführer von Hochrisiko-KI-Systemen

Sehr viel weitergehender sind die Anforderungen an KI-Systeme (»AI system«), insbesondere an solche mit hohem Risiko gemäß Art. 6 des AI Acts. Bevor ein Hochrisiko-KI-System in den Verkehr gebracht wird, stellen Einführer dieser Systeme unter anderem sicher, dass ein Konformitätsbewertungsverfahren durchgeführt und die technische Dokumentation erstellt werden und ihre Namen, ihre eingetragenen Handelsnamen oder ihre eingetragene Handelsmarke und Kontaktanschrift beigefügt sind.

Verpflichtungen für Händler von Hochrisiko-KI-Systemen

Bevor Händler ein Hochrisiko-KI-System auf dem Markt bereitstellen, überprüfen sie unter anderem, ob das Hochrisiko-KI-System mit der erforderlichen CE-Konformitätskennzeichnung versehen ist, ob die erforderliche Dokumentation und Gebrauchsan-

weisung beigefügt sind und ob der Anbieter bzw. gegebenenfalls der Einführer des Systems die Verpflichtungen für Hochrisiko-KI-Systeme erfüllt hat.

Verpflichtungen für Anbieter von Hochrisiko-KI-Systemen

Anbieter von Hochrisiko-KI-Systemen unterliegen umfangreichen Verpflichtungen nach dem AI Act. Händler, Einführer, Betreiber oder sonstige Dritte gelten nach dem AI Act als Anbieter eines Hochrisiko-KI-Systems, wenn sie:

1. ihren Namen oder ihr Markenzeichen auf ein Hochrisiko-KI-System setzen, das bereits in Verkehr gebracht oder in Betrieb genommen wurde,
2. eine wesentliche Änderung an einem Hochrisiko-KI-System vornehmen, das bereits in Verkehr gebracht oder in Betrieb genommen wurde (und es weiterhin ein Hochrisiko-KI-System bleibt),
3. ein KI-System, einschließlich eines KI-Systems für allgemeine Zwecke, das nicht als Hochrisiko-KI-System eingestuft wurde und bereits in Verkehr gebracht oder in Betrieb genommen wurde, so wesentlich verändern, dass das KI-System zu einem Hochrisiko-KI-System wird, oder
4. eine wesentliche Änderung des KI-Systems für hohe Risiken vornehmen.

In den Punkten 1–3 gilt der Anbieter, der das KI-System ursprünglich in Verkehr gebracht oder in Betrieb genommen hat, nicht mehr als Anbieter. Der ursprüngliche Anbieter ist somit nicht mehr verantwortlich und soll dem neuen Anbieter lediglich die technische Dokumentation und alle relevanten und vernünftigerweise zu erwartenden Informationen und Fähigkeiten des KI-Systems, den technischen Zugang oder sonstige Unterstützung auf der Grundlage des allgemein anerkannten Stands der Technik zur Verfügung, die für die Erfüllung der in dieser Verordnung festgelegten Verpflichtungen erforderlich sind.

Dies gilt auch für Anbieter von KI-Systemen mit allgemeinem Verwendungszweck (»General Purpose AI system«), wenn diese direkt in ein Hochrisiko-KI-System integriert sind.

2.2 Checkliste und Vorüberlegungen vor der Beschaffung von KI (IT Procurement)

Vor der Beschaffung von KI-Lösungen sollten vielfältige technische, kommerzielle und rechtliche Bewertungen und Festlegungen der internen und externen Anforderungen an die zu beschaffende KI erfolgen.

1. Empfehlenswert ist zunächst die **Erstellung einer generellen KI-Strategie in einem partizipativen Abstimmungsprozess mit allen relevanten Stakeholdern** (d. h. Management, Leitungen von Fachabteilungen, KI-Fachleuten, Datenschutz, Informationssicherheit, Rechtsabteilung, Compliance, Betriebsrat, betroffene Beschäftigte etc.) zur Dokumentation, in welcher Weise die Nutzung von KI-Anwendungen zur Erreichung der betrieblichen Ziele beitragen soll.

2. Unerlässlich ist die **Festlegung der wesentlichen technischen, kommerziellen und rechtlichen Anforderungen** an die zu beschaffende KI als »IT Procurement-Projekt« unter Beteiligung der relevanten Abteilungen und Stakeholder.

3. **Entscheidung, ob** die Leistung **als AI as a Service** (»AIaaS« – vergleichbar mit Software as a Service/SaaS) bezogen werden soll, **oder als »on premise«-Installation**.

   Die Entscheidung zwischen AIaaS und on premise-Installation, also einer Installation im Rechenzentrum des Unternehmens oder bei einem Dienstleister des Unternehmens, ist von entscheidender Bedeutung. Sie hat direkte Auswirkungen auf die Skalierbarkeit und Flexibilität von KI-Lösungen. AIaaS ermöglicht Unternehmen den Zugang zu leistungsstarken KI-Algorithmen und Ressourcen ohne die Notwendigkeit eigener Infrastruktur, insbesondere spezieller Serverprozessoren in Form von Grafikprozessoren, sog. »GPUs«. Dies ist besonders vorteilhaft für Unternehmen mit begrenzten Ressourcen oder einem kurzfristigen Bedarf an KI-Funktionalität. Auf der anderen Seite kann die Entscheidung für ein on premise-installiertes KI-Produkt eine bessere Anpassung an spezifische geschäftliche Anforderungen ermöglichen, was zu einer höheren Effektivität und Präzision führen kann. Dennoch müssen Organisationen bei dieser Wahl auch die langfristigen Kosten, Wartungsaufwände und die Fähigkeit zur Skalierung im Blick behalten, um die optimale KI-Strategie für ihre individuellen Anforderungen zu bestimmen.

4. Daher ist die **Prüfung von strategischen Erwägungen** wie der Nutzung von (vielleicht schon vorhandener) eigener Infrastruktur oder der eines Drittanbieters (z. B. Cloud) und ggf. entstehender Abhängigkeit von externen Dienstleistern z. B. durch Offenlegung von Know-how oder Daten, sowie Folgekosten (z. B. Updates, Wartung, Wechsel des externen Dienstleisters zu späteren Zeitpunkten) sehr entscheidend. Darüber hinaus spielen hier auch Aspekte des Schutzes personenbezogener Daten und der Betriebs- und Geschäftsgeheimnisschutz eine wesentliche Rolle.

5. Die **Festlegung der wichtigsten rechtlichen »Must Have«- oder »No Go«-Konditionen**, insbesondere Haftung, Datenschutz, Datennutzung, Gewährleistung Einhaltung regulatorischer Anforderungen, Rechten an geistigem Eigentum und der Nutzung der von KI verwendeten und generierten Daten und Regelungen für Implementierung, Verfügbarkeit, Skalierbarkeit, Support und Wartung samt SLA (Service Level Agreement), ist von entscheidender Bedeutung für den sachgerechten Vergleich und für die Risiken der ausgewählten KI.

6. Dabei ist insbesondere die **Festlegung des erforderlichen sachlichen und räumlichen Umfangs der Lizenzierung bzw. Nutzungsrechte** an KI, der von KI verwendeten internen oder externen Daten (»Input«) und den dadurch generierten Daten (»Output«) relevant. Hier ist auch die Nutzung der eingegebenen (Trainings-)Daten durch den KI-Provider zu berücksichtigen. Diese sollte bzw. kann ggf. durch eine vertragliche Vereinbarung oder die Wahl einer bestimmten Lizenz ausgeschlossen werden, wo nötig. Weiteres besonderes Augenmerk sollte auf die Nutzung von R&D- bzw. Entwicklungsdaten durch die KI sowie auf die Nutzung in produktiver Umgebung (Software und Hardware) gelegt werden.

7. Vor allem aus technischer und kommerzieller Sicht ist frühzeitig zu bewerten, ob die **Implementation bzw. Integration der KI** intern erfolgt oder ob dafür ein externer IT-Servicepartner nötig oder sinnvoll ist.

8. Gleiches gilt für die **Bewertung, ob der Einkauf der KI über externe IT-Servicepartner** (Systemintegrator, Reseller oder Kooperationspartner der KI-Anbieter) **vorteilhaft ist**.

9. **Rechtlich** sind insbesondere folgende Punkte zu prüfen:
   - **Die Möglichkeit und Sinnhaftigkeit der (teilweisen) Nutzung von eigenen Einkaufsbedingungen oder Vertragsbedingungen**, die meist allerdings nicht dem komplexen Vertragsgegenstand eines KI-Systems gerecht werden,
   - **Anbieter-AGB bzw. Vertragsbedingungen**, sowie
   - **Bewertung und Auswahl des bevorzugten anwendbaren Rechts für Beschaffung von KI, um ggf. für Besteller vorteilhaftere Jurisdiktionen zu nutzen** (die z. B. weniger Haftungsausschlüsse erlauben).

10. **Risikoabwägung**

    Elementar für die Auswahl ist eine »State of the Art«-Risikoabwägung. Im Rahmen der Risikoabwägung gilt es bei KI neue Herausforderungen zu beachten, sodass bisherige Prozesse des Risikoassessments ggf. überarbeitet werden müssen. Das BSI nennt in seinem Bericht zur Lage der IT-Sicherheit in Deutschland 2023 vor allem die Nutzung der KI für Cyberangriffe, das Ausnutzen von Schwachstellen in der KI, Herausforderung im Training und vor allem das umfassende Wissen der KI bei einer weiten Einbindung in bestehende Strukturen und Datenbanken als neue Bedrohungen.[2]

2 https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/Publikationen/Lageberichte/Lagebericht2023.pdf?__blob=publicationFile&v=6 S. 40ff
Guidelines for secure AI system development (ncsc.gov.uk)

11. Die Risikobewertung sollte u. a. die vom Fraunhofer IAO empfohlenen Punkte[3] berücksichtigen:

- Welche Datensätze werden für Entwicklung bzw. das Training der KI genutzt? Können oder sollten Trainingsdaten ggf. gesondert auf dem Markt für Trainingsdatensätze beschafft werden?
- Welche Datensätze werden im anschließenden Praxisbetrieb verarbeitet?
- Welche rechtlichen Rahmenbedingungen müssen dabei beachtet werden?
- Ist es für die Zwecke der jeweiligen Anwendung ausreichend, anonyme Daten zu verwenden?
- Wer erhält Zugriff auf die Daten? Wie sind betriebliche Datenbestände gegen Missbrauch oder Betriebsspionage geschützt? Müssen die Risiken für personenbezogene Daten vorab in einer Datenschutz-Folgenabschätzung geprüft und dokumentiert werden? Wer führt diese durch? Wie werden Mitbestimmungsrechte hinsichtlich des Schutzes personenbezogener Daten wahrgenommen? Wie werden die Kriterien der Datensparsamkeit und Zweckbindung der Datennutzung umgesetzt?
- Vor allem sind aber die vertraglichen und gesetzlichen Haftungsrisiken aus der konkret geplanten Nutzung der KI für interne Zwecke und/oder Leistungen für Kundinnen und Kunden im Detail zu bewerten und soweit möglich angemessen durch vertragliche Gewährleistungs-, Haftungs- und Freistellungsregeln abzusichern (siehe Ziffer 4.3). Dabei dürfen mögliche direkte und indirekte (Vermögens-)Schäden beim Kunden oder Dritten aufgrund der Nutzung durch von der KI generierten Daten nicht außer gelassen werden.
- Nutzung von bereitgestellten Filtern.
- KI-Governance.
- Es ist weiterhin zu beachten, dass die unter 3. genannten rechtlichen Aspekte bereits im Rahmen der Beschaffung berücksichtigt werden sollten, da im eingekauften Produkt enthaltene Rechtsmängel sich auch auf die Einsatzphase auswirken würden.

Da KI-Technologie fast immer in Form von Software implementiert wird, gelten die Überlegungen und Anforderungen, die bei der Beschaffung von Software zu beachten sind, grundsätzlich entsprechend. Wird die Software »as a Service« (AIaaS, s. oben) angeboten, gilt es zu beachten, dass der Anbieter des generativen KI-Systems neben der »reinen« Software auch eine Infrastruktur bereitstellt, die für die Nutzung des konkreten KI-Systems erforderlich ist. Dabei bedarf es einer ausreichenden Kontrolle über die genutzten KI-Systeme der Anbieter durch den Kunden, um etwa sicherzustellen, dass gesetzliche oder interne Anforderungen eingehalten werden. Volle technische Kontrolle und insbesondere auch Souveränität über die eigenen Daten, die in die KI fließen sollen, hat der Kunde nur bei einer on premise-Lösung.

Aufgrund der Vergleichbarkeit zu SaaS-Lösungen, kann für die vertragliche Gestaltung und Prüfung der Beschaffung von KI-Systemen grundsätzlich auf die entsprechenden Anforderungen zu SaaS-Verträgen zurückgegriffen werden. Insbesondere sind im Rahmen der vertraglichen Prüfung ausreichende Nutzungsrechte und Service-Level für die Verfügbarkeit zu vereinbaren, die dem Zweck der beabsichtigten Anwendungsbereiche des jeweiligen KI-Systems im Unternehmen entsprechen. In Verträgen über KI-Systeme sollte zudem die Verantwortlichkeit für Herkunft und Qualität von Trainingsdaten eindeutig festgelegt sein.

3 Fraunhofer IAO, Leitfaden zu Strategie und Wandel für den KI-Einsatz, Leitfaden zur Durchführung von KI-Projekten (fraunhofer.de) https://www.digital.iao.fraunhofer.de/de/publikationen/Leitfaden_DurchfuehrungVonKI-Projekten.html, zuletzt besucht am 12. Februar 2024

Darüber hinaus sollte es klare Regelungen dafür geben, wenn die vertragsgemäße Nutzung des KI-Systems durch den Kunden gegen regulatorische Vorgaben, Gesetze und/oder Rechte Dritter verstößt. Hier sollten etwa Klauseln zur Freistellung von solchen Ansprüchen Dritter durch den Anbieter des KI-Systems enthalten sein. Schließlich sollte ausdrücklich vereinbart werden, in welchem Umfang und in welcher Form die in dem KI-System verarbeitete Kundendaten vom Anbieter des KI-Systems (ggf. zu eigenen Zwecken) genutzt werden dürfen. Je nachdem, wie sensibel und kritisch diese Kundendaten sind, sollten sie durch ausdrückliche vertragliche Regelungen vor einer vertragswidrigen Nutzung geschützt werden.

Die spezifischen Risiken, die mit der Nutzung eines generativen KI-System einhergehen können, sollten schon im Rahmen des Beschaffungsprozess ausreichend geprüft werden. Daher sollte die beabsichtige Nutzung des jeweiligen KI-Systems insbesondere von Experten aus dem Datenschutz, der Informationssicherheit, der Rechtsabteilung und Compliance, dahingehend geprüft werden, ob ein rechtskonformer Einsatz möglich ist und welche Risiken ggf. bestehen. Bei dieser Prüfung sind die von den Anbietern offerierten unterschiedlichen Lizenzmodelle (z. B. Versionen speziell für Unternehmen anstatt von allgemein offen zugänglichen Versionen) sowie unter Umständen bereitgestellte Sicherheitsmechanismen, wie etwa Filter oder der Ausschluss bestimmter Datenverarbeitungen, zu berücksichtigen.

Wegen der Komplexität einer solchen interdisziplinären Prüfung und auch wegen der zunehmenden gesetzlichen Anforderungen, z. B. durch den AI Act mit einem ggf. einzurichtenden eigenen Risikomanagementsystem, empfiehlt sich die Implementierung einer KI-Governance, die in den Beschaffungsprozess einzubinden ist. Diese Funktion kann dann sowohl rechtliche als auch technische Anforderungen an die Nutzung von KI-Systemen im Unternehmen bündeln, deren Einhaltung überwachen und darüber hinaus als Kompetenzstelle für interne und externe Ansprechpartner dienen.

3 Rechtliche Aspekte in der Einsatzphase

3.1 Datenverarbeitung nach DSGVO

3.1.1 Einleitung, Allgemeines und Anwendungsbereich

Die Datenschutz-Grundverordnung (DSGVO) und das Bundesdatenschutzgesetz (BDSG) sind zwei wesentliche rechtliche Rahmenwerke, die den Einsatz von Künstlicher Intelligenz in Unternehmen in der Europäischen Union und in Deutschland regeln, sofern die KI personenbezogene Daten verarbeitet. Sie sind anwendbar, wenn Unternehmen als Verantwortliche oder Auftragsverarbeiter personenbezogene Daten im Rahmen der KI-Nutzung verarbeiten. Dabei ist irrelevant, ob die KI selbstständig Entscheidungen trifft oder als unterstützendes Werkzeug dient.
Im Gegensatz zu KI-Systemen für Profiling, Diagnostik etc., die grundsätzlich (sensible) personenbezogene Daten verarbeiten und in der Regel keine generative KI darstellen, werden beim Einsatz generativer KI personenbezogene Daten weniger häufig unmittelbar verarbeitet. Ein Anwendungsbeispiel für eine solche unmittelbare Verwendung ist das Account-Management.[4]

Im folgenden Abschnitt werden auch Bezüge zu Anwendungsfällen jenseits von generativer KI – bspw. diskriminativer KI – hergestellt. Dies soll den Blick trotz des gesetzten Rahmens dieses Leitfadens weiten und für Problembereiche innerhalb des Datenschutzrechtes sensibilisieren, die sich bei KI außerhalb des Anwendungsbereiches der generativen KI ergeben. Dies betrifft vor allem Bereiche des Profilings und generell Bereiche, wo Entscheidungen autonom bzw. maschinengestützt getroffen werden.

Beispiele für eine breitere Verarbeitung personenbezogener Daten auch im Rahmen von generativer KI sind die Erstellung von Texten mit personenbezogenem Inhalt oder das Generieren von Bildern und Videos auf Basis von Original-Material mit Personenbezug. In diesen Fällen gelten die Regelungen der DSGVO und des BDSG.

Ein Unternehmen, das mittels KI personenbezogene Daten verarbeitet, gilt als »Verantwortlicher« im Sinne der DSGVO, wenn es Entscheidungen über die Zwecke und Mittel der Datenverarbeitung trifft. Die Verarbeitung umfasst jede Operation, die mit personenbezogenen Daten durchgeführt wird – von der Erhebung über die Speicherung und Analyse bis hin zur Löschung.

4 Zwar sind in gängigen Modellen generativer KI selbst personenbezogene Informationen nicht als Klardatum enthalten. Allerdings können solche personenbezogenen Informationen in Form von Wahrscheinlichkeiten bzw. Gewichtungen, auf deren Basis diese Modelle arbeiten, repräsentiert sein. Diese Wahrscheinlichkeiten repräsentieren letztlich die Trainingsdaten, aus denen sie errechnet wurden, und somit potenziell auch jeden Personenbezug, der in diesen Trainingsdaten enthalten ist – vergleichbar mit einer Pseudonymisierung, bei der allerdings die KI selbst die Mittel hat, den Personenbezug wiederherzustellen. Zumindest bei allen KI-Modellen, die mit CommonCrawl oder vergleichbaren Datensätzen trainiert wurden, dürfte ein Personenbezug in den Trainingsdaten nicht auszuschließen sein. In Einzelfällen »erinnert« sich das Modell an die personenbezogene Information. Zu vertiefteren Datenschutzfraen konsultieren Sie bitte unseren Praxisleitfaden KI & Datenschutz, im Erscheinen.

Bei der Nutzung von KI kann es zu Konstellationen kommen, in denen das Unternehmen entweder als Auftragsverarbeiter oder in gemeinsamer Verantwortung mit dem KI-Hersteller agiert:

Bei der Auftragsverarbeitung verarbeitet das Unternehmen Daten im Auftrag eines anderen Unternehmens (dem Auftraggeber). Die KI dient als Werkzeug zur Erledigung spezifischer Aufgaben. Die Verantwortlichkeit für die Einhaltung der Datenschutzvorschriften liegt primär beim Auftraggeber.

Wenn das Unternehmen und der KI-Hersteller gemeinsam Entscheidungen über die Verarbeitung personenbezogener Daten treffen, spricht man von gemeinsamer Verantwortlichkeit. Beide Parteien müssen in diesem Fall die Einhaltung der DSGVO sicherstellen und dies in einer Vereinbarung dokumentieren.

In einigen Fällen kann es zur getrennten Verantwortung kommen, z. B. wenn der Anbieter der KI-Daten für Zwecke, die vom Auftraggeber weder bestimmt noch mit ihm abgesprochen sind, verarbeitet.

Eine Herausforderung kann im Falle der gemeinsamen Verantwortlichkeit darin liegen, die Rechte und Pflichten und damit auch die Haftung zwischen dem KI-Hersteller und dem einsetzenden Unternehmen für beide Parteien akzeptabel aufzuteilen.

Die wichtigsten Rechtsgrundlagen für die Verarbeitung personenbezogener Daten durch KI im Unternehmenskontext sind:

- Einwilligung (Art. 6 Abs. 1 lit. a DSGVO): Die betroffene Person muss ihre Einwilligung zur Datenverarbeitung geben. Diese muss freiwillig, bestimmt und informiert abgegeben werden.
- Vertragserfüllung (Art. 6 Abs. 1 lit. b DSGVO): Die Datenverarbeitung ist notwendig, um einen Vertrag zu erfüllen, in dem die betroffene Person Partei ist oder um einen Vertrag auf Anfrage der betroffenen Person zu schließen.
- Rechtliche Verpflichtung (Art. 6 Abs. 1 lit. c DSGVO): Die Verarbeitung ist notwendig, um einer rechtlichen Verpflichtung nachzukommen.[5]
- Berechtigte Interessen (Art. 6 Abs. 1 lit. f DSGVO): Die Verarbeitung ist zur Wahrung berechtigter Interessen des Unternehmens oder eines Dritten notwendig, sofern nicht die Interessen oder Grundrechte und Grundfreiheiten der betroffenen Person überwiegen.

Dabei ist es wichtig zu beachten, dass es für die Erhebung und Nutzung personenbezogener Daten zum Training einer KI und der Verarbeitung dieser Daten durch die KI jeweils einer – gegebenenfalls unterschiedlichen – Rechtsgrundlage bedarf. Es sollte geprüft werden, ob die Datennutzung sachgerechterweise von vorneherein auf aggregierte Daten ohne Personenbezug beschränkt werden kann, oder eine Verwendung anonymer Daten ausreicht. Dies gilt erst recht mit Blick auf die sich teilweise widersprechenden Vorgaben der DSGVO und des Data Acts bzgl. der Nutzung von durch verbundene Produkte oder Dienstleistungen generierte Daten, inklusive von Betriebsgeheimnissen. Hier dürfte eine Vermischung von personenbezogenen Daten nach der DSGVO,

5 Bislang besteht jedoch noch keine rechtliche Pflicht zum Einsatz von KI.

IoT-Daten nach dem Data Act und vertraulicher interner Daten (R&D, Produktionsdaten etc.) wohl zu vermeiden sein, um komplexe Folgeprobleme zu vermeiden (siehe dazu auch Ziffer 3.1.2).

Wenn personenbezogene Daten in Staaten außerhalb des Geltungsbereiches der DSGVO übermittelt werden, muss sichergestellt sein, dass trotzdem ein den Anforderungen der DSGVO entsprechendes Schutzniveau gewährleistet ist, und eine der Grundlagen aus Art. 45 ff. DSGVO, z. B. ein Angemessenheitsbeschluss der EU-Kommission oder zwischen Datenexporteur- und Importeur abgeschlossene Standarddatenschutzklausel, vorliegt.

Neben den Bestimmungen der DSGVO sind im Daten- und KI-Kontext auch die Bestimmungen des Data Acts sowie des Data Governance Acts zu beachten, auf die im Rahmen dieses Leitfadens jedoch nicht vertieft eingegangen werden soll.

3.1.2 Pflichten im Einzelnen

Zweckbindung

Die Datenschutz-Grundverordnung legt strenge Regeln für die Verarbeitung personenbezogener Daten fest. Der Grundsatz der Zweckbindung spielt eine wesentliche Rolle, besonders im Kontext der Künstlichen Intelligenz.

Gemäß der DSGVO müssen Daten für festgelegte, eindeutige und legitime Zwecke erhoben werden. Das bedeutet, dass bei der Nutzung von KI-Systemen die Daten nur für den spezifischen Zweck verarbeitet werden dürfen, für den sie gesammelt wurden.

Wenn KI-Systeme Daten verarbeiten und Ergebnisse liefern, müssen diese Ergebnisse im Einklang mit dem ursprünglichen Zweck stehen. Nach Erreichen dieser Ergebnisse müssen die Daten gelöscht werden. Eine Weiterverarbeitung zu einem anderen Zweck erfordert in der Regel eine neue Rechtsgrundlage.

Beim Training von KI-Modellen mit personenbezogenen Daten muss sichergestellt werden, dass die Datenverarbeitung den Zwecken entspricht, für die die Daten ursprünglich erhoben wurden. Darüber hinaus müssen Maßnahmen wie Pseudonymisierung oder Anonymisierung ergriffen werden, um die Privatsphäre der Betroffenen zu schützen. Gegebenenfalls können Daten zu Trainingszwecken auf Basis einer Einwilligung als Rechtsgrundlage verarbeitet werden. Dies ist zu empfehlen, falls keine ausreichende Pseudonymisierung möglich ist und Anonymisierung dem Zweck entgegensteht. Eine Pseudonymisierung ist nur dann möglich, wenn die Zuordnung von Pseudonymen und Klarnamen nach der Verarbeitung durch und mittels KI noch verlässlich erhalten bleibt.

Aktuell etablierte KI-Modelle verwenden häufig Datensets als Teil der Trainingsdaten, für die pauschal enorme Mengen im Internet zugänglicher Daten aus zahlreichen Quellen gesammelt werden (z. B. CommonCrawl und WebText). Für alle diese Daten muss gewährleistet sein, dass eine Weiterverarbeitung als Trainingsdaten für die KI entweder durch den ursprünglichen Verarbeitungszweck abgedeckt ist oder eine neue Rechtsgrundlage vorliegt.

Vermeidung der Entstehung besonderer Art personenbezogener Daten iSd Art. 9 DSGVO

Wenn besondere Kategorien personenbezogener Daten im Sinne des Art. 9 Abs. 1 DSGVO Gegenstand der Verarbeitung sind, hat der verantwortliche Datenverarbeiter erhöhte Schutzanforderungen umzusetzen, damit die Risiken für den Schutz der Rechte und Freiheiten natürlicher Personen minimal bleiben.

Die im Bereich der Verfahren der Künstlichen Intelligenz durchgeführten umfangreichen Datenverarbeitungen sind dahingehend zu beurteilen, ob aus den personenbezogenen Daten im Laufe der Anwendung und Weiterentwicklung der KI Informationen abgeleitet werden können, die besondere Kategorien personenbezogener Daten nach Art. 9 Abs. 1 DSGVO sein können.

Dies kann z. B. der Fall sein, wenn aus personenbezogenen Daten Schlüsse auf besondere personenbezogene Daten abgeleitet werden können. Siehe dazu EuGH:

»Art. 9 Abs. 1 DSGVO ist dahingehend auszulegen, dass die Veröffentlichung personenbezogener Daten, die geeignet sind, die sexuelle Orientierung einer natürlichen Person indirekt zu offenbaren, auf der Website der Behörde, die für die Entgegennahme und die inhaltliche Kontrolle von Erklärungen über private Interessen zuständig ist, eine Verarbeitung besonderer Kategorien personenbezogener Daten im Sinne dieser Bestimmung darstellt.« (EuGH v. 1.8.2022 - C-184/20)

Sollte es in einem KI-System möglich sein, aus vorhandenen Daten indirekt besondere personenbezogene Daten (nach Art. 9 Abs.1 DSGVO) abzuleiten, gelten auch für diese besonderen personenbezogenen Daten erhöhte Anforderungen zum Schutz der natürlichen Personen, d. h. insbesondere erhöhte Anforderungen in Bezug auf die rechtlichen Grundlagen zur Verarbeitung der Daten (Art. 9 DSGVO).

Ferner sind nach Art. 10 des AI Acts bei sogenannten Hochrisiko-KI-Systemen die Risiken bei der Nutzung von besonderen Kategorien personenbezogener Daten »angemessene Vorkehrungen für den Schutz der Grundrechte und Grundfreiheiten natürlicher Personen« zu treffen, »wozu auch technische Beschränkungen einer Weiterverwendung und modernste Sicherheits- und Datenschutzmaßnahmen wie Pseudonymisierung oder Verschlüsselung gehören« (Art. 10 Abs. 5 des Entwurfs der EU KI-Verordnung).

Daraus lässt sich schließen, dass bei den umfangreichen Verarbeitungen von Daten im Rahmen des Trainings und der Anwendung von KI-Systemen zu prüfen ist, ob durch die Anwendung besondere Kategorien personenbezogener Daten (Art. 9 DSGVO) entstehen können. Dann wären für das Training und für die Anwendung eines KI-Systems die genannten erhöhten Anforderungen zum Schutz dieser besonderen personenbezogenen Daten umzusetzen. Zum aktuellen Stand der Technik dürfte dies bei Anwendungen von generativer KI eher weniger der Fall sein, dafür aber bei anderen KI-Technologien, insb. der diskriminativen.[6]

6 Siehe zu den technologieunabhängigen vertieften Datenschutzausführungen den datenschutzspezifischen Praxisleitfaden KI & Datenschutz, im Erscheinen.

Informationspflichten (Transparenz)

Einer der Hauptgrundsätze der DSGVO ist die Transparenz. Unternehmen müssen Personen, deren Daten sie verarbeiten, umfassend darüber informieren. Dies schließt Informationen über den Zweck der Datenverarbeitung, die Rechtsgrundlage, die Dauer der Speicherung und die Rechte der betroffenen Personen ein. Besonders wichtig sind dabei die Artikel 13 Abs. 2 lit. f und 14 Abs. 2 lit. g, die sich auf automatisierte Entscheidungen, einschließlich Profiling, konzentrieren. Sie verlangen, dass Unternehmen die Betroffenen über die Existenz solcher Verarbeitungsmechanismen, deren Logik sowie die Bedeutung und die beabsichtigten Auswirkungen informieren.

Die Integration von KI in Geschäftsprozesse bringt allerdings Herausforderungen mit sich. Eine der größten ist dabei die Transparenz bei selbstlernenden Systemen. Diese Systeme sind dynamisch. Ihre Entscheidungsfindungsprozesse können komplex und schwer nachvollziehbar sein. Dies erschwert die Einhaltung der Transparenzvorschriften der DSGVO.

Ein weiteres Problem ergibt sich, wenn Unternehmen KI-Lösungen einsetzen, die sie nicht selbst entwickelt haben. In solchen Fällen sind sie von den Datenschutzpraktiken Dritter abhängig, was zusätzliche Risiken in Bezug auf die Einhaltung der DSGVO birgt. Häufig fehlt es auch an Transparenz, sodass der Verpflichtung zur Information der Betroffenen über die Verarbeitung ihrer personenbezogenen Daten möglicherweise nicht vollständig nachgekommen wird.

Ein proaktiver Ansatz zur Einhaltung der DSGVO-Vorschriften ist das Konzept »Privacy by Design«. Es bedeutet, dass Datenschutzmaßnahmen von Anfang an in die Entwicklung und Implementierung von KI-Systemen integriert werden. Dies umfasst die Minimierung der Datenerfassung, die Sicherstellung der Transparenz und das Einbetten von Kontrollmöglichkeiten für die Nutzerinnen und Nutzer.

Die Rechenschaftspflicht ist ein weiterer wesentlicher Aspekt der DSGVO, der eng mit dem Einsatz von Künstlicher Intelligenz in Unternehmen verknüpft ist. Diese Pflicht erfordert von Unternehmen nicht nur die Einhaltung der Datenschutzvorschriften, sondern auch den Nachweis, dass geeignete Maßnahmen und Verfahren implementiert wurden, um die Compliance mit der DSGVO zu gewährleisten.

Im Kontext der KI bedeutet dies, dass Unternehmen nicht nur transparent über ihre Datenverarbeitungspraktiken informieren und diese an die DSGVO anpassen müssen, sondern auch belegen müssen, dass ihre KI-Systeme den Datenschutzvorschriften entsprechen. Dies beinhaltet die Dokumentation der Datenverarbeitungsprozesse, die Durchführung von Datenschutz-Folgenabschätzungen, insbesondere bei risikoreichen KI-Anwendungen, sowie die Implementierung und Überwachung von Datenschutzmaßnahmen.

Die Rechenschaftspflicht ist besonders herausfordernd im Umgang mit KI, da die Technologie oft komplexe und dynamische Datenverarbeitungsprozesse umfasst. Unternehmen müssen daher sicherstellen, dass ihre KI-Systeme nicht nur zum Zeit-

punkt der Implementierung, sondern auch über ihren gesamten Lebenszyklus hinweg den Datenschutzanforderungen entsprechen. Dies erfordert eine kontinuierliche Überwachung und Anpassung der Systeme sowie eine klare Dokumentation aller Änderungen und Entscheidungen.

Eine weitere Herausforderung liegt bei der Verwendung von KI in der Frage, ob und inwieweit es als Voraussetzung für einen ausreichenden Grad an Transparenz erforderlich ist, den Algorithmus und/oder die Trainingsdaten offenzulegen. Im Falle des Einsatzes eines proprietären KI-Systems eines externen Anbieters stünde dem typischerweise das Interesse des Herstellers entgegen, diese für die Leistungsfähigkeit der KI entscheidenden Faktoren gerade nicht offenzulegen. Es entsteht somit ein »Blackbox«-Problem, da für die Verarbeitung wesentliche Aspekte der KI gegebenenfalls nicht ausreichend transparent dargestellt werden können.

Recht auf Berichtigung & Löschen

Das Recht auf Berichtigung (Art. 16 DSGVO) gibt Individuen die Möglichkeit, die Korrektur unrichtiger oder unvollständiger sie betreffender personenbezogener Daten zu fordern. Dies bedeutet für KI-Systeme, dass sie über Mechanismen verfügen müssen, die eine schnelle und präzise Anpassung von Daten ermöglichen, sobald ein Fehler festgestellt wird.

Das Recht auf Löschung, auch bekannt als »Recht auf Vergessenwerden« (Art. 17 DSGVO), verpflichtet zur Löschung personenbezogener Daten unter bestimmten Umständen, beispielsweise wenn die Daten für die ursprünglichen Zwecke nicht mehr notwendig sind oder die betroffene Person ihre Einwilligung widerruft. In Bezug auf KI stellt dies sicher, dass Systeme in der Lage sind, Daten effizient und vollständig zu entfernen, wenn dies erforderlich ist.

Ein weiterer wichtiger Aspekt im Kontext von KI und Datenschutz ist das Datamanagement. Vor der Nutzung von Daten in der Produktion müssen diese validiert und verifiziert werden. Dies gewährleistet, dass KI-Systeme mit zuverlässigen und akkuraten Informationen arbeiten. Dies gilt nicht nur für die Eingangsdaten, sondern auch für Daten, die während des Produktionsprozesses generiert werden.

Eine besondere Herausforderung im Umgang mit KI und Datenschutz ist das Konzept des »Unlearning«. Es geht darum, wie KI-Systeme Daten »vergessen« oder aus ihren Modellen entfernen können, insbesondere im Hinblick auf die Anforderungen des Rechts auf Löschung. Dies ist ein komplexer Prozess, da KI-Modelle dazu neigen, Informationen tief in ihren Strukturen zu verarbeiten und zu speichern. Es erfordert daher fortschrittliche Techniken, um sicherzustellen, dass die betreffenden Daten vollständig und effektiv aus den Modellen entfernt werden, ohne die Integrität oder Leistungsfähigkeit der KI zu beeinträchtigen.

Sonstige Betroffenenrechte

Die DSGVO legt einen besonderen Fokus auf den Schutz personenbezogener Daten und räumt Betroffenen spezifische Rechte ein. Diese Rechte sind auch im Fall der Verarbeitung personenbezogener Daten durch KI von großer Bedeutung, wobei ihre Umsetzung besondere Herausforderungen mit sich bringt.

Eine der Hauptforderungen der DSGVO ist die Transparenz in der Verarbeitung personenbezogener Daten. Im KI-Kontext wird dies durch die Komplexität oder unvollständige Informationen zur Datenverarbeitung bei proprietären Systemen (Blackbox) und die oft intransparenten Entscheidungsprozesse erschwert. KI-Systeme können so programmiert sein, dass sie selbstlernend sind und Entscheidungen auf der Grundlage von Algorithmen treffen, die für Außenstehende schwer nachvollziehbar sind. Dies wirft Fragen hinsichtlich der Erklärbarkeit und Nachvollziehbarkeit von KI-Entscheidungen auf, die für die Wahrung der Transparenzpflicht essenziell sind.

Ein weiteres Recht ist das Auskunftsrecht der betroffenen Person. Gemäß der DSGVO haben Individuen das Recht zu erfahren, ob und wie ihre persönlichen Daten verarbeitet werden. Bei KI-Anwendungen kann es schwierig sein, präzise Auskünfte zu geben, insbesondere wenn es um komplexe Datenverarbeitungsvorgänge geht. Unternehmen müssen daher sicherstellen, dass sie auch bei komplexen KI-Operationen in der Lage sind, klare und verständliche Informationen über die Datenverarbeitung zu liefern.

Ein zentrales Element der DSGVO ist auch das Recht auf Widerruf einer Einwilligung und der Widerspruch gegen die Datenverarbeitung (Art. 21). Sofern ein Betroffener seine Einwilligung widerruft oder der Datenverarbeitung widerspricht, müssen Unternehmen in der Lage sein, die Verarbeitung dieser Daten unverzüglich zu stoppen. Dies kann insbesondere bei fortgeschrittenen KI-Systemen, die auf großen Datenmengen trainiert wurden, eine Herausforderung darstellen.

Datenminimierung & Speicherbegrenzung

Der Grundsatz der Datenminimierung besagt, dass nur so viele personenbezogene Daten verarbeitet werden sollen, wie unbedingt nötig sind, um den festgelegten Zweck zu erfüllen. Im Kontext generativer KI kann dies herausfordernd sein. Die Systeme sind oft auf große Datenmengen angewiesen, um effektiv zu lernen und präzise zu sein. Hier muss ein Gleichgewicht gefunden werden zwischen dem Bedarf an umfangreichen Daten für das Training von KI-Modellen und dem Schutz der Privatsphäre.

Der Grundsatz der Speicherbegrenzung verlangt, dass personenbezogene Daten nur so lange gespeichert werden, wie sie für den Zweck ihrer Verarbeitung benötigt werden. Im KI-Bereich ist es oft schwierig, den genauen Zeitraum zu bestimmen, für den Daten benötigt werden, insbesondere bei lernenden Systemen, die kontinuierlich auf Daten zugreifen. Es ist wichtig, klare Kriterien für die Aufbewahrungsdauer zu definieren und die Daten zu löschen, sobald sie ihren Zweck erfüllt haben.

Datensicherheit

Die DSGVO stellt hohe Anforderungen an die Datensicherheit. Artikel 32 verlangt von Unternehmen, Schutzmaßnahmen nach dem aktuellen Stand der Technik umzusetzen. Dies beinhaltet eine fortlaufende Überwachung und Anpassung der Sicherheits- und Datenschutzpraktiken von KI-Systemen. Zudem fordert die DSGVO Transparenz und Rechenschaft, insbesondere im Hinblick auf die Rechtfertigung der Datenverarbeitung. Ein besonderes Problem ist die »Richtigkeit der Daten«, da KI-Systeme auf Basis fehlerhafter Daten irreführende Ergebnisse liefern können. Um die DSGVO-Anforderungen zu erfüllen, müssen Unternehmen ihre KI-Systeme kontinuierlich überwachen und validieren.

Vorbehalt menschlicher Entscheidung

Die DSGVO setzt mit Art. 22 eine wichtige Begrenzung für automatisierte Entscheidungsfindungssysteme, einschließlich Profiling, fest.[7] Dieser Artikel beschränkt die Verwendung solcher Systeme, wenn deren Entscheidungen rechtliche Wirkungen haben oder die betroffene Person erheblich beeinträchtigen. Es wird klargestellt, dass Entscheidungen, die signifikante Auswirkungen auf Einzelpersonen haben, nicht ausschließlich auf automatisierten Verarbeitungsprozessen basieren dürfen. Die Definition von »automatisierten Entscheidungen« umfasst Entscheidungen, die ohne menschliches Eingreifen getroffen werden und auf algorithmischer Datenverarbeitung basieren. Für das Verbot gelten jedoch bestimmte Voraussetzungen: Die Entscheidung muss rechtliche Wirkungen entfalten oder die Person erheblich beeinträchtigen. Es existieren aber Ausnahmen von diesem Verbot, wenn beispielsweise die automatisierte Entscheidung für den Abschluss oder die Erfüllung eines Vertrags notwendig ist, die betroffene Person ausdrücklich eingewilligt hat oder die automatisierte Entscheidung durch das Recht der EU oder der Mitgliedstaaten erlaubt ist, sofern dieses Recht angemessene Maßnahmen zum Schutz der Rechte und Freiheiten sowie der berechtigten Interessen der betroffenen Person vorsieht. Praktisch relevant dürfte wohl nur die Einwilligung sein.

Des Weiteren ist eine Unterscheidung zwischen entscheidungsunterstützenden Systemen und selbstentscheidenden Systemen wichtig. Während entscheidungsunterstützende Systeme lediglich Daten, Analysen oder Empfehlungen für menschliche Entscheidungsträger liefern, treffen selbstentscheidende Systeme Entscheidungen ganz ohne menschliches Eingreifen. In der DSGVO ist insbesondere die Nutzung der selbstentscheidenden Systeme reglementiert, da diese ein hohes Potenzial haben, erhebliche Auswirkungen auf Einzelpersonen auszuüben. Ziel des Artikels 22 der DSGVO ist es, die potenziellen Risiken der Automatisierung für die Grundrechte und Freiheiten von Einzelpersonen zu begrenzen und sicherzustellen, dass Entscheidungen, die erhebliche Auswirkungen haben, eine menschliche Überprüfung beinhalten.

7 KI gestützte Entscheidungsfindungssysteme stellen in der Regel keine generative KI dar, dennoch gilt es die Bestimmungen von Art. 22 DSGVO zu beachten, insbesondere, wenn Ergebnisse generativer KI für automatisierte Entscheidungen genutzt werden.

Datenschutz-Folgenabschätzung

Die Durchführung einer Datenschutz-Folgenabschätzung ist gemäß Art. 35 DSGVO erforderlich, wenn eine Verarbeitung personenbezogener Daten voraussichtlich zu einem hohen Risiko für die Rechte und Freiheiten der betroffenen Personen führt. Dies kann insbesondere beim Einsatz neuer Technologien, wozu KI-Anwendungen im Regelfall zählen dürften, der Fall sein.

In der Datenschutz-Folgenabschätzung sind die Verarbeitungsvorgänge und die Risiken der Verarbeitung für Betroffene sowie die getroffenen technisch-organisatorischen sowie andere risikominimierende Maßnahmen zu beschreiben und zu bewerten. Wie schon bei den Anforderungen bezüglich Informationspflichten und Transparenz liegt auch hier eine besondere Hürde in der Komplexität und Dynamik von KI-Systemen, die bereits eine Beschreibung der Verarbeitung, aber insbesondere auch eine Bewertung der Risiken, erschweren können.

Grundsätzlich nehmen die deutschen Aufsichtsbehörden an, dass der Einsatz von generativer KI mit einem hohen Risiko verbunden ist. Verbleibt unter Berücksichtigung aller risikominimierenden Maßnahmen in der Bewertung ein hohes Risiko, muss gemäß Art. 36 DSGVO die zuständige Aufsichtsbehörde konsultiert werden.

3.1.3 Praxishinweise

Anonymisierung im Kontext der KI

Anonymität liegt gemäß Erwägungsgrund 26 DSGVO vor, wenn die Identifizierbarkeit einer natürlichen Person nicht oder nur unter dem Einsatz von nach allgemeinem Ermessen unwahrscheinlichen Mitteln möglich ist. Dabei sind alle Mittel zu berücksichtigen, die vernünftigerweise zur Identifikation genutzt werden könnten, unabhängig davon, ob sie technischer, rechtlicher oder sonstiger Natur sind. Als Maßstab, ob es wahrscheinlich ist, dass bestimmte Mittel zur Identifizierung genutzt werden oder nicht, sind insbesondere die erforderlichen Kosten, Fähigkeiten sowie der benötigte Zeit- und Arbeitsaufwand relevant.

Es ist nicht erforderlich, dass sich alle zur Identifizierung einer Person erforderlichen Informationen in den Händen der jeweils betrachteten Stelle befinden, sondern es reicht aus, dass ein realistischerweise nutzbarer Weg vorhanden ist, die Informationen von einer anderen Stelle zu erlangen.

Der Status der Anonymität muss regelmäßig überprüft werden, da zum Beispiel aufgrund neuer technologischer Entwicklungen eine Identifikation gegebenenfalls zu einem späteren Zeitpunkt leichter möglich sein kann als zum Zeitpunkt der ursprünglichen Bewertung.

Unter dem Begriff Anonymisierung versteht man Verfahren, mit denen personenbezogene Daten dahingehend verändert werden, dass sie die Identifikation einer natürlichen Person nicht mehr oder nur noch mit dem oben beschriebenen unverhältnismäßigen Aufwand ermöglichen. Anony-

misierungsverfahren arbeiten entweder auf Basis von Randomisierung oder Generalisierung. Die eingesetzten Verfahren sollten stets dem Stand der Technik entsprechen.
Auch der Vorgang der Anonymisierung stellt eine Verarbeitung personenbezogener Daten im Sinne der DSGVO dar und erfordert somit eine Rechtsgrundlage.

Für KI liegt im Kontext der Anonymisierung eine große Herausforderung im Umfang des ihr verfügbaren »Hintergrundwissens«, 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.

Zusammenarbeit mit Dienstleistern

Die Zusammenarbeit mit Dienstleistern beim Einsatz von KI kann sich insbesondere dahingehend beschwerlich gestalten, dass die gesetzlich geforderten Kontrollmöglichkeiten gegenüber Dienstleistern häufig ihre Grenzen darin finden dürften, dass es typischerweise nicht den Interessen der Hersteller von KI-Modellen entspricht, für die Leistungsfähigkeit der KI entscheidende Faktoren wie den Algorithmus oder die Trainingsdaten- und Methoden offenzulegen. Solange dies auch nicht z. B. durch Zertifizierungen o. Ä. reguliert ist, entsteht somit ein »Blackbox«-Problem, durch das wesentliche Aspekte der KI gegebenenfalls nicht hinreichend kontrolliert werden können, um der Rechenschaftspflicht der DSGVO Genüge zu tun.

3.2 IT-Sicherheit

Effektive Cybersicherheit (= IT-Sicherheit) ist insbesondere für Hochrisiko-KI-Systeme essenziell und erfordert u. a. auch durch den Gesetzgeber (vgl. AI Act, Art. 15) sowohl effektive Schutzmaßnahmen als auch hinreichende Nutzerinformationen über das jeweils realisierte Cybersicherheitsschutzlevel (vgl. AI Act, Art. 13).

Dabei unterscheiden sich die meisten KI-Systeme bzgl. Cybersicherheitsrisiken und den dafür anwendbaren Schutzmaßnahmen **nicht** wesentlich von klassischen IT-Systemen bzw. klassischen IT-Produkten. Folglich betrachtet u. a. die aktuell entstehende, horizontale Regulierung zur Produkt-Cybersicherheit in der EU (»Cyber Resilience Act«, CRA) die Anforderungen zur Cybersicherheit von Hochrisiko-KI-Systemen aus Artikel 15 des AI Acts als erfüllt, sofern das KI-System die grundlegenden Cybersicherheitsanforderungen des CRA erfüllt (vgl. Art. 8 CRA).

Dennoch gibt es einige KI-spezifische Angriffspunkte (»AI specific vulnerabilities«), wenngleich diese nicht zwangsweise zur Cybersicherheit zugeordnet werden. Im Folgenden ein kurzer Überblick der bekanntesten KI-spezifischen Angriffspunkte.

Datenvergiftung (»Data Poisoning«): Bei diesem Angriff werden schädliche, falsche oder irreführende Daten ins Trainingsset eingeschleust, um die Ergebnisse der KI zu manipulieren. Ziel des Angreifers ist die Erzeugung falscher oder tendenziöser Ergebnisse.

Da sich dieser Angriff letztlich auf einen unzureichenden Schutz der Zugriffsberechtigungen bzw. die Integrität oder Authentizität der Trainingsdaten zurückzuführen lässt, kann und sollte man diesen Bedrohungen mit klassischen Schutzmaßnahmen der Cybersicherheit (z. B. Zugriffskontrolle, Signaturen etc.) effektiv begegnen.

Eingabeangriffe (»Input Attacks« auch »Evasion Attacks« oder »Adversarial Examples«): Angreifer versuchen, durch speziell präparierte Eingaben (wie Textprompts) dem KI-System Antworten zu entlocken, die vom Anbieter des KI-Systems nicht vorgesehen oder erlaubt sind.

Da diese Angriffspunkte v.a. in der (unzureichenden) Funktionalität des KI-Systems begründet sind, sind hier v.a. Maßnahmen zur Verbesserung der Funktionalität, Qualität und Robustheit angebracht.

Modell-Extraktion (»Model Extraction Attack«): Angreifer extrahieren Informationen über die Architektur oder die Parameter eines KI-Modells, um ein ähnliches Modell zu erzeugen. Damit könnten sie dann ein eigenes KI-System trainieren oder das Originalmodell gezielt kompromittieren.

Bei diesem KI-spezifischen Angriffspunkt handelt es sich im Wesentlichen um (i. d. R. legales; vgl. § 3 Abs. 1, GeschGehG) Reverse-Engineering und ist damit nur unzureichend via Cybersicherheit direkt adressierbar, sondern eher über funktionale Einschränkungen (z. B. begrenzte Abrufe pro Zeiteinheit).

Denial-of-Service (DoS): Bei diesem Angriff geht es darum, das KI-System durch eine Flut von Anfragen zu überlasten, sodass es für legitime Nutzer unzugänglich wird.

Als Gegenmaßnahmen werden empfohlen:

- Robustes Training: Durch die Verwendung von Techniken wie »Out-of-Distribution Detection« oder »Adversarial Training« kann das Modell resistent gegen Datenvergiftung gemacht werden.
- Zugriffskontrollen und Verschlüsselung: Strenge Zugriffskontrollen und Verschlüsselungsmechanismen können dabei helfen, unbefugten Zugriff auf das Modell und seine Parameter zu verhindern, was die Gefahr einer Modell-Extraktion mindert.
- Input-Sanitization: Durch die Überprüfung aller Eingabedaten auf Anomalien oder manipulierte Inhalte können viele Arten von Eingabeangriffen abgewehrt werden. Für diese Überprüfung kann wiederum KI eingesetzt werden.
- Differential Privacy: Mit Techniken der Differential Privacy wird versucht, die Identifizierbarkeit von Einzelpersonen in den Daten zu minimieren. So werden Reidentifikationsangriffe erheblich erschwert.
- Rate Limiting: Eine einfache, aber effektive Methode zur Abwehr von DoS-Angriffen ist die Begrenzung der Anzahl der Anfragen an das System pro Zeiteinheit.
- Regelmäßige Überprüfung und Updates: Eine ständige Überwachung der Systemaktivitäten und regelmäßige Aktualisierungen sind unerlässlich, um auf neue Bedrohungsvektoren reagieren zu können.

3.3 Schutzrechte

3.3.1 Urheberrecht

3.3.1.1 Schutzfähigkeit des Outputs

Schutzfähigkeit als Werk

Gemäß § 2 Abs. 2 UrhG sind lediglich Werke persönlicher geistiger Schöpfung urheberrechtlich geschützt. Das aktuelle Verständnis des Urheberrechts setzt einen direkten Bezug zwischen Mensch und geistiger Schöpfung voraus. Daher sind maschinell erzeugte Werke vom Urheberrechtsschutz ausgenommen. Dies basiert auf der Annahme, dass nur ein Mensch eine persönlich-geistige Schöpfung hervorbringen kann. In diesem Sinne sind Werke, die von Künstlicher Intelligenz generiert wurden, nicht schutzfähig, da ihnen der Bezug zu einem menschlichen Urheber fehlt.

Ausnahmsweise können KI-Erzeugnisse urheberrechtlichen Schutz genießen, wenn ein Mensch ein KI-System lediglich als Werkzeug verwendet, die gestalterischen Entscheidungen allerdings selbst trifft (sog. »KI-assisted works«). Ausschlaggebend ist der Einfluss des Nutzers auf das Erzeugnis. Ein urheberrechtlicher Schutz kann daher dann in Erwägung gezogen werden, wenn der menschliche Nutzer dem KI-System derart präzise Anweisungen gibt, dass die endgültige Gestaltung des Werkes bereits festgelegt ist und die KI lediglich diese Gestaltung umsetzt. Ein entsprechend kreativer Einfluss des Nutzenden auf der Input-Ebene ist zumindest denkbar, wenn die Eingabeinformationen oder Prompts so konkret ausgestaltet sind, dass in ihnen eine persönliche geistige Schöpfung gesehen werden kann. Denn in diesem Fall nutzt der Mensch den kreativen Gestaltungsspielraum beim Schaffen des Werkes selbst aus und verwendet die KI nur als technisches Hilfsmittel. Allerdings werden die Eingabeinformationen oder Prompts in der Regel eher eine Idee oder ein Konzept liefern, auf deren Basis die KI dann arbeitet, was alleine jedoch nicht für die Annahme eines maßgeblichen Einflusses ausreichen würde. Abzustellen ist auf den konkreten Einzelfall.

In eine ähnliche Richtung argumentiert das United States Copyright Office: Es erklärt in seiner Richtlinie »Copyright Registration Guidance: Works Containing Material Generated by Artificial Intelligence«, dass nur Werke, die das Produkt menschlicher Kreativität sind, urheberrechtlich geschützt werden können. Es lehnt die Registrierung von Werken ab, die ausschließlich von einer KI-Technologie ohne menschlichen Beitrag erzeugt wurden. Es akzeptiert jedoch die Registrierung von Werken, die sowohl menschliche als auch KI-generierte Elemente enthalten, sofern die menschlichen Elemente ausreichend originell und kreativ sind.

Demgegenüber gibt es einige Länder, die den Programmierer der Künstlichen Intelligenz als Urheber sehen. Das sind insbesondere, aber nicht abschließend, Großbritannien (Section 9 (3) Copyright, Designs and Patents Act 1988), Neuseeland (Section 5 (2) lit. a Copyright Act 1994) und Irland (Section 21 lit. f Copyright and Related Rights Act 2000). In allen drei zitierten Gesetzen findet sich folgende Formulierung: *»the author shall be taken to be the person by whom the*

arrangements necessary for the creation of the work are undertaken«. Unter »arrangements necessary for the creation« ist die Erstellung der Künstlichen Intelligenz, also des Algorithmus, zu verstehen.

Insgesamt kann man also festhalten, dass an dem von KI generierten Output in der Regel kein Urheberrecht entsteht. Ein urheberrechtlicher Schutz kann jedoch durch eine hinreichend kreative Bearbeitung des Outputs erreicht werden. Dies bedeutet, dass der Output einer KI zur urheberrechtlichen Schutzfähigkeit gelangen kann, wenn er von einer natürlichen Person so bearbeitet wird, dass ihm ein eigener schöpferischer Gehalt zukommt. Soweit die KI lediglich als Werkzeug verwendet wird, ist es für den Verwender notwendig den Schaffungsprozess hinreichend zu dokumentieren. Denn nach allgemeinen Grundsätzen hat derjenige den urheberrechtlichen Schutz darzulegen und zu beweisen, der sich auf diesen beruft. Soweit darauf abgestellt wird, dass Künstliche Intelligenz als Werkzeug eingesetzt wurde, muss also der Verwender seinen eigenen Beitrag darlegen und beweisen. Da bisher menschliche Schöpfungen der Regelfall waren, wurde diese meist ungeprüft unterstellt. Dieser Nachweis für einen eigenen menschlichen schöpferischen Beitrag unter Nutzung einer KI als Werkzeug dürfte künftig deutlich schwieriger werden, spätestens wenn KI-Erzeugnisse für bestimmte Ergebnisse eher Regel statt Ausnahme sind. Dann bedarf es neuer bzw. veränderter Lösungen zum Nachweis der persönlichen geistigen Schöpfungen.

Schutzfähigkeit über Leistungsschutzrechte

Das Urheberrecht bietet neben dem Werkschutz auch den Schutz verwandter Schutzrechte, den sogenannten Leistungsschutzrechten gemäß den §§ 70ff. UrhG.

Die Voraussetzungen für den Schutz von Leistungsschutzrechten nach UrhG variieren stark. Sie sind, anders als das Schutzrecht für Werke nach § 2 Abs. 2 UrhG, nicht an die Erfüllung einer persönlichen geistigen Schöpfung geknüpft, sondern zielen auf »Leistungen anderer Art« ab, welche »der schöpferischen Leistung des Urhebers ähnlich sind oder in Zusammenhang mit den Werken der Urheber erbracht werden« (siehe amtliche Begründung zum Entwurf des Urheberrechtsgesetzes, Bundestagsdrucksache BT-Drucksache 4/270 vom 23. März 1962, S. 33 f.).

Welches Leistungsschutzrecht im konkreten Fall in Betracht kommt, hängt von der Art des generierten Outputs ab. Denkbar ist mit Blick auf KI generierten Output das Leistungsschutzrecht des Tonträgerherstellers nach § 85 f. UrhG, der Laufbildschutz nach § 95 UrhG, sowie in einem gewissen Rahmen das Recht des Datenbankherstellers nach § 87a – § 87e UrhG.

Soweit der Output einer Künstlichen Intelligenz in Tonfolgen besteht, so können diese potenziell unter das Tonträgerherstellerrecht fallen. Dieses Recht dient einem Investitionsschutz, geschützt wird die wirtschaftliche, technische und organisatorische Leistung (BGH, Urteil vom 20. 11. 2008 - I ZR 112/06 GRUR 2009, 403, 404). Eine persönliche geistige Schöpfung ist dagegen nicht erforderlich. Daher könnten auch KI-generierte Tonfolgen unter den Schutzbereich fallen, sobald diese erstmals fixiert werden.

Vergleichbar verhält es sich beim Recht des Laufbildschutzes nach § 95 UrhG. Auch dieses Recht dient dem Schutz der wirtschaftlichen und organisatorischen Verantwortung (BGH Urt. v. 6. 2.

2014 – I ZR 86/12, GRUR 2014, 363). Geschützt wird eine Abfolge von Bild- und ggf. Tonfolgen, durch die der Eindruck eines Bewegtbildes entsteht, ohne dass eine persönliche geistige Schöpfung vorliegen muss. Soweit eine KI einen derartigen Film als Output liefert, ist ein Laufbildschutz möglich.

Sowohl das Recht des Tonträgerherstellers als auch das Schutzrecht für Laufbilder stünden dann demjenigen zu, der die wirtschaftliche und organisatorische Verantwortung für die Erstellung des Outputs trägt. Dies wäre regelmäßig der Programmierer der Künstlichen Intelligenz bzw. das Unternehmen, das die Künstliche Intelligenz betreibt.

Das Datenbankherstellerrecht der §§ 87a ff. UrhG schützt nach Art und Umfang wesentliche Investitionen in die Herstellung, Überprüfung oder Darstellung einer Datenbank. Gemäß § 87a Abs. 2 UrhG ist derjenige der Rechtsinhaber, der diese Investitionen getätigt hat. Wichtig ist, dass es sich bei der Datenbank im urheberrechtlichen Sinne um eine Sammlung von Werken, Daten oder anderen unabhängigen Elementen handeln muss, die systematisch oder methodisch angeordnet und einzeln mithilfe elektronischer Mittel oder auf andere Weise zugänglich sind und deren Beschaffung, Überprüfung oder Darstellung eine nach Art oder Umfang wesentliche Investition erfordert. Die Elemente gelten zunächst als unabhängig, wenn sie einen eigenständigen Informationsgehalt aufweisen. Es muss möglich sein, die Elemente zu trennen, ohne, dass sie durch die Trennung ihren Informationsgehalt verlieren. Stellt das KI-generierte Ergebnis also eine Datenbank im Sinne der §§ 87a ff. UrhG dar, kommt das Leistungsschutzrecht des Datenbankherstellers in Betracht. Entscheidend ist dabei, dass dem Rechtsinhaber nach § 87b UrhG nur das ausschließliche Recht zur Vervielfältigung, Verbreitung und öffentlichen Wiedergabe der gesamten Datenbank oder eines nach Art oder Umfang wesentlichen Teils der Datenbank zusteht. Einzelne Datenbankelemente oder mehrere unterhalb der Wesentlichkeitsschranke sind vom Ausschließlichkeitsrecht nicht erfasst.

Regelmäßig nicht in Betracht dürfte hingegen das Schutzrecht für Lichtbildhersteller kommen. Die Rechtsprechung geht davon aus, dass das Leistungsschutzrecht des Lichtbildherstellers nach §§ 72 ff. UrhG ein Mindestmaß an persönlicher (wenn auch nicht schöpferischer) geistiger Leistung erfordert. Bei einem von Künstlicher Intelligenz geschaffenen und auf Algorithmen basierendem Output dürfte es hieran mangels persönlicher Leistung fehlen. Etwas anderes könnte nur dann in Betracht kommen, wenn ein Mensch einen bestimmenden Einfluss auf einen durch KI durchgeführten Prozess zur Erstellung von Lichtbildern hat.

Die Frage, ob von KI generierte journalistische Erzeugnisse unter das neue Recht des Presseverlegers fallen, ist aktuell unklar und in der juristischen Literatur umstritten. Teilweise wird darauf abgestellt, dass nur ein aus journalistischen Beiträgen zusammengestelltes Online-Angebot geschützt wird und keine Zusammenstellung computergenerierter Produkte. Hier wird man die weiteren Entwicklungen abwarten und auf Entscheidungen der Gerichte warten müssen.
Wegen des lückenhaften Schutzes sollte je nach Einzelfall über alternative Schutzkonzepte wie vertragliche Regelungen oder Kopierschutzmechanismen nachgedacht werden.

3.3.1.2 Haftung durch Eingabeinformation

Neben den selbst formulierten Handlungsanweisungen können Eingabeinformationen und Prompts auch geschützte Inhalte Dritter enthalten. Besonders relevant wird dies, wenn in den Eingabeinformationen oder Prompts ein geschütztes Werk direkt wiedergegeben wird, zum Beispiel in der Form: »Fasse den [geschützten Text XYZ] zusammen«, wobei der Text in die Eingabeinformationen bzw. den Prompt kopiert wird.

Es ist wichtig zu beachten, dass die zustimmungsfreie Nutzung urheberrechtlich geschützter Werke eine urheberrechtliche Verletzungshandlung darstellt, soweit keine urheberrechtliche Schranke vorliegt. Betreibt man ein neuronales Netzwerk nicht auf eigenständiger Hardware, die vom Internet isoliert ist, muss man davon ausgehen, dass man eine Vervielfältigung auf den Servern des KI-Anbieters veranlasst. Diese Server löschen nicht in jedem Fall die Eingabeinformationen bzw. Prompts, sondern können sie zur Optimierung der KI verwenden, was urheberrechtlich relevant sein kann.

Die urheberrechtlichen Schranken des § 44a UrhG (vorübergehende Vervielfältigung) und § 53 UrhG (Privatkopie) dürften in diesem Kontext nicht anwendbar sein. § 44a UrhG erlaubt die vorübergehende Vervielfältigung nur dann, wenn sie keine eigenständige wirtschaftliche Bedeutung hat. Die Vervielfältigung in KI-Servern ist jedoch nicht nur vorübergehend, sondern kann für Trainings- und Optimierungszwecke der KI längerfristig genutzt werden. § 53 UrhG erlaubt die Nutzung nur für den privaten Kreis und für den eigenen Gebrauch. Jedoch findet die Nutzung urheberrechtlich geschützter Werke in einer kommerziellen, öffentlichen Umgebung statt, was über den privaten Gebrauch hinausgeht.

Die faktischen Haftungsrisiken sind in diesem Zusammenhang jedoch zunächst als gering einzuschätzen. Das liegt daran, dass Informationen darüber, wer Vervielfältigungen auf den Servern der KI-Anbieter veranlasst, in der Regel nicht nach außen dringen. Allerdings ist besondere Vorsicht geboten, da durch solche Handlungsanweisungen die Wahrscheinlichkeit von Rechtsverletzungen durch den Output der KI erhöht wird (siehe dazu weiter unten).

3.3.1.3 Relevanz der Urheberrechte an KI-Trainingsdaten auf Hersteller- und Anwenderseite

Die Betrachtung des Herstellerhandelns im Kontext der Künstlichen Intelligenz (KI) ist essenziell, da Handlungen der Hersteller direkt Einfluss auf die Anwendungsphase der KI haben können. Ein zentraler Aspekt dabei ist die Vervielfältigungshandlung während des Trainingsprozesses. In einem ersten Schritt wird häufig eine Sammlung von urheberrechtlich geschützten Werken erstellt, was die Vervielfältigungsrechte gemäß § 16 UrhG berührt. Dies umfasst das Sammeln von Trainingsdaten und das Training der KI, bei dem die Werke im Sinne des § 16 UrhG vervielfältigt werden. Auch wenn ein KI-System Werke zur Kenntnis nimmt und daraus Informationen zieht, ist dies urheberrechtlich relevant, da das Speichern der Trainingsdaten in einem Korpus eine Vervielfältigung darstellt.

Eine zentrale Rolle beim Training vom KI-Systemen spielt § 44b UrhG. § 44b UrhG erlaubt die Vervielfältigung bei der Erstellung von Datensammlungen zum KI-Training unter den folgenden Voraussetzungen:

1. **Rechtmäßige Zugänglichkeit des Trainingsmaterials**: Nach § 44b Abs. 2 S. 1 UrhG müssen die verwendeten Werke rechtmäßig zugänglich sein, was der Fall ist, wenn der Nutzer es ohne Rechtsverstoß abrufen kann.

2. **Löschpflicht nach Abschluss des Trainings**: Gemäß § 44b Abs. 2 S. 2 UrhG müssen Vervielfältigungen nach Beendigung des Trainings gelöscht werden. Der Trainingskorpus muss also nach Abschluss des konkreten KI-Projekts vernichtet werden.

3. **Opt-out des Rechtsinhabers**: § 44b Abs. 3 UrhG ermöglicht Rechtsinhabern, die Nutzung ihrer Werke für Text- und Data-Mining durch einen maschinenlesbaren Opt-out (Nutzungsvorbehalt) auszuschließen. Dies bedeutet, dass die Schranke des § 44b UrhG nach § 44b Abs. 3 UrhG nur dann eingreift, wenn der Rechtsinhaber sich die Nutzung nicht vorbehalten hat. In Bezug auf im Internet veröffentlichte Werke verlangt § 44b Abs. 3 S. 2 UrhG, dass der Opt-out in maschinenlesbarer Form erklärt werden muss.

Soweit die vorstehenden Voraussetzungen erfüllt sind, ist die Vervielfältigung von Werken für die automatisierte Analyse von KI-Systemen zulässig. Die deutsche Bundesregierung hat zuletzt in einer parlamentarischen Anfrage ebenso wie Regelungen des AI Acts die Anwendbarkeit des § 44b UrhG, der seinen Ursprung in der EU DSM-Richtlinie hat, bestätigt.

Es existieren dennoch Meinungen, dass die gegenwärtige Rechtslage die Nutzung der Text- und Data-Mining-Schranke des § 44b UrhG als Rechtfertigungsbasis eventuell nicht unterstützt. Diese Ansicht basiert auf der Argumentation, dass die Verwendung reiner KI-Trainingsdaten lediglich als vorbereitende Maßnahme für spätere Analysen fungiert und somit nur indirekt zur Informationsgewinnung beiträgt. Es wird betont, dass, solange der primäre Zweck nicht die unmittelbare Informationsgewinnung ist – wie es bei der Datenanalyse der Fall wäre –, diese Vorgehensweise nicht von der genannten Norm abgedeckt ist.

In Deutschland ist daher ein rechtssicherer Weg über die Lizenzierung der zum Training genutzten Werke zu empfehlen.

Das Schutzlandprinzip im Urheberrecht ermöglicht es grundsätzlich, neuronale Netzwerke in Ländern mit geringeren Urheberrechtsstandards zu trainieren und anschließend als fertiges Produkt in Deutschland anzubieten, wodurch die Einschränkungen des § 44b UrhG praktisch umgangen werden können. Erwägungsgrund 60j des AI Acts nimmt diesen Umstand auf und stellt klar, dass ein außerhalb der EU trainiertes KI-System nur in der EU angeboten werden darf, wenn beim Training unionsrechtliche Urheberrechtsvorgaben beachtet wurden. Erwägungsgründe dienen der Auslegungshilfe und bilden keine direkte Anspruchsgrundlage. Es wird sich zeigen müssen, wie dieser Konflikt aufzulösen ist.

Für Anwender ergeben sich aus einem rechtswidrigen Training nur dann Probleme, wenn im Output Vervielfältigungen urheberrechtlich geschützter Werke auftreten. Dies kann vorkommen,

wenn die KI Trainingsdaten ganz oder teilweise im Output reproduziert, ein Phänomen, das als »Erinnern« bezeichnet wird. Die Forschung ist sich uneinig, wie häufig dies der Fall sein kann, aber das Risiko bleibt bestehen.

Zudem besteht die Frage, ob Vervielfältigungen der trainierten Werke im neuronalen Netzwerk vorliegen, was zu einer »Infizierung« des KI-Modells führen könnte. Die herrschende Meinung sieht die Informationsgewinnung aus geschützten Werken und die Anpassung der Gewichtungswerte des neuronalen Netzwerks nicht als eine solche Vervielfältigung an. Es gibt jedoch auch die Ansicht, dass die in das neuronale Netzwerk eingeflossenen Parameter eine Vervielfältigung darstellen könnten, was erhebliche Folgen hätte, da ein gezieltes Entfernen dieser Parameter oft nicht möglich ist.

Bei der Fortbildung der KI im Live-Betrieb muss zwischen Training mit eigenen und fremden Daten unterschieden werden. Bei fremden Daten sind die Grenzen des § 44b UrhG zu beachten (s. o.). Bei eigenen Daten müssen diese Grenzen des Urheberrechts nicht berücksichtigt werden,[8] sofern die entsprechenden Rechte für das Training vorliegen, was in der Regel vertragliche Regelungen mit Mitarbeitenden beinhaltet. Das Training stellt wahrscheinlich eine eigenständige Nutzungsart dar, die wirtschaftlich selbstständig verwertbar ist und im Zweifel vertraglich eingeräumt werden muss.

3.3.1.4 Rechte der KI-Hersteller am generierten Output

Die Entwickler der eingesetzten KI erhalten ganz regelmäßig kein Urheberrecht, wenn Dritte mit der KI Erzeugnisse generieren. Allenfalls und ganz ausnahmsweise, wenn der Programmierer bereits bei der Entwicklung des Programms substanzielle Gestaltungsentscheidungen zu dem konkreten Erzeugnis getroffen hat, kommt ein urheberrechtlicher Schutz zugunsten des Programmierers in Betracht. Das wird bei KI-Systemen regelmäßig nicht der Fall sein.

Es kann sein, dass die KI-Hersteller Rechte am Output durch die Nutzungsbedingungen zu erlangen suchen, oder zumindest den Einsatz der Erzeugnisse einschränken. Die Wirksamkeit dieser Regelungen ist dann im Einzelfall zu prüfen (etwa im Licht des AGB-Rechts) und nur schuldrechtlicher Natur.

3.3.1.5 Urheberrechtliche Haftung auf Anwenderseite

Ein relevantes Beispiel für die immaterialgüterrechtliche Haftung auf Anwenderseite ist der Einsatz von KI-generierten Erzeugnissen, wie Fotos, für Werbemaßnahmen. Dabei ist zunächst festzuhalten, dass der künstlerische Stil, ähnlich wie abstrakte Ideen, nicht unter den Schutz des Urheberrechts fällt. Jedoch können Urheberrechtsverletzungen entstehen, wenn der von der KI generierte Output Teile eines urheberrechtlich geschützten Werkes enthält, die für das Training der KI verwendet wurden. In diesem Zusammenhang ist zu beachten, dass die Haftung für solche

8 Andere Grenzen, die sich bspw. aus Persönlichkeitsrechten, dem Recht am eigenen Bild oder der eigenen Stimme sowie sonstigen personenbezogenen Daten ergeben, gelten jedoch weiterhin und sich zu beachten.

Vervielfältigungen unabhängig davon besteht, ob das Training der KI mit dem urheberrechtlich geschützten Material legal erfolgte.

Ein wesentlicher Maßstab für die Feststellung einer Urheberrechtsverletzung ist die Wiedererkennbarkeit des vorbestehenden, geschützten Werkes im KI-Output im Rahmen einer Gesamtbetrachtung. Hierbei ist entscheidend, dass nur jene Teile des vorbestehenden Werkes berücksichtigt werden, die tatsächlich urheberrechtlich geschützt sind. Eine Einzelfallbetrachtung ist daher unerlässlich.

Das Risiko von Urheberrechtsverletzungen kann durch bestimmte Maßnahmen reduziert werden. Eine Möglichkeit besteht darin, dass Nutzer in ihren Prompts nicht explizit geschützte Werke nennen. Weiterhin könnte in einigen Fällen die Pastiche-Schranke des § 51a UrhG anwendbar sein, allerdings ist der genaue Inhalt dieses Begriffs und die Reichweite der Schranke derzeit noch nicht abschließend geklärt.

Im Falle einer festgestellten Urheberrechtsverletzung stehen dem Rechteinhaber jedenfalls Unterlassungsansprüche zu. Zudem können, bei Nachweis von Verschulden, Schadensersatzansprüche geltend gemacht werden. Dabei ist zu beachten, dass im Urheberrecht bereits eine niedrigschwellige Fahrlässigkeit für die Begründung der Haftung ausreichend sein kann.

3.3.2 Know-how-Schutz und Geschäftsgeheimnisschutz

Bei der Anwendung von KI im Unternehmen sollte der Schutz von Geschäftsgeheimnissen iSd Geschäftsgeheimnisschutzgesetzes (GeschGehG) beachtet werden. Hierbei bestehen sowohl Chancen als auch Risiken: Zum einen kann der Geschäftsgeheimnisschutz den Urheberschutz an KI-Modellen, Trainingsdaten und Prompts sowie an KI-generiertem Output ergänzen bzw. etwaige Schutzlücken schließen. Zum anderen dürfen bestehende Geschäftsgeheimnisse nicht durch einen unbedarften Umgang bei der Eingabe in KI-Anwendungen gefährdet werden.

An Geschäftsgeheimnissen bestehen verkehrsfähige Schutzrechte, an denen der Berechtigte Lizenzen einräumen kann und die (unter Beachtung der Schutzvoraussetzungen) grundsätzlich eine umfassende Verwertung erlauben. Daraus resultiert ein mitunter erheblicher wirtschaftlicher Wert. Davon abzugrenzen ist der uneinheitlich verwendete Begriff des Know-hows. Dieser bezeichnet kein gewerbliches Schutzrecht, sondern beschreibt lediglich den faktischen Vorteil einer geheimen Information, mittels derer ein Unternehmen sich von Wettbewerbern abheben kann.

Der Geschäftsgeheimnisschutz wird in Deutschland durch das auf einer europäischen Richtlinie beruhende Gesetz zum Schutz von Geschäftsgeheimnissen gewährleistet. Eine Information wird nach dessen § 2 zum Geschäftsgeheimnis, wenn sie (i) geheim und deshalb von wirtschaftlichem Wert ist, (ii) sie Gegenstand von angemessenen Geheimhaltungsmaßnahmen ist und (iii) soweit ein berechtigtes Interesse an ihrer Geheimhaltung besteht. Damit hat das GeschGehG in inhaltlicher Hinsicht einen sehr breiten Anwendungsbereich.

Dies hat zur Folge, dass jede Kategorie von Informationen dem Geheimnisschutz zugänglich ist, wenn und soweit sie die vorgenannten Voraussetzungen erfüllt. Hierunter können mit Blick auf KI-Anwendungen beispielsweise trainierte und untrainierte KI-Modelle, Algorithmen, Trainingsdaten sowie durch KI-Anwendung generierter und anderweitig nicht schutzfähiger Output von KI-Anwendungen (grafische Darstellungen, Video- und Musiksequenzen, Software und Computerprogramme sowie technische (Produkt-)Gestaltungen wie CAD oder Halbleiterdesigns) fallen. Die inhaltlichen Beschränkungen anderer Rechte des Geistigen Eigentums, wie z. B. die fehlende persönliche, geistige Schöpfung iSd Urheberrechts, bestehen hier nicht. Denn anders als das Urheberrecht verlangt das GeschGehG gerade keinen menschlichen Beitrag bei der Entstehung des Geschäftsgeheimnisses, sodass rechtlicher Schutz auch in Abwesenheit eines menschlichen Urhebers in Betracht kommt.

Geheim ist eine Information, wenn sie weder insgesamt noch in der genauen Zusammensetzung ihrer Bestandteile Dritten allgemein bekannt oder ohne Weiteres zugänglich ist. Es liegt in der Verantwortung des Rechtsinhabers, dies sicherzustellen und mit angemessenen Geheimhaltungsmaßnahmen (wie dem Abschluss von Verschwiegenheitsvereinbarungen insbesondere mit Kunden, Geschäftspartnern und Angestellten, der Beschränkung des Zugriffs auf die geschützte Information nach dem Need-to-know-Prinzip, IT-Sicherheitsmaßnahmen- und Schulungen) dauerhaft aufrechtzuerhalten.

3.3.2.1 Schutz von KI-Modellen nach dem GeschGehG

KI-Modelle sind nach herrschender Auffassung wohl nur beschränkt dem Schutz als Computerprogramm nach § 69a UrhG zugänglich, nämlich soweit sie das Ergebnis einer eigenen geistigen Schöpfung ihres Urhebers sind. Hierbei ist zu unterscheiden. Während der dem KI-Modell zugrundeliegende Algorithmus bei hinreichender Schöpfungshöhe sowie die Einbettung des KI-Modells in einen Application-Layer als Computerprogramm angesehen werden können, ist das KI-Modell hinsichtlich der im Wege des Training-Prozesses ausgeprägten Gewichtungen der neuronalen Verknüpfungen kein Computerprogramm. Denn diese Gewichtungen sind automatisiert durch den Trainings-Prozess entstanden. Sie sind gewissermaßen durch die KI selbst programmiert und daher dem urheberrechtlichen Computerprogrammschutz nicht zugänglich.

Diese Gewichtungen sind jedoch für die Arbeitsweise des KI-Modells entscheidend und stellen damit dessen wirtschaftlichen Wert dar. Da diese nicht Teil des geschützten Computerprogramms sind, ist ihr Schutz durch entsprechende Geheimhaltungspflichten sicherzustellen. Es sind bereits Fälle bekannt geworden, in denen durch entsprechendes Prompt-Engineering die Gewichtungen eines KI-Modells rekonstruiert werden konnten. So finden sich oftmals in den Nutzungsbedingungen von KI-Anbietern Klauseln, die das gezielte Reverse Engineering des KI-Modells iSd § 3 Abs. 1 Nr. 2 GeschGehG sowie sonstige Handlungen, die auf dessen Rekonstruktion abzielen, untersagen. Hierdurch soll die Extraktion der Gewichtungen zum Zweck des Nachbaus des KI-Modells unterbunden werden.

3.3.2.2 Schutz von KI-generiertem Output nach GeschGehG

Ob und inwieweit KI-generierter Output dem Geschäftsgeheimnisschutz zugänglich ist, hängt von der Art und Weise des Einsatzes der KI ab. Entscheidend für die Bewertung als »geheim« ist, ob ein Dritter auf den Output zugreift oder zugreifen kann. Nutzt der Anwender eine selbst entwickelte (proprietäre) KI – sind also KI-Anwender- und Anbieter personenidentisch – oder nutzt der Anwender zwar die KI eines Drittanbieters, aber betreibt diese lokal auf seiner Infrastruktur (on premise, s. oben), dann kann der Output geheim sein. Entscheidend hierfür ist dann lediglich das Informationsmanagement im Unternehmen des KI-Anwenders. Dass der Inhalt ursprünglich durch eine KI erzeugt wurde, ist für dessen Schutzfähigkeit unerheblich.

Nutzt der Anwender die KI dagegen als Cloud-Produkt des Drittanbieters (AI as a Service – AIaaS, s. oben), ist jedenfalls faktisch eine Zugangsmöglichkeit des Anbieters eröffnet. In diesem Fall kommt es auf die Nutzungsbedingungen des Anbieters an: Verpflichtet sich dieser zu Vertraulichkeit oder gestatten sie ihm den Zugriff auf den Output etwa zu Zwecken der Produktentwicklung und Qualitätskontrolle? Nur im ersten Fall besteht Geschäftsgeheimnisschutz am Output. Im sogenannten freemium-Vertrieb machen KI-Anbieter die Geheimhaltung teilweise schlicht davon abhängig, ob der Anwender eine kostenfreie oder kostenpflichtige Version der KI verwendet.

Aus Sicht des Anwenders empfiehlt sich hier unbedingt die genaue Prüfung der Nutzungsbedingungen. Sichert der Anbieter die Vertraulichkeit des Outputs zu, sodass dieser nur »für die Augen des Nutzers« bestimmt ist, kann der generierte Output als Geschäftsgeheimnis des Anwenders betrachtet werden. Sofern die Geheimhaltungskette über den gesamten Produktvertrieb aufrechterhalten werden kann, können auf Grundlage des Geschäftsgeheimnisschutzes auch tragfähige Lizenzmodelle entwickelt werden.

3.3.2.3 Prompts als Geschäftsgeheimnisse

Auch Prompts können Geschäftsgeheimnisse darstellen oder enthalten. Da die Entwicklung und Gestaltung von Prompts mitunter ganz erheblichen zeitlichen und personellen Aufwand erfordern können, erscheint die Schutzfähigkeit auch interessengerecht. Aus der Praxis sind zudem Fälle bekannt, dass Prompts exklusiv für Kunden entwickelt werden und daher auch ein schutzwürdiges Wirtschaftsgut darstellen.

Wie schon bei den Ausführungen zum Schutz von KI-generiertem Output nach dem GeschGehG ist auch hier die Art und Weise des KI-Betriebs (proprietär, on premise oder AIaaS, s. dazu genauer oben) entscheidend. Daher sind unbedingt die einschlägigen Nutzungsbedingungen genau in den Blick zu nehmen. Behält sich der Anbieter darin das Recht vor, seine KI mit den Prompts und Input-Daten der Anwender zu trainieren, resultiert daraus ein erhebliches Risiko für die enthaltenen Geschäftsgeheimnisse. Teilweise finden sich in den Nutzungsbedingungen von KI-Anbietern sogar Regelungen, nach denen sich diese die einfachen, weltweiten Nutzungsrechte an Prompts einräumen lassen, um diese auch über die Trainingszwecke hinaus weiterzuverwerten.

Das Risiko des Geheimnisverlusts gilt aber nicht nur im Verhältnis zum KI-Anbieter. Darüber hinaus kann nicht ausgeschlossen werden, dass die KI die zum Training verwendeten Informationen oder Teile davon gegenüber Dritten als Output wieder offenlegt. Hierdurch riskiert der Anwender den Verlust bestehender Geschäftsgeheimnisse, wenn diese zunächst als (Teil eines) Prompt in eine cloudbasierte KI (AIaaS) eingegeben und damit z. B. zu Trainingszwecken in das KI-Modell aufgenommen werden. So wurde z. B. der Fall bekannt, in dem die Softwareentwickler eines IT-Unternehmens Quellcode zur Kontrolle von programmiertechnischen Fehlern in eine KI-Anwendung eingespeist haben. Hierdurch wurde das Risiko begründet, dass dieser Quellcode – je nach konkreter Gestaltung der KI-Anwendung – bei entsprechender Ansteuerung des KI-Modells von diesem auch gegenüber unbefugten Dritten reproduziert wird.

3.3.3 Weitere Schutzrechte

3.3.3.1 Marken

Aus der Nutzerperspektive sind Marken für die Eingabe (Prompt) in eine KI-Anwendung und den Output relevant. Für das Training einer KI-Anwendung mit Daten, die Marken enthalten, gelten die Ausführungen zu Ziffer 3.3.1. Im Trainingsprozess wird die digitale Repräsentation der Marke, aber nicht die Marke in ihrer Markenfunktion genutzt.

Die Markenverwendung für einen Prompt ist zunächst eine maschinenlesbare Bestimmung der Marke. Eine Markennutzung liegt vor, wenn durch die Bestimmung der Marke ein Output geschaffen wird, der diese Marke wiedergibt. Allerdings: anders als z. B. bei der Buchung eines Keywords führt die Wiedergabe der Marke im KI-Output noch nicht zu einer Nutzung im geschäftlichen Verkehr. Im Fall der Keywords wird die Anzeige einer elektronischen Werbung bewirkt (Google Ad). Der KI-Output bedarf erst einer weiteren Verwendung, anhand der sich die Rechtmäßigkeit der Markennutzung beurteilt.

Zu berücksichtigen ist, dass die KI auch ohne Angabe einer Marke im Prompt, eine Marke in einem Output wiedergeben kann. Dies ist insbesondere bei bekannten Marken der Fall, die z. B. nahezu repräsentativ für eine bestimmte Warengattung sind. Viele dieser Marken sind nicht nur unter einem Bildzeichen oder Wort eingetragen, sondern sie können auch z. B. in einer Produktform geschützt sein. Zudem gibt es Farben oder auch Muster, die dem Markenschutz unterstehen. Der Katalog der Markenformen ist nicht abgeschlossen. Häufig sind diese »besonderen« Marken auch bekannte Marken, die insbesondere gegen Rufausbeutung geschützt sind.

Vor der Nutzung eines KI-Outputs muss daher geklärt werden, ob sich darin fremde Marken finden. KI-Output, der erkennbar eine fremde geschützte Marke wiedergibt, kann im geschäftlichen Verkehr nicht ohne Zustimmung des Inhabers für die unter der Marke geschützten Waren und Dienste verwendet werden. Sollte es sich um eine bekannte Marke handeln, ist diese vor rufausbeutenden oder rufschädigenden Handlungen geschützt. Zudem kann sich der Inhaber einer bekannten Marke gegen Verwässerung wenden.

KI-Output, der eine humorvolle, kritische oder kreative Auseinandersetzung mit der Marke vermittelt, kann nach Abwägung im Einzelfall eine zulässige Form der Meinungsäußerung oder der künstlerischen Freiheit sein, sodass die Rechte des Markeninhabers zurücktreten. Die Verwendung von Outputs, die die Marke trivialisieren oder schlicht entwerten, bleibt rechtswidrig.

Der Output, den KI-Systeme generieren, kann insbesondere als Wort- oder Bildmarke markenrechtlich geschützt werden. Der Schutz einer Marke entsteht in der Regel durch deren Eintragung in das zuständige Markenregister, wobei es auch den Sonderfall des Schutzes durch Verkehrsgeltung gibt. Bei der Überlegung, KI-generierte Inhalte als Marke anzumelden, sollte im Vorfeld gründlich geprüft werden, ob bereits ältere Markenrechte bestehen, um mögliche Konflikte mit Rechteinhabern zu vermeiden.

3.3.3.2 Designs

KI-Output kann, ob ausdrücklich in einem Prompt vorgegeben oder nicht, ein geschütztes Design wiedergeben. Ein Design wird dann unerlaubt nachgeahmt, wenn es bei dem informierten Benutzer keinen anderen Gesamteindruck erweckt als das geschützte Design.

Eine Abbildung eines geschützten Designs kann auch eine unerlaubte Nachahmung sein, wenn sie die wesentlichen Merkmale des geschützten Designs übernimmt oder nur unwesentliche Abweichungen aufweist. Dies gilt auch für eigenständige Gestaltungen, sofern diese an den charakteristischen Elementen des Designs orientiert wurde.

KI-Output, der ein geschütztes Design wiedergibt, kann nicht ohne Zustimmung des Inhabers verwendet werden. Allerdings ist eine Wiedergabe zulässig sein, sollte die Verwendung des KI-Outputs in einem bestimmten Kontext erfolgen, wie z. B. für Lehrzwecke oder als Bildzitat.

3.3.3.3 Persönlichkeitsrechte

KI-Output kann, ob ausdrücklich über einen Prompt eingegeben oder aufgrund der Datensätze des KI-Tools generiert, Personen oder auch Merkmale von Personen, z. B. die Stimme, erkennbar machen.

Die Verbreitung des KI-Outputs, der eine Person identifizierend beinhaltet, kann das Recht auf informationelle Selbstbestimmung, das die Privatsphäre und die Ehre einer Person schützt, verletzen, sofern die Verbreitung gegen ihren Willen erfolgt. Das Allgemeine Persönlichkeitsrecht schützt z. B. auch die stimmliche Identität einer Person. Darüber können Rechte am eigenen Bild betroffen sein, wenn es an der Einwilligung der Person fehlt.

Sollte eine Person über ihre Darbietungen Eingang in einem Prompt oder KI-Output finden, gelten die Schutzinstrumente des Urheber- und Leistungsschutzrechts (s. o. Ziffer 3.3.1) sowie des Datenschutzrechts (s. o. Ziffer 3.1.1) und anderer die Persönlichkeitsrechte betreffender Gesetze. Im Einzelfall kann, unter Abwägung widerstreitender grundgesetzlicher sonstiger Interessen, die Verwendung des KI-Outputs gestattet sein.

3.4 Haftungsrechtliche Aspekte

3.4.1 Typische Risiken beim KI-Einsatz

Der Einsatz generativer KI in Unternehmen birgt eine Reihe von Risiken, die sowohl operationelle als auch rechtliche Implikationen haben können. Zu den häufigsten Herausforderungen zählen Fehler und Mängel in den Trainingsdaten, die zu unzuverlässigen oder irreführenden Ergebnissen in der Einsatzphase führen können, wie Vorurteile, Ungenauigkeiten, ungenaue, unethische und schädliche Inhalte wie unbeabsichtigte Diskriminierung, die durch verzerrte Daten oder Algorithmen entstehen kann. Darüber hinaus können KI-Systeme fehlerhafte Schlussfolgerungen ziehen oder aufgrund ungeeigneter Algorithmen und Entscheidungslogiken in spezifischen Anwendungskontexten versagen.

In Produkten und Dienstleistungen eingebettete bzw. verwendete fehlerhafte Ergebnisse können zu Unzufriedenheit der Kundinnen und Kunden, Vertragsbruch, (unverhältnismäßiger) Haftung und Vertragsstrafen, Geschäfts- und Umsatzverlusten, Reputationsschäden, Bußgeldern, Problemen in der Lieferkette usw. führen.

Sicherheitslücken und Manipulationsanfälligkeit stellen weitere kritische Risiken dar, ebenso wie potenzielle Hardwarefehler, die die Funktionalität der KI beeinträchtigen können. Zudem müssen Unternehmen die möglichen Rechte Dritter am Output der KI berücksichtigen, um Urheberrechtsverletzungen oder andere rechtliche Konflikte zu vermeiden.

Zuletzt bestehen Compliance-relevante Risiken durch Nichteinhaltung der gesetzlichen und/oder vertraglichen Verpflichtungen, Verstoß gegen Anforderungen an Dokumentation, Nachvollziehbarkeit, Transparenz und Rechenschaftspflicht und Risiken des Third-Party-Managements (bei Nutzung von fremder KI).

3.4.2 Haftung des Anwenders nach derzeitigen Haftungsregelungen

3.4.2.1 Vertragliche Haftung

Gegenüber dem KI-Hersteller

Bei der vertraglichen Haftung gegenüber dem Hersteller von KI-Systemen spielen mögliche Pflichtverletzungen eine zentrale Rolle. Unternehmen, die KI-Lösungen einsetzen, müssen die Vertragsbedingungen genau prüfen, insbesondere im Hinblick auf Garantien, Leistungszusicherungen und Verantwortlichkeiten. Eine Pflichtverlet-

zung kann vorliegen, wenn die KI nicht wie vereinbart funktioniert, beispielsweise aufgrund von Fehlern in der Software oder unzureichender Leistungsfähigkeit.

In vielen Fällen versuchen Hersteller, ihre Haftung durch vertragliche Begrenzungen zu minimieren. Dies kann Haftungsausschlüsse für bestimmte Arten von Schäden oder Haftungsobergrenzen umfassen. Unternehmen sollten diese Bedingungen sorgfältig bewerten und bei Bedarf Verhandlungen führen, um einen angemessenen Schutz ihrer Interessen zu gewährleisten.

Gegenüber den Endnutzern bzw. Kundinnen und Kunden (z. B. bei Angebot eines Chatbots, den Kundinnen und Kunden »bedienen«)

Im Verhältnis zum Endnutzer oder Kunden, beispielsweise beim Einsatz eines Chatbots, können ebenfalls vertragliche Haftungsfragen aufkommen. Mögliche Pflichtverletzungen in diesem Kontext umfassen die Nichterfüllung von Leistungszusagen, wie die Genauigkeit oder Zuverlässigkeit der durch den Chatbot generierten Antworten.

Haftungsbegrenzungen gegenüber Kundinnen und Kunden gestalten sich oft komplex, insbesondere angesichts der strengen Vorschriften des Verbraucherschutzrechts. Unternehmen, die eine Begrenzung ihrer Haftung anstreben, müssen gewährleisten, dass ihre Haftungsbeschränkungen nicht nur rechtlich durchsetzbar sind, sondern auch den geltenden rechtlichen Anforderungen entsprechen. Dies schließt eine sorgfältige Überprüfung und Gestaltung der Allgemeinen Geschäftsbedingungen (AGB) ein, um sicherzustellen, dass diese klar, verständlich und nicht unangemessen benachteiligend für den Verbraucher sind.

Ein konkretes Beispiel für eine Haftungssituation könnte die Richtigkeit der generierten Inhalte sein. Falsche oder irreführende Informationen, die durch einen Chatbot bereitgestellt werden, könnten zu Schäden oder Verlusten für den Nutzer führen. Zudem können Rechtsverletzungen entstehen, wenn der Chatbot geschützte Inhalte Dritter ohne entsprechende Lizenz verwendet. In solchen Fällen könnten Unternehmen sowohl gegenüber ihren Kunden als auch gegenüber den Rechteinhabern haftbar gemacht werden.

3.4.2.2 Deliktsrechtliche Haftung

Verschuldensabhängige Haftung

Wesentliche Anspruchsgrundlage der deliktsrechtlichen verschuldensabhängigen Haftung ist § 823 Abs. 1 BGB. Hiernach ist der zurechenbare Verursacher nur verantwortlich, wenn er sich pflichtwidrig verhalten hat, was sich insbesondere aus der Verletzung von Verkehrssicherungspflichten ergibt. In Bezug auf KI-Systeme können nicht nur die Hersteller oder Entwickler, also die sog. Produzenten, sondern auch der Anbieter oder Anwender als zurechenbarer Verursacher in Betracht kommen. In der Praxis bedeutet dies, dass je nach Rolle und Einfluss auf das Produkt oder die Dienstleistung unterschiedliche Verkehrssicherungspflichten gelten können. Diese Pflichten zielen darauf ab, Schäden für Personen oder Eigentum zu vermeiden und die Sicherheit zu gewährleisten.

Für den **Hersteller** sind die Verkehrssicherungspflichten besonders umfassend. Sie umfassen Konstruktions-, Fabrikations-, Instruktions-, Produktbeobachtungs- und Gefahrabwendungspflichten. Der Hersteller ist verantwortlich dafür, dass das Produkt sicher entworfen, hergestellt und mit den notwendigen Informationen und Warnungen versehen wird.

Anbieter eines Produktes oder einer Dienstleistung, die beispielsweise KI-Systeme nutzen oder integrieren, können ebenfalls Verkehrssicherungspflichten unterliegen. Diese umfassen üblicherweise die Pflicht, sicherzustellen, dass die angebotenen Produkte oder Dienstleistungen keine Gefahr für die Nutzer oder Dritte darstellen. Auch müssen Anbieter über eventuelle Risiken informieren und bei bekannt gewordenen Gefahren reagieren.

Anwender oder **Nutzer** eines Produktes, besonders im Bereich der KI, können ebenfalls bestimmte Pflichten haben, insbesondere wenn sie die KI-Systeme in einem Kontext nutzen, der Risiken für andere mit sich bringt. Hier können Pflichten wie die sachgemäße Anwendung gemäß den Anweisungen des Herstellers oder Anbieters und die Berücksichtigung von Sicherheitshinweisen relevant sein. Bei Anwendern, die das Produkt in einem professionellen oder kommerziellen Kontext nutzen, können die Verkehrssicherungspflichten umfangreicher sein, insbesondere wenn sie das Produkt modifizieren oder in einer Weise verwenden, die über den ursprünglichen Anwendungsbereich hinausgeht.

Eine Haftung gemäß § 823 Abs. 1 BGB kommt nur dann in Betracht, wenn dem Anspruchsgegner ein schuldhafter Verstoß gegen die Verkehrssicherungspflichten nachgewiesen werden kann. Der relevante Sorgfaltsmaßstab wird durch die spezifischen Umstände des Einzelfalls bestimmt. Grundsätzlich hat der Geschädigte sämtliche anspruchsbegründende Tatsachen, also objektiven Tatbestand, Verschulden, Schaden und Kausalität, zu beweisen. Diese Beweislastverteilung gilt gegenüber dem Anbieter und Anwender der KI.

Sonderfall: Hersteller und Entwickler (Produzenten)

Für verschuldensabhängige Ansprüche aus § 823 I BGB gegenüber Herstellern und Entwicklern aus § 823 I BGB, die sog. Produzentenhaftung, sieht die Rechtsprechung eine Beweislastumkehr vor. Da der Geschädigte selbst wenig Einsicht in den Betrieb des Schädigers hat und er daher schwierig seinen Anspruch darlegen und beweisen kann, muss der Schädiger hier beweisen, dass ihn kein Verschulden trifft, er also keine Verkehrspflicht verletzt hat. Der Geschädigte muss lediglich nachweisen, dass sein Schaden durch einen Produktfehler verursacht wurde, der im Organisations- und Gefahrenbereich des Herstellers entstanden ist.

Neben den bereits dargestellten Verkehrspflichten wie Konstruktions-, Fabrikations- und Instruktionspflichten, obliegt den Herstellern und Entwicklern als Produzenten insbesondere eine fortlaufende Produktbeobachtungspflicht. Diese Pflicht basiert auf der Annahme, dass der Produzent aufgrund seiner Nähe zum Produkt und technischen Expertise am besten geeignet ist, relevante Informationen über sicherheitsrelevante Eigenschaften seiner Produkte zu sammeln und zu analysieren.

Die aus der Produktbeobachtungspflicht resultierenden Reaktionspflichten können variieren, sind aber häufig durch Warnpflichten gekennzeichnet. Der Umfang und Inhalt dieser Reaktionspflich-

ten hängen maßgeblich von der Größe der potenziellen Gefahr und der Effektivität möglicher Gegenmaßnahmen ab. Die genaue Ausgestaltung der Produktbeobachtungspflicht ist jedoch nicht immer eindeutig und Gegenstand fortlaufender juristischer Diskussionen.

Verschuldensunabhängige Haftung

Produkthaftung nach dem Produkthaftungsgesetz (ProdHaftG) eine besondere Form der verschuldensunabhängigen Gefährdungshaftung. Im Gegensatz zur verschuldensabhängigen Haftung ist ein vorwerfbarer Verstoß gegen Sorgfaltspflichten keine Voraussetzung für die Haftung nach dem ProdHaftG.

Haftungsgrundlage

Die Produkthaftung kommt in Betracht, wenn die folgenden Voraussetzungen erfüllt sind: Verletzung eines geschützten Rechtsguts (Tötung einer Person, Verletzung von Körper oder Gesundheit, Beschädigung einer Sache)

- durch ein fehlerhaftes Produkt
- mit daraus resultierendem (finanziellen) Schaden
- sowie kein Vorliegen einer gesetzlichen Ausnahme aus § 1 Abs. 2, Abs. 3 ProdHaftG.

Schäden am fehlerhaften Produkt selbst oder reine Vermögensschäden, die nicht unmittelbare Folge der Rechtsgutsverletzung sind, fallen nicht unter die Produkthaftung. Potenzielle Haftungsgründe können Konstruktions-, Fabrikations- oder Instruktionsfehler sein.

Anwendung auf KI-Systeme

Nach dem Produkthaftungsgesetz (ProdHaftG) wird unter einem Produkt traditionell eine »bewegliche«, also eine körperliche Sache verstanden. In der rechtlichen Diskussion ist allerdings die Einordnung von Software als Produkt im Sinne des § 2 ProdHaftG, insbesondere wenn sie als eigenständige (»standalone«) Software oder getrennt von der Hardware vertrieben wird, umstritten. Die herrschende Meinung tendiert jedoch dazu, auch Software als »Produkt« anzusehen. Weniger Schwierigkeiten bereitet die Beurteilung bei Hardware-Software-Systemen. In diesem Fall ist es gleichgültig, ob ein Fehler der Hardware vorliegt, oder ein Fehler der mit dieser verbundenen Steuerungssoftware (»embedded software«). Das System wird als ein Gesamtprodukt betrachtet.

Fehlerhaftigkeit eines KI-Systems

Nach § 3 ProdHaftG ist ein Produkt fehlerhaft, wenn es nicht die Sicherheit bietet, die berechtigterweise zu erwarten ist. Das einzuhaltende Sicherheitsniveau richtet sich nach dem Stand von Wissenschaft und Technik zu dem Zeitpunkt des Inverkehrbringens. Haftungsrelevante Fehler in den Ausgabeergebnissen von KI-Systemen können sich nicht nur aus einer unzureichenden Programmierung der verwendeten Algorithmen ergeben, sondern auch aus unvollständigen oder für den angestrebten Lernzweck ungeeigneten Trainingsdaten. Aufgrund der Komplexität der Entscheidungsprozesse innerhalb eines KI-Systems ist es in der Regel schwierig, die Fehlerhaftigkeit nachzuweisen.

Beschränkung der Herstellerhaftung

Die Haftung des Herstellers ist auf Fehler beschränkt, die zum Zeitpunkt des Inverkehrbringens des Produkts vorlagen. Sollte ein Produkt nach seiner Markteinführung weiterentwickelt werden, so obliegt dem Hersteller keine Haftung für Mängel, die erst im Rahmen dieser Weiterentwicklungen entstehen. Dies gilt ebenso für Änderungen in den Sicherheitsstandards: Der Hersteller ist verpflichtet, die zum Zeitpunkt des Inverkehrbringens eines neuen Produkts geltenden Sicherheitsstandards einzuhalten, jedoch nicht für Produkte, die bereits auf dem Markt sind.

Im Kontext von KI-Systemen ist der Hersteller nach dem Produkthaftungsrecht nicht haftbar für Fehlfunktionen, die durch nachträgliche Veränderungen im Einsatz des Systems entstehen. Solche Veränderungen können beispielsweise eine Umprogrammierung durch den Nutzer, den Einsatz des Systems außerhalb des vorgesehenen Anwendungsbereichs oder die Verwendung ungeeigneter Datenbestände umfassen. Andererseits bleibt der Hersteller haftbar, wenn ein Fehler auf eine unzureichende Programmierung der Lernalgorithmen zurückgeführt werden kann und sich dieser Fehler erst nach dem Inverkehrbringen des Produkts manifestiert.

Ferner kann ein Fehlgebrauch eines KI-Systems oder eine Falscheingabe von Informationen durch den Anspruchsteller zu einer Anspruchsreduzierung wegen Mitverschuldens führen.

Beweislast

In Fällen der Produkthaftung nach dem ProdHaftG liegt die Beweislast für das Vorliegen eines Produktfehlers, den entstandenen Schaden und die Kausalität zwischen dem Fehler und dem Schaden beim geschädigten Anspruchsteller. Dies bedeutet konkret, dass der Geschädigte nachweisen muss, dass der Schaden durch einen Mangel des Produkts verursacht wurde.

Wichtig zu beachten ist, dass für die Haftung im Rahmen des ProdHaftG kein Verschulden des Herstellers erforderlich ist. Das heißt, ein vorwerfbarer Verstoß gegen eine Sorgfaltspflicht muss nicht vorliegen, um eine Haftung zu begründen. Es findet auch keine Beweislastumkehr statt, wie sie die Rechtsprechung bei der Produzentenhaftung vorsieht.

Allerdings hat der Hersteller hier ebenfalls die Möglichkeit, sich von einer Produkthaftung zu entlasten. Dies kann erfolgen, indem er belegt, dass kein haftungsrelevanter Fehler innerhalb seiner Verantwortungssphäre begangen wurde. Zudem ist nach § 1 Abs. 2 Ziff. 5 ProdHaftG eine Haftung ausgeschlossen, wenn der Fehler zum Zeitpunkt des Inverkehrbringens des Produkts nach dem damaligen Stand von Wissenschaft und Technik nicht erkennbar war (sog. Einwand des Entwicklungsfehlers).

3.4.3 Maßnahmen gegen Haftungsrisiken

Um Haftungsrisiken im Zusammenhang mit dem Einsatz von KI-Systemen zu minimieren, können verschiedene Maßnahmen ergriffen werden.

In einem ersten Schritt sollten die relevanten Haftungsrisiken, die sich aus dem geplanten Einsatz von KI für interne Zwecke und/oder Dienstleistungen für Kunden ergeben, detailliert evaluiert werden. Hierbei sollten mögliche direkte und indirekte (Vermögens-) Schäden beim Kunden oder Dritten durch die Nutzung der von der KI generierten Daten nicht außer Acht gelassen werden. Eine Absicherung ist durch vertragliche Gewährleistungs-, Haftungs- und Freistellungsregelungen möglich und empfehlenswert.

Eine weitere Möglichkeit besteht in der gesellschaftsrechtlichen Auslagerung des KI-Systems in eine Betreibergesellschaft mit haftungsbeschränkter Rechtsform, z. B. eine GmbH. Darüber hinaus empfiehlt sich der Abschluss entsprechender Versicherungen.

Eine weitere präventive Maßnahme ist die vertragliche Haftungsbeschränkung im gesetzlich zulässigen Umfang – auch im Arbeitsverhältnis. Dies schafft einen klaren Haftungsrahmen. Die Konkretisierung der Leistungen und des Leistungsumfangs in Verträgen trägt ebenfalls zur Risikominimierung bei.

3.4.4 Geplante Haftungsregelungen für KI-Anwender

Die oben dargelegten Ausführungen verdeutlichen, dass auf nationaler Ebene bereits umfassende gesetzliche Haftungsregelungen für den Einsatz von KI existieren. Parallel dazu werden auf europäischer Ebene zusätzliche Regelungen entwickelt, um den Umgang mit KI weiter zu normieren und zu präzisieren. Im Fokus der aktuellen Diskussionen stehen dabei nicht nur der AI Act, sondern auch die KI-Haftungsrichtlinie und die neu überarbeitete Produkthaftungsrichtlinie.

NLF & AI Act: »Produktsicherheitsrecht«

Das New Legislative Framework (NLF) stellt einen zentralen Pfeiler im Bereich des europäischen Produktsicherheitsrechts dar und spielt auch im Kontext des AI Acts eine entscheidende Rolle. Es definiert grundlegende Anforderungen an KI-Systeme, die Aspekte wie Risikomanagement, Transparenz, Robustheit, IT-Sicherheit und die menschliche Überwachung der Künstlichen Intelligenz umfassen. Entsprechend des risikobasierten Ansatzes des NLF werden KI-Systeme in vier Kategorien eingeteilt, angefangen bei risikofreien bis hin zu Hochrisiko-KI-Systemen, wobei für jede Kategorie spezifische Regulierungsmaßnahmen vorgesehen sind.

Innerhalb des AI Acts (siehe dazu im Detail weiter oben) werden verschiedene Rollen im Umgang mit KI-Systemen definiert. Dazu gehören die Provider, also die Entwickler oder Hersteller von KI-Systemen, die für die Einhaltung der gesetzlichen Vorgaben verantwortlich sind, sowie die User, also die Anwender von KI-Systemen, die je nach Risikokategorie bestimmten Auflagen folgen müssen.

Neue Pflichten, die mit dem AI Act einhergehen, sind insbesondere für Hersteller von Hochrisiko-KI-Systemen relevant. Sie müssen nun unter anderem die Konformität mit harmonisierten europäischen Normen sicherstellen und eine CE-Kennzeichnung anbringen, die die Übereinstimmung mit den relevanten rechtlichen Anforderungen signalisiert. Außerdem müssen sämtliche Sicherheitskomponenten, die in Produkte integriert werden, die unter das NLF fallen, den Anforderungen des AI Acts entsprechen.

Die Haftungsregelungen im AI Act zielen darauf ab, die Verantwortlichkeit für Schäden, die durch KI-Systeme verursacht werden, klar zu regeln. Dies ist besonders bei Hochrisiko-KI-Systemen von Bedeutung, wo strenge Anforderungen sicherstellen sollen, dass diese Systeme keine Schäden verursachen, die die Sicherheit von Verbrauchern oder die öffentliche Ordnung gefährden.
Der risikobasierte Ansatz des AI Acts ist für Anwender generativer KI von besonderer Bedeutung, da er gewährleistet, dass die von diesen Systemen ausgehenden Risiken adäquat eingeschätzt und reguliert werden. Anwender müssen die Risikokategorie ihres KI-Systems verstehen und entsprechend den gesetzlichen Anforderungen handeln, was insbesondere für Hochrisiko-Systeme bedeutende Implikationen hat.

Neue KI-Haftungsrichtlinie & neue Produkthaftungsrichtlinie

Die geplante KI-Haftungsrichtlinie soll nur auf außervertragliche, verschuldensabhängige Schadensersatzansprüche anwendbar sein und wird daher vor allem im bestehenden Deliktsrecht zu Veränderungen führen. Eine Ausstrahlung auf vertragliche Ansprüche ist explizit nicht vorgesehen. In der Praxis könnten jedoch Spannungen entstehen, da Anspruchsteller möglicherweise versuchen, Beweise, die sie unter Art. 3 (Offenlegung von Beweismitteln) der KI-Haftungsrichtlinie gewonnen haben, auch für vertragliche Ansprüche zu verwenden. Denn das deutsche Zivilprozessrecht kennt im Grundsatz kein Institut, das eine Verwertung von Beweismitteln für vertragliche Ansprüche verwehrt.

Nach der geplanten KI-Haftungsrichtlinie soll der Anspruchsteller (beispielsweise ein KI-Anwender) das Recht haben, Informationen über ein Hochrisiko-KI-System zu erhalten, welches verdächtigt wird, einen Schaden verursacht zu haben. Diese Richtlinie zielt also auf die Normierung eines Auskunftsanspruchs gegen Anbieter und Betreiber von Hochrisiko-KI-Systemen ab. Das Hauptziel ist es, Anspruchsstellern effektive Werkzeuge zur Verfügung zu stellen, um potenziell haftende Personen zu ermitteln und einschlägige Beweismittel für einen Anspruch zu erhalten. Darüber hinaus soll die Richtlinie einen zusätzlichen Anreiz bieten, die Anforderungen des AI Acts bezüglich der Dokumentation und Aufzeichnung von Informationen zu erfüllen. Wird die Vorlage der geforderten Informationen versäumt, so wird eine Verletzung der Sorgfaltspflicht widerleglich vermutet. Daher ist eine sorgfältige und umfassende Dokumentation notwendig, um sich effektiv gegen Ansprüche zu schützen. Zudem wird nach der KI-Haftungsrichtlinie ein Kausalzusammenhang zwischen dem Sorgfaltspflichtverstoß und der Ausgabe des KI-Systems widerleglich vermutet. Für Schadensersatzansprüchen gegen Anbieter von Hochrisiko-KI-Systemen erfordert diese Kausalitätsvermutung, dass der Anspruchsteller nachweisen muss, dass der Anbieter bestimmte Vorgaben der KI-VO nicht erfüllt hat. Diese Vermutung setzt voraus, dass ein möglicher ursächlicher Zusammenhang zwischen dem Sorgfaltspflichtverstoß und dem Ausgabeverhalten des KI-Systems besteht und dass ebendieses Ausgabeverhalten ursächlich für den entstandenen Schaden war.

Zu erwähnen ist in diesem Zusammenhang auch die neue (geplante) Produkthaftungsrichtlinie. Diese Richtlinie überschneidet sich in Teilen mit der KI-Haftungsrichtlinie, insbesondere hinsichtlich der Schadensersatzansprüche gegen Hersteller. Auch in der Produkthaftungsrichtlinie ist ein Auskunftsanspruch sowie eine widerlegbare Vermutung vorgesehen, was eine methodische Übereinstimmung mit der KI-Haftungsrichtlinie darstellt. Ein weiterer bemerkenswerter Punkt dieser Richtlinie ist die ausdrückliche Einbeziehung von Software in den Produktbegriff.

3.5 Arbeitsrechtliche Aspekte

Einer der wichtigsten und zugleich komplexesten Bereiche, in denen KI eingesetzt wird, ist die Arbeitswelt. KI-Technologien revolutionieren nicht nur die Art und Weise, wie Arbeit ausgeführt wird, sondern werfen auch grundlegende Fragen zum Arbeitsrecht auf. Der folgende Abschnitt zielt darauf ab, die arbeitsrechtlichen Aspekte im Zusammenhang mit der Anwendung von KI zu beleuchten und einen umfassenden Überblick über die Herausforderungen und Chancen zu geben, die sich aus dem Einsatz von KI am Arbeitsplatz ergeben.

3.5.1 Individualarbeitsrecht

Die Integration von KI in den Arbeitsalltag bringt neue Herausforderungen für das Individualarbeitsrecht mit sich. Das Direktionsrecht des Arbeitgebers berechtigt ihn, seine Arbeitnehmerinnen und Arbeitnehmer zu verpflichten, KI oder bestimmte KI-Systeme als Arbeitsmittel zu verwenden. Das Weisungsrecht beinhaltet nicht nur das Ob der Nutzung, sondern auch das Wie, also welche Verhaltensregeln- und Verbote diesbezüglich zu beachten sind. Wenn Arbeitnehmerinnen und Arbeitnehmer gegen die Weisungen verstoßen, stellt dies eine Verletzung arbeitsvertraglicher Pflichten dar. In solchen Fällen ist der Nachweis des Verstoßes durch den Arbeitgeber entscheidend. Dieser ist gerade bei externen KI-Systemen, auf die der Arbeitgeber keinen Zugriff hat, sehr schwierig zu erbringen. Die Reaktion auf eine Pflichtverletzung kann eine Abmahnung sein, und in schwerwiegenderen Fällen oder bei wiederholten Verstößen ist sogar eine verhaltensbedingte Kündigung möglich.

Ein weiterer wichtiger Aspekt ist der Schutz von Geschäftsgeheimnissen. KI-Systeme, die große Mengen sensibler Daten verarbeiten, erfordern besondere Sicherheitsmaßnahmen. Dies gilt umso mehr, wenn die KI die Eingaben zu Trainingszwecken verarbeitet und die Eingaben als Output bei unternehmensfremden Nutzerinnen und Nutzern der KI erscheinen können. Arbeitnehmerinnen und Arbeitnehmer sind verpflichtet, Geschäftsgeheimnisse zu schützen, was den verantwortungsvollen Umgang mit KI-Systemen einschließt, die Zugang zu diesen Informationen haben.

Schutz vor Diskriminierungen durch KI (AGG)

- Arbeitgeber können beim Einsatz von KI gegen die Vorgaben des Allgemeinen Gleichbehandlungsgesetzes (AGG) verstoßen und somit verpflichtet sein, betroffenen Beschäftigten eine angemessene Entschädigung und/oder Schadensersatz zu zahlen, vgl. § 15 AGG.
- Gem. § 1 AGG ist Ziel dieses Gesetzes, Benachteiligungen aus Gründen der Rasse oder wegen der ethnischen Herkunft, des Geschlechts, der Religion oder Weltan-

schauung, einer Behinderung, des Alters oder der sexuellen Identität zu verhindern oder zu beseitigen. Auch wenn der Anwendungsbereich des AGG nicht auf das Arbeitsrecht beschränkt ist, so spielen dennoch arbeitsrechtliche Streitigkeiten im Zusammenhang mit dem AGG eine wesentliche Rolle.

- Voraussetzung für eine Haftung des Arbeitgebers nach den Vorgaben des AGG ist, dass er Beschäftigte (vgl. § 6 AGG) ist, wegen eines der in § 1 AGG genannten Merkmalen mittelbar oder unmittelbar benachteiligt oder (sexuell) belästigt. Hierbei gilt die Anweisung zur Benachteiligung einer Person aus einem in § 1 AGG genannten Grundes ebenfalls als Benachteiligung (vgl. § 3 Abs.5 S.1 AGG).
- Eine unmittelbare Benachteiligung liegt vor, wenn eine (konkrete) Person wegen eines in § 1 genannten Grundes eine weniger günstige Behandlung erfährt, als eine andere Person in einer vergleichbaren Situation erfährt, erfahren hat oder erfahren würde.
- Werden KI-Systeme unzureichend programmiert oder haben sie eine ungenügende Datengrundlage, aus der sie dann logische Schlüsse im Rahmen des Selbstlernprozesses ziehen, kann eine unmittelbare oder mittelbare Diskriminierung eines Bewerbers oder eines Arbeitnehmers wegen eines in § 1 AGG genannten Merkmals zu vermuten sein, auch wenn dies vom Betreiber gar nicht intendiert war. Die Betreiber von KI-Systemen, also hier die Arbeitgeber, müssen deshalb darauf achten, wenn sie solche Systeme im Personalmanagement und im Bewerbungsprozess einsetzen, dass Entscheidungen der KI-Systeme nicht gegen das Benachteiligungsverbot verstoßen.

3.5.2 Betriebsverfassungsgesetz

Bei der Einführung und Anwendung von KI-Systemen sind die Mitbestimmungs- und Beteiligungsrechte des Betriebsrats zu beachten.

§ 87 Abs. 1 Nr. 1 BetrVG bezieht sich auf das Mitbestimmungsrecht bei »Ordnung des Betriebs und des Verhaltens der Arbeitnehmer im Betrieb«. Dies kann auch die Einführung und Anwendung von KI-Systemen umfassen, sofern diese Systeme einen Einfluss auf die Arbeitsorganisation oder das Arbeitsverhalten der Mitarbeiterinnen und Mitarbeiter haben. Beispielsweise kann der Betriebsrat mitbestimmen, wenn eine KI-Lösung eingesetzt wird, um Arbeitsabläufe zu optimieren, die Arbeitsorganisation zu verändern oder das Verhalten der Beschäftigten am Arbeitsplatz zu beeinflussen. Dies könnte der Fall sein, wenn KI zur Steuerung von Arbeitsprozessen, zur Verteilung von Arbeitsaufgaben oder zur Überwachung und Steuerung von Arbeitsabläufen verwendet wird.

Ein Schlüsselelement ist das Mitbestimmungsrecht nach § 87 Abs. 1 Nr. 6 BetrVG. Dieses gewährt dem Betriebsrat ein Mitbestimmungsrecht bei der Einführung und Anwendung von Technologien, die potenziell zur Überwachung von Verhalten oder Leistung der Arbeitnehmer dienen können. KI-Systeme fallen oft unter diese Kategorie, da sie regelmäßig Daten sammeln und zu analysieren, die Rückschlüsse auf die Leistung oder das Verhalten der Arbeitnehmer zulassen. Entscheidend ist hierbei die objektive Eignung der Technologie zur Überwachung, unabhängig von der Absicht des Arbeitgebers.

Des Weiteren sind die Unterrichtungs- und Beratungsrechte nach § 90 Abs. 1 Nr. 3 BetrVG von Bedeutung. Diese verpflichten den Arbeitgeber zur frühzeitigen Einbindung des Betriebsrats in die Planungsprozesse für den Einsatz von KI, sodass dieser seine Sichtweisen und Bedenken einbringen kann.

Das Recht zur Hinzuziehung eines Sachverständigen nach § 80 Abs. 3 S. 2 BetrVG ist ebenfalls von großer Wichtigkeit. Aufgrund der Komplexität von KI-Technologien kann der Betriebsrat Expertinnen und Experten hinzuziehen, um ein umfassendes Verständnis der Implikationen und Potenziale der Technologie zu gewinnen. Während der Gesetzgeber das Ob der Hinzuziehung anerkennt, bleibt der Umfang der Hinzuziehung offen.

Schließlich regelt § 95 Abs. 2a BetrVG eine Zustimmungsbedürftigkeit bei der Anwendung von KI-generierten Auswahlrichtlinien. Dies bedeutet, dass der Einsatz von KI-Systemen zur Generierung von Richtlinien, beispielsweise für Beförderungen oder Versetzungen, der Zustimmung des Betriebsrats bedarf.

Wenn ein Arbeitgeber plant, Maßnahmen durchzuführen, die den Betrieb wesentlich verändern, wie in § 111 BetrVG beschrieben, muss er dies frühzeitig mit dem Betriebsrat im Rahmen eines Interessenausgleichs besprechen. Die Einführung von Künstlicher Intelligenz kann unter bestimmten Umständen als solch eine Betriebsänderung angesehen werden.

Die Implementierung von KI kann zu Effizienzsteigerungen führen, die möglicherweise einen geringeren Bedarf an Arbeitskräften nach sich ziehen. Dies kann von der Reduzierung der Arbeitsplätze bis hin zur Schließung ganzer Betriebe oder wesentlicher Teile davon reichen. Ein solcher Personalabbau könnte schon eine betriebsändernde Maßnahme nach § 111 Abs. 1 S. 3 Nr. 1 BetrVG sein. Allerdings werden solche Rationalisierungsmaßnahmen oft erst nach der Entscheidung zur Einführung von KI getroffen, welche selbst eine eigenständige Betriebsänderung sein kann.

Die Einführung von KI in einem Unternehmen kann wesentliche Veränderungen in den Betriebsanlagen und Arbeitsmethoden mit sich bringen. Diese Veränderungen können Betriebsänderungen nach § 111 S. 3 Nr. 4 und 5 BetrVG darstellen. Es ist wichtig zu untersuchen, ob der Arbeitgeber grundlegend neue Arbeitsmethoden oder Produktionsverfahren einführt, indem er entweder die Art und Weise, wie menschliche Arbeitskraft eingesetzt wird, oder das technische Verfahren ändert. Jedoch stellt nicht jede Verbesserung der Methoden und Verfahren eine Betriebsänderung dar. Es muss eine wesentliche Änderung sein, bei der eine neue Methode die bisherige ersetzt. Ob eine Änderung durch KI-Einsatz als wesentlich angesehen wird, ist umstritten. Einige Fachmeinungen betonen qualitative Aspekte und verlangen einen deutlichen technologischen Fortschritt. Bei Unsicherheiten kann auch die Anzahl der betroffenen Mitarbeitenden als Indiz herangezogen werden.

Zusammengefasst ist es entscheidend, die spezifischen Umstände jedes Einzelfalls zu berücksichtigen und die Auswirkungen der Maßnahme auf Arbeitsplätze und Arbeitsbedingungen im jeweiligen Unternehmen genau zu analysieren.

Falls Arbeitsplatzverluste trotz des Interessenausgleichs unvermeidlich sind, kommt der Sozialplan gemäß § 112 Abs. 1 S. 2 BetrVG zum Einsatz. Dieser regelt den Ausgleich oder die Milderung der

wirtschaftlichen Nachteile für die betroffenen Beschäftigten, die durch die Betriebsänderung entstehen, beispielsweise durch Abfindungen, Umschulungen oder andere Unterstützungsformen.

3.5.3 Betriebssicherheitsverordnung

Die Betriebssicherheitsverordnung (BetrSichV) in Deutschland spielt eine zentrale Rolle bei der Gewährleistung von Sicherheit und Gesundheitsschutz bei der Verwendung von Arbeitsmitteln- und Anlagen. Im Kontext der fortschreitenden Integration von Künstlicher Intelligenz in industrielle und gewerbliche Prozesse ergeben sich neue Herausforderungen und Chancen, die es im Rahmen der BetrSichV zu betrachten gilt.

Bei der Implementierung von KI in Arbeitsprozessen ist es wichtig, potenzielle Gefahren zu identifizieren und geeignete Maßnahmen zur Minimierung dieser Risiken zu ergreifen.

Zum einen muss die physische Sicherheit der Mitarbeiterinnen und Mitarbeiter berücksichtigt werden. Dies bezieht sich insbesondere auf Arbeitsplätze, an denen Menschen direkt mit KI-gesteuerten Maschinen interagieren. Es muss sichergestellt werden, dass diese Systeme sicher sind und nicht zu Verletzungen führen. Dazu gehört die Überprüfung der Maschinensicherheit, die Einhaltung von Sicherheitsabständen und die Implementierung von Notfallstoppsystemen.

Daneben sind die psychologischen Auswirkungen der KI-Nutzung zu beachten. Die Angst vor Arbeitsplatzverlust, der Umgang mit überwachenden KI-Systemen oder die ständige Interaktion mit KI kann Stress und psychische Belastungen für die Mitarbeitenden bedeuten. Hier ist es wichtig, Schulungen anzubieten, Transparenz zu schaffen und Unterstützung für die Beschäftigten zu gewährleisten.

3.5.4 Arbeitsschutz bzw. Anwendung von KI und Datenschutz

Der Einsatz von KI-Technologien zur Bearbeitung personenbezogener Beschäftigtendaten erfordert ein Gleichgewicht zwischen technologischen Innovationen und dem Schutz der Privatsphäre der beteiligten Mitarbeitenden. Dieser Abschnitt hebt die wesentlichen Herausforderungen im Datenschutz hervor, die mit der Verwendung von KI einhergehen (siehe zu den allgemeinen datenschutzrechtlich relevanten Bestimmungen oben ausführlich). Sobald ein Arbeitgeber personenbezogene Beschäftigtendaten mittels KI verarbeitet, werden sowohl die Datenschutzgrundverordnung als auch das Bundesdatenschutzgesetz relevant.

Rechtsgrundlage für die Verarbeitung

Die Nutzung von KI zur Datenverarbeitung von Mitarbeitenden, einschließlich Bewerberinnen und Bewerbern gemäß § 26 Abs. 8 Nr. 2 BDSG, muss auf eine Verarbeitungsgrundlage gestützt werden. Alle Grundlagen in Art. 6, 9 DSGVO und § 26 BDSG, mit Ausnahme der Einwilligung, verlangen die Erforderlichkeit der Datenverarbeitung. Insbesondere in den Bereichen der Persönlichkeitsanalyse und Emotionserkennung ist diese Erforderlichkeit sorgfältig zu prüfen, was eine Abwägung zwischen den Persönlichkeitsrechten der Mitarbeiterin bzw. des Mitarbeiters und den Verarbeitungsinteressen des Arbeitgebers voraussetzt.

Sofern eine Betriebsvereinbarung die Verarbeitungsgrundlage darstellt, ist sorgfältig darauf zu achten, dass durch diese nicht das Datenschutzniveau der DSGVO abgesenkt wird. Insofern bietet es sich an, auch hier einen Erforderlichkeitsmaßstab für die Datenverarbeitung anzulegen.

Eine Datenverarbeitung kann auch mit der Einwilligung des oder der Beschäftigten gemäß § 26 Abs. 2 S. 1 BDSG erfolgen, vorausgesetzt, sie ist insbesondere freiwillig erteilt. Besonders in Bewerbungsverfahren kann die Freiwilligkeit herausfordernd sein, wenn Bewerberinnen und Bewerber Nachteile bei der Ablehnung der Einwilligung fürchten. Dem kann entgegengewirkt werden, wenn traditionelle Bewerbungsverfahren ohne KI-Einsatz eine gleichwertige Option sind. Wirtschaftliche Vorteile für Beschäftigte, wie etwa das kostenlose Erhalten eines durch KI erstellten Persönlichkeitsprofils, können auch eine Rolle spielen.

Rechte der Betroffenen

Gemäß Art. 12 ff. DSGVO haben betroffene Mitarbeitende bestimmte Rechte, wenn ihre Daten mittels KI verarbeitet werden. Arbeitgeber müssen Mitarbeitende über die Datenverarbeitung umfassend informieren, nicht benötigte Daten löschen und unrichtige Daten korrigieren. Bei automatisierten Entscheidungen, einschließlich Profiling, gemäß Art. 13 Abs. 2 lit. f und Art. 14 Abs. 2 lit. g DSGVO bestehen besondere Informationspflichten. Arbeitgeber müssen über die Logik, Tragweite und die Auswirkungen der Datenverarbeitung informieren, was bei komplexen KI-Systemen eine Herausforderung darstellen kann.

Verbot automatisierter Entscheidungen

Art. 22 Abs. 1 DSGVO verbietet automatisierte Entscheidungen, die rechtliche Wirkungen haben oder die betroffene Person erheblich beeinträchtigen. KI darf demnach keine finalen Entscheidungen über Einstellungen, Beförderungen, Kündigungen oder Abmahnungen treffen. Ausnahmen von diesem Verbot sind in Art. 22 Abs. 2 DSGVO geregelt, allerdings nur unter strengen Voraussetzungen.

Datenschutz-Folgenabschätzung

Der Einsatz von KI unter Verarbeitung von Beschäftigtendaten kann eine Datenschutz-Folgenabschätzung gemäß Art. 35 DSGVO erforderlich machen. Hierbei sind die Verarbeitungsprozesse, Risiken für betroffene Personen und risikomindernde Maßnahmen zu analysieren und zu bewerten. Die Komplexität und Dynamik von KI-Systemen stellen dabei eine besondere Herausforderung dar, sowohl in Bezug auf die Erfüllung von Informationspflichten und Transparenz als auch bei der Bewertung der Verarbeitungsvorgänge und Risiken.

3.5.5 Erstellung einer unternehmensinternen Richtlinie zur Nutzung generativer KI

Die Einführung einer internen Richtlinie zum Umgang mit generativen KI-Tools in einem Unternehmen ist aus mehreren Gründen sinnvoll. Zum Ersten bietet sie die Möglichkeit, die Anwenderinnen und Anwender innerhalb des Unternehmens über Möglichkeiten und Anwendung dieser Tools zu informieren und ihnen ein Basisverständnis über deren Funktionsweise zu verschaffen. Hierbei sollte sie den Anwendenden auch ein Gefühl dafür geben, was ein solches Tool zu leisten vermag und wo andererseits dessen Grenzen liegen. Ein weiterer Zweck ist die Festlegung und innerbetriebliche Umsetzung von verbindlichen Regeln und Verhaltensvorgaben. Die Richtlinie sollte dabei nicht nur die einschlägigen, rechtlichen Anforderungen berücksichtigen, sondern idealerweise auch auf die konkreten, unternehmensrelevanten Anwendungsfälle eingehen. Entscheidend dabei ist, dass die Anwenderinnen und Anwender als Adressaten der Richtlinie in die Lage versetzt werden, die Regelungen in der täglichen Praxis auch umsetzen zu können. Auf diese Weise lassen sich rechtliche Risiken für das Unternehmen effektiv minimieren.

Nicht zu unterschätzen ist auch die vertrauensbildende Funktion einer solchen Richtlinie. So kann sie Vertrauen fördern, indem sie zur Transparenz in Hinblick auf die Nutzung von KI-Tools beiträgt und signalisiert, dass das Unternehmen nicht nur vertrauensvoll mit solchen Tools umgeht, sondern sich auch der damit verbundenen Risiken bewusst ist. Dieses Vertrauen geht einher mit ethischen Überlegungen. Je mehr sich das Unternehmen mit den potenziellen Auswirkungen der Nutzung von KI-Tools sowohl auf die Gesellschaft als auch einzelne Stakeholder auseinandersetzt, desto besser kann es den Risiken und Herausforderungen von KI (wie zum Beispiel Missbrauch, Manipulation und Diskriminierung) begegnen. Auch auf ESG-Ziele eines Unternehmens kann die Richtlinie einzahlen, indem sie die Anwenderinnen und Anwender dazu anleitet und befähigt, Auswirkungen z. B. auf die Umwelt oder die Gesellschaft bei der Nutzung von KI-Tools zu berücksichtigen.

Vor diesem Hintergrund sollte die Richtlinie inhaltlich sowohl auf das Unternehmen als auch auf die Anwendenen der KI-Tools ausgerichtet sein. Vor allem folgende Fragestellen können dabei adressiert werden:

- Welche KI-Tools und Applikationen sind vom Unternehmen zu welchem Zweck zur Nutzung durch Mitarbeiterinnen und Mitarbeiter freigegeben?
- Unter welchen Bedingungen dürfen darüber hinaus andere KI-Tools und Applikationen genutzt werden?
- Ist die Privatnutzung gestattet oder ausgeschlossen?
- Was ist bei der Erstellung der Eingabe bzw. der »Prompts« zu beachten?
- Welche Schulungsangebote gibt es?
- Gibt es Vorgaben zur Dokumentation? Zum Beispiel, welche Eingaben für den generierten Content genutzt wurden bzw. wie der Prozess war, um zur endgültigen Eingabe zu kommen bzw. welchen menschlichen Beitrag zum generierten Content geleistet wurde.

- Welche Vertraulichkeitsklassen sind für die Eingabe von Informationen zugelassen?
- In welchen Grenzen dürfen ggf. personenbezogenen Daten für die Eingabe genutzt werden?
- Wie ist die Ausgabe zu prüfen, ggf. zu überarbeiten und wie darf sie weiterverwendet werden?
- Wie ist die Ausgabe auf Copyright und IP-Verletzungen zu prüfen?
- Wie ist die Ausgabe für die Weiterverwendung als KI-generierter Content zu kennzeichnen?
- Welche Rechte werden an der Ausgabe erworben und können diese an Dritte übertragen werden?

Angesichts des rapiden technischen Fortschrittes sowie der ebenso zu erwartenden regulatorischen und gerichtlichen Entwicklungen zu rechtlichen Fragen, ist es essenziell, dass diese Entwicklungen fortlaufend verfolgt werden und relevante Entwicklungen in der Richtlinie verarbeitet werden.

Eine Möglichkeit, die Richtlinie den jeweiligen Anwenderinnen und Anwendern näherzubringen, ist, entsprechende Schulungen (ggf. elektronisch) zum Inhalt der Richtlinie anzubieten. Zudem kann die Teilnahme an einer solchen Schulung auch zur Voraussetzung gemacht werden, dass die Mitarbeitenden entsprechende generative KI-Tools zur Verfügung gestellt bekommt.

3.6 Verträge und Willenserklärungen

In einer Welt, in der Künstliche Intelligenz zunehmend in verschiedenen Bereichen eingesetzt wird, ist es wichtig, ihre rechtliche Stellung zu verstehen. Ein zentraler Aspekt ist, dass KI keine eigene Rechts- und Geschäftsfähigkeit besitzt. Dies bedeutet, dass sie nicht als eigenständige rechtliche Einheit betrachtet wird und somit keine Verträge abschließen oder rechtlich verbindliche Entscheidungen treffen kann.

In Bezug auf die Arbeitserbringung gilt das Prinzip der Höchstpersönlichkeit. Dies bedeutet, dass bestimmte Aufgaben und Dienstleistungen, die eine persönliche Leistung erfordern, nicht an eine KI delegiert werden können. Beispielsweise können medizinische oder rechtliche Beratungen, die individuelles Fachwissen und menschliches Urteilsvermögen erfordern, nicht allein durch KI erfolgen. In diesen Bereichen unterstützt KI zwar, aber die endgültige Verantwortung und Entscheidung liegt beim Menschen.

Trotz ihrer fehlenden Rechtsfähigkeit kann KI in der Rolle eines Vertreters, Boten oder Erfüllungsgehilfen agieren. Als Vertreter oder Bote kann KI im Namen einer Person oder eines Unternehmens Informationen übermitteln oder spezifische Anweisungen ausführen. Allerdings liegt die rechtliche Verantwortung weiterhin bei der Person oder dem Unternehmen, das die KI einsetzt. Als Erfüllungsgehilfe kann KI in bestimmten Prozessen assistieren, wie etwa bei der Datenauswertung oder der Automatisierung von Routineaufgaben. Auch hier bleibt die Haftung bei demjenigen, der die KI betreibt.

Es ist entscheidend zu verstehen, dass trotz der fortschreitenden Entwicklung und der zunehmenden Autonomie von KI-Systemen die rechtliche Verantwortung immer bei Menschen oder juristischen Personen liegt. Dies stellt sicher, dass trotz des Einsatzes von KI die menschliche Verantwortung und Kontrolle gewahrt bleiben, was insbesondere in rechtlichen und ethischen Fragen von großer Bedeutung ist.

4 Ethische Aspekte beim Einsatz von generativer KI

4.1 Verhältnis Ethik – Recht

KI-Technologien haben weitreichende Auswirkungen nicht nur im Alltag, sondern auch in der Unternehmenswelt. Sie verbessern Entscheidungsprozesse, steigern die Produktivität und ermöglichen eine bislang unerreichte Effizienz. Allerdings bringt dieser Fortschritt auch Herausforderungen mit sich, insbesondere in Bezug auf einen verantwortungsvollen und ethischen Einsatz von KI. Für Unternehmen ist es entscheidend, durch den verantwortungsvollen Einsatz von KI die Akzeptanz und das Vertrauen sowohl der Mitarbeitenden als auch das der Kundschaft zu erhalten und zu stärken.

4.2 Ethik beim Einsatz von KI

Obwohl es keine einheitlichen Prinzipien für die ethische Bewertung von KI gibt, existieren zahlreiche Ansätze in der gesellschaftlichen und wissenschaftlichen Ethik-Diskussion. Beispielsweise hat die EU-Kommission in ihren Ethik-Richtlinien von 2019 vier Grundprinzipien für eine KI-Ethik festgelegt: »Respekt der menschlichen Selbstbestimmung«, »Schadensvorbeugung«, »Fairness« und »Erklärbarkeit«. Der Deutsche Ethikrat betonte in seiner Stellungnahme vom März 2023 die Wichtigkeit, dass KI den menschlichen Verstand, Selbstbestimmung und Verantwortung nicht untergräbt. Zahlreiche führende Tech-Unternehmen haben sich auf freiwillige Schutzmaßnahmen geeinigt, um die Risiken von KI zu mindern.

Eine umfassende Risikobewertung ist vor der Implementierung von KI-Lösungen in Unternehmensprozessen unerlässlich, um negative Auswirkungen zu minimieren und einen reibungslosen Übergang zu den neuen Technologien zu gewährleisten. Besondere Aufmerksamkeit erfordert die Transparenz im Einsatz von KI, um das sogenannte Blackbox-Phänomen zu vermeiden. Die Funktionen und Entscheidungsprozesse der KI sollten klar und verständlich sein, um Vertrauen zu schaffen und Nachvollziehbarkeit zu ermöglichen.

Die Fairness des Einsatzes von KI ist ebenfalls von zentraler Bedeutung. KI-Systeme müssen frei von Vorurteilen und Diskriminierung sein und zur Förderung von Chancengleichheit beitragen. Dies erfordert eine sorgfältige Gestaltung der Systeme und eine bewusste Auswahl und Überprüfung der Trainingsdaten.

Ein weiterer wichtiger Aspekt ist das Prinzip der menschlichen Verantwortlichkeit. Entscheidungen, die von KI-Systemen getroffen werden, müssen stets von Menschen überwacht und verantwortet werden (»human in the loop« oder »human oversight«). Klare Kommunikations- und Berichtsprozesse sind hierfür essenziell.

Die Autonomie des Menschen sollte stets respektiert werden, wobei KI als Werkzeug zu betrachten ist, das dem Menschen dient. Mitarbeiter sollten die Funktionen und Entscheidungen von KI-Systemen verstehen und kritisch hinterfragen können. Schulungen und Weiterbildungen sind hierfür unerlässlich.

Um den ethischen Anforderungen gerecht zu werden, sollten Unternehmen Strukturen im Bereich Corporate Digital Responsibility etablieren, um eine unternehmensübergreifende KI-Strategie zu entwickeln. Die Berücksichtigung ethischer Aspekte bei der Anwendung von KI ist für Unternehmen essenziell, um Vertrauen zu stärken und sich Wettbewerbsvorteile zu sichern.

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.

Bitkom e.V.

Albrechtstraße 10
10117 Berlin
T 030 27576-0
bitkom@bitkom.org

bitkom.org

bitkom