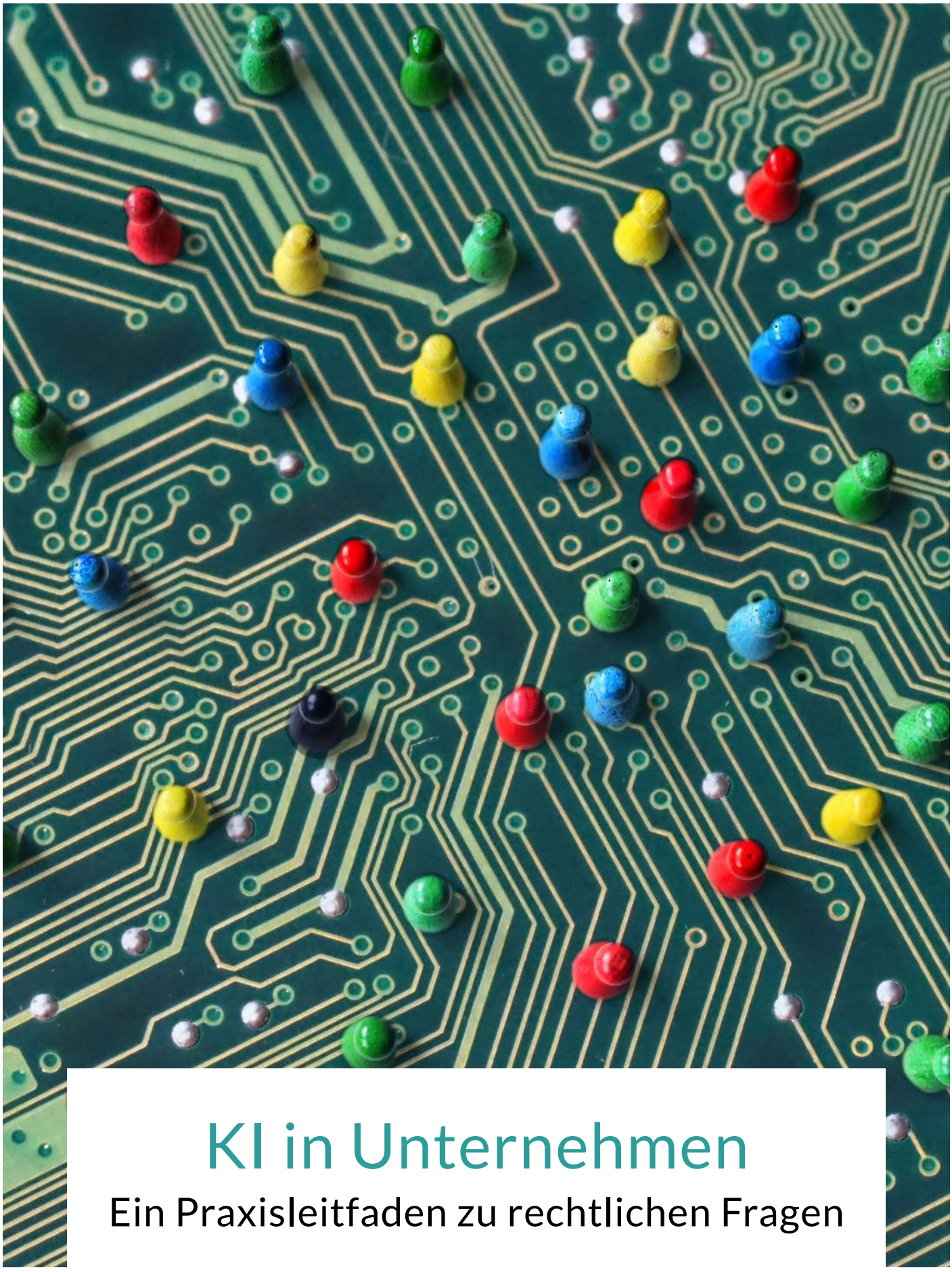

KI in Unternehmen

Ein Praxisleitfaden zu rechtlichen Fragen

iRIGHTS law

BertelsmannStiftung

KI in Unternehmen

Ein Praxisleitfaden zu rechtlichen Fragen

Dr. Till Kreutzer,
Prof. Dr. Per Christiansen

Vorwort

In den letzten Jahren haben wir uns im Zuge unseres Projektes zur betrieblichen digitalen Transformation von Unternehmen mit der Frage befasst, in welcher Weise sich die Digitalisierung auf unsere Art des Arbeitens und Wirtschaftens auswirkt. Im Fokus stand dabei vorwiegend die Kombination von technischem und kulturellem Wandel, in dem sich die Digitalisierung widerspiegelt. Ausgegangen sind wir dabei stets von intuitiv handhabbaren und aktuell weit verbreiteten digitalen Werkzeugen, die sich auf unsere Art des Arbeitens auswirken. Es ging deshalb unter anderem um die Weiterentwicklung des Kommunikationsverhaltens, der persönlichen Vereinbarkeit zwischen Arbeit und Aspekten des Privatlebens, aber auch von Werten, von Fragen / Aspekten der Gesundheit und der sozialen Sicherung. Des Weiteren beschäftigte uns das Spannungsverhältnis zwischen neuen Arbeitsmethoden und dem bestehenden arbeitsrechtlichen Rahmen sowie vor allem auch die Veränderung der Örtlichkeit(en), die wir ganz persönlich mit unserer Arbeit in Verbindung bringen. Das heißt aber auch, dass schon bisher eine gewisse Spannung zwischen dem technisch sowie kulturell Machbaren und dem regulatorisch Möglichen existierte.

Speziell das Machbare erfährt aber nun eine immer stärkere Dynamisierung durch die expansive Nutzung aller denkbaren Formen von Daten. Die offensive Weiterentwicklung von Künstlicher Intelligenz zur Datengenerierung, Datennutzung, Datenanalyse und Datenvermarktung kommt nun auch im Tätigkeitsfeld der kleinen und mittleren Unternehmen (KMU) an. Dies ist nicht zuletzt dem Umstand geschuldet, dass der finanzielle Aufwand sowie die Anforderungen an die Kompetenzen zur Nutzung dieser Werkzeuge in den letzten Jahren signifikant abgenommen haben. Dabei ist es jedoch gerade für KMU schwierig, die technische, juristische und datenschutzrelevante Expertise vor Ort im eigenen Haus vorzuhalten, um die Rahmenbedingungen des Einsatzes dieser mächtigen Werkzeuge mit letzter Sicherheit einschätzen zu können. Nun möchten wir mit dieser Publikation eine handlungsorientierte Unterstützung anbieten.

Die Publikation soll Orientierung bieten in zentralen rechtlichen Fragen der Anwendung Künstlicher Intelligenz (KI). Sie richtet sich an Projektverantwortliche und Praktiker:innen in KMU, die Systeme speziell des Maschinellen Lernens in die betrieblichen Abläufe einfügen wollen. Eine rechtliche Grundorientierung ist wichtig, weil viele rechtliche Aspekte möglichst frühzeitig im Projekt adressiert und damit die Weichen richtig gestellt werden müssen. Bei der Themenauswahl und der Art und Weise der Darstellung haben wir auf gute Allgemeinverständlichkeit mit Erläuterungen von notwendigen Fachbegriffen geachtet.

Nach einer Beschreibung des Untersuchungsthemas (Was meinen wir mit KI?) beschäftigen wir uns zunächst mit Rechtsfragen beim Design und bei der Beschaffung von KI-Technologien. Hier geht es vor allem um die juristische Bewertung verschiedener Entstehungs- und Beschaffungsstrategien wie Eigenentwicklungen oder dem Einkauf von Standardlösungen. Die Darstellung folgt sodann den Projektphasen der Beschaffung, des Systemdesigns, des Trainierens mit Trainingsdaten und des späteren Einsatzes im Live-Betrieb.

In all diesen Phasen spielen sowohl das Immaterialgüterrecht (Wem gehört die KI? Wann darf ich Inhalte für Trainingszwecke verwenden? Wem gehören die Arbeitsergebnisse einer KI?) als auch das Datenschutzrecht (Welche Anforderungen bestehen an das System bei Umgang mit personenbezogenen Daten? Welche Daten darf ich für Trainingszwecke verwenden? Wie vermeide ich Diskriminierungen durch das System?) eine zentrale Rolle. Flankierend sind im betrieblichen Einsatz auch Haftungsrisiken und arbeitsrechtliche Aspekte zu bedenken. Letztere sind in größerer Tiefe jedoch weiteren Publikationen vorbehalten.

Die Lektüre mag verständlicherweise so manche Leser:innen angesichts der Komplexität und des Facettenreichtums des Themas sowie der vielen noch unklaren Fragen beunruhigend erscheinen. Die Anzahl der „legal requirements“ für Maschinelles Lernen ist überaus hoch, vor allem im Datenschutzrecht. Aber viele Vorgaben führen im Ergebnis zu einer besseren Produktqualität und helfen, spezifische Risiken für den kommerziellen Einsatz von KI zu vermeiden. Die Rechtsunsicherheit in diesem Feld bedeutet nicht, dass man nicht handeln darf oder sollte. Im Gegenteil, sie sollte ein Impuls für frühzeitiges und vorausschauendes Handeln und damit Teil des Risikomanagements sein. Positiv betrachtet: Rechtsunsicherheit eröffnet Spielräume für Innovation.

Die Publikation ersetzt keine rechtliche Beratung im Einzelfall. Sofern jedoch Künstliche Intelligenz im Betriebskontext eingesetzt werden soll, sei es im Rahmen des Recruitings oder auch zur Produktentwicklung oder -herstellung, so findet man hier eine „Landkarte der Entscheidungshilfen“. Welche Überlegungen sind anzustellen, wenn ein algorithmisches System im eigenen Betrieb programmiert wird im Gegensatz dazu, wenn dieses eingekauft wird? Was hat das für Konsequenzen für die nächsten Schritte und zum Beispiel den Kompetenzaufbau im Betrieb? Was bedeutet das für die Haftung des Unternehmens? All dies will wohl überlegt sein. Kennt man die wichtigen Punkte, ist eine Orientierung möglich und die Entscheidung kann getroffen werden.

Wir hoffen, zum besseren Verständnis und zur Entscheidungsfindung einen Beitrag zu leisten.

Gütersloh,
Dr. Ole Wintermann,
Birgit Wintermann

Inhalt

1 Einführung

Künstliche Intelligenz (KI) und vor allem maschinelles Lernen gewinnen rasant Bedeutung in wirtschaftlichen Wertschöpfungsketten. Für eigentlich jedes Unternehmen stellt sich die Frage, ob die eigenen Produkte, Produktions- oder Dienstleistungsprozesse jetzt oder mittelfristig mit maschinellem Lernen kundenfreundlicher, effizienter und kostengünstiger gestaltet werden könnten.

In technischer Hinsicht sind Lösungen des maschinellen Lernens überraschend leicht verfügbar. Manche Lösungen können fertig entwickelt lizenziert werden, oftmals in einem „Pay per use"-Vergütungsmodell. Viele große Cloud-Anbieter wie Amazon, IBM, Google und Microsoft bieten vorkonfigurierte Lösungen im Baukastensystem an, die einerseits Tools für Standardaufgaben des maschinellen Lernens bereithalten (Mustererkennung, Bilderkennung, Sprachanalyse, Data-Mining, „Predictive Analytics" usw.), andererseits komplementäre Dienstleistungen wie Cloudspeicher, „Data Warehouses" und Rechenkapazität bis hin zu einer Komplettlösung beinhalten. Selbst wenn man eine eigene Lösung auf eigenen Systemen anstrebt, sind die zugrunde liegenden Softwareprinzipien und Algorithmen oft schon seit Jahrzehnten bekannt und die notwendigen Technologien für Entwickler:innen leicht verfügbar.

Weit weniger trivial ist hingegen der Umgang mit den Daten, die durch Methoden des maschinellen Lernens verarbeitet werden. Maschinelles Lernen arbeitet mit erheblichen modellierten Datenmengen und produziert neue Daten, die kommerziell genutzt werden können. Denkt man sich maschinelles Lernen in Phasen, angefangen von der Konzeption der Lösung über das Anlernen / Trainieren des Systems bis hin zu dessen Einsatz, so gibt es für alle Phasen rechtliche (und kostenrelevante) Rahmenbedingungen, die im Einzelnen davon abhängen, welche Art von Daten verarbeitet und erzeugt werden. Verarbeitet das System des maschinellen Lernens beispielsweise personenbezogene Daten (wie Namen, Kontaktdaten, Gesundheitsinformationen, biometrische Merkmale), gelten substanzielle Anfor-

Künstliche Intelligenz (KI)-Systeme sind von Menschen entwickelte Software- oder Hardwaresysteme, die in der physischen oder digitalen Dimension agieren, indem sie ihre Umgebung durch Datenerfassung wahrnehmen und die gesammelten strukturierten oder unstrukturierten Daten interpretieren. Durch die Interpretation bzw. die Verarbeitung der aus diesen Daten abgeleiteten Informationen treffen sie die Entscheidung über die besten Maßnahmen zur Erreichung des vorgegebenen Ziels (Europäische Kommission 2019a: 6; frei übersetzt).

derungen aus der Datenschutz-Grundverordnung (DSGVO; Europäische Union 2016), die es nachweisbar zu erfüllen gilt. Werden Inhalte (wie z. B. Texte oder Musik) genutzt, die womöglich urheberrechtlich geschützt sind, können (kostenpflichtige) Lizenzen einzuholen sein. Auch auf der Output-Seite eines Systems stellen sich Fragen, etwa wem die Arbeitsergebnisse einer Künstlichen Intelligenz gehören und wer unter welchen Umständen für Fehlentscheidungen des Systems haftet.

Die rechtliche Behandlung der Daten ist bereits auf der Ebene der Konzeption einer Lösung des maschinellen Lernens mit allen Anforderungen vollständig mitzudenken. Für ein IT-Management mag dies wenig nach agilem Arbeiten klingen. Jedoch würde es einen Managementfehler darstellen, (datenschutz-)rechtliche Expertise erst spät oder nur partiell in ein Vorhaben für maschinelles Lernen einzubinden, weil sich im schlimmsten Fall Versäumnisse nicht mehr korrigieren lassen und man die Unverwertbarkeit der Arbeitsergebnisse oder bei massiven Datenschutzverletzungen hohe Bußgelder riskiert. Umgekehrt zielen viele rechtliche Vorgaben darauf ab, dass das System auf einer sicheren IT-Grundlage diskriminierungs- und fehlerfrei mit einer hohen Akzeptanz der Betroffenen arbeitet, was aus Unternehmensperspektive nur erstrebenswert ist.

Der nachfolgende Leitfaden ist eine Handreichung über die rechtlichen Rahmenbedingungen, die es bei Einführung und Betrieb einer Lösung für maschinelles Lernen in betrieblichen Kontexten zu beachten gilt. Er kann nicht die notwendige Rechtsberatung im Einzelfall ersetzen, sondern ist als erste Orientierung für zu bearbeitende Gesichtspunkte und als praktische Hilfestellung gedacht, und zwar sowohl für diejenigen, die unmittelbar den Einsatz solcher Lösungen planen, als auch diejenigen, die vielleicht jetzt schon strategisch die Weichen für einen späteren möglichen Einsatz richtigstellen wollen.

1.1 Was ist maschinelles Lernen?

Als maschinelles Lernen („machine learning“) werden IT-Verfahren bezeichnet, in denen ein System durch das Wiederholen einer bestimmten Aufgabe (sog. Trainieren) lernt, seine Aufgabe immer besser zu bewältigen.

Zwei Gesichtspunkte sind hierfür typisch: Zum einen sind den zugrunde liegenden Algorithmen in Abgrenzung zu einem normalen Softwareprogramm die einzelnen Schritte zur Lösung der gestellten Aufgabe nicht von vornherein vorgegeben, sondern das System muss Lösungswege ausprobieren und erlernen. Es erkennt Zusammenhänge und Muster zwischen Eingangs- und Ausgangsgrößen in den Daten und übersetzt das Erlernte in eigenes Systemverhalten. Zum anderen verbessern sich die Arbeitsergebnisse mit zunehmender Erfahrung. Dies wiederum erklärt den Datenhunger von maschinellem Lernen.

ABBILDUNG 1:
KI klassifiziert Husky fälschlicherweise als Wolf
Quelle: Fischer 2018

Ein fast schon klassisches Beispiel ist das von Bauckhage (Fischer 2018): Ein System bekam die Aufgabe, auf Fotos Wölfe von Hunden zu unterscheiden – besonders bei Huskys keine leichte Aufgabe.

Das System wurde trainiert, indem es auf eine große Menge von Fotos mit Wölfen oder Hunden angewendet wurde und bei jedem Erkennungsversuch Feedback erhielt, ob es seine Aufgabe richtig, wenigstens besser oder falsch gelöst hatte. Anhand des Feedbacks verfeinerte das System seine Erkennungsfähigkeiten laufend und verbesserte die Erkennungsrate. Das funktioniert in der Praxis erstaunlich gut. Dabei lässt sich dieses Beispiel auch ohne Weiteres in den Unternehmenskontext übertragen, indem man Hunde und Wölfe z. B. durch Bauteil A und Bauteil B ersetzt.

An diesem Beispiel zeigt sich gut die Bedeutung von Trainingsdaten. Die Erkennungsrate und damit der Nutzen des Systems hängt ganz wesentlich von der Anzahl und der Qualität der Trainingsdaten und Trainingsvorgänge ab. Dabei genügt es nicht, eine große Menge von Fotos zu beschaffen. Die Trainingsdaten sollten möglichst fehlerfrei und für die vorgesehene Aufgabe repräsentativ sein. Finden sich in dem Foto-Pool auch Bilder, die andere Motive als Hunde oder Wölfe zeigen, kann das maschinelle Lernen hierdurch korrumpiert werden. In Abhängigkeit von dem konkreten Lernverfahren müssen die Fotos für das Trainieren des Systems zusätzlich aufbereitet werden (Labeling), damit ein Feedback an das System korrekt und maschinenlesbar gegeben werden kann, ob das Foto nun einen Hund oder einen Wolf zeigt.

Auch eine dritte Eigenschaft von maschinellem Lernen zeigt sich an diesem Beispiel: das sog. Blackbox-Phänomen. Das System entschied, ob es auf einem Foto einen Hund oder einen Wolf erkennt. Aber es kommunizierte ohne weitere Vorkehrungen nicht, was die Gründe für diese Entscheidung waren. Es zeigte nicht den

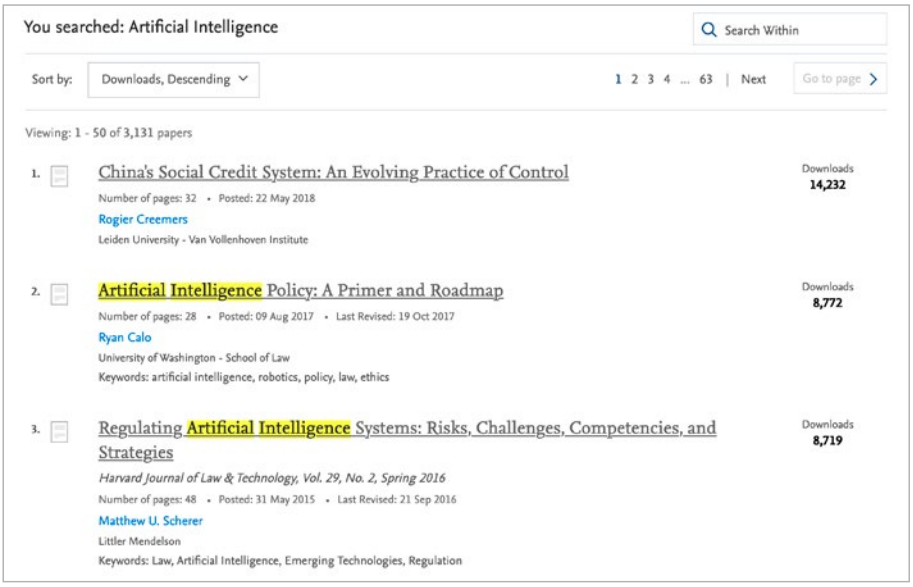

ABBILDUNG 2:
Ein Beispiel für „Explainable AI": Suchergebnisseite einer Literatursuche – Das System erklärt mit einer Markierung, weshalb ein bestimmter Suchtreffer in dem Suchergebnis aufgelistet wird
Quelle: https://www.ssrn.com/index.cfm/en/ –> Suchbegriff „Artificial Intelligence"

Lösungsweg, den es sich selbst entwickelt hat, sondern lediglich das Ergebnis (Blackbox). Diese Intransparenz ist aus vielen Gründen nicht gewünscht und der Forschungszweig „Explainable AI" oder „XAI" versucht, maschinelles Lernen zunehmend nachvollziehbarer zu gestalten. In vorliegendem Hunde-Wölfe-Beispiel hat sich gezeigt, dass das betreffende System Hunde und Wölfe ganz überraschend nicht anhand von Eigenschaften der Tiere (Fellfarbe, Augenfarbe usw.), sondern anhand der Hintergründe bzw. des Kontextes unterschieden hat (sog. Clever-Hans-Strategie). Waldartige Hintergründe und Schnee sprechen für Wolf, Garten und Straßen für Hunde. Das System traf also Entscheidungen zwischen Hunden und Wölfen, ohne diese selbst unterscheiden zu können. So amüsant und geschmeidig eine solche Ausweichstrategie des Systems wirkt, so macht sie ein System beispielsweise im medizinischen Kontext jedoch völlig unbrauchbar.

Die Clever-Hans-Strategie geht zurück auf ein Pferd, den klugen Hans, welches um 1900 aufgrund seiner angeblichen Rechenkünste „als wissenschaftliche Sensation" galt. Allerdings konnte der Kluge Hans „in etwa 90 Prozent der Fälle die richtige Antwort aus der Reaktion der Fragsteller ableiten", nicht etwa aufgrund seiner Rechenkünste. Ähnliche Lösungsstrategien könne auch bei KI-Systemen beobachtet werden, indem z. B. Bilder „vorwiegend anhand des Kontextes" klassifiziert werden, so wie im Hund-Wolf-Beispiel. (Fraunhofer-Institut für Nachrichtentechnik und Heinrich-Hertz-Institut 2019)

Das Funktionsprinzip des maschinellen Lernens gilt nicht nur für Bilderkennung, sondern für eine Vielzahl von Aufgabenstellungen mit beliebigen Daten. Die Verfahren des maschinellen Lernens können verfeinert werden, z. B. indem man Prinzipien anwendet, nach denen das menschliche Gehirn strukturiert ist (neuronale Netze und „deep learning"). Man kann maschinelles Lernen mit Sensoren zu kognitiven Systemen verbinden und in der Robotik mit physischen Handlungsmöglichkeiten ausstatten – wie etwa bei autonomen Fahrzeugen.

Bei alledem sind aber die Limitierungen nicht außer Acht zu lassen, die maschinelles Lernen mit allen anderen Formen der künstlichen Intelligenz teilt: Maschinelles Lernen ist nicht intelligent verstanden im Sinne einer menschlichen Intelligenz. Es sieht von außen nur so aus. Eine Maschine versteht nicht, was ein Hund ist. Wenn eine Maschine Hund in „dog" übersetzt, versteht sie den Sinn des Textes nicht. Wenn man bei einem digitalen Sprachassistenten Futter für einen Hund bestellt, versteht auch dieser nicht den Sinn der Worte, sondern gleicht den Klang lediglich mit gespeicherten Daten ab. Eine Maschine versteht auch keine kausalen Zusammenhänge, warum etwas geschieht, sondern sieht nur zahlenmäßige Korre-

lationen. Künstlicher Intelligenz fehlen Komponenten, die es verhindern, dass sie menschlichem Verhalten in allen Punkten gleicht, etwa Bewusstsein, Emotionen, semantisches Verständnis, überhaupt die Vielfalt der zu „Intelligenz“ gehörenden Fähigkeiten.

Stand heute ist maschinelles Lernen daher entgegen der Science-Fiction-Reputation in so manchen Medien nicht mehr und nicht weniger als ein Verfahren, spezifisch gestellte Aufgaben anhand von großen Datenmengen besonders effizient zu lösen.

1.2 Wie kann KI in der Unternehmenspraxis eingesetzt werden?

Es mangelt nicht an Studien, die das kommerzielle Potenzial maschinellen Lernens aufzeigen (bislang allerdings ohne Berücksichtigung der Effekte der Covid-19-Pandemie). Eine Studie des McKinsey Global Institute (2018) schätzt, das Gebiet der KI werde bis 2030 einen globalen jährlichen Wachstumsschub für das Bruttoinlandsprodukt (BIP) in Höhe von durchschnittlich 1,2 Prozentpunkten auslösen, was bei Weitem die Wachstumsschübe von Dampfmaschine und Industrierobotern übertreffe. Für Deutschland wird ein durch KI verursachtes Wachstum in Höhe von 1,3 Prozentpunkten prognostiziert. Damit lägen die prognostizierten durchschnittlichen Wachstumseffekte durch KI für Deutschland in der Größenordnung von ca. 50 Milliarden Euro jährlich. Zu einer ähnlichen Einschätzung gelangt eine Studie von PWC (2018), die ebenfalls eine BIP-Steigerung von 1,2 Prozentpunkten prognostiziert. In Deutschland könnten dabei 65 Prozent aller Jobs zumindest teilweise durch KI-Systeme ergänzt, unterstützt oder automatisiert werden. Das Wachstum werde zum einen von wirtschaftsstarken Branchen (Kraftfahrzeuge, Maschinenbau, Chemie und Elektronik / Elektrotechnik) getragen, die von der Anwendung spezieller KI-Lösungen profitieren. Zum anderen spielen Start-ups, die KI-basierte Produkte und Dienstleistungen entwickeln, eine Rolle.

Nach einer repräsentativen Umfrage des Branchenverbandes Bitkom (2020a) sehen zwar 73 Prozent der befragten Unternehmen KI als wichtigste Zukunftstechnologie an. Jedoch hat 2020 nur jedes siebte Unternehmen in KI-Technologie investiert.

Einen Überblick über die aktuelle Nutzung von KI in Deutschland verschafft die KI-Landkarte der Plattform Lernende Systeme (2019).

Eine Studie von Schmeiss und Friederici (2020) untersuchte die KI-Start-ups in Deutschland und kommt zu dem Ergebnis, die überwiegende Anzahl von Start-ups entwickele nicht vollständig neue Wertschöpfungsstrukturen und Produkte für neu identifizierte Kundenbedürfnisse („KI as a Solution“), sondern ziele vielmehr auf die Optimierung einzelner Aufgaben in bestehenden Wertschöpfungsprozessen ab („KI as a Service“). Bestehende Strukturen würden Start-ups mit letzteren

Anwendungsbeispiele für maschinelles Lernen im Unternehmenskontext

Produktdesign

- Ein Modehersteller lässt sich von einer KI, die mit Modefotos aus vergangenen Jahrzehnten trainiert wurde, eine neue Kollektion entwerfen.
- Ein Start-up entwickelt eine KI, mit der anhand von Kameras der Zustand des Wassers in der Fischzucht beurteilt und besser gesteuert werden kann.
- Ein Medienhaus lässt sich von einer KI journalistische Artikel schreiben (Roboterjournalismus).
- Ein Rückversicherer lässt durch eine KI globale Nachrichtenquellen auf Schadensfälle auswerten und kann auf diese Weise Versicherungstarife gestalten („early loss detection").
- Ein KFZ-Hersteller entwickelt ein autonom fahrendes Fahrzeug.
- Ein IT-Konzern entwickelt Spracherkennungssysteme wie Siri, Alexa oder Cortana.
- Ein Maschinenbauer entwickelt Bauteile durch Simulationen (digitaler Zwilling).

Fertigungsprozesse / Logistik

- Eine KI optimiert die Supply Chain und / oder die Transportwege.
- Eine KI optimiert die Produktionsprozesse und reduziert Stau, möglicherweise auch unter Personaleinsparung.
- Eine KI lernt für die Produktion Objekte zu erkennen (Griff in die Kiste durch Roboter).

Qualitätskontrolle

- Eine KI erkennt Fehlerzustände in Produktionsprozessen und entwickelt selbstständig Strategien zur Feinjustierung der Maschinenparameter.
- Mit einer KI wird die Leistungsfähigkeit einer visuellen Inspektion von Produktionsprozessen gesteigert.
- Eine KI plant eine vorausschauende Wartung von Maschinen („predictive maintanance").

Marketing, Vertrieb, Customer-Relationship-Management

- Eine KI segmentiert Kundengruppen passgenau, analysiert Kundenverhalten und prognostiziert die Nachfrage.
- Eine KI optimiert die Werbekampagne.
- Eine KI berät Kunden und Kundinnen bei der Auswahl von Produktvarianten und erkennt Upsell-Potenziale.
- Eine KI steuert eine dynamische Preisgestaltung.
- Eine KI erzeugt mehr oder weniger selbsttätig Texte für Werbebroschüren und andere Marketingunterlagen.

Kundenservice

- Kunden und Kundinnen werden im ersten Kontakt von einem Chatbot / digitalen Assistenten / Serviceroboter betreut.
- Eine KI beantwortet einfache Kundenanfragen automatisch und schafft damit den Kundenberater:innen mehr Zeit für komplexere Anliegen.
- Ein:e Hersteller:in von Computer-Games analysiert zeitnah die Online-Communities zu den eigenen Produkten und kann auf Stimmungen reagieren.

Controlling

- Eine KI bucht Zahlungsein- und -ausgänge.
- Eine KI erstellt automatisierte Forecasts.

Sicherheit / Compliance

- Eine KI erkennt frühzeitig Cyberangriffe.
- Eine KI erkennt strafbare und / oder jugendgefährdende Inhalte in eigenen Internetdiensten.
- Eine KI deckt unternehmensinterne Compliance-Verstöße auf („fraud detection").

Human Resources

- Eine KI trifft eine Bewerbervorauswahl.
- Eine KI unterstützt das People Management durch Prognosen über Talententwicklungen und zu erwartende Kündigungen.

Geschäftsmodellen bei Monetarisierbarkeit und Verfügbarkeit von nutzbaren Daten begünstigen.

Die Anwendungen und die Anwendungsmöglichkeiten für maschinelles Lernen sind so vielfältig wie die Daten, die verfügbar sind. Dies ist in zweierlei Hinsicht zu verstehen: Zum einen eignet sich maschinelles Lernen im Prinzip für alle Formen von Daten, von simplen Maschinen- / Sensordaten bis hin zu komplexer Sprachanalyse. Zum anderen wird maschinelles Lernen auf die Bereiche limitiert, in denen ausreichend Daten in angemessener Qualität verfügbar sind.

1.3 Wie unterstützt die Politik?

Die Politik hat das kommerzielle Potenzial des maschinellen Lernens erkannt und fördert sowohl die Entwicklung von KI-Technologien (in einem globalen Wettbewerb mit China und den USA) als auch die Verfügbarkeit von nutzbaren Daten. Die Strategie der EU-Kommission zur Förderung von maschinellem Lernen ist niederlegt in einem Strategiepapier (Europäische Kommission 2018a) und einem koordinierten Aktionsplan (Europäische Kommission 2018b), kürzlich ergänzt durch das White Paper „On Artificial Intelligence – A European approach to excellence and trust“ (Europäische Kommission 2020a) sowie die korrespondierende Communication: „A European strategy for data“ (Europäische Kommission 2020b), beide vom 19.2.2020. Hervorzuheben ist, dass die Kommission das Problem der Knappheit von Trainingsdaten adressiert. Erwähnenswert sind überdies die Verlautbarungen der EU Kommisssion – High-Level Expert Group on Artificial Intelligence (2020c).

Es ist zu erwarten, dass diese Strategien in Zukunft sowohl in gesetzgeberischer Aktivität als auch in Förderprogrammen resultieren werden. Gleiches gilt auf nationaler Ebene für die bereits 2018 verabschiedete KI-Strategie der Bundesregierung, die es sich zum Ziel gesetzt hat, eine „europäische Antwort auf datenbasierte Geschäftsmodelle“ (Bundesregierung 2018: 9) zu finden, die den hiesigen Werte- und Sozialstrukturen entspricht und gleichzeitig exportfähig ist.

2 Rechtliche Aspekte bei Design / Beschaffung einer Lösung mit maschinellem Lernen

2.1 Eigenentwicklung oder Beschaffung

Bei der Einführung von KI-Technologien in Unternehmen können verschiedene Ansätze verfolgt werden. Denkbar ist es, sie selbst zu entwickeln bzw. im Auftrag nach den eigenen Vorgaben von einem Dritten entwickeln zu lassen (Eigenentwicklung). Alternativ können Standardlösungen eingekauft und lizenziert werden (Beschaffung, Einkauf).

Welche Strategie vorzugswürdig ist, hängt von verschiedenen Faktoren ab, wie z. B. den Kosten, der Verfügbarkeit von Marktangeboten, der durch die KI zu lösenden Aufgaben, der Leistungsfähigkeit des Unternehmens bei der eigenen Entwicklung von Software usw. Häufig wird der Erwerb einer auf dem Markt erhältlichen Lösung vorzugswürdig sein. Angesichts der Vielfalt denkbarer Konstellationen erscheinen pauschale Ratschläge zu solchen strategischen Aspekten jedoch wenig hilfreich.

Entsprechend sollen im Folgenden lediglich einige generelle rechtliche Unterschiede zwischen Eigenentwicklungs- und Beschaffungslösungen aufgezeigt werden.

2.2 Rechte an der KI

Ein wesentlicher Unterschied zwischen der Eigen- bzw. Auftragsentwicklung und dem Einkauf von Standardlösungen liegt darin, wem Rechte an der KI (nicht zu verwechseln mit den Rechten an den Erzeugnissen der KI) zustehen. Bei der Eigenentwicklung liegen die Rechte im Zweifel im eigenen Unternehmen, bei eingekauften Lösungen müssen sie, ebenso wie das Produkt selbst, vertraglich erworben – lizenziert – werden.

Softwarelösungen im KI-Bereich können in verschiedener Hinsicht rechtlich geschützt sein. Innovative (neue) Verfahren oder durch KI-Technologien gesteuerte Produkte können patentfähig sein. Das Patentrecht schützt zwar keine reinen Rechenregeln bzw. Handlungsanweisungen (Algorithmen) oder Software als solche (also den Code, dieser kann urheberrechtlich geschützt sein). Implementierungen von Algorithmen oder Software in technische Erzeugnisse können jedoch durchaus patentierbar sein. Für den Patentschutz ist eine Erteilung des Schutzrechts durch die Patentbehörden (in Deutschland das Deutsche Patent- und Markenamt, DPMA) erforderlich.

Auch ein Urheberrechtsschutz kommt in Betracht. Das Urheberrecht bedarf, anders als das Patentrecht, keiner Formalitäten wie Registrierungen oder Erteilungsbeschlüsse. Es entsteht an schutzfähigen Inhalten (Werken) automatisch durch deren Erschaffung. Urheberrechtlich geschützt sind u. a. Computerprogramme, Softwaredokumentationen, Bilder, Musik, Filme und andere Inhalte. Das Urheberrecht bezieht sich auf den konkreten Inhalt. In Bezug auf Software ist beispielsweise der Programmcode urheberrechtlich schutzfähig, nicht jedoch die Algorithmen, die durch den Code implementiert werden, Konzepte oder die abstrakten Funktionen des Programms. Die Schutzanforderungen des Urheberrechts für Computerprogramme sind äußerst gering. Soweit es sich um eine eigene Schöpfung – also etwa nicht um eine bloße Kopie des Programms – handelt, die nicht vollständig banal ist oder nur aus wenigen Zeilen Code besteht, wird die Software urheberrechtlich geschützt sein. Ist dies der Fall, darf sie von anderen nicht ohne Zustimmung (Lizenz) verwendet werden.

Algorithmen als solche sind zwar weder urheber- noch patentrechtlich schutzfähig. Sie können jedoch unter gewissen Umständen als Geschäftsgeheimnisse nach dem Gesetz zum Schutz von Geschäftsgeheimnissen (GeschGehG) vor unbefugter Nutzung oder Veröffentlichung geschützt sein (➔ 3.1.4.3). Geschäftsgeheimnisse sind, vereinfacht ausgedrückt, Informationen, die nicht allgemein bekannt (also geheim) sind, die mit entsprechenden Maßnahmen bewusst geheim gehalten werden und bei denen ein berechtigtes Interesse an der Geheimhaltung besteht. Diese Aspekte werden auf intern entwickelte Algorithmen für KI, die nicht öffentlich bekannt gemacht werden, häufig zutreffen. Ist dies der Fall, dürfen sie nicht unbefugt, also ohne Zustimmung des Inhabers, erlangt, genutzt oder veröffentlicht werden.

2.2.1 Inhaber:innen von Rechten an der KI bei Eigen- und Auftragsentwicklung

Wird die KI-Technologie intern (also: durch eigene Mitarbeiter:innen / Angestellte) entwickelt, stehen die Rechte hieran dem Unternehmen zu. Entsprechende Regelungen für diese Rechtezuordnung finden sich im Urheberrechtsgesetz (UrhG) und im Gesetz über Arbeitnehmererfindungen (ArbnErfG). Etwaig als Geschäftsgeheimnisse geschützte Algorithmen stehen in aller Regel ebenfalls dem Unternehmen zu. Kurzum: Werden KI-Technologien inhouse entwickelt, stehen die Rechte hieran von vornherein dem eigenen Unternehmen zu.

Werden freie Mitarbeiter:innen oder Subunternehmer:innen / Auftragsentwickler:innen eingesetzt, sind die gesetzlichen Regelungen über die Zuordnung der Rechte weniger eindeutig. Mit ihnen sollten vertragliche Vereinbarungen geschlossen werden, in denen die Inhaberschaft an etwaigen Schutzrechten für ihre Arbeitserzeugnisse geregelt werden.

Bei Auftragsentwicklungen (z. B. durch einen dienstleistendes Unternehmen) hängt die Frage der Rechtezuordnung von den vertraglichen Vereinbarungen ab. In solchen Konstellationen werden Softwareentwicklungs- und Lizenzverträge geschlossen, in denen geregelt ist, wem welche Rechte in welchem Umfang zustehen. Will sich die:der Auftraggeber:in umfassende Verwertungsrechte an urheberrechtlich geschütztem Programmcode oder patentrechtlich schützenswerten Produkten und Verfahren sichern, muss er / sie sich diese Rechte exklusiv sowie zeitlich und räumlich unbegrenzt einräumen lassen. Solche Verträge sind meist sehr komplex, daher empfiehlt es sich, sie von Juristen und Juristinnen entwickeln und ggf. verhandeln zu lassen, die auf diesem Gebiet Erfahrung haben.

2.2.2 Inhaber:innen von Rechten an der KI beim Einkauf von Standardlösungen

Werden dagegen Standardlösungen beschafft, können exklusive Rechte nicht erworben werden. Ansonsten könnten Anbieter:innen sie schließlich keinen anderen Kunden und Kundinnen mehr überlassen. Die Erwerbenden erhalten in solchen Konstellationen lediglich nicht exklusive Rechte zur Nutzung der jeweiligen Lösung. Auch wird den Erwerbenden in der Regel kein Zugriff auf den Quellcode der Software gewährt, sodass sie auf die bloße Anwendung / den Einsatz derselben beschränkt ist. Eigene Verbesserungen, „Bugfixes“ oder individuelle Anpassungen sind den Erwerbern und Erwerberinnen nicht möglich. Hierfür sind sie auf die Anbieter:innen angewiesen. Damit kann das Problem des „vendor lock-in“ entstehen. Geht ein:e Anbieter:in beispielsweise insolvent, ist die eingekaufte Lösung unter Umständen nicht mehr dauerhaft nutzbar (weil sie von der:dem Anbieter:in nicht mehr weiterentwickelt wird). Auf eine andere KI-Lösung umzustellen, kann erhebliche Probleme und Kosten nach sich ziehen.

Standardlösungen können als Produkt oder als Dienst zur Verfügung gestellt werden. Als Produkt wird Software in ausführbarer Form überlassen und es wird über die Nutzung ein Lizenzvertrag geschlossen. Die Lizenznehmer:innen erhalten hier eigene Rechte zur Ausführung und ggf. weitere Nutzungen des Programms auf eigenen Rechnern oder Servern („on premises"). Ein großer Trend in der Softwarebranche ist daneben das so genannte „Software as a Service"-Modell (SaaS), häufig auch als Cloud-Computing bezeichnet. Bei diesem Modell erwerben Verwender:innen nicht das Recht, eine Software auf eigenen Systemen („on premises") zu nutzen, sondern Zugriffs- und Nutzungsbefugnisse für einen Softwaredienst. Rechtlich ist diese Variante der Miete ähnlich, während die klassische Softwareüberlassung dem Kauf entspricht.

Die beiden Modelle können für die Kunden und Kundinnen Vor- und Nachteile haben. Sie unterscheiden sich rechtlich und praktisch in mancher Hinsicht.
Ein Beispiel: Bei der Softwareüberlassung kaufen Erwerber:innen ein Produkt. Es gehört ihnen, sie können es auf eigenen Systemen verwenden und ggf. weiterveräußern. Somit sind sie u. a. von den Systemen der Anbieter:innen unabhängig, können für Back-ups selbst sorgen usw. Bei der Servicevariante erwerben Kunden und Kundinnen dagegen nur Zugriff- und Nutzungsbefugnisse. Die Abhängigkeit von den Anbietern und Anbieterinnen ist dabei (noch) größer, beispielsweise in Bezug auf die Verfügbarkeit der Software bzw. des Dienstes. Wird der Dienst etwa eingestellt, entfällt die Nutzungsmöglichkeit unmittelbar. Dafür müssen die Erwerber:innen hier keine Rechnerkapazitäten für die Anwendung bereitstellen, sie müssen sich weder um die Wartung oder die IT-Sicherheit des Systems kümmern.

Welche Rechte und Pflichten genau bestehen, sollte gerade bei Outsourcing-Strategien im KI-Kontext präzise durch Verträge geregelt werden. Insbesondere in Bezug auf SaaS-Lösungen sind die gesetzlichen Regelungen, etwa zur Leistungsdefinition, Haftung, oder Gewährleistung, äußerst vage. Verträge dieser Art sind schon typenmäßig schwer zuzuordnen, sie können als Miet-, Dienst- oder auch Werkverträge eingeordnet werden und in der Regel wird es sich um Mischverträge unterschiedlicher Typen handeln. Da die gesetzlichen Regeln über Kauf-, Werk- und Dienstverträge sehr unterschiedlich sind, sollte man sich auf sie nicht verlassen und durch die Gestaltung präziser Verträge eine eigene Rechtsgrundlage schaffen. Hierbei sollte vor allem auf gute Leistungsbeschreibungen geachtet werden. Diese definieren nicht nur, was die KI-Anbieter:innen schulden, sondern im Umkehrschluss auch, was eine nicht genügende, mangelhafte, Leistung ist. Hiervon hängen wiederum die Ansprüche der Auftraggeber:innen ab, sollte nicht ordnungsgemäß geleistet werden. Kurzum: Präzise Leistungsbeschreibungen sind wichtig, um die Rechte und Pflichten von Diensleistenden und Kunden und Kundinnen bzw. Auftraggebern und -geberinnen und Auftragnehmern und -nehmerinnen festzulegen. Um sie zu erstellen, ist in der Regel sowohl juristisches als auch technisches Know-how unerlässlich.

Bedenke: Vertragliche Regelungen sind wichtig, um das gemeinsame Verständnis der Vertragsparteien zu formulieren und zu dokumentieren. Darüber hinaus müssen sie im Streitfall auch von Dritten – Anwälten und Anwältinnen, Gerichten, Sachverständigen – interpretiert und ausgelegt werden. Es reicht also nicht aus, wenn die Parteien schon wissen, was gemeint ist. Allgemeinverständlichkeit ist ebenso wichtig wie technische oder juristische Präzision.

2.2.3 Vor- und Nachteile bei Eigenentwicklungen oder Erwerb von KI-Technologien

Wie gesagt sind abstrakte Ratschläge für oder gegen verschiedene Beschaffungsstrategien nicht möglich. Was sinnvoll ist, hängt von den Umständen und Anforderungen im jeweiligen Fall ab. Anstelle dessen stellen wir im Anschluss einige Faktoren gegenüber, die strategisch von Bedeutung sind. Naturgemäß muss hierbei stark pauschalisiert werden. Jeder Aspekt kann im konkreten Fall anders liegen, auch sind neben den hier berücksichtigten Standardkonstellationen gemischte Formen des Softwareerwerbs möglich (z. B. Erwerb von Standardsoftware mit Vereinbarungen über individuelle Weiterentwicklungen und Anpassungen). Auch werden die genannten Faktoren nicht bewertet. Ob sie jeweils relevant, vor- oder nachteilig sind, ist anhand der konkreten Konstellation abzuwägen (siehe Tabelle 1).

2.3 Datenschutzrechtliche Anforderungen und technisch-organisatorische Maßnahmen bei der Verarbeitung von personenbezogenen Daten

Führt das Vorhaben des maschinellen Lernens zu einer Verarbeitung von personenbezogenen Daten, so gilt das Datenschutzrecht mit zahlreichen zu erfüllenden Anforderungen. Personenbezogene Daten sind alle solche Daten, die sich auf Personen beziehen oder auf Personen beziehbar sind, beispielsweise Namen / Benutzernamen, Bankdaten, Gesundheitsdaten, Positionsdaten, Kaufverhalten, Werbe-Identifiyer, IP-Adressen (➔ 3.1.1).

Wenn personenbezogene Daten verarbeitet werden sollen, ist es alternativlos, mit datenschutzrechtlicher Expertise die gesamte Lösung durchzuplanen. Die Datenschutz-Grundverordnung (DSGVO) und das Bundesdatenschutzgesetz (BDSG) stellen nicht nur Anforderungen auf, unter welchen Umständen personenbezogene Daten überhaupt verarbeitet werden dürfen, sondern auch Anforderungen, wie eine solche Verarbeitung auf den IT-Systemen zu organisieren und abzusichern ist. All dies und die damit verbundenen Kosten sind bei der Konzeption zu berücksichtigen.

Die gesetzlichen Vorgaben sind dabei ziemlich abstrakt und vage. So gelten die allgemeinen Grundsätze der Datenverarbeitung (Art. 5 DSGVO): Rechtmäßigkeit der Verarbeitung, Fairness, Transparenz, Zurechenbarkeit, Zweckbindung, Datenminimierung, Richtigkeit, Begrenzung der Speicherdauer, Integrität und Vertraulichkeit. Diese Grundsätze müssen von den verantwortlichen Personen durch frühzeitig geplante technische und organisatorische Maßnahmen umgesetzt werden („Privacy by Design“ – Art. 25 DSGVO). Aber was bedeutet dies konkret? Die gesetzlichen Aussagen können operativ nur sinnvoll umgesetzt werden, wenn sie konkretisiert und ausgefüllt werden. Eine Konkretisierung des einschlägigen

TABELLE 1: **Eigenentwicklung vs. Standardlösung**

Faktor	Eigenentwicklung (inhouse, Auftragsentwicklung)	Standardlösung (Softwareprodukt, Service)
Schutzrechte (Patent, Urheberrecht, Geschäftsgeheimnis)	▮ eigene Rechte	▮ nur Nutzungsbefugnis
Laufende Kosten	▮ Kosten für Weiterentwicklung, Anpassung und Wartung (keine laufenden Lizenzgebühren)	▮ Lizenzgebühren, ggf. Kosten für Support und Wartung, Updategebühren etc.
Vermögenswert	▮ eigenes Asset (Vermögensgut), eigene Schutzrechte ▮ kann an Dritte lizenziert oder veräußert werden	▮ nur Nutzungsbefugnis, kein eigener Vermögenswert (erzeugt Kosten, schafft kein nachhaltiges eigenes Vermögen) ▮ keine Veräußerungs- oder Lizenzierungsmöglichkeiten (außer Veräußerung der jeweiligen Kopie bei Softwareüberlassung als Gebrauchtsoftware)
Autonomie	▮ „Sourcecode" steht zur Verfügung, eigene Veränderungen, „Bugfixes", Weiterentwicklungen möglich ▮ kein „vendor lock-in", kein Insolvenzrisiko Dritter ▮ keine zwingende Nutzung fremder Systeme ▮ Daten und Arbeitsergebnisse stehen unmittelbar und nur intern zur Verfügung	▮ kein Zugriff auf den „Sourcecode", nur Nutzung möglich ▮ Insolvenzrisiko, starke Bindung an Vertragspartner:in ▮ keine Autonomie hinsichtlich der verwendeten Systeme ▮ Daten und Arbeitsergebnisse liegen bei SAAS- / Cloud-Lösungen auf fremden Systemen (Datenschutz / -sicherheit!)
Haftung / Gewährleistung	▮ bei Inhouse-Entwicklung: reine Eigenverantwortung, keine Vertragspartner:innen für Haftung und Gewährleistung ▮ bei Fremdentwicklung (Beschaffung): Gewährleistung, Haftung aufseiten der Entwickler:innen (vertraglich und gesetzlich definiert, Verhandlungssache)	▮ Haftungs- und Gewährleistungsansprüche (vertraglich und gesetzlich definiert)
Wartung, Anpassung, Weiterentwicklung	▮ eigene Verantwortung	▮ bei SaaS, Cloud: in der Regel Betriebs- und Wartungspflicht der Anbieter:innen ▮ bei Softwareüberlassung: abhängig von Anbieterbedingungen (bzw. Verhandlungssache)
Support	▮ bei Inhouse-Entwicklung: eigene Verantwortung ▮ bei Fremdentwicklung: Vereinbarungssache	▮ zumeist durch Anbieter:innen geleistet ▮ Supportlevel kann oft variiert werden (Kostenfrage)
Verfügbarkeit	▮ Nutzung auf eigenen Servern / Rechnern / Rechenzentren möglich (eigene Verantwortung)	▮ bei Softwareüberlassung: Nutzung auf eigenen Servern / Rechner / Rechenzentren möglich (eigene Verantwortung) ▮ bei SaaS, Cloud: hängt von Anbietern und Anbieterinnen ab; ▮ hohe Verfügbarkeitsraten oft gegen Entgelt möglich, schützen aber nur gegen Schäden (nicht gegen die faktischen Nachteile eines Systemausfalls)
Möglichkeit der Weitergabe / Unterlizenzierung, z. B. an verbundene Unternehmen (Konzern, Konsortien usw.)	▮ bei Inhouse-Entwicklungen: keine Einschränkungen ▮ bei Fremdentwicklungen: Vereinbarungssache	▮ in der Regel keine frei verfügbaren Rechte zum Teilen / zur Weitergabe
Entwicklungsdauer und -aufwand	▮ erhebliche Entwicklungszeiten bei eigenen Entwicklungen ▮ erhebliche eigene Ressourcen und Know-how bei Inhouse-Entwicklungen erforderlich ▮ Gefahr des Scheiterns von IT-Projekten	▮ steht unmittelbar bereit, Zeitaufwand entsteht allenfalls durch Einkaufsprozedere (Vertragsverhandlungen etc.) ▮ Ressourcen werden nur für Implementierung / Bedienung, nicht für Projektentwicklung und -durchführung benötigt ▮ kein Projektmanagement nötig, keine / geringe Gefahr des Scheiterns des Projekts
Datenschutzrechtliche Verantwortlichkeit	▮ eigene datenschutzrechtliche Verantwortlichkeit ▮ angemessene technisch-organisatorische Maßnahmen bei der Verarbeitung von personenbezogenen Daten notwendig	▮ eigene datenschutzrechtliche Verantwortlichkeit ▮ Abgrenzung zu den Verantwortlichkeiten der Anbieter:innen notwendig ▮ Vorkehrungen notwendig, dass Anbieter:innen datenschutzrechtlich korrekt handeln ▮ Anbieter:innen können technisch-organisatorische Maßnahmen bereitstellen

Quelle: eigene Darstellung

Datenschutzrechts gibt es bislang kaum, weil sowohl das Themenfeld des maschinellen Lernens als auch die DSGVO als solche sehr jung sind und bislang auch keine Rechtsprechung hierzu existiert. Diese Rechtslage erschwert ein einfaches Abarbeiten von Vorgaben in der Konzeptionsphase, kann auf der anderen Seite für innovative Anwendungen auch Chancen und Gestaltungsoptionen eröffnen.

Die europäischen Datenschutzbehörden bemühen sich, durch Veröffentlichungen klarzustellen, wie die Rechtslage von ihnen ausgelegt wird. Diese Dokumente sind keine Gesetze, sondern Stellungnahmen, teilweise (fast schon politische) Diskussionspapiere, die im Prinzip nicht rechtlich bindend sind. Man mag mit guten Gründen abweichende Lösungen finden und diese ggf. mit der Aufsicht abstimmen. Dennoch bieten diese Dokumente eine gute Hilfestellung und eine Orientierung, welche Gesichtspunkte für die (sich koordinierenden) europäischen Datenschutzbehörden in ihrer Aufsichtstätigkeit wichtig sind.

Maßgeblich sind auf bisherigem Stand:

- Das Diskussionspapier der britischen Datenschutzbehörde von 2017 „Big data, artificial intelligence, machine learning and data protection“ (ico 2017).
- Ein Report der Norwegischen Datenschutzbehörde (Datatilsynet 2018) „Articificial Intelligence and Privacy“ vom Januar 2018.
- Die ICDPPC-Erklärung „Declaration on ethics and data protection in artificial intelligence“ vom 23. Oktober 2018 (ICDPPC 2018).
- Die Entschließung „Hambacher Erklärung zur Künstlichen Intelligenz“ der Datenschutzkonferenz (DSK 2019a) vom 3. April 2019.
- Das Positionspapier der DSK (2019b) vom 6. November 2019 zu empfohlenen technischen und organisatorischen Maßnahmen bei der Entwicklung und dem Betrieb von KI-Systemen.
- Die Guidance „Explaining decisions made with AI“ (ico 2020a) des britischen Information Commissioner’s Office und des Alan Turing Institute vom 20. Mai 2020.
- Die „Guidance on AI and data protection“ (ico 2020b) des britischen Information Commissioner’s Office vom 30. Juli 2020.

Im Folgenden wird ein Überblick über die grundsätzlich zu bedenkenden Arbeitspakete gegeben, die aber in Abhängigkeit von der konkreten Ausgestaltung der KI-Lösung individuell zu prüfen und abzuarbeiten sind.

2.3.1 Festlegung der datenschutzrechtlichen Verantwortlichkeiten

Ein Design einer KI-Lösung beginnt mit der Festlegung der datenschutzrechtlichen Verantwortlichkeiten. Dies gilt zum einen im Außenverhältnis. Wer datenschutzrechtlich Verantwortliche:r ist, ist rechenschaftspflichtig für die Einhaltung des Datenschutzrechts (Art. 5 II DSGVO), insbesondere für die Rechtmäßigkeit der Datenverarbeitung, die Einhaltung der oben dargestellten Grundsätze der Datenverarbeitung (Art. 5 DSGVO), die Erfüllung der Betroffenenrechte (Art. 12 ff.

DSGVO) sowie die Sicherheit der Datenverarbeitung (Art. 32 DSGVO). Grundsätzlich ist Verantwortliche:r, wer allein oder gemeinsam mit anderen über die Zwecke und Mittel der Datenverarbeitung entscheidet, also typischerweise die:der Verwender:in. Entwickler:innen sind, wenn sie nicht mit personenbezogenen Daten arbeiten, keine datenschutzrechtlichen Verantwortlichen. Verwendet man dienstleistende / cloudanbietende Unternehmen, kommt eine gemeinsame Verantwortlichkeit (Art. 26 DSGVO) oder eine Auftragsdatenverarbeitung in Betracht, bei der man sich die Einhaltung des Datenschutzrechts durch die Dienstleistenden vertraglich garantieren lassen muss (➔ 2.3.10). Ähnlich wichtig ist eine Abgrenzung der Verantwortlichkeiten, wenn man ein System bezieht, das vortrainiert wurde und mit unternehmensinternen Daten nur weiter spezialisiert werden soll. Die Zuweisung der Verantwortlichkeiten ist aber auch im betrieblichen Innenverhältnis geboten, weil ein ordnungsgemäßes Delegieren Projektarbeit erleichtert und zu den allgemeinen Sorgfaltspflichten einer Geschäftsführung gehört.

2.3.2 Festlegung der zu verwendenden Daten sowie der Zwecke und Mittel der Datenverarbeitung

Im nächsten Schritt muss die Verarbeitung der personenbezogenen Daten geplant werden. Sowohl für Trainingsdaten als auch für personenbezogene Daten im späteren Einsatz des Systems (z. B. die Nutzerdaten eines digitalen Assistenten) muss zunächst festgelegt und dokumentiert werden, zu welchen Zwecken diese Daten erhoben und genutzt werden sollen bzw., wenn die Daten bereits vorhanden sind, zu welchen ursprünglichen Zwecken diese erhoben worden sind. Das ist von erheblicher Bedeutung, da personenbezogene Daten grundsätzlich nur zu den zuvor festgelegten Zwecken verwendet werden dürfen (Art. 5 I b DSGVO) und spätere Zweckänderungen nur in engen Grenzen möglich sind (Art. 6 IV DSGVO). Die Zweckbestimmung hat damit große kommerzielle Relevanz. Können bereits existierende Daten für ein Training der KI eingesetzt werden? Sind spätere Änderungen am System, Einsatzzweck und Geschäftsmodell möglich? Aus Unternehmensperspektive empfiehlt sich zum Erhalt einer Flexibilität bei der Datennutzung eine möglichst weite Zweckbestimmung bis an die Grenze dessen, was von den Datenschutzbehörden als inhaltslose Hohlphrase (z. B. „Verbesserung der Nutzererfahrung“ oder „zukünftige Forschung“) nicht mehr akzeptiert wird.

Im zweiten Schritt sind die zu verwendenden Daten im Detail zu spezifizieren, also die Kategorien der Daten, deren Herkunft, Menge und Eigenschaften.

Im dritten Schritt sind die Mittel der Datenverarbeitung festzulegen. Dazu sind im Prinzip das gesamte System und der darauf laufende Datenverarbeitungsvorgang (einschließlich eines Löschkonzepts für nicht mehr benötigte Daten) zu spezifizieren. Das Positionspapier der Datenschutzkonferenz (DSK 2019b) listet zahlreiche Aspekte auf, wie beispielsweise: Welches Lernmodell für maschinelles Lernen soll gewählt werden? Welche Ergebnisse des Systems sollen als angemessen und korrekt oder als unerwünscht gelten? Was sind die Eingabe- und Ausgabeparameter? Wie soll die Zweckbindung gesichert werden? In welcher Weise findet eine

Interaktion zwischen der KI und Menschen statt? Soll das System im Live-Betrieb weiter lernen? Auf welchen Systemen soll das maschinelle Lernen stattfinden (lokal oder serverbasiert)? An welche Stellen werden Daten weitergegeben? Das Positionspapier (a. a. O.: 19) enthält für diese Gesichtspunkte eine Checkliste.

Im Zuge der Planung der Datenverarbeitung stellt sich natürlich auch die Frage, ob und in welchem Umfang die personenbezogenen Daten überhaupt rechtmäßig erhoben und genutzt werden dürfen (Art. 6 DSGVO). Die Anforderungen hierfür werden detailliert unter ➔ 3.1 erörtert, sind aber zwingend schon in der Konzeptionsphase zu klären, weil es sinnlos wäre, ein System zu realisieren für Daten, die nicht verwendet werden dürfen.

In der Planung des Systems des maschinellen Lernens ist schließlich zu berücksichtigen, dass ein solches System nicht ohne menschliche Letztentscheidung eingesetzt werden darf, wenn die Entscheidungen für die Betroffenen rechtliche Wirkungen entfalten oder ähnlich einschneidend sind (Art. 22 DSGVO, ➔ 4.2.3).

2.3.3 Datenschützende Modellierung von Trainingsdaten

Ein potenzielles Konfliktfeld zeigt sich bei den Vorgaben der Datenschutzkonferenz (DSK) zur Gewinnung und Verwendung von Trainingsdaten (siehe auch ➔ 3.1). Hier enthält das Positionspapier (DSK 2019b) zahlreiche Vorgaben, die die Verwendung von personenbezogenen Trainingsdaten wegen des Grundsatzes der Datenminimierung (Art. 5 I c DSGVO) auf das notwendige Maß beschränken sollen, ohne an Repräsentativität bzw. statistischer Akkuratheit für die zugedachte Aufgabe zu verlieren. Das steht in einem gewissen Widerspruch zu der Funktionsweise von maschinellem Lernen, dessen Ergebnisse mit zunehmender Datenmenge und -qualität im Prinzip immer besser werden.

Im Kern wird verlangt, die Veredelung von personenbezogenen Rohdaten bis hin zu einem Trainingsdaten-Set genau zu spezifizieren in Bezug auf Menge, Zweck, Herkunft, Bereinigungsverfahren (Normalisierung, Standardisierung, Komplettierung, Fehlerbereinigung, Fehlertestverfahren) und weitere aufgelistete Aspekte. Da die Nachvollziehbarkeit des Trainingsvorgangs ein zentraler Prüfungsgesichtspunkt der Aufsicht ist, müssen die verwendeten Roh- und Trainingsdaten inventarisiert und gegen unbeabsichtigte Vermengungen, Veränderungen oder Abflüsse gesichert werden. Der Report der norwegischen Datenschutzbehörde schlägt zudem Verfahren des maschinellen Lernens vor, die mit weniger Trainingsdaten auskommen („Generative Adversial Networks“ bzw. synthetische Daten, „Matrix Capsules“) oder die die personenbezogenen Daten in der lokalen Kontrolle der Betroffenen belassen („Federated Learning“). Hinter allem steht im Grunde wieder der „Privacy by Design“-Gedanke: Die Modellierung der Trainingsdaten soll schon im Ausgangspunkt datenschutzoptimierend durchgeplant werden.

2.3.4 Methoden zur Optimierung des Datenschutzes, insbesondere Pseudonymisierung

Ein für das Design einer Lösung maschinellen Lernens wichtiger Aspekt ist der Einsatz von Methoden, die den Datenschutz optimieren ohne den Einsatzzweck des Systems zu gefährden. Die Datenschutzaufsicht wird stets hinterfragen, ob das maschinelle Lernen (sowohl bei Trainingsdaten als auch bei Daten im späteren Einsatz) nicht in einer Weise hätte realisiert werden können, die die Ziele des Datenschutzes besser verwirklicht. Hierbei ist beispielsweise zu fragen, ob man anonymisierte, pseudonymisierte oder synthetische Daten hätte verwenden und wann welche Daten / Zwischenergebnisse frühstmöglich gelöscht werden können, ob die Daten hinreichend gegen unbefugten Zugriff gesichert sind und ob aus den Datensätzen alle Bestandteile bereinigt werden, die für die Zwecke des Systems überhaupt nicht gebraucht werden. Letzteres setzt voraus, dass man im Vorfeld Hypothesen für die Entscheidungsfindung definiert und daran gemessen irrelevante Daten aus den Rohdatensätzen eliminiert. Eine weitere typische Datenschutzoptimierung wäre es, Trainingsvorgänge und den späteren Live-Betrieb auf den lokalen Endgeräten der Betroffenen ablaufen zu lassen, ohne die personenbezogenen Daten im Netz zu transportieren; eine Variante hiervon ist das oben erwähnte „Federated Learning".

Beispiel: Zur Optimierung einer Software, die demographische Strukturen und Entwicklungen in bestimmten Wohngebieten berechnet, werden Alter, Herkunft, vorherige Wohnorte usw. der Bewohner:innen relevant sein. Deren Namen und Telefonnummern spielen dagegen für die Fragestellung keine Rolle. Entsprechend ließe sich die Anwendung mit Datensätzen trainieren, die bestimmte personenbezogene Daten gar nicht (mehr) enthalten.

Eine Anonymisierung erfolgt übrigens nicht allein schon dadurch, dass man – so die verbreitete Fehleinschätzung – die Namen aus Datensätzen löscht. Eine echte Anonymisierung, die eine Repersonalisierung praktisch ausschließt, erfolgt über mathematisch komplexe, anerkannte Anonymisierungsverfahren. Jedenfalls nach Auffassung des Bundesdatenschutzbeauftragten bedarf es überdies auch einer ausdrücklichen Rechtsgrundlage für den Anonymisierungsvorgang. Dennoch mag sich der Aufwand einer Anonymisierung gegenüber den sonst zu erfüllenden substanziellen datenschutzrechtlichen Anforderungen effizienter darstellen. Ist eine Anonymisierung nicht möglich, sind die Daten wenigstens zu pseudonymisieren, was datenschutzrechtlich aus dem Grundsatz der Datenminimierung folgt (Art. 5 I c DSGVO). Dem Gebot der Pseudonymisierung ist nicht Genüge getan, wenn Daten eines Betroffenen in einer großen Masse weiterer Daten scheinbar verschwinden. Vielmehr müssen die Identifikationsmerkmale durch Kennzeichen ersetzt werden, die eine Feststellung der Identität des Betroffenen ohne Kenntnis der Zuordnungsvorschrift ausschließen. Weiterführungen kann auf die Bitkom-Handreichung „Anonymisierung und Pseudonymisierung von Daten für Projekte des maschinellen Lernens" verwiesen werden (Bitkom 2020b).

Eine laufende Überarbeitung der Anforderungen und eine (Weiter)Entwicklung von geeigneten Verfahren, Methoden und Tools zur Optimierung des Datenschutzniveaus ist zu erwarten. Die Forschungszweige hierzu nennen sich „Privacy Preserving Machine Learnings" (PPML) oder „Privacy Enhancing Technologies" (PET). Um nur ein Beispiel zu nennen: Das Konzept „Transfer Learning" versucht den Trainingszustand eines Systems vollständig auf ein weiteres System zu übertragen und so die Notwendigkeit eines eigenen Trainingsvorgangs mit den damit verbun-

denen Datenverarbeitungsvorgängen zu ersparen. Andere Stichworte in diesem Zusammenhang sind „Differential Privacy", bei der den Trainingsdaten statistische Störungen hinzugefügt werden, oder „Homomorphic Encryption", bei der das System nicht mit Klar-, sondern mit verschlüsselten Daten rechnet. Beim Design einer Lösung des maschinellen Lernens ist es empfehlenswert, sich einen Überblick über den aktuellen Stand der Technik an datenschützenden Verfahren und Methoden zu verschaffen.

2.3.5 Methoden zur Verhinderung von Diskriminierungen

Beispiele für „machine bias"
Eine Jobempfehlungsmaschine bevorzugt Männer gegenüber Frauen.

Ein System des „Predictive Policing" prognostiziert eine höhere Verbrechenswahrscheinlichkeit bei People of Color.

Ein Lieferservice spart bestimmte Gebiete wegen eines durchschnittlichen niedrigen Haushaltseinkommens aus.

Weiter Beispiele finden sich unter: Friedman und Nissenbaum 1996; Wikipedia 2020

Ein bekanntes Problem des maschinellen Lernens ist der sog. „machine bias". Die Praxis kennt mittlerweile etliche Beispiele, in denen die Arbeitsergebnisse eines maschinellen Lernens bestimmte Personen unbeabsichtigt rechtswidrig diskriminieren.

Diese Problematik betrifft nicht alle Fälle von maschinellem Lernen, sondern nur solche Systeme, die Auswirkungen auf Menschen haben (rechtliche Entscheidungen, Profile usw.). Die Datenschutzbehörden sehen in der Verhinderung von Diskriminierungen durch KI-Systeme einen zentralen Schwerpunkt ihrer gesetzlichen Aufgaben in diesem Themenfeld. Das mag überraschen, hat aber mit dem Schutz von Grundrechten vor den Folgen fehlerhafter Datenverarbeitung und dem Grundsatz der Datenverarbeitung nach Treu und Glauben (Art. 5 I a DSGVO) zu tun.

Wenn man über Diskriminierung spricht, wird man unterscheiden müssen: Bestimmte Diskriminierungen sind rechtlich (durch das Allgemeine Gleichstellungsgesetz, AGG) verboten. Diese Verbote betreffen nur ganz partiell bestimmte Fallgruppen und sind lückenhaft. Andere Diskriminierungen oder jedenfalls Ungleichbehandlungen sind rechtlich nicht verboten, aber sozial unerwünscht oder produkt-/reputationsschädigend, beispielsweise eine Benachteiligung wegen Übergewichts oder des optischen Erscheinungsbildes einer Person. Beim Design eines Systems ist es daher notwendig festzulegen, welche Ungleichbehandlungen über die ohnehin verbotenen hinausgehend als nicht tolerabel gelten sollen und sich nicht auf die Prognosen des Systems auswirken dürfen. Man kann zu dem Entschluss gelangen, das System solle in der Realität vorhandene Ungleichbehandlungen (z. B. zwischen den Geschlechtern) nicht einfach abbilden, sondern gezielt zu neutralisieren versuchen. Eine solche Fragestellung kann durchaus eine ganz grundsätzliche Wertedebatte auslösen.

Rechtlich verboten sind nach dem AGG im Zusammenhang mit Arbeitsverhältnissen Benachteiligungen aus Gründen der Rasse, ethnischen Herkunft, des Geschlechts, der Religion, einer Behinderung, des Alters oder der sexuellen Identität, sofern für die Benachteiligung kein rechtfertigender Grund besteht. Im allgemeinen Zivilrecht (und damit im allgemeinen Unternehmensverkehr) sind solche Benachteiligungen wegen der Vertragsfreiheit im Prinzip zulässig. Ausnahmen gelten für auch im Zivilrecht stets verbotene Benachteiligungen aufgrund von

Rasse oder ethnischer Herkunft sowie für Massengeschäfte und Versicherungsverträge, für die wegen deren Breitenwirkung die gleichen Benachteiligungsverbote wie im Arbeitsverhältnis gelten. Massengeschäfte sind solche, die typischerweise ohne Ansehen der Person zu vergleichbaren Bedingungen in einer Vielzahl von Fällen zustande kommen (§ 19 AGG). Wird maschinelles Lernen im Unternehmenskontext beispielsweise für Vertrieb oder Customer Service eingesetzt, können die Benachteiligungsverbote bei Massengeschäften einschlägig sein.

Diskriminierungen können nicht nur unmittelbar erfolgen, wenn dediziert an die genannten Kriterien angeknüpft wird, sondern auch mittelbar, wenn an andere Kriterien als die im Gesetz genannten angeknüpft wird, aber die Auswirkungen eine Gruppe diskriminierend treffen (§ 3 Abs. 2 AGG). Beispielsweise ist eine Schlechterstellung von Teilzeitkräften eine mittelbare Diskriminierung von Frauen, weil die meisten Teilzeitkräfte Frauen sind. Derartige mittelbare Diskriminierungen sind äußerst schwer im Vorfeld zu erkennen und auch im AGG-Recht immer für eine Überraschung gut.

Kommt es zu Diskriminierungen, entsteht für die Anwender:innen ein Reputations- und Haftungsrisiko, weil bei Vorliegen von hinreichenden Indizien die Beweislast umgekehrt wird: Die Anwender:innen müssen beweisen, dass sie nicht diskriminiert haben (§ 22 AGG).

Ungeklärt ist bislang, ob sich die Datenschutzbehörden in ihrer Aufsichtstätigkeit nur auf die im AGG genannten Kriterien beziehen werden, oder – mit einem Verweis auf die Fairness von Datenverarbeitungen – sich für weitere mögliche Themenfelder für Diskriminierungen zuständig sehen. Für Letzteres spricht Erwägungsgrund 71 der DSGVO, der anders über das AGG hinausgehend für das Profiling auch Diskriminierungen wegen politischer Meinung, Religion oder Weltanschauung, Gewerkschaftszugehörigkeit und genetischer Anlagen oder Gesundheitszustand anspricht und dabei klarstellt, dass diese Aufzählung nicht abschließend ist.

Zu Diskriminierungen bei Einsatz von maschinellem Lernen kann es aus verschiedenen Gründen kommen, vorrangig bei einem „machine bias“ in den Trainingsdaten, aber z. B. auch in den späteren Verarbeitungsvorgängen durch diskriminierende Zielvorstellungen oder fehlerhafte Gestaltung der Entscheidungskriterien. Wenn die Trainingsdaten nicht repräsentativ für die zugedachte Aufgabe oder in sich fehlerhaft sind, lernt das System auf fehlerhafter Grundlage – fehlerhafte Resultate sind dann keine Überraschung. Die Auswahl der Daten, deren Kategorisierung, Priorisierung und Modellierung kann dazu führen, dass eine Bevölkerungsgruppe unterrepräsentiert ist oder sich kulturelle, soziale oder institutionelle Aspekte in das System einschleifen und die künstliche Intelligenz dadurch einzelne Gruppen benachteiligt, ohne dass hierfür ein rechtfertigender Grund vorliegt. Ein „machine bias“ kann sehr subtil und in den Trainingsdaten als solchen vorab kaum zu erkennen sein.

Die Verlautbarungen der Datenschutzbehörden betonen, ein System des maschinellen Lernens stets mit dem Ziel zu bewerten, auch verdeckte Diskriminierungen durch konkrete technische oder sonstige Maßnahmen zuverlässig auszuschließen. Erforderlich ist ein Qualitätsmanagement, welches Prüfroutinen und insbesondere mathematische und statistische Verfahren beinhaltet, welche unmittelbare oder mittelbare Diskriminierungen von natürlichen Personen in Bezug auf die rechtlich relevanten Kriterien ausschließen. Was das konkret bedeutet, ist bislang nicht vollständig geklärt und hängt auch von dem konkreten Verfahren des maschinellen Lernens ab. Der Stand der Technik und der entsprechenden Prüfroutinen entwickelt sich rapide. Man wird aber in jedem Fall aufzeigen müssen, ob und wie man bei der Modellierung der Trainingsdaten antizipierte Schieflagen bereinigt hat. Ein weiterer Aspekt ist die Herstellung von Transparenz der Entscheidungsvorgänge (➔ 2.3.6). Schließlich sind laufende Stichproben und Testverfahren erforderlich, die sich zum einen auf die Diskriminierungsverbote nach dem AGG, zum anderen auf die ggf. vorab definierten weiteren nicht akzeptablen Benachteiligungen beziehen. Diese Tests müssen sowohl die Trainingsvorgänge als auch den späteren Live-Betrieb abdecken. Werden Proxy-Variablen verwendet, ist besonderes Augenmerk darauf zu richten, ob sich hierdurch nicht ein „bias" einschleicht. In jedem Falle wird man zur Vermeidung von mittelbaren Diskriminierungen prüfen müssen, ob verbotene Anknüpfungspunkte nicht durch hochkorrelierende Ersatzvariablen ersetzt wurden, etwa eine Anknüpfung an das Geschlecht durch eine Anknüpfung an Körpergewicht und Vorname.

Die Prüfungsanforderungen steigen, wenn sich das System im Einsatz laufend anhand von Nutzerdaten weiterentwickelt. Werden Daten aus dem laufenden Einsatz nicht vorab geprüft und bearbeitet, kann sich leicht ein „machine bias" einschleifen. Die Datenschutzkonferenz verlangt ausdrücklich eine Risikoüberwachung zum Schutz vor Diskriminierungen auch während eines Live-Betriebs. Die „Guidance on AI an data protection" des Information Commissioner's Office (ico 2020b) enthält hilfreiche Handreichungen hierzu.

Quelle: Pixabay/StockSnap https://pixabay.com/de/photos/menschen-mann-lesung-zeitung-2566677/. (Pixabay-Lizenz: https://pixabay.com/de/service/license/) – Montage von D. Ehlers (Überschrift eingefügt)

Beispiel: Ein Chatbot, der rassistische Sprache übernahm – zugleich ein PR-Gau

- „Twitter-Nutzer machen Chatbot zur Rassistin" (https://www.zeit.de/digital/internet/2016-03/microsoft-tay-chatbot-twitter-rassistisch)
- „Rassistischer Chat-Roboter: Mit falschen Werten bombardiert" (https://www.sueddeutsche.de/digital/microsoft-programm-tay-rassistischer-chat-roboter-mit-falschen-werten-bombardiert-1.2928421)
- „Tay, Microsoft's AI chatbot, gets a crash course in racism from Twitter" (https://www.theguardian.com/technology/2016/mar/24/tay-microsofts-ai-chatbot-gets-a-crash-course-in-racism-from-twitter)
- „Vom Hipster-Mädchen zum Hitler-Bot" (https://www.spiegel.de/netzwelt/web/microsoft-twitter-bot-tay-vom-hipstermaedchen-zum-hitlerbot-a-1084038.html)
- „Zum Nazi und Sexisten in 24 Stunden" (https://www.faz.net/aktuell/wirtschaft/netzwirtschaft/microsofts-bot-tay-wird-durch-nutzer-zum-nazi-und-sexist-14144019.html)

In Bezug auf die Verantwortlichkeiten für Diskriminierungen durch KI sind rechtlich noch viele Fragen ungeklärt. Die Ursache für einen „machine bias“ ist oftmals nicht eine technische bzw. technisch zu adressierende, sondern eine gesellschaftliche, die in den Daten lediglich so abgebildet ist, wie sie auch real existiert. In welchem Umfang kann man Hersteller:innen einer KI dafür in die Pflicht nehmen, aus vorhandenen Daten eine idealtypische Gesellschaft herauszumodellieren, wie immer die auch aussehen mag und wer auch immer bestimmt, wie diese auszusehen hat? Pragmatisch wird man aber sagen können: Jedenfalls Diskriminierungen, die durch das AGG oder die DSGVO verboten sind, können nicht legal von einer KI vorgenommen werden. Entsprechend müssen präventive Maßnahmen zur Verhinderung sowie ein zeitnahes Monitoring mit entsprechenden Reaktionsmöglichkeiten aufgezeigt werden.

2.3.6 Methoden zur Herstellung von Transparenz / Überwindung des Blackbox-Phänomens

Ein weiterer zentraler Punkt ist die Herstellung von Transparenz. Transparenz ist eine Grundbedingung für Nachvollziehbarkeit. Transparenz ist sowohl für die Betroffenen der verwendeten Daten (Art. 5, 12, 13, 14 DSGVO) als auch für die Aufsichtsbehörden herzustellen (Art. 5 II DSGVO), nach Auffassung der Datenschutzkonferenz für die Aufsicht in größerer Tiefe. Transparenz ist eine Grundbedingung dafür, dass die Rechtmäßigkeit der Datenverarbeitung überprüft und Rechte der Betroffenen geltend gemacht werden können. Dies verdeutlicht den Stellenwert, den Transparenz im Datenschutzrecht einnimmt. Von der Pflicht zur Herstellung von Transparenz gegenüber den Betroffenen gibt es nur sehr wenige Ausnahmen, etwa wenn die Betroffenen schon von der Datenverarbeitung wissen oder wenn sich die Erteilung der Information als unmöglich erweisen oder einen unverhältnismäßig hohen Aufwand erfordern würde (Art. 14 V DSGVO).

Transparenz ist dabei ein vager Begriff, der in diesem Kontext auch anders verstanden werden kann, nämlich als Grundbedingung für die Akzeptanz eines Produkts in der Zielgruppe oder im Markt. Dieser Vertrauensaspekt ist im Unternehmenskontext ebenfalls wichtig. Wenn man sich das eingangs erwähnte Beispiel (➔ 1.1) der Aufgabe vergegenwärtigt, auf Fotos Wölfe von Hunden zu unterscheiden, so liegt auf der Hand: Man vertraut diesem System viel mehr, wenn man nachvollziehen kann, dass es Wölfe und Hunde wirklich aufgrund von deren Eigenschaften und nicht aufgrund eines Bildhintergrundes unterscheidet. Welche Faktoren Transparenz und Vertrauen in algorithmische Entscheidungen bei den Betroffenen fördern, wird z. B. in der Arbeit von Beining (2019) untersucht, auf die verwiesen wird. Hilfreiche Hinweise finden sich auch in der Handreichung vom Information Commissioner’s Office und Alan Turing Institute (ico 2020a).

Im juristischen Sinne ist Transparenz enger gefasst und nur bezogen auf die Informationen, die nach der DSGVO bereitzustellen sind. Die Pflichtinformationen gegenüber Betroffenen sind in Art. 13, 14 DSGVO aufgelistet, beispielsweise

der Zweck der Datenverarbeitung, Speicherdauer, Kontaktdaten der Verantwortlichen, Empfänger:innen der Daten usw. Grob gesagt muss man den Betroffenen alle Informationen zur Verfügung stellen, damit diese den Datenverarbeitungsvorgang verstehen und sinnvoll über die Verwendung / Preisgabe ihrer personenbezogenen Daten entscheiden und ggf. ihre Rechte geltend machen können, und der Aufsicht sämtliche Informationen, die sie zur Erfüllung ihrer Aufsichtspflicht und der Prüfung der Rechtmäßigkeit der Datenverarbeitung benötigen. Aber was heißt dies genau im Kontext des maschinellen Lernens?

Die Datenschutzkonferenz stellt in der Hambacher Erklärung (DSK 2019a: 3) eine Maximalforderung auf. So heißt es: „Personenbezogene Daten müssen in einer für die betroffene Person nachvollziehbaren Weise verarbeitet werden (Art. 5 Abs. 1 lit. a DSGVO). Dies erfordert insbesondere eine transparente Verarbeitung, bei der die Informationen über den Prozess der Verarbeitung und ggf. auch über die verwendeten Trainingsdaten leicht zugänglich und verständlich sind (Art. 12 DSGVO). Entscheidungen, die auf Grundlage des Einsatzes von KI-Systemen erfolgen, müssen nachvollziehbar und erklärbar sein. Es genügt nicht, dass das Ergebnis erklärbar ist, darüber hinaus muss die Nachvollziehbarkeit im Hinblick auf die Prozesse und das Zustandekommen von Entscheidungen gewährleistet sein. Nach der DSGVO ist dafür auch über die involvierte Logik ausreichend aufzuklären. Diese Transparenzanforderungen sind fortwährend zu erfüllen, wenn KI-Systeme zur Verarbeitung von personenbezogenen Daten eingesetzt werden. Es gilt die Rechenschaftspflicht des Verantwortlichen (Art. 5 Abs. 2 DSGVO)."

In der Praxis wird diese Forderung erhebliche Probleme verursachen. Es muss die involvierte Logik aufgezeigt werden. Bei Systemen des maschinellen Lernens auf aktuellem Stand ist es oftmals nicht oder nur mit erheblichem Aufwand möglich nachzuvollziehen, wie das System konkret auf seine Entscheidung gekommen ist (Blackbox-Phänomen). In der Forschung werden allerdings zunehmend Verfahren entwickelt, die den Entscheidungsprozess offenlegen („Explainable AI" oder XAI) oder nachträglich ermitteln (z. B. „Blackbox Tinkering", „Local Interpretable Model-Agnostic Explanations", LIME; „Layer-wise Relevance Propagation", LRP; „Spectral Relevanz Analysis", SpRay u. a.). Als Anwender:in / Hersteller:in wird man sich also nicht auf die Position zurückziehen können, eine Transparenz sei technisch unmöglich. Vielmehr muss man beim Design eines Systems die notwendige Transparenz auf dem Stand der Technik mit einplanen, idealerweise den Algorithmus von Beginn an mit Transparenzfunktionen entwickeln und überdies verfügbare Techniken einsetzen. Erwägenswert erscheint auch der in der Wissenschaft diskutierte Ansatz der „counterfactual explanations" von Wachter, Mittelstadt und Russel (2018). In diesem Ansatz wird Transparenz nicht durch Offenlegung der internen Logik, sondern dadurch hergestellt, dass erklärt wird, welche Umstände zu einer anderen Entscheidung geführt hätten. Beispiel: „Ihr Kreditantrag wurde abgelehnt, weil Ihr Jahreseinkommen bei 40.000 Euro lag. Würde Ihr Jahreseinkommen über 50.000 Euro liegen, würde dem Antrag stattgegeben werden." Es leuchtet ein, dass derartige Informationen für die Betroffenen oft viel relevanter als die involvierte Logik des Systems sind.

Wie detailliert die Informationen über die Datenverarbeitung sein müssen, ist noch nicht vollständig geklärt. Hier zeigt sich ein typischer Konflikt aus der Praxis: Datenschutzbehörden fordern typischerweise möglichst weitgehende Transparenz. Dies mag aber nicht im Interesse eines Unternehmens liegen, welches die Kunden und Kundinnen nicht verschrecken oder überfordern möchte. Womöglich verfügen Anwender:innen in Beschaffungsszenarien über Dienstleistende / Cloudanbietende überhaupt nicht über die notwendigen Informationen, die die Dienstleistenden für sich als Geschäftsgeheimnis behandelt. Nicht zuletzt kann die Offenlegung der Entscheidungslogik leicht missbraucht werden, weil man mit Kenntnis der Funktionsweise das System für eigene Zwecke kapern oder korrumpieren könnte. Erklären muss man aber jedenfalls das logische Grundprinzip und die Grundannahmen (Entscheidungsquellen und deren Relevanz sowie die wesentlichen Strukturmerkmale der Entscheidungsfindung). Krafft und Zweig (2019) haben aufgezeigt, dass eine sinnvolle Informationstiefe abhängig von Kontext und Risikograden des Systems ist; manche Systeme sind gefährlicher als andere. Normalerweise ist es nicht erforderlich, den eigentlichen Algorithmus oder den Quellcode offenzulegen. Dies liegt auf der Linie der Rechtsprechung zur Schufa, welche ebenfalls nicht ihren (geschäftsrelevanten) Scoring-Mechanismus offenlegen musste. Als Faustformel für die Praxis sollte man sich vor Augen halten, dass gegenüber Betroffenen zu detaillierte Informationsfülle das Gegenteil von Transparenz bewirken kann. Man sollte in Privacy Policies und dergleichen auf Verständlichkeit achten, was eine Beschränkung auf die wesentlichen Informationen in den gesetzlich vorgeschriebenen Pflichtangaben notwendig macht. Transparenz gegenüber der Aufsicht hingegen kann technische Expertise unterstellen und die Aufsicht wird sicher nach technischen Dokumenten wie Protokollen, Logs und spezifischen Dokumentationen fragen.

Transparenz eines Systems bezieht sich nach dem Verständnis der Datenschutzbehörden nicht nur retrospektiv auf bereits vollzogene Datenverarbeitungsvorgänge, sondern auch prospektiv auf zukünftige Datenverarbeitungsvorgänge. Das wird bei selbstlernenden Systemen häufig eine Herausforderung darstellen, und zwar technisch sowie in der Anwendung zur Verhinderung von Missbrauchsszenarien. Die detaillierten Anforderungen der Datenschutzbehörden an die Planung eines Systems einschließlich der Definition von Wissensdomänen, der Kartierung der Datenquellen, den Monitoring-Anforderungen usw. dienen auch dazu, unter dem Transparenzaspekt Prognosen für das Systemverhalten zu ermöglichen.

Schließlich stellt sich die Frage, auf welchem kommunikativen Weg Informationspflichten gegenüber den Betroffenen sinnvoll erfüllt werden können. Typischerweise erfolgt dies in Privacy Policies / Datenschutzerklärungen sowie Einwilligungstexten. Bei vielen Anwendungen, etwa bei Chatbots oder dem Einsatz von digitalen Assistenten, die sich im laufenden Betrieb weiterentwickeln, sind die Anforderungen an den Kommunikationsmodus noch nicht vollständig geklärt. In der Praxis kann man sich an aktuellen Empfehlungen der Branchenverbände orientieren, etwa dem Leitfaden „Machine Learning und die Transparenzanforderungen der DSGVO“ (Bitkom 2018)

Die Bedeutung von Transparenz und Diskriminierungsschutz kann nicht genug betont werden. Neben den datenschutzrechtlichen Komponenten könnte zukünftig auch eine haftungsrechtliche treten. Aufgrund einer Resolution des Europäischen Parlaments (2020) von Anfang Februar 2020 sollen KI-Systeme nachvollziehbare und diskriminierungsfreie Algorithmen anwenden, andernfalls sollen die Betroffenen die Verantwortlichen für Nachteile haftbar machen können.

2.3.7 Erfüllung von Rechten der Betroffenen

Den von einer Datenverarbeitung betroffenen Personen stehen umfangreiche Rechte zur Verfügung, insbesondere Auskunfts- und Löschungsansprüche (Art. 15 ff. DSGVO). Organisatorisch ist sicherzustellen, dass berechtigte Ansprüche tatsächlich erfüllt werden können. Häufiger Praxisfehler: Die Kontaktangaben, die man im Rahmen der Informationspflichten angegeben hat, funktionieren nicht bzw. die angegebenen E-Mail-Postfächer werden nicht abgerufen. Für Details kann auf die Guidance des Information Commissioner's Office (ico 2020a) verwiesen werden. Nachvollziehbarerweise geht die Datenschutzaufsicht Beschwerden von Betroffenen, ihre Rechte würden nicht erfüllt werden, regelmäßig nach.

2.3.8 Datenschutz-Folgenabschätzung

Sollen personenbezogene Daten für maschinelles Lernen verarbeitet werden, ist in der Regel zuvor eine Datenschutz-Folgenabschätzung („Privacy Impact Assessment", PIA) durchzuführen (Art. 35 DSGVO) und zu dokumentieren. Eine solche müssen Verantwortliche immer dann durchführen, wenn eine Form der Verarbeitung, insbesondere bei neuen Technologien, aufgrund der Art, des Umfangs, der Umstände und der Zwecke der Verarbeitung voraussichtlich ein hohes Risiko für die Rechte und Freiheiten natürlicher Personen zur Folge hat. Nach Auffassung der Datenschutzkonferenz ist dies der Fall bei dem „Einsatz von künstlicher Intelligenz zur Verarbeitung personenbezogener Daten zur Steuerung der Interaktion mit den Betroffenen oder zur Bewertung persönlicher Aspekte der betroffenen Person" (Bayerisches Landesamt für Datenschutzaufsicht 2018b: 3). Wesentlich weiter geht dabei noch der regulatorische Ansatz des Information Commissioner's Office (ico 2020a), der im Sinne einer „risk based regulation" die Risikoanalyse in einer Datenschutzfolgenabschätzung gewissermaßen zum Kern des gesamten Systemdesigns erklärt. Der Umgang mit dem Problem der Diskriminierung durch „machine bias" (➔ 2.3.5) sollte in der Folgenabschätzung adressiert werden. Nützliche Hilfestellungen für eine Risikoanalyse gibt auch die „Assessment List for Trustworthy Artificial Intelligence" (EU-Kommission – High Level Expert Group on Artificial Intelligence 2020c). Die deutschen Datenschutzbehörden stellen allgemeine Muster und Anleitungen für eine Datenschutz-Folgenabschätzung zur Verfügung (DSK 2017).

2.3.9 Maßnahmen der Datensicherheit / IT-Sicherheit

Der Umgang mit personenbezogenen Daten erfordert zu dokumentierende technisch-organisatorische Maßnahmen zur Datensicherheit (Art. 24, 32 DSGVO) auf dem Stand der Technik mit dem Ziel eines „dem Risiko angemessenen Schutzniveaus". Hierzu gibt es allgemeine Empfehlungen des Bundesamtes für Sicherheit in der Informationstechnik (2020a) (Grundschutz), die uneingeschränkt neben den speziellen Verlautbarungen zur Künstlichen Intelligenz gelten wie für jedes andere IT-System auch. Orientieren kann man sich auch am Standard-Datenschutzmodell der Datenschutzkonferenz (DSK 2020). Was konkret zu tun ist, ist gewissermaßen eine Abwägungsentscheidung und hängt von der Art der verwendeten Daten und dem Risiko für die Betroffenen ab. Man benötigt also ein eigenes individuelles Konzept, für das man aber weitgehend auf Standards zurückgreifen kann. Für besonders sensible Systeme ist eine Zertifizierung nach ISO 27001 (Bundesamt für Sicherheit in Informationstechnik 2020b) erwägenswert. Soweit für das betreffende System eine umfassende und detaillierte Risikoanalyse angezeigt ist, kann man sich an den ISO-Normen 31000 und 27005 sowie dem BSI-Standard 200-3 orientieren.

Überblicksartig sind die in § 64 BDSG aufgelisteten Punkte darzustellen:

- Pseudonymisierung und Verschlüsselung von personenbezogenen Daten
- Sicherstellung der Vertraulichkeit, also Zutrittskontrolle (z. B. abschließbare Büroräume), Zugangskontrolle (z. B. Log-in mit Benutzername und Passwort, Virenschutz) sowie Zugriffskontrolle (z. B. Rollen- und Berechtigungskonzepte) und Trennungskontrolle (Maßnahmen, die eine Vermischung von zu unterschiedlichen Zwecken erhobener Daten verhindern). Insbesondere darf das System nur durch Befugte trainiert, genutzt und überwacht werden.
- Sicherstellung der Integrität, also Verfälschungsschutz der Daten durch Eingabe- und Weitergabekontrolle oder digitale Signaturen
- Sicherstellung der Verfügbarkeit (z. B. Back-ups)
- Sicherstellung der Belastbarkeit (z. B. Firewalls)
- Verfahren zur regelmäßigen Überprüfung, Bewertung und Evaluierung

Für die meisten Unternehmen sollten derartige Maßnahmen ohnehin Standard sein. Darüber hinaus sind gesondert die spezifischen Risiken der Systeme mit maschinellem Lernen zu adressieren, etwa der Verlust von großen Datenmengen, Integritätsrisiken durch Verwendung von Drittsoftware oder auch das Risiko der vorsätzlichen Kompromittierung des Systems (vgl. hierzu ico 2020b). Viele KI-Systeme laufen separiert von der Unternehmens-IT auf virtuellen Maschinen.

Stellungnahme eines Anwalts der umstrittenen Gesichtserkennungssoftware Clearview nach einem erheblichen Hack-Angriff auf das System: Datenpannen seien nun einmal „Teil des Lebens im 21. Jahrhundert". Quelle: Brühl 2020

2.3.10 Auftragsverarbeitung, Übermittlungen in Drittländer, insbesondere Cloud-Computing

Eine Lösung mit maschinellem Lernen wird in der Praxis selten eine Stand-Alone-Lösung eines Unternehmens vollständig mit eigener IT sein. Vielmehr wird man sich partiell oder vollständig Dienstleistungsunternehmen bedienen und diesen im Zuge der Datenverarbeitung auch personenbezogene Daten übertragen, z. B. cloudanbietende Unternehmen. Dann aber sind nach allgemeinen datenschutzrechtlichen Regelungen die Voraussetzungen hierfür zu schaffen:

Werden personenbezogene Daten durch Dienstleistende verarbeitet, ist dies (sofern nicht ausnahmsweise eine gemeinsame Verantwortlichkeit vorliegt) eine Auftragsverarbeitung. Eine solche ist nur zulässig, wenn die dienstleistende Unternehmen „hinreichend Garantien dafür bieten, dass geeignete technische und organisatorische Maßnahmen so durchgeführt werden, dass die Verarbeitung im Einklang mit den Anforderungen dieser Verordnung erfolgt und den Schutz der Rechte der betroffenen Personen gewährleistet“ (§ 28 DSGVO). Hier gilt im Grunde dasselbe wie für die eigenen technisch-organisatorischen Maßnahmen: Die Anforderungen steigen mit dem Risiko für die Betroffenen. Man muss zeigen können, dass die Dienstleistenden in einer Weise mit den Daten umgehen, die nach dem Stand der Technik gegenüber den Risiken der Datenverarbeitung angemessen ist. Verschlüsselung ist bei Massen von personenbezogenen Daten praktisch Pflicht. Bei sensiblen Daten kann die Abwägung sogar dazu führen, dass ein Outsourcing bzw. die Nutzung von Clouds nicht zulässig ist. Eine sorgfältige Auswahl der Dienstleistenden / Cloudanbietenden liegt dabei im Eigeninteresse. Das Outsourcing von Verarbeitungsprozessen entlastet nämlich nicht davon, für die Einhaltung des Datenschutzes auch bei Fehlern der Dienstleistenden selbst zu haften. In der Praxis sind Gütesiegel und Zertifikate für datenschutzrechtliche Compliance eine gute Orientierungshilfe.

Eine Auftragsverarbeitung erfordert es überdies zwingend, die Dienstleistenden mit einem Vertrag zur Auftragsverarbeitung auf die Einhaltung des Datenschutzrechts zu verpflichten. Das Bayerische Landesamt für Datenschutzaufsicht (2018a) sowie Branchenverbände wie Bitkom (2017) stellen Muster mit den gesetzlichen Mindestanforderungen zur Verfügung. Viele dienstleistende Unternehmen erlauben mittlerweile einen unproblematischen Vertragsschluss online. Bei Anbietern und Anbieterinnen im Massengeschäft sollte geprüft werden, ob mit den bereitgestellten Verträgen wirklich die Anforderungen der DSGVO und die notwendigen technisch-organisatorischen Maßnahmen bindend versprochen werden, was z. B. in Bezug auf Kontrollrechte oder die richtige Löschung von Daten problematisch sein kann.

Wenn personenbezogene Daten in Drittländer außerhalb der EU übertragen werden sollen, ist ein zusätzlicher Aspekt zu beachten. Das ist z. B. der Fall bei der Nutzung von Clouds von US-Anbietern und -Anbieterinnen auf Rechenzentren in den USA. Im Kern erlaubt die DSGVO einen Export personenbezogener Daten nur

dann, wenn die Daten im Zielland ähnlich sicher geschützt sind wie in der EU, entweder ganz generell oder aufgrund von vertraglichen Regeln. Hier stellt die DSGVO verschiedene Optionen zur Verfügung (Einwilligung der Betroffenen, sog. EU-Standardvertragsklauseln, Binding Corporate Rules, Verhaltensregeln nach Art. 46 DSGVO).

Eine Sonderregelung galt mit dem sog. Privacy Shield für die USA, einem Mechanismus, bei dem US-Unternehmen pragmatisch durch eine Erklärung in den USA die Einhaltung des EU-Datenschutzniveaus bestätigen konnten. Dieser Mechanismus wurde im Juli 2020 vom EuGH für unwirksam erklärt (InfoCuria 2020). Auf aktuellem Stand ist die Rechtslage für einen Datenexport in die USA ungeklärt. Wegen des überragenden Praxisbedürfnisses für einen Datenaustausch mit den USA ist zu erwarten, dass zeitnah Lösungen entwickelt werden.

2.3.11 Dokumentationen

„Wer schreibt, der bleibt" ist eine alte Anwaltsweisheit. Dies gilt auch im Datenschutz. Als Verantwortliche:r muss man letztlich beweisen können, dass man DSGVO-konform handelt. Dazu sind saubere Dokumentationen von allen oben erläuterten Schritten notwendig.

Die DSGVO leitet hierzu an, indem es Verfahrensverzeichnisse verlangt (Art. 30 DSGVO), für die die Datenschutzbehörden Muster (Der Hessische Beauftragte für Datenschutz und Informationsfreiheit 2020) zur Verfügung stellen. In diesen Verzeichnissen sind auch die technisch-organisatorischen Maßnahmen zu dokumentieren. Ebenso sind die Verträge über Auftragsverarbeitung und die Datenschutz-Folgenabschätzung verfügbar zu halten. Was sich bürokratisch anhört, ist tatsächlich ein Hilfsmittel, um das eigene Vorhaben sauber planen und durchführen zu können. Wer solche Verzeichnisse selbst einmal anlegt, wird leicht die Erfahrung machen, dass sich dabei zahlreiche Optimierungspotenziale auftun. Ein besonderes Augenmerk sollte dabei auf die Dokumentation all der Maßnahmen gelegt werden, die die oben dargestellten kritischen Punkte für maschinelles Lernen betreffen. Auch sollte die Dokumentation laufende Reviews und ggf. Anpassungen hergeben.

2.3.12 Betriebliche Datenschutzbeauftragte

Schließlich ist daran zu denken, dass unter Umständen ein:e (interne:r oder externe:r) betriebliche:r Datenschutzbeauftragte:r bestellt werden muss. Das ist z. B. der Fall, wenn die Kerntätigkeit der Verantwortlichen im Umgang mit sensiblen Daten besteht (Art. 37 DSGVO), also beispielsweise der Umgang mit Gesundheits- oder biometrischen Daten, oder wenn Verarbeitungsvorgänge stattfinden, für die eine Datenschutz-Folgenabschätzung vorgeschrieben ist (➔ 2.3.8). Letzteres ist

nach Auffassung der Datenschutzaufsicht bei maschinellem Lernen mit personenbezogenen Daten typischerweise der Fall. Ebenso ist eine Bestellung erforderlich, wenn im Unternehmen insgesamt (also nicht nur für die KI, sondern z. B. auch Personalabteilung usw.) in der Regel mindestens 20 Personen ständig mit der automatisierten Verarbeitung personenbezogener Daten beschäftigt werden (§ 38 BDSG). Angesichts der zahlreichen und komplexen datenschutzrechtlichen Vorgaben ist eine lösungsorientierte und fachkompetente Persönlichkeit in dieser Rolle ein Erfolgsfaktor.

3 Rechtliche Aspekte des Trainierens mit Daten

Menge und Qualität der Trainingsdaten in Bezug auf die zu erfüllende Aufgabe sind Schlüsselfaktoren für die geplanten Arbeitsergebnisse eines Systems mit maschinellem Lernen. Trainings- und die zugrundeliegenden Rohdaten sind allerdings oftmals ein knappes Gut. Nicht selten stehen schlicht nicht genügend geeignete Daten für ein Vorhaben zur Verfügung. Aus unternehmerischer Perspektive ist maschinelles Lernen idealerweise mit einer vorausschauenden und mit entsprechenden Ressourcen und (datenschutzrechtlichen) Verantwortlichkeiten unterlegten *Datenstrategie* zu verknüpfen, wie zur Unterstützung des eigenen Geschäftsmodells wertvolle Datenpools aufgebaut und fremdbezogen werden können und wie die eigenen Pools gegenüber Wettbewerbern zu schützen sind. Letztere Frage wird bei Dienstleistungsunternehmen für maschinelles Lernen relevant, die unter Umständen den Trainingsstand des Systems aufgrund der Unternehmensdaten in Zukunft weiterverwenden.

3.1 Welche Daten darf man für das Trainieren verwenden?

Im Folgenden geht es um die recht komplexe Frage, welche Roh- und Trainingsdaten unter welchen Voraussetzungen überhaupt für maschinelles Lernen verwendet werden dürfen. Um einem häufigen Missverständnis vorzubeugen: Man kann nicht ohne Weiteres Daten aus dem Internet oder den sozialen Medien abgreifen, um damit eigene Systeme zu trainieren. Gleiches gilt für Daten von Geschäftspartnern und Geschäftspartnerinnen. Wie man welche Daten verwenden kann, hängt rechtlich davon ab, um was für Daten es sich handelt.

3.1.1 Verschiedene Kategorien von Daten

Im Überblick muss man folgende Arten von Daten differenzieren: nicht personenbezogene Daten, personenbezogene Daten und angereicherte Inhaltsdaten.

Nicht personenbezogene Daten sind solche Informationen, die sich nicht auf eine identifizierte oder identifizierbare Person beziehen. Mit anderen Worten: Solche Informationen, die schlicht nicht mehr einem Individuum zugeordnet werden können, auch nicht, wenn man sie mit anderen Daten kombiniert und so womöglich Baustein zu Baustein fügt. Hierunter fallen Daten, die nichts mit Personen zu tun haben (z. B. Wetterdaten) und sog. Maschinendaten, also technische Informationen wie Messwerte, Zählerstände, statistische Daten und Sensordaten aus digitalisierten Prozessen. Das Potenzial solcher Maschinendaten ist für das maschinelle Lernen und die unternehmerische Wertschöpfung enorm und gilt gerade im Mittelstand als bei Weitem noch nicht ausgereizt (VDMA 2019). Ebenso fallen in die Kategorie der nicht personenbezogenen Daten vollständig anonymisierte Daten. Das sind ursprünglich personenbezogene Daten, bei denen durch Aggregation und

ABBILDUNG 3:
Arten von Daten
Quelle: Bertelsmann Stiftung

Nicht personenbezogene Daten:

nicht auf identifizierte oder identifizierbare natürliche Personen beziehbar

Beispiele: Wetterdaten, Maschinendaten, Messwerte, Zählerstände, statistische Daten, Sensordaten, etc.; vollständig anonymisierte Daten, synthetische Daten

Verwendung: unproblematisch, Datenschutzrecht und DSGVO gelten nicht; ggfs. gelten Beschränkungen in Bezug auf vertragliche Nutzung, Leistungsschutz-/Datenbankrechte, Schutz von Geschäftsgeheimnissen

Personenbezogene Daten

auf identifizierte oder identifizierbare natürliche Personen beziehbar; direkte oder indirekte Zuordnung zu einem Individuum möglich, z. B. in Kombination mit weiteren irgendwo verfügbaren Daten

Beispiele: Name, Adresse, Geburtsdatum, Geschlecht, Fingerabdruck, Steuer-ID, Kontonummer, Standortdaten der Person, KFZ-Kennzeichen, Kapitalvermögen, Kundenkonto, Inhalte in Social-Media-Profilen, IP-Adresse im Serverlog einer Website

Verwendung: Datenschutzrecht und DSGVO ist zu beachten, vertragliche Beschränkungen können ebenfalls gelten.

Inhalte:

durch Immaterialgüterrechte geschützte Werke

Beispiele: Texte, Musik, Graphiken, Fotos, Videos

Verwendung: Das anwendbare Immaterialgüterrecht ist zu beachten, etwa Urheber- oder Leistungsschutzrechte. Know-how-Schutz und vertragliche Beschränkungen können ebenfalls gelten. Enthalten die Werke zugleich personenbezogene Daten wie z. B. bei einem Foto einer realen Person, kommt im Prinzip zugleich das Datenschutzrecht zur Anwendung sowie ggf. andere Persönlichkeitsrechte.

anerkannte Anonymisierungsverfahren der Personenbezug nachträglich entfällt, z. B. Nutzungsstatistiken. Schließlich sind auch sog. synthetische Daten nicht personenbezogenen, also Daten, die sich gar nicht auf real existierende Personen beziehen, aber in ihrer Qualität mit echten personenbezogenen Daten vergleichbar sind (sog. „deep fakes").

Die Verwendung von nicht personenbezogenen Daten in den aufgezeigten Varianten für das maschinelle Lernen ist rechtlich in weiten Zügen unproblematisch erlaubt. Für nicht personenbezogene Daten gelten das Datenschutzrecht und insbesondere die DSGVO nicht, was operativ vieles im Vergleich zur Nutzung von personenbezogenen Daten erleichtert (➔ 3.1.2). Gehören die Maschinendaten im Falle einer Fremdbeschaffung einem anderen Unternehmen, können allerdings vertragliche Nutzungsbeschränkungen (z. B. Non Disclosure Agreements, NDA), Leistungsschutz- / Datenbankrechte sowie der Schutz von Geschäftsgeheimnissen betroffen sein. (➔ 3.1.4 ff.).

Personenbezogene Daten sind demgegenüber solche Informationen, die sich auf identifizierte oder identifizierbare natürliche Personen beziehen. Damit sind solche Informationen gemeint, die man einem Menschen zuordnen kann, und zwar entweder direkt oder auch in der Kombination mit weiteren irgendwo verfügbaren Daten. Letzterer Aspekt ist wichtig und bei der Konzeption der Datenverarbeitung zu bedenken: Daten in der Hand von Anwendenden können auch dann personenbezogen sein, wenn eine Kombination mit anderen Daten, die womöglich in der Hand einer ganz anderen Stelle liegen, die Identifizierung von Personen möglich macht. Dies gilt, sofern die Verantwortlichen (diese Einschränkung kann im Einzelfall für bestimmte Daten relevant werden) auf diese Daten irgendwie zugreifen können, und sei es mit rechtlichen Mitteln.

Sobald ein Personenbezug vorliegt, gilt das Datenschutzrecht. Für die Verwendung von personenbezogenen Daten für das maschinelle Lernen sind die DSGVO und deren ausführende Gesetze die zentrale Regelungsmaterie. Dies gilt auch für den praktisch relevanten Fall, Nutzerdaten auf irgendeine Weise aus dem Internet zu beschaffen (➔ 3.1.3).

Als **Inhalte** werden hier demgegenüber angereicherte Informationen bzw. Werke bezeichnet, die ggf. über das Datenschutzrecht hinaus in anderer Weise rechtlich geschützt sind. Ein gutes Beispiel sind Texte oder Videos. Soll anhand von Texten oder Videos maschinell gelernt werden, ist zu beachten, dass solche Werke durch Urheber- / Leistungsschutzrechte geschützt sein können und damit unter Umständen lizenziert werden müssen. Da auch Inhalte aus Sicht einer KI nichts anderes als Daten sind, aber rechtlich meist besonderem Schutz durch Immaterialgüterrechte oder Persönlichkeitsrechte unterliegen, werden sie hier gesondert behandelt. Inhalte sind in manchen Fällen gleichzeitig personenbezogene Daten (z. B. ein Video, das eine reale Person zeigt), in anderen Fällen nicht (z. B. ein Stück Programmcode). In ersteren Fällen kommt es im Prinzip zu einer Doppelung von anwendbaren Regelungsregimen; zu prüfen ist dann sowohl eine datenschutzrechtliche Zulässigkeit als auch eine Zulässigkeit unter immaterialgüterrechtlichen Aspekten (➔ 3.1.4).

3.1.2 Überblick: Datenschutzrechtliche Grundlagen für die Nutzung personenbezogener Daten zu Trainingszwecken

Sollen für das maschinelle Lernen personenbezogene Daten verwendet werden, muss sich im Prinzip für jedes einzelne Datum zeigen lassen, dass es nach den Vorgaben der DSGVO rechtmäßig erhoben wurde und jetzt im Rahmen des Einsatzes für die künstliche Intelligenz rechtmäßig verarbeitet wird. Anderenfalls ist die Nutzung der Daten (bußgeldbewehrt) verboten. Die hier behandelten Fragen sind typischerweise schon bei der Konzeption eines Systems zu bedenken (➔ 2.3.2).

Die DSGVO sieht in Art. 6 I eine Reihe von Möglichkeiten vor, personenbezogene Daten zu erheben bzw. zu verarbeiten, etwa Einwilligungen der Betroffenen (Art. 6 I 1 a DSGVO), die Verarbeitung zu Zwecken der Vertragserfüllung (Art. 6 I 1 b Var 1 DSGVO) oder die Nutzung aufgrund eines überwiegenden „berechtigten Interesses“ (Art. 6 I 1 f DSGVO). Diese rechtlichen Optionen sind kontextabhängig und müssen im konkreten Anwendungsfall für alle Kategorien von geplanten Trainingsdaten vorab geprüft werden. Die Einwilligung als Rechtsgrundlage ist aus Unternehmenssicht entgegen verbreiteter Ansicht selten die beste Option, weil Einwilligungen jederzeit widerrufen werden können und dann zu klären wäre, wie sich der Widerruf auf das System auswirkt.

Geht es darum, Roh- / Trainingsdaten durch eine erstmalige Erhebung neu zu gewinnen, legt man den Zweck dieser Datenverarbeitung (Nutzung als Trainingsdaten und Nutzung im Rahmen der später geplanten Anwendungsszenarien) fest und begründet dann die Datenverarbeitung aufgrund einer oder mehrerer der in Art. 6 I DSGVO genannten Rechtsgrundlagen.

Sollen allerdings bereits vorhandene personenbezogene Daten für Trainingszwecke gewissermaßen zweitverwertet werden – dies dürfte der Normalfall sein –, dann kommt es, bevor man eine Rechtsgrundlage für die Verarbeitung identifiziert, zunächst auf den Zweck an, zu dem diese Daten ursprünglich erhoben wurden. Für personenbezogene Daten gilt nämlich der Grundsatz der Zweckbindung der Daten (Art. 5 I b DSGVO). Wurden diese Daten bereits ursprünglich auch für Trainingszwecke auf dem konkreten System oder den Zweck des geplanten Einsatzes erhoben, ist ein Trainieren unproblematisch. Dies gilt auch für Daten, die von den Betroffenen frei von jeglicher Zweckbindung in die Public Domain veröffentlicht wurden.

Wurden die Daten jedoch ursprünglich für ganz andere Zwecke erhoben, dürfen diese Daten nur dann für ein Trainieren des Systems verwendet werden, wenn die Voraussetzungen für eine Zweckänderung nach Art. 6 IV DSGVO vorliegen. Dies wiederum kommt nur in Betracht, wenn entweder eine wirksame Einwilligung der Betroffenen vorliegt, es durch eine Spezialvorschrift gestattet ist, oder – dies ist der wichtigste Fall – der neue Trainingszweck nicht inkompatibel mit dem ursprünglichen Verarbeitungszweck ist. Letzteres setzt eine komplizierte Prüfung im Einzelfall voraus, bei der eine Reihe abstrakter Kriterien zu beachten ist: eine mögliche Verbindung bzw. Sachzusammenhang zwischen dem ursprünglichen Erhebungszweck und dem Trainingszweck / dem Zweck des Systems; ein möglicher Zusam-

menhang, in dem die Daten erhoben wurden; mögliche Folgen der Datenverarbeitung sowie das Vorhandensein von Garantien zum Schutz der Betroffenen, insbesondere Anonymisierung, Pseudonymisierung und Verschlüsselung. Hier wird deutlich: Die engen, komplizierten und rechtsunsicheren Zulassungsvoraussetzungen für eine Zweckänderung stellen die zentrale Hürde für Big-Data-Anwendungen und maschinelles Lernen aus bestehenden Datenpools dar. Umgekehrt kann eine Zweckänderung leichter gerechtfertigt werden, wenn man auf technisch-organisatorischer Ebene sorgfältig und datenschützend mit den Daten umgeht.

Möglicherweise kommt für das konkrete System auch die besondere Ausnahme der DSGVO für im öffentlichen Interesse liegende statistische Zwecke, Archivzwecke, wissenschaftliche oder historische Zwecke (Art. 5 I b, 89 DSGVO) in Betracht. In diesen besonderen Fällen wird eine Zweckkompatibilität gesetzlich vermutet, was die Nutzung von bereits erhobenen Daten für solche Zwecke wesentlich erleichtert.

Die Zweckbindung der Trainingsdaten bedeutet auch, dass man bei Änderungen der Aufgabenstellung des Systems die bisherigen Datenbestände nur dann weiter nutzen kann, wenn sich die Änderungen im Rahmen des geplanten Zwecks bewegen oder die Voraussetzungen für eine Zweckänderung vorliegen. Hat man beispielsweise einen Trainingsdatenbestand für ein System zur Empfehlung von Musik (mit personenbezogenen Daten) aufgebaut, darf dieser Bestand nicht für ein System zur Identifizierung der politischen Tendenz genutzt werden, auch wenn dies technisch ohne Weiteres möglich wäre. Würde sich ein Trainingsvorgang außerhalb des ursprünglichen Zwecks bewegen und scheidet eine Zweckänderung aus, bleibt nur die Überlegung, ob das Training neu auf Basis einer Nutzungsbefugnis nach Art. 6 I DSGVO gerechtfertigt werden könnte.

Eine Anonymisierung von personenbezogenen Daten, um aus diesen Trainingsdaten zu gewinnen, ist allerdings zulässig, weil dies im Ergebnis den Datenschutz fördert. In vielen Fällen, in denen die Nutzung von Daten aus dem Internet an dem Zweckbindungsgrundsatz scheitert, könnte eine (fachgerechte) Anonymisierung einen Lösungsweg darstellen. Gegenwärtig wird kontrovers diskutiert, ob eine Anonymisierung ebenfalls einer ausdrücklichen Rechtsgrundlage bedarf, was die Lage kompliziert macht und eigentlich widersinnig ist. Hier ist zu empfehlen, den Stand der Diskussion in den aktuellen Verlautbarungen der Datenschutzbehörden zu verfolgen.

Ist die Nutzung von Daten aus dem Netz ausnahmsweise mit dem Zweckbindungsgrundsatz kompatibel, ist umstritten, ob man dann rechtlich zusätzlich auch noch eine Rechtsgrundlage nach Art. 6 I DSGVO benötigt (so die herrschende Auffassung), oder ob mit einer Zweckänderung nach Art. 6 IV bereits alle Voraussetzungen für die Nutzung der Daten erfüllt sind. Wer sichergehen will, stützt sich immer auch auf eine Rechtsgrundlage nach Art. 6 I DSGVO.

Bislang kaum erörtert ist das Problem der Fehler in massenhaften Datensammlungen. Bei enormen Datenmengen ist es praktisch nicht auszuschließen, dass sich Ausreißer in den Datensätzen befinden, die nicht datenschutzkonform sind. Dies erkennt auch die Datenschutzaufsicht an. Selbst bei sorgfältigem Handeln kann

mal ein Fehler passieren. Datenschutz bewirkt letztlich eine Organisationsverantwortung. Es geht weniger darum, dass es nicht zu einzelnen Fehlern in der Datenverarbeitung kommt, sondern vielmehr darum, dass der Datenschutz konzeptionell gewährleistet ist und die Verantwortlichen sorgfältige Prozesse der Auswahl, Bereinigung und Kontrolle nachweisen können. Ganz ähnlich werden Ausreißer im Produkthaftungsrecht und im Arbeitsrecht behandelt. Für das Restrisiko von Ausreißern kann man ähnlich wie im Massen-Lizenzgeschäft verfahren und eine angemessene Rückstellung für Ansprüche von Betroffenen bilden.

Wenn personenbezogene Daten verwendet werden, dann löst dies unabhängig von den Fragen der Zweckbindung und der Rechtsgrundlage für die Datenverarbeitung eine Reihe von weiteren Pflichten aus, die bereits beim Design eines Systems zu berücksichtigen sind, also Transparenzpflichten, Pflichten zur Erfüllung der Betroffenenrechte und Pflichten zum Einsatz angemessener technisch-organisatorischer Maßnahmen (➔ 2.3), insbesondere auch für den Prozess der Veredelung von Rohdaten zu einem Trainingsdatensatz (➔ 2.3.3).

Besonders sensibel ist die Verarbeitung personenbezogener Daten besonderer Kategorien, aus denen die rassische und ethnische Herkunft, politische Meinungen, religiöse oder weltanschauliche Überzeugungen oder die Gewerkschaftszugehörigkeit hervorgehen, ebenso genetische oder biometrische Daten, Gesundheitsdaten oder Daten zu Sexualleben / sexueller Orientierung. Für solche Daten erhöht die DSGVO an verschiedenen Stellen die Anforderungen an eine rechtmäßige Nutzung (vgl. Art. 9 DSGVO). Bei solchen Daten ist das Risiko der Datenverarbeitung für die Betroffenen besonders hoch. Entsprechend wird eine Datenschutzaufsicht besonderen Augenmerk auf den rechtmäßigen Umgang mit solchen Daten legen. Für die Praxis ist zu raten, sich bei personenbezogen Daten besonderer Kategorien überdies an den einschlägigen Ethik-Grundsätzen für künstliche Intelligenz (Europäische Kommission 2019b) zu orientieren, etwa bei Systemen der Gesichtserkennung.

3.1.3 Nutzerdaten aus dem Internet und aus sozialen Netzen

Praktisch relevant ist natürlich die Frage, wie und in welchem Umfang Nutzerdaten aus dem Internet und aus sozialen Medien genutzt werden können. Auch hier muss sich nach den Vorgaben der DSGVO zunächst einmal eine Rechtsgrundlage aufzeigen lassen, die die konkrete Datenerhebung und -nutzung vollumfänglich rechtfertigt. Zudem sind der Zweckbindungsgrundsatz und unter Umständen die Nutzungsbedingungen der Plattformen / sozialen Netze zu beachten. Dies gilt sowohl dann, wenn man mit einem eigenen „Crawler“ Daten aus dem Netz erhebt, als auch dann, wenn man mit Anbietenden von Plattformen oder sozialen Netzen vertragliche Vereinbarungen über einen Datenzugriff schließt und diesen über Schnittstellen, Apps, „Fanpages“ usw. realisiert.

3.1.3.1 Einwilligungslösungen

Eine Option besteht darin, Einwilligungen der Betroffenen einzuholen. Mit diesem Mittel generieren US-amerikanische und chinesische Digitalkonzerne riesige eigene Datenpools, indem sie, insbesondere im Business-to-Consumer-Geschäft (B2C), Zugang zu den digitalen Dienstleistungen gegen eine Einwilligung in die Nutzung der personenbezogenen Daten der Kunden und Kundinnen zur Verfügung stellen. Die Einwilligung erklären diese regelmäßig bei Registrierung für den Dienst. Die Möglichkeit, eine eigene Einwilligung einzuholen, besteht theoretisch gesehen immer, und sei es durch individuelle Abfragen. Es ist stets eine Überlegung wert, ob man einen eigenen Datenpool mit den Daten der eigenen Kunden und Kundinnen aufbauen kann. Praktisch gesehen stehen Einwilligungslösungen allerdings vor großen Problemen. Plattformbetreiber:innen und soziale Netze gestatten das Abfragen von Daten normalerweise nicht ohne Weiteres und sichern sich rechtlich und durch technische Maßnahmen ab. Unzählige Webseitenbetreiber:innen um Erlaubnis zu ersuchen, ihre Daten nutzen und auswerten zu dürfen, wird in der Regel am Aufwand scheitern. Zu beachten ist dabei auch, dass die Anforderungen an eine rechtskonforme Einwilligung hoch sind: Die Nutzer:innen müssen eine auf angemessenen Informationen beruhende, willensmängelfreie, unmissverständliche und nachträglich beweisbare Einwilligungserklärung abgeben. In Allgemeinen Geschäftsbedingungen (AGB) versteckte Einwilligungserklärungen in einer Sprache, die mehr verschleiert als aufklärt, genügen nicht, auch wenn man dies in der Praxis immer mal wieder sieht. Einwilligungen können überdies jederzeit widerrufen werden.

Davon unabhängig ist es vorstellbar, dass Plattformanbieter:innen oder soziale Netze von den eigenen Kunden und Kundinnen die Einwilligung einholen, die in dem eigenen Angebot erzeugten Daten an Dritte weitergeben zu dürfen, etwa zur Verwendung als Trainingsdaten. Nach den aufgezeigten Grundsätzen setzt dies voraus, dass die Nutzer:innen vor Abgabe der Einwilligungserklärung erfahren, an wen konkret welche Daten zu welchem Zweck geliefert werden. Eine pauschale Zustimmung zu einer Datenweitergabe an beliebige Dritte ist demnach nicht möglich, wohl aber eine individuelle Abfrage der Datenweitergabe z. B. bei der Verwendung von Schnittstellen, Apps oder Authentifizierungsverfahren.

3.1.3.2 Lösungen aufgrund eines überwiegenden berechtigten Interesses

Alternativ zu einer Einwilligungslösung kommt als Rechtfertigung für einen Bezug von Trainingsdaten aus dem Internet oder sozialen Medien der Rechtfertigungsgrund „überwiegendes berechtigtes Interesse“ in Betracht. Dieser Rechtfertigungsgrund (Art. 6 I f DSGVO) ist für die Rechtsanwendung schwer zu handhaben, weil er hochabstrakt ist, eine Abwägung beinhaltet und bislang wenig ausfüllende Präzedenzen zur Verfügung stehen. Aber wegen der praktischen Schwierigkeiten bei Einwilligungslösungen ist er von großer Bedeutung. Im ersten Schritt muss man aufzeigen können, ein berechtigtes Interesse daran zu haben, personenbezogene

Daten als Trainingsdaten für maschinelles Lernen zu verwenden. Das wäre bei einem Training für ein legitimes kommerzielles System normalerweise anzunehmen. In einem zweiten Schritt muss man zeigen, dass die Verarbeitung der Daten für das berechtigte Interesse erforderlich ist. Auch dies dürfte selten eine Hürde darstellen. Gleichwohl sieht man auch an dieser Voraussetzung, dass Datenbestandteile in den Datensätzen, die man für das Trainieren nicht benötigt, nicht erhoben und verarbeitet werden dürfen. In einem dritten und entscheidenden Schritt muss dann aber gegen die (Grund)Rechte der betroffenen Personen abgewogen werden. Das Privatheitsinteresse der Betroffenen darf nicht überwiegen. An diesem Punkt beginnt die argumentative Arbeit, die letztlich von den Rahmenbedingungen und Einzelfallumständen des betreffenden Systems abhängt. Aber bestimmte Leitlinien lassen sich geben. Auch ist davon auszugehen, dass die Datenschutzbehörden und Gerichte diese Vorschrift immer weiter konkretisieren werden.

Für ein überwiegendes Interesse an der Verwendung für maschinelles Lernen spricht z. B.:

- Die personenbezogenen Daten wurden von den Betroffenen mit der Zielsetzung veröffentlicht, dass jede:r die Daten sehen / nutzen kann.
- Die Betroffenen haben die Daten selbst der Öffentlichkeit (und nicht nur einem begrenzten Nutzerkreis in sozialen Medien) zugänglich gemacht.
- Die beabsichtigte Nutzung für das maschinelle Lernen steht in einem engen Zusammenhang mit dem Zweck, zu dem die Daten ursprünglich erhoben wurden.
- Das beabsichtigte maschinelle Lernen hat einen Gemeinwohlnutzen.
- Die Daten werden unverzüglich pseudonymisiert.

Für ein überwiegendes Privatheitsinteresse der Betroffenen (und damit gegen die Zulässigkeit einer Verwendung) spricht:

- Die Daten über die Betroffenen wurden rechtswidrig erhoben.
- Die Datensätze wurden nicht angemessen bereinigt und pseudonymisiert, enthalten für das Training unerhebliche weitere Daten und verletzen das Gebot der Datensparsamkeit.
- Betroffene haben in irgendeiner Weise signalisiert, dass sie mit der Verwendung ihrer Daten nicht einverstanden sind (z. B. Beschränkung des Zugriffs, Opt-out aus Suchfunktionen, strenge Privacy-Settings).
- Die personenbezogenen Daten wurden nicht von den Betroffenen in das Netz eingestellt und diese haben ein berechtigtes Löschungsinteresse (z. B. persönlichkeitsrechtsverletzende Paparazzi-Aufnahmen).
- Es handelt sich um sensible Daten besonderer Kategorien (Art. 9 DSGVO), z. B. biometrische Daten für eine Gesichtserkennung. Die Verarbeitung sensibler Daten ist allerdings nicht grundsätzlich verboten, sondern unterliegt nur erhöhten Schutzanforderungen, was auch in der Abwägung zu berücksichtigen ist.
- Die Gewinnung der Trainingsdaten würde getarnt oder unter Täuschung erfolgen (wie z. B. im Cambrigde-Analytica-Skandal durch eine App mit dem verharmlosenden Titel „thisisyourdigitallife“ oder durch ein Fake-Profil).

- Die Gewinnung der Trainingsdaten würde den Allgemeinen Geschäftsbedingungen der Anbieter:innen von Plattformen / sozialen Netzen widersprechen, die oftmals eine unautorisierte Datenerhebung auf den eigenen Plattformen untersagen.
- Die Daten sind nicht allgemein öffentlich zugänglich und sollen zu einem anderen Zweck genutzt werden als dem, zu dem sie ursprünglich erhoben wurden, werden also für ein maschinelles Lernen aus ihrem ursprünglichen Kontext entfernt (und auch eine ausnahmsweise Zweckänderung nach Art. 6 IV DSGVO und § 24 BDSG kommt nicht in Betracht).
- Es bestehen hohe Risiken einer Diskriminierung, des Identitätsdiebstahls / -betrugs, finanzieller Verluste oder einer Rufschädigung.

Überdies lassen sich für eine Abwägung auch Argumente gewinnen aus dem Umfang der Trainingsdaten, deren Detailliertheit, aus den Auswirkungen auf die Betroffenen (beim Trainieren einer KI typischerweise sehr gering) und aus den Maßnahmen, die zur Verhinderung von Diskriminierungen bei dem Training getroffen wurden (➔ 2.3.5).

3.1.3.3 Zweckänderungen

Auch für Daten aus dem Netz gilt der Zweckbindungsgrundsatz (➔ 3.1.2). Dies stellt kein Problem dar, wenn die Betroffenen die Daten frei im Netz veröffentlicht haben, denn dann wurden die Daten gerade nicht für einen bestimmten Zweck gewidmet. Dann kann man sie auch – vorbehaltlich der übrigen datenschutzrechtlichen und immaterialgüterrechtlichen Anforderungen – für Trainingszwecke nutzen. Anders sieht es aus, wenn die Daten einer Zweckbestimmung unterliegen, beispielsweise, wenn sie in einer geschlossenen Nutzergruppe / einem sozialen Netz zu einem bestimmten Thema veröffentlicht wurden. Eine Verwendung solcher Daten ist nur zulässig, wenn die Voraussetzungen für eine Zweckänderung nach Art. 6 IV DSGVO vorliegen oder wenn die Erhebung dieser Daten über eine Rechtsgrundlage nach Art. 6 I DSGVO neu legitimiert wird.

Es wird deutlich: Die Gewinnung von Trainingsdaten aus dem Netz ist oftmals auch ohne Einwilligung möglich, erfordert aber die gründliche und dokumentierte Entwicklung eines Konzepts. In vielen Fällen dürfte der Weg über eine Anonymisierung oder synthetische Daten die Lösung mit dem geringsten Aufwand darstellen.

3.1.4 IP, Schutzrechte und Know-how-Schutz bei der Nutzung von fremden Inhalten zu Trainingszwecken

Rechtliche Beschränkungen der Nutzung von Inhalten zu Trainingszwecken können sich zudem aus Immaterialgüterrechten, Know-how-Schutz oder Verträgen ergeben. Ob z. B. Urheber- oder Leistungsschutzrechte zu rechtlichen Beschränkungen führen, hängt zunächst von der Herkunft der Inhalte ab (in diesem Zusam-

menhang werden die Trainingsdaten als Inhalte, Content oder Trainingsinhalte bezeichnet, um eine Abgrenzung zu den anderen Datenformen, wie Rohdaten etc., zu verdeutlichen, ➔ Abbildung 3). Handelt es sich um selbst erzeugte Inhalte, werden die notwendigen (Nutzungs)Rechte in aller Regel vorhanden sein. Gegebenenfalls ist ratsam, dies in den Verträgen mit Angestellten und Mitarbeitenden zu regeln, die an der Erzeugung / Schaffung des Contents beteiligt sind.

Bei einer KI-bezogenen Nutzung von fremden Datenbeständen (an denen keine eigenen Rechte bestehen), die Inhalte enthalten, sind schutzrechtliche Aspekte generell zu beachten. Ob die Verwendung von Trainingsdaten diesbezüglichen Beschränkungen unterliegt, hängt neben ihrer Herkunft vor allem davon ab, um welche Art von Inhalten es sich handelt. Anders als reine Fakten und Informationen (Daten im eigentlichen Sinn) können Inhalte wie Texte, Musik, Software u. a. durch Immaterialgüterrechte wie das Urheberrecht oder Leistungsschutzrechte geschützt sein. Sollen sie zum Trainieren verwendet werden, bedarf es hierfür einer Erlaubnis (vertragliche Gestattung = Lizenz) oder einer gesetzlichen Befugnis.

3.1.4.1 Urheberrecht und Leistungsschutzrechte an Trainingsinhalten

Urheberrecht und verwandte Schutzrechte (synonym mit: Leistungsschutzrechte) schützen kreative Gestaltungen und hiermit zusammenhängende Leistungen gegen ungefragte Verwendung. Das Urheberrecht schützt „persönliche geistige Schöpfungen“ (= Werke). Damit gemeint sind menschliche kreative Äußerungen (Gestaltungen) wie Texte, Fotos oder auch Computerprogramme.

Leistungsschutzrechte schützen dagegen verschiedene Arten von Leistungen, die in der Regel in der Vermittlung oder Aufbereitung von urheberrechtlich geschützten Werken liegen. Beispielsweise werden Tonträger- und Filmherstellerrechte für die Produktion von Musik und Filmen gewährt.

Das **Datenbankherstellerrecht** schützt die Investitionen der Hersteller:innen von Datenbanken.

Das **Lichtbildrecht** ist dem Urheberrecht sehr ähnlich. Es existiert neben dem Urheberrecht an Fotografien, das nur an in gewissem Maß kreativen Aufnahmen entsteht. Anders als das Urheberrecht ist das Lichtbildrecht unabhängig von qualitativen Anforderungen. Es gilt daher für solche Fotografien, die das für ein Urheberrecht notwendige Maß an Individualität bzw. Kreativität nicht erreichen (wie z. B. simple Schnappschüsse).

Das Urheberrecht und die Leistungsschutzrechte entstehen automatisch. Anders als beispielsweise das Patentrecht erfordern sie also keinen formalen Akt, müssen nicht beantragt oder bewilligt werden. Sie entstehen durch Erbringung der Werkschöpfung oder jeweiligen Leistung ohne weiteres Zutun.

Urheberrechtsschutz

Das Urheberrecht schützt keine Fakten oder Informationen, sondern nur menschliche kreative Schöpfungen. Daten und Informationen sind keine Schöpfungen. Sie werden nicht geschaffen, sondern existieren einfach. Dementsprechend gibt es keinen urheberrechtlichen Schutz von Rohdaten. Erst wenn diese Daten weiterverarbeitet – beispielsweise in einem Text verschriftlicht – werden, kommt ein Urheberrechtsschutz in Betracht. Dabei sind die Anforderungen an die urheberrechtliche Schutzfähigkeit (die sog. Schöpfungshöhe) generell sehr gering. Selbst banale und kurze Texte, einfache Computerprogramme, schlichte Popmusikstücke oder bereits kleine Filmausschnitte sind meist urheberrechtlich geschützt. Ist das der Fall, dürfen sie nur genutzt – z. B. für eine Datenanalyse auf den eigenen Server kopiert – werden, wenn hierfür eine vertragliche oder gesetzliche Gestattung vorliegt.

„Das beste Feature dabei ist, dass der Computer nicht einfach nur das reproduzieren kann, was du ihm vorgibst oder was Bach bereits geschrieben hat. Dieses Netzwerk ist in der Lage, den Prozess zu generalisieren. Musik kann in Interaktion mit dem Computer generiert werden. Ich gebe etwa eine Melodie ein und Deep Bach komponiert die drei Begleitstimmen und Harmonien dazu. So wird Komposition auch für Menschen zugänglich, die nicht den theoretischen Hintergrund haben." (berlinvalley.com 2018).

Das Urheberrecht endet 70 Jahre nach dem Tod des Autors. Entsprechend können Werke, bei denen diese Schutzdauer abgelaufen ist, frei verwendet werden. Dies hat beispielsweise das Projekt „DeepBach" (berlinvalley.com 2018) ermöglicht, in dem eine KI trainiert wurde, Bach-artige Werke zu komponieren.

Leistungsschutzrechte

Trainingsinhalte können auch durch Leistungsschutzrechte geschützt sein. Eine Musikaufnahme beispielsweise ist – neben dem Urheberrecht an Komposition und Text für die Musik als solche – durch das Tonträgerherstellerrecht geschützt und darf nur unter bestimmten Umständen genutzt werden. Andere Leistungsschutzrechte gelten für einfache Fotografien, Film- oder Videosequenzen und vieles mehr.

Die Verwendung von Daten aller Art (gleich, ob es sich um Inhalte oder um Rohdaten handelt), die in Datenbanken gespeichert sind, kann durch das Datenbankherstellerrecht reglementiert sein. Das Datenbankherstellerrecht schützt nicht die Daten selbst, sondern lediglich die Datenbank. Einzelne Daten aus Datenbanken zu kopieren, verstößt daher nicht gegen Datenbankherstellerrechte. Der Abgriff der Daten wird erst relevant, wenn wesentliche Teile des Datenbestandes übernommen werden. Das ist nicht nur dann der Fall, wenn ein quantitativer Großteil des Datenbestandes kopiert wird. Da das Datenbankherstellerrecht dem Investitionsschutz dient, schützt es auch vor Entnahmen / Übernahmen, die in qualitativer Hinsicht wesentlich sind. Enthält eine Datenbank also beispielsweise vor allem öffentlich verfügbare Informationen und darüber hinaus einen kleinen Anteil schwer zu beschaffender oder aufwendig erzeugter Analysedaten, kann auch die Übernahme der letzteren „wesentlich" sein, obwohl sie quantitativ nur einen kleinen Teil des Datenbestandes ausmachen.

Auch eine sukzessive Entnahme jeweils unwesentlicher Teile kann gegen das Datenbankherstellerrecht verstoßen. Greift ein „Crawler" beispielsweise bei jedem Zugriff nur zwei Datensätze ab, liegt hierin jeweils keine wesentliche Entnahme. Werden jedoch kleine Datenmengen systematisch durch wiederholte Zugriffe entnommen, kann auch dies sukzessiv zu einer wesentlichen Entnahme führen.

Schutzrechte bei frei und öffentlich verfügbaren Inhalten oder Datenbanken

Ob geschützte Inhalte von ihrem Rechteinhaber frei verfügbar gemacht werden (z. B. im Internet), ändert grundsätzlich nichts an ihrem Schutz. Keineswegs können Internetinhalte ohne Weiteres frei verwendet werden, nur weil sie ohne Zugangsbeschränkungen im Netz stehen. Gleiches gilt für Online-Datenbanken. Auch bei im Netz veröffentlichten und kostenlos nutzbaren Inhalten und Datenbanken gelten die oben genannten Regeln. Wenn sich Rechteinhaber entscheiden, ihre Errungenschaften frei ins Netz zu stellen, verzichten sie hierdurch nicht auf ihre Rechte.

Befugnisse zur Nutzung von durch Urheber- oder Leistungsschutzrechte geschützte Inhalte zu Trainingszwecken

Verträge und Lizenzen

Texte, Fotos, Tonaufnahmen oder andere geschützte Inhalte dürfen nur im Rahmen von Datenanalysen oder Trainingsprogrammen genutzt werden, wenn hierfür eine Befugnis besteht. Nutzen bedeutet aus Sicht des Urheberrechts und der Leistungsschutzrechte, dass die Daten kopiert, öffentlich verfügbar gemacht oder weitergegeben werden. Rein flüchtige Kopien, wie sie bei der Nutzung digitaler Inhalte stets im Arbeitsspeicher des Nutzerrechners entstehen, sind unbedenklich (sie sind gesetzlich gestattet). Werden geschützte Werke oder Datenbanken jedoch dauerhaft zu Analysezwecken kopiert und gespeichert, greift das Schutzrecht und es bedarf einer gesetzlichen oder vertraglichen Gestattung. Das bedeutet, dass zwar nicht jede Verwendung von geschützten Inhalten zum „machine learning" urheberrechtlich / leistungsschutzrechtlich relevant ist. Gehen die Maßnahmen jedoch mit Kopien und Speicherungen einher, sind solche Rechte zu beachten. Das ist z. B. immer der Fall, wenn die Trainingsinhalte auf einem Server, etwa in der Cloud oder einem „Data Warehouse", dauerhaft gespeichert werden.

Zulässig sind Nutzungen, wenn sie entweder per Gesetz erlaubt sind oder vom Rechteinhaber gestattet wurden. Einen Vertrag abzuschließen (eine Lizenz einzuholen) hat den Vorteil der Rechtssicherheit. Stimmt die:der Inhaber:in der Daten zu, kann man sie nach den ausgehandelten vertraglichen Regeln nutzen. Je nach Datenquelle wird dieser Weg jedoch faktisch nicht immer möglich sein. Insbesondere bei der Erschließung, Analyse und Auswertung von großen Onlinedatenbeständen (beispielsweise von frei zugänglichen Onlineinhalten) scheitern vertragliche Maßnahmen häufig an der Masse der Rechteinhaber, die man für eine Einzelrechteklärung kontaktieren müsste. Wer beispielsweise Tausende von frei zugänglichen Nachrichtenartikeln nutzen will, müsste unter Umständen Tausende Verträge abschließen.

Die notwendigen Befugnisse können auch durch genannte freie bzw. offene Lizenzen erlangt werden. Auch hierbei handelt es sich um Verträge, durch die Nutzungsrechte eingeräumt werden (Lizenzverträge). Solche Lizenzen gibt es in vielerlei Form, beispielsweise für Software (sog. „Open Source"-Software) oder für Inhalte wie Bilder oder Texte („Open Content"). „Open Content"- und „Open Source"-Lizenzen gestatten es allen Interessierten, die hierunter lizenzierten Artefakte zu nutzen, ohne eine individuelle Gestattung einzuholen. Die Nutzung unterliegt vertrag-

ABBILDUNG 4:
Creative Commons

Namensnennung 4.0 International

Namensnennung – Share Alike 4.0 International

Namensnennung – Keine Bearbeitungen 4.0 International

Namensnennung – Nicht kommerziell 4.0 International

Namensnennung – Nicht kommerziell – Share Alike 4.0 International

Namensnennung – Nicht kommerziell – Keine Bearbeitungen 4.0 International

Quelle: https://de.creativecommons.net/was-ist-cc/

lich über die Lizenz festgelegten und rechtlich verbindlichen Regeln, die es zu beachten gilt. Im Internet finden sich Millionen von Bildern, Texten, Videos und Computerprogrammen, die unter solchen Lizenzen verfügbar gemacht werden. Solche Inhalte können in aller Regel zu Trainingszwecken oder zur Datenanalyse verwendet werden, da die offenen Lizenzen solche und die meisten anderen Nutzungen gestatten.

Gerade für Datensammlungen werden oft freie Lizenzen eingesetzt. In diesem Fall werden die Daten als „Open Data" oder „Open Research Data" gekennzeichnet, indem sie unter einer „Open Data"-Lizenz öffentlich zugänglich gemacht werden. Hierdurch sollen Datenmengen für jeden zugänglich und nutzbar gemacht werden, u. a. um Innovationen zu fördern. Sie können nach den für die jeweilige Datenbank geltenden Lizenzbestimmungen verwendet werden. Für „Open Data" gibt es spezielle offene Lizenzen, wie die „Open Database License".

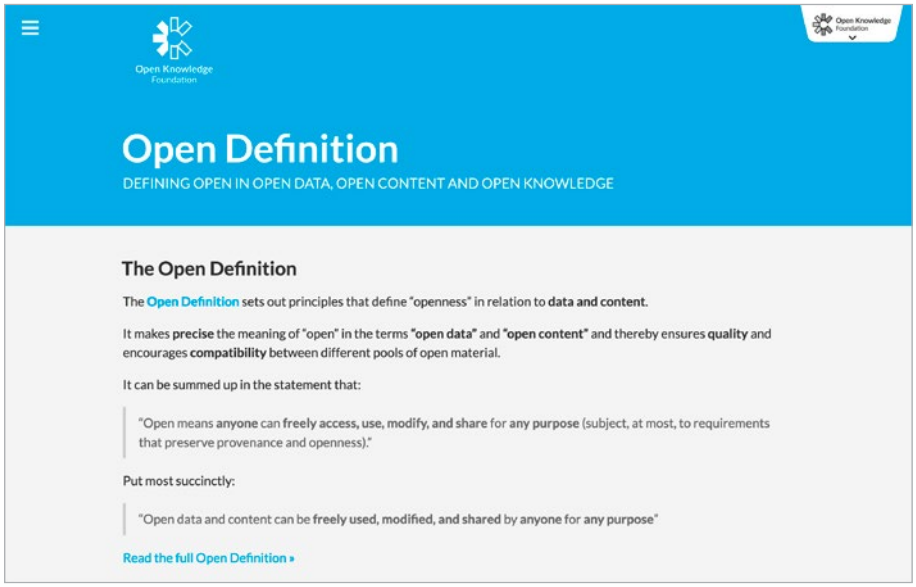

ABBILDUNG 5:
„Open Data"
Quelle: https://opendefinition.org/

Gesetzliche Nutzungsbefugnisse (Schrankenbestimmungen)

Sind urheber- oder leistungsschutzrechtlich geschützte Trainingsinhalte dagegen weder durch offene Lizenzen zur Nutzung freigegeben noch können hierüber Verträge geschlossen werden, dürfen sie nur unter besonderen Umständen nach den gesetzlichen Regeln kopiert und / oder weitergegeben werden. Das deutsche Gesetz (das Urheberrechtsgesetz, UrhG) sieht für manche Arten von Nutzungen Ausnahmen vor. Diese sog. Schrankenbestimmungen machen es möglich, geschütztes Material zu verwenden, ohne hierfür eine individuelle Erlaubnis einzuholen. Sie gelten stets nur für bestimmte Zwecke, z. T. auch nur für bestimmte Nutzergruppen (es gibt beispielsweise spezielle Regelungen für die Nutzung im Unterricht und zu Forschungszwecken, auf die sich nur öffentliche Bildungs- und Forschungseinrichtungen berufen können, die keine kommerziellen Zwecke verfolgen).

Die automatisierte Auswertung von urheberrechtlich geschütztem Material zu wissenschaftlichen Zwecken ist unter gewissen Umständen gestattet (nach einer Schrankenbestimmung zu Text- und Data-Mining, § 60d UrhG). Eine andere Regelung gestattet die Nutzung von geschütztem Material in gewissen Grenzen zu Forschungszwecken (§ 60c UrhG).

Diese gesetzlichen Nutzungserlaubnisse sind (bislang) auf Nutzungen zu nicht kommerziellen Zwecken beschränkt. Unternehmensbezogene Verwendungszwecke werden hierdurch derzeit nicht gestattet. Allerdings werden sie im Rahmen einer Neuregelung erweitert werden, die – aufgrund der Vorgaben einer EU-Richtlinie – bis Juni 2021 umgesetzt werden soll. Eine neue Schrankenbestimmung wird dann auch für (urheberrechtlich relevantes) Text- und Data-Mining und Datenanalysen gelten, die zu kommerziellen Zwecken vorgenommen werden. Wie die Regelung im Detail aussehen wird, wird sich allerdings erst im Gesetzgebungsverfahren zeigen, das 2020 begonnen hat.

Referentenentwurf des Bundesministeriums der Justiz und für Verbraucherschutz zu Text- und Data-Mining (TDM) im Urheberrechtsgesetz (UrhG) vom 2.9.2020

§ 44b Text und Data Mining

(1) Text und Data Mining ist die automatisierte Analyse von einzelnen oder mehreren digitalen oder digitalisierten Werken, um daraus Informationen insbesondere über Muster, Trends und Korrelationen zu gewinnen.

(2) Zulässig sind Vervielfältigungen von rechtmäßig zugänglichen Werken für das Text und Data Mining. Die Vervielfältigungen sind zu löschen, wenn sie für das Text und Data Mining nicht mehr erforderlich sind.

(3) Nutzungen nach Absatz 2 Satz 1 sind nur zulässig, wenn der Rechtsinhaber sich diese nicht vorbehalten hat. Ein Nutzungsvorbehalt bei online veröffentlichten Werken ist nur dann wirksam, wenn er in maschinenlesbarer Form erfolgt.

§ 60d Text und Data Mining für Zwecke der wissenschaftlichen Forschung

(1) Vervielfältigungen für Text und Data Mining (§ 44b Absatz 1 und Absatz 2 Satz 1) für Zwecke der wissenschaftlichen Forschung sind nach Maßgabe der nachfolgenden Bestimmungen zulässig.

(2) Zu Vervielfältigungen berechtigt sind Forschungsorganisationen. Forschungsorganisationen sind Hochschulen, Forschungsinstitute oder sonstige Einrichtungen, die wissenschaftliche Forschung betreiben, sofern sie

1. nicht kommerzielle Zwecke verfolgen,
2. sämtliche Gewinne in die wissenschaftliche Forschung reinvestieren oder
3. im Rahmen eines staatlich anerkannten Auftrags im öffentlichen Interesse tätig sind.

Nicht nach Satz 1 berechtigt sind Forschungsorganisationen, die mit einem privaten Unternehmen zusammenarbeiten, das einen bestimmenden Einfluss auf die Forschungsorganisation und einen bevorzugten Zugang zu den Ergebnissen der wissenschaftlichen Forschung hat.

(3) Zu Vervielfältigungen berechtigt sind ferner
   1. öffentlich zugängliche Bibliotheken, Archive, Einrichtungen im Bereich des Film- oder Tonerbes und öffentlich zugängliche Museen sowie
   2. einzelne Forscher, sofern sie nicht kommerzielle Zwecke verfolgen.

(4) Berechtigte nach den Absätzen 2 und 3, die nicht kommerzielle Zwecke verfolgen, dürfen Vervielfältigungen nach Absatz 1 einem bestimmt abgegrenzten Kreis von Personen für deren gemeinsame wissenschaftliche Forschung sowie einzelnen Dritten zur Überprüfung der Qualität wissenschaftlicher Forschung öffentlich zugänglich machen. Sobald die gemeinsame wissenschaftliche Forschung oder die Überprüfung der Qualität wissenschaftlicher Forschung abgeschlossen ist, ist die öffentliche Zugänglichmachung zu beenden.

(5) Berechtigte nach Absatz 2 Satz 1 und Absatz 3 Nummer 1 dürfen Vervielfältigungen nach Absatz 1 mit angemessenen Sicherheitsvorkehrungen gegen unbefugte Benutzung aufbewahren, solange sie für Zwecke der wissenschaftlichen Forschung oder zur Überprüfung wissenschaftlicher Erkenntnisse erforderlich sind.

(6) Rechtsinhaber sind befugt, erforderliche Maßnahmen zu ergreifen, um zu verhindern, dass die Sicherheit und Integrität ihrer Netze und Datenbanken durch Text und Data Mining gemäß Absatz 1 gefährdet werden.

Quelle: Bundesministerium der Justiz und für Verbraucherschutz 2020

3.1.4.2 Marken- und Designrechte

Marken- und Designrechte können an geschützten Produktgestaltungen oder -bezeichnungen, Firmennamen und vielerlei anderen *Assets* bestehen. Beispielsweise würde eine Bilddatenbank mit 3-D-Modellen von Kraftfahrzeugen oder Mobiltelefonen viele geschützte Marken und Designs enthalten.

Selbst wenn das jedoch der Fall ist, wäre eine Auswertung der Datenbank zu Trainings- oder Analysezwecken marken- und designrechtlich generell unbedenklich. Diese gewerblichen Schutzrechte zielen nicht darauf ab, jede nur denkbare Nutzung der geschützten Bezeichnungen oder Gestaltungen von einer Erlaubnis des Rechteinhabers abhängig zu machen. Vielmehr schützt das Markenrecht vor allem vor Produktfälschungen und Verwechslungen von Anbietenden, Produkten und Dienstleistungen. So können Inhaber:innen eines Markenrechts beispielsweise gegen die Verwendung einer ähnlichen – verwechslungsfähigen – Produktbezeichnung verwendet vorgehen. Mit einem Designrecht können Plagiate verboten werden, deren Gestaltung dem eigenen Produkt sehr ähnelt.

Auch wenn es noch nicht abschließend gerichtlich geklärt ist, kann man derzeit davon ausgehen, dass durch reine Speicher- und Analyseprozesse weder Marken- noch Designrechte verletzt werden. Solche Maßnahmen greifen nicht in den Schutzbereich dieser Rechte ein. Sie richten sich nicht an den Markt und dienen nicht der Nutzung der Marken und Designs im Wettbewerb. Schon gar nicht kommt es hierbei zu Täuschungen der Verbraucher:innen.

3.1.4.3 Know-how- und Geheimnisschutz

Know-how ist als solches nicht geschützt. Es ist ein weit verbreiteter Irrglaube, dass eine Art Eigentumsrecht an (exklusivem) Wissen besteht. Der Know-how-Schutz ist vielmehr als Geheimnisschutz ausgestaltet (siehe das Gesetz zum Schutz von Geschäftsgeheimnissen, GeschGehG). Dieser erfasst nur Geschäftsgeheimnisse (und nicht jede Art von Geheimnis und schon gar nicht jede Art von Wissen). Zudem greift der Geheimnisschutz nur, wenn die Inhaber:innen der Information diese mit „angemessenen Geheimhaltungsmaßnahmen" sichern und solange die Informationen und das Wissen tatsächlich geheim, d. h. nicht öffentlich bekannt sind. Im Übrigen unterliegt der Geschäftsgeheimnisschutz einer Vielzahl von Ausnahmen und Einschränkungen (z. B. bei Nutzungen im öffentlichen Interesse durch Whistleblower).

Definition von Geschäftsgeheimnissen in § 2 GeschGehG
Geschäftsgeheimnis ist eine Information,
a) die weder insgesamt noch in der genauen Anordnung und Zusammensetzung ihrer Bestandteile den Personen in den Kreisen, die üblicherweise mit dieser Art von Informationen umgehen, allgemein bekannt oder ohne Weiteres zugänglich ist und daher von wirtschaftlichem Wert ist und
b) die Gegenstand von den Umständen nach angemessenen Geheimhaltungsmaßnahmen durch ihren rechtmäßigen Inhaber ist und
c) bei der ein berechtigtes Interesse an der Geheimhaltung besteht.

Wer sich auf den Geheimnisschutz berufen will, sollte vor allem die genannten Geheimhaltungsmaßnahmen sorgfältig dokumentieren, regelmäßig überprüfen und ggf. optimieren. Im Streitfall sind Inhaber:innen der Information insofern beweispflichtig. Kann der Beweis über die Maßnahmen nicht geführt werden, besteht auch kein Anspruch auf Geheimnisschutz. Mit Geheimhaltungsmaßnahmen sind beispielsweise IT-Sicherheitskonzepte, Zugangs- oder Zugriffsbeschränkungen oder auch Geheimhaltungsvereinbarungen (solche reichen allein nicht aus) gemeint. Welche Maßnahmen in diesem Sinne angemessen sind, hängt vom jeweiligen Einzelfall ab. Hierbei sind verschiedene Indikatoren zu berücksichtigen, beispielsweise die Art der Informationen oder die Leistungsfähigkeit der Inhaber:innen.

Verstöße gegen den Geschäftsgeheimnisschutz durch Datenauswertungen oder KI-Trainingsmaßnahmen kommen daher nur in Betracht, wenn die Daten per se widerrechtlich erlangt wurden. Mit durch einen Hackerangriff aus einem fremden System gestohlenen Daten dürfen KI-Systeme nicht trainiert werden. Ebenso wenig mit geheimen Daten, die ehemalige Mitarbeiter:innen von Konkurrenten und Konkurrentinnen abgegriffen und ihren neuen Arbeitgebern und Arbeitgeberinnen zur Verfügung stellen.

3.1.4.4 Verträge

Unabhängig von der Frage, ob die zu verwendenden Daten durch Immaterialgüter- oder andere Rechte geschützt sind, kann ihre Nutzung auch vertraglichen Einschränkungen unterliegen. Bietet beispielsweise jemand eine Datenbank mit Forschungsdaten gegen Entgelt an, wird deren Nutzung vertraglichen Vereinbarungen unterliegen. Ob die Daten Rechtsschutz genießen, spielt letztlich keine Rolle, wenn man eine Nutzungsvereinbarung abschließt, denn diese ist per es einzuhalten. Ihre Verwendung unterliegt dann den jeweiligen vertraglichen Regelungen.

3.2 Was ist in der Trainingsphase für das System zu beachten?

In der Trainingsphase werden im Prinzip die Rohdaten zu einem Trainingsdatensatz hinreichender Menge und Qualität modelliert und anschließend dem Algorithmus oder dem neuronalen Netz zur Verarbeitung geliefert. Für diese Vorgänge ist bei personenbezogenen Daten ein umfassendes Konzept einschließlich technisch-organisatorischer Maßnahmen erforderlich, das beim Design des Systems festgelegt wurde (➔ 2.3) und nun ausgeführt wird.

Für die Trainingsphase ist das Hauptaugenmerk darauf zu richten, dass von diesem Konzept nicht abgewichen wird, oder, wenn Änderungen auftreten, diese begründet und dokumentiert werden. Die Versuchung mag mitunter groß sein, zur Verbesserung des Systems Datenquellen und -sätze nachzuschieben, auf neue Ideen oder Entdeckungen zu reagieren oder mit dem System etwas Neues auszuprobieren. Genau an einem solchen Verhalten setzt (bei personenbezogenen Daten) allerdings die Datenschutzaufsicht an. Ein zentraler Ansatzpunkt für die Kontrolle von Systemen ist die Nachvollziehbarkeit des Trainingsvorgangs. Die Datenschutzkonferenz verlangt in Bezug auf personenbezogene Daten eine laufende Überwachung und Protokollierung des Trainingsvorgangs mit Angabe der Fehlerrate (Differenz zwischen Soll- und Ist-Ergebnissen), soweit dies möglich ist, anderenfalls durch andere Verifikationsverfahren, wie z B. Testläufe mit präparierten Testdatensätzen („Blackbox Tinkering“). Auf diese Weise soll auch bei Blackboxes eine Nachvollziehbarkeit hergestellt werden. Daher ist dann aber auch eine unkontrollierte Vermengung von Datensätzen, eine unbeabsichtigte Veränderung oder ein Abfluss von Roh- oder Trainingsdaten ebenso wie eine unzulässige Zweckänderung von Daten zu vermeiden. Es mag dem unternehmerischen Denkmuster eines Start-ups entsprechen, zu experimentieren, was die Maschine sonst noch so an kommerziell verwertbaren Erkenntnissen hervorbringen könnte. Aber wenn man solche potenziellen Einsatzgebiete nicht von Beginn an mit eingeplant hat, muss man (bei personenbezogenen Daten) zuvor noch einmal die gesamte Schleife einer datenschutzrechtlichen Prüfung durchlaufen.

Um es klar zu sagen: Es ist keinesfalls verboten, ein System zu verändern und neue Ideen zu entwickeln. Nur muss dabei sorgsam und rechtmäßig mit personenbezogenen Daten umgegangen werden. Das Gleiche kann in Bezug auf Inhalte gelten, die immaterialgüterrechtlich geschützt sind. Hat man beispielsweise einen Datenbestand für einen ganz bestimmten Zweck und nur zur einmaligen Nutzung lizenziert, müssen Veränderungen des Nutzungszwecks oder Mehrfachnutzungen möglicherweise nachlizenziert werden.

Es wird deutlich: Gründliche und umfassende Dokumentation in der Trainingsphase ist der Schlüssel für den zu erbringenden Nachweis, ein datenschutzkonformes und diskriminierungsfreies System zu betreiben. Auch für das Lizenzmanagement kann dies erforderlich sein.

4 Rechtliche Aspekte in der Einsatzphase

4.1 IP, Schutzrechte und Know-how-Schutz an KI-Erzeugnissen

Die Arbeitsergebnisse von KI-Technologien (der „Output“) können gegen unautorisierte Verwendung durch Dritte geschützt sein. Infrage kommen hierfür verschiedene IP-Rechte (Intellectual Property Rights, IPR = Immaterialgüterrechte) und Know-how-Schutz. Auch kann man sich unter Umständen mit Verträgen gegen unerwünschte Nutzungen schützen. Grundsätzliche Informationen zu den verschiedenen Schutzrechten finden sich oben in ➔ 3.1.4.

ABBILDUNG 6:
Beispiel Edmond de Bellamy

Quelle: Artificial intelligence software (2018). „Portrait of Edmond de Belamy“. https://www.christies.com/img/LotImages/2018/NYR/2018_NYR_16388_0363_000(edmond_de_belamy_from_la_famille_de_belamy).jpg (Download 29.10.2020), Public Domain

Ob KI-Ergebnisse konkret geschützt sind, hängt oft von den Details des jeweiligen Falles ab. Dabei variieren sowohl die technischen Umstände als auch der Grad menschlicher Einflussnahme auf den Output in der Regel erheblich. Solche Unterschiede wirken sich meist auf die rechtliche Beurteilung aus.

4.1.1 Urheberrecht und Leistungsschutzrechte am Output

KI-Technologien können mittlerweile Bilder malen, Musik komponieren und vortragen. Sie können Computerprogramme schreiben und Designs gestalten.

Solche Outputs könnten durch Urheber- oder Leistungsschutzrechte geschützt sein. Erst durch einen solchen Rechtsschutz werden sie verkehrsfähig (können

lizenziert und verkauft werden) und sind gegen die ungefragte Übernahme und Weiterverwendung geschützt. Sind keine IPR gegeben, sind die Inhalte gemeinfrei und können völlig legal einfach übernommen und weiterverwendet werden. Rechtsschutz ist dann nur sehr eingeschränkt über Verträge möglich.

Die Frage, ob an KI-generierten Schöpfungen Urheber- oder Leistungsschutzrechte entstehen können, wird derzeit viel diskutiert. Hierbei stellt sich ganz generell die Frage der Rechtszuordnung. Wenn solche Rechte entstehen, wem würden sie zustehen? Der entwickelnden Person des automatisierten Systems oder Algorithmus? Dessen Arbeitgeber:in? Der KI selbst?

Gegen eine Zuordnung von Rechten an eine KI spricht ein grundlegender Aspekt. Maschinen oder Algorithmen haben keine Rechte und können keine Rechte besitzen. Sie können in eigenem Namen keine Verträge schließen, nicht verklagt werden oder Straftaten begehen. Juristisch ausgedrückt: Sie sind nicht rechtsfähig. Diese rechtliche Tatsache wird sich jedenfalls solange nicht ändern, wie es keine Künstlichen Intelligenzen gibt, die wirklich intelligente, eigenständige Persönlichkeiten darstellen.

Auch eine Zuordnung von Immaterialgüterrechten (v. a. Urheber- und Leistungsschutzrechten) an die Entwickler:innen der KI, deren Arbeitgeber:innen oder die Eigentümer:innen des Systems wirft erhebliche Schwierigkeiten mit der geltenden Rechtsordnung auf. Solche Rechte werden in aller Regel denjenigen zugeordnet, die das Werk geschaffen (beim Urheberrecht) oder die geschützte Leistung erbracht haben (bei Leistungsschutzrechten). Programmierer:innen eines Selbstlernprogramms sind jedoch nicht Schöpfer:innen der Ergebnisse, die dieses erzeugen (ebenso wenig wie die Entwickler:innen von Microsoft Schöpfer:innen von in MS Word geschriebenen Texten sind). Ihnen Urheber- oder Leistungsschutzrechte am Output zuzuordnen ist daher schwer zu begründen. Das wäre, als würde man dem Herstellern bzw. den Herstellerinnen von Kameras mit Autofokus Rechte an den hiermit erstellten Aufnahmen gewähren. Hierzu jedoch im Einzelnen.

4.1.1.1 Urheberrechtsschutz

Schon heute malen KI-Technologien Bilder, schreiben Bücher und komponieren Lieder, mehr oder weniger vollständig autonom. Fraglich ist, ob es sich hierbei um „Schöpfungen“ handelt, die urheberrechtlich geschützt sein können.

Für die urheberrechtliche Schutzfähigkeit von KI-erzeugten Daten und Inhalten ist es entscheidend, ob und inwiefern sie auf die gestalterische Tätigkeit einer menschlichen Person zurückzuführen sind. Das Urheberrecht ist ein personenbezogenes Schutzrecht. Es gilt nur für persönliche geistige Schöpfungen eines Menschen. Weder fallen rein technische Gestaltungen noch solche von Tieren unter das Urheberrecht.

ABBILDUNG 7:
Affen-Selfie von Naruto

Quelle: Selfie eines Schopfaffen (Macaca nigra) in Nord-Sulawesi (Indonesien), einer Makakenart, der die Kamera des Fotografen David Slater auf sich selbst gerichtet und den Auslöser betätigt hat. https://commons.wikimedia.org/wiki/File:Macaca_nigra_self-portrait_(rotated_and_cropped).jpg#filehistory (Download 29.10.2020), Public Domain

Hat also eine KI beispielsweise eine Graphik gestaltet, einen Text geschrieben oder ein Musikstück komponiert, bestehen hieran keine Urheberrechte. Reine KI-Werke sind also nicht urheberrechtsfähig. Das kann anders sein, wenn der Output durch menschliche Tätigkeit bearbeitet wird, sofern diese Tätigkeit schöpferisch ist und ein Mindestmaß an Originalität erreicht wurde. Schreibt beispielsweise eine KI einen längeren Nachrichtentext, der dann von einem bzw. einer Redakteur:in umfangreich stilistisch und sprachlich überarbeitet wird, kann die:der Redakteur:in an der editierten Fassung ein Urheberrecht besitzen.

Das bedeutet: Je autonomer die KI bei solchen Schöpfungen arbeitet, desto unwahrscheinlicher ist ein urheberrechtlicher Schutz des Outputs.

4.1.1.2 Leistungsschutz

Ob der Output von KI-Technologien durch Leistungsschutzrechte geschützt sein kann, hängt davon ab, um welche Art von Arbeitsergebnis es sich handelt. Grundsätzlich können (zumindest manche) Leistungsschutzrechte auch an nicht menschlichen Erzeugnissen entstehen. Die Anforderungen der verwandten Schutzrechte sind sehr unterschiedlich.

Fotos

Aus technischer Sicht ist es möglich, dass mit Methoden des maschinellen Lernens Fotografien, Bildkollagen oder andere technische Formen von Abbildungen selbsttätig erzeugt werden. Beispielsweise können Künstliche Intelligenzen mit aus Bildern bestehenden Datenpools trainiert und dadurch in die Lage versetzt werden, eigene Bilder zu erzeugen. Voraussetzung hierfür ist lediglich, dass die für die Neukomposition erforderlichen Inhalte vorhanden sind. Für solche nicht schöpferischen Abbildungen, die keine menschlichen, eigenschöpferischen Züge aufweisen, käme allenfalls das Lichtbildrecht (➔ 3.1.4.1) in Betracht. Indes: Auch, wenn das Lichtbildrecht nicht voraussetzt, dass sich menschliche Kreativität in der Abbildung manifestiert, kommt es bei reinen KI-Erzeugnissen nicht zur Anwendung. Nach ganz herrschender Meinung in der Rechtsliteratur und der Rechtsprechung wird das Lichtbildrecht nur für menschliche Leistungen gewährt. Entsprechend genießen weder von Tieren (wie dem schwarzen Makaken Naruto, ➔ Abbildung 7) noch von Maschinen erzeugte Fotos und sonstige Abbildungen Lichtbildschutz.

Das Gleiche gilt im Übrigen für Videos oder Filme. Sie können zwar – auch wenn es sich nicht um schöpferische Leistungen, sondern um simple Filmchen handelt – ähnlich wie einfache Fotos geschützt sein (nach dem Laufbildrecht, § 95 UrhG). Auch dies wird jedoch nicht für rein technische Erzeugnisse gewährt.

Musik- und Filmproduktionen

Musikproduktionen werden herkömmlich durch das Tonträgerherstellerrecht geschützt (s. §§ 85, 86 UrhG). Anders als das Urheberrecht an der Komposition

bezieht sich dieses Leistungsschutzrecht auf die technisch-organisatorische Leistung, Töne und akustische Werke erstmalig auf einem Tonträger festzuhalten. Neben herkömmlichen Musikproduktionen (Aufnahmen, Erstellung eines Masters) fallen hierunter auch alle anderen Arten von Tonaufnahmen. Ob der auf Tonträger verkörperte Ton selbst schutzfähig ist, es sich also z. B. um die Darbietung eines urheberrechtlich geschützten Werks wie eines Musikstücks handelt, ist irrelevant. Auch an Aufnahmen von Vogelstimmen oder Geräuschen können Tonträgerherstellerrechte bestehen. „Schutzgegenstand ist die im Tonträger verkörperte Herstellerleistung als immaterielles Gut" heißt es in der Gesetzesbegründung. Tonträgerhersteller ist diejenige Person, die die organisatorische Hoheit über die Aufnahme besitzt, also einerseits die maßgeblichen technischen und wirtschaftlichen Leistungen erbringt und andererseits, soweit erforderlich, das Rechtemanagement durchführt (insbesondere die Verträge mit Kunstschaffenden etc. abschließt).

Ein Unternehmen oder eine natürliche Person, die mit Methoden des maschinellen Lernens Musik produziert, abmischt und auf einen Master kopiert, kann also Tonträgerherstellerrechte an der entstandenen Aufnahme haben. Ob die durch die KI eingespielte Musik selbst geschützt ist, ist unerheblich. Ist die KI jedoch gleichzeitig Interpret und Produzent, organisiert sie also eigenständig den Prozess von der Tonerzeugung bis zur fertigen Aufnahme, dürfte ein Tonträgerherstellerrecht nicht in Betracht kommen. Tonträgerhersteller:in kann nur eine natürliche oder juristische Person sein (Art. 3c des Internationalen Abkommens über den Schutz der ausübenden Kunstschaffenden, der Hersteller von Tonträgern und der Sendeunternehmen vom 26.10.1961, sog. Rom-Abkommen). Wenn es keine solche Person gibt, die die organisatorische Hoheit über den Produktionsprozess hatte, gibt es auch kein Tonträgerherstellerrecht. Eine KI ist keine Person im Rechtssinne.

Für Filmproduktionen gilt dasselbe, lediglich mit dem Unterschied, dass es nicht um die „auf den Ton beschränkte Festlegung der Töne einer Darbietung oder anderer Töne" (so Art. 3b des Rom-Abkommens), sondern um die Festlegung von visuellen und / oder audiovisuellen Leistungen geht.

ABBILDUNG 8:
„Proto"
Quelle: hollyherndon.com/proto 2019

Künstlerische Darbietungen wie Gesang, Schauspiel oder Tanz

Auf Holly Herndons Album „Proto" singt u. a. eine KI namens „Spawn" (thefader.com 2019; youtube 2018). Besteht an deren Gesangseinlagen ein Schutzrecht oder könnten sie von jedem ohne rechtliche Probleme, z. B. bei einer Live-Darbietung, aufgezeichnet, gesampelt und kopiert werden (an dieser Stelle geht es nur um den Gesang als solchen, nicht um die Rechte an der Tonaufnahme des Gesangs, s. o.).

An den Leistungen von Interpreten und Interpretinnen bestehen grundsätzlich „Rechte der ausübenden Künstler". In Art. 3a des Rom-Abkommens heißt es: „Für die Zwecke dieses Abkommens versteht man unter: «Ausübenden Künstlern» die Schauspieler, Sänger, Musiker, Tänzer und anderen Personen, die Werke der Literatur oder der Kunst aufführen, singen, vortragen, vorlesen, spielen oder auf irgendeine andere Weise darbieten". „Ausübender Künstler im Sinne dieses Gesetzes ist, wer ein Werk oder eine Ausdrucksform der Volkskunst aufführt, singt, spielt oder

auf eine andere Weise darbietet oder an einer solchen Darbietung künstlerisch mitwirkt", sagt § 73 UrhG. Mit diesen Definitionen wird klargestellt, dass nur die Darbietungen von natürlichen Personen mit dem Leistungsschutzrecht der ausübenden Künstler:innen belohnt werden. Da Künstliche Intelligenzen keine Personen, sondern technische Konstrukte sind, kommt ein Schutz ihrer Gesangseinlagen oder anderen Darbietungen (etwa: rein computergenerierte schauspielerische Leistungen) nicht in Betracht.

Datenbankherstellerrecht

Technologien des maschinellen Lernens können Datenbanken nicht nur auslesen, sondern sie auch mehr oder weniger eigenständig konzipieren, organisieren und befüllen. Erbringt eine Person oder ein Unternehmen hierfür wesentliche Investitionen, entstehen im Zweifel Datenbankherstellerrechte, die der Person oder dem Unternehmen zustehen. Ob die Datenbank durch menschliche Leistungen oder rein automatisiert erzeugt wurde, ist unerheblich. Allein entscheidend sind die hierfür aufgewendeten Investitionen. Nur wenn diese zu gering, zu unwesentlich sind, kommt ein solches Schutzrecht nicht in Betracht.

Leistungsschutzrecht für Presseverlage

Das Leistungsschutzrecht für Presseverlage wird in vorliegendem Kontext z. B. relevant, wenn Technologien des maschinellen Lernens vollautomatisch Inhalte für Presseerzeugnisse erstellen. Hierzu ist bereits heute ein Trend zu beobachten. So hat beispielsweise die Nachrichtenagentur Associated Press Frankreich eine umfangreiche Strategie zur (u. a.) automatisierten Nachrichtenerzeugung aufgesetzt (AP o. J.). Von der Inhaltserstellung über Redigate bis zur Auswahl und Zusammenstellung könnten Presseerzeugnisse (jedenfalls in manchen Ausprägungen) zukünftig in hohem Maß computergeneriert produziert werden.

Ob solche Erzeugnisse dem Leistungsschutzrecht für Presseverleger zugänglich wären, ist noch schwer einzuschätzen. Dieses Recht schützt Presseverleger:innen vor ungefragten Verwendungen ihrer Publikationen. In Bezug auf dieses noch sehr neue Schutzrecht sind viele Fragen ungeklärt. Zwar gab es in Deutschland zwischen 2013 und 2019 bereits ein solches Leistungsschutzrecht. Es wurde jedoch durch den Europäischen Gerichtshof für unzulässig erklärt (iRights info 2019). Erst im Zuge der 2019 verabschiedeten Richtlinie zum Urheberrecht im digitalen Binnenmarkt wird es erneut eingeführt werden (vermutlich 2021). Entsprechend kann über dessen Details noch wenig gesagt werden.

Im Zweifel wird es auch hier – ähnlich dem Datenbankherstellerrecht – weniger darauf ankommen, ob die Inhalte des Presseerzeugnisses von Menschen oder Maschinen erstellt bzw. ob das Presseerzeugnis automatisiert oder manuell kuratiert wird. Entscheidend scheint es nach derzeitigem Stand eher zu sein, ob in die jeweilige Presseveröffentlichung nennenswerte Investitionen geflossen sind und ob diese von einer natürlichen oder juristischen Person erbracht wurden. Wäre dies der Fall, kämen die o. g. Gedanken zum Datenbankherstellerrecht zur Anwen-

dung: Ein Presseverlag könnte solche Rechte erlangen, auch wenn die Inhalte des Presseerzeugnisses von einer KI erstellt wurden, soweit er die organisatorische Hoheit über die Publikation ausübt und hierin investiert. Die KI selbst wird auch bei diesem Recht nicht als Rechteinhaberin (also als Presseverlegerin) angesehen werden können.

4.1.2 Patentrecht

KI-Technologien können – und werden – als solche patentiert werden. Gleiches gilt im Grundsatz für Erfindungen, bei deren Entstehung KI-Technologien als Hilfsmittel und zur Unterstützung eingesetzt werden („computer aided inventions"). Anders sieht es jedoch bei rein von KI erfundenen Technologien aus („computer generated inventions"). Wie das Urheberrecht ist das Patentrecht grundsätzlich personenbezogen. Erfindungen sind nach geltender Rechtslage nur dann geschützt, wenn sie von einem Menschen stammen. „Das Recht auf das Patent hat der Erfinder oder sein Rechtsnachfolger" heißt es in § 6 PatentG. Das heißt im Umkehrschluss, dass kein Recht auf ein Patent entsteht, wenn es keine:n Erfinder:in als menschliche Person gibt. Ergo können auf rein computergenerierte technische Erfindungen keine Patente erteilt werden (Europäisches Patentamt 2020). Dies wird momentan jedenfalls in den meisten Rechtsordnungen der Welt gelten.

Möglich ist, dass sich die diesbezügliche Rechtslage irgendwann ändert. Die für gewerbliche Schutzrechte zuständige UN-Organisation, die „World Intellectual Property Organisation" (WIPO), hat bereits erkannt, dass die Schutzfähigkeit von KI-Erzeugnissen (durch Urheber-, Patent- oder andere Schutzrechte) ein regulatives Thema mit erheblicher Bedeutung ist. Immerhin wird davon ausgegangen, dass Immaterialgüterrechte, wie das Patentrecht, erhebliche Anreize für Investitionen und Innovationen setzen. Rechtspolitisch wird daher z. B. eruiert, ob zukünftig auch rein computergenerierte Erfindungen patentfähig sein sollten, was gesetzliche Änderungen erfordern würde. Die WIPO hat daher Anfang 2020 eine umfangreiche Konsultation (WIPO 2019) über diese und andere Fragen durchgeführt (Issue 1 und 2 der Konsultation).

4.1.3 Geheimnisschutz

Erzeugen Algorithmen Know-how oder Geschäftsgeheimnisse, können diese nach den Regeln des Gesetzes zum Schutz von Geschäftsgeheimnissen (GeschGehG) gegen die unbefugte Verwendung geschützt sein. Wie solche Geschäftsgeheimnisse entstehen (menschlich, maschinengeneriert), ist hierfür unerheblich (siehe Näheres hierzu oben ➔ 3.1.4.3.).

4.2 Datenschutzrechtliche Vorgaben bei personenbezogenen Output-Daten

In der Einsatzphase gibt es einige typische – wenn man so will – datenschutzrechtliche Sollbruchstellen:

4.2.1 Zweckbindung bei personenbezogenen Output-Daten

Unbeschadet der soeben dargestellten Rechtslage im Immaterialgüterrecht gelten für die Output-Daten bzw. Arbeitsergebnisse des Systems, wenn und soweit es sich dabei um personenbezogene Daten handelt, die Vorgaben der DSGVO. Dies wiederum ist relevant für die Kommerzialisierbarkeit der Arbeitsergebnisse. Es gilt wieder der Zweckbindungsgrundsatz, sodass die Arbeitsergebnisse nur im Rahmen des für sie vorgesehenen Zwecks (oder bei ausnahmsweise zulässiger Zweckänderung) verwendet werden dürfen. Ein instruktives Beispiel der Datenschutzkonferenz (DSK 2019b): Eine Musikempfehlungsmaschine und deren Output darf nicht dazu verwendet werden, politische Profile der Nutzer:innen zu verkaufen. Ebenso müssen für eine Weiterverarbeitung der Arbeitsergebnisse eine Rechtsgrundlage in der DSGVO gefunden und die datenschutzrechtlichen flankierenden Organisationspflichten erfüllt werden.

4.2.2 Technisch-organisatorische Maßnahmen und laufendes Monitoring

Die den Risiken angemessenen technisch-organisatorischen Maßnahmen beim Einsatz des Systems wurden beim Design des Systems spezifiziert und müssen nun ausgeführt werden (➔ 2.3). Besonders wichtig ist hier aber die laufende Überwachung des Systems, gerade wenn es im Live-Betrieb weiterlernt. Ein laufendes Monitoring des Systems liegt im Eigeninteresse eines Unternehmens. Aber auch die Datenschutzaufsicht verlangt Maßnahmen, mit denen das System in ausreichendem Maße auf Fehlentscheidungen und vor allem rechtswidrige Diskriminierungen überprüft wird. In simpelsten Falle sind dies wiederholte Testläufe mit Testdaten. Aber die Technologien zur Qualitätskontrolle von Systemen maschinellen Lernens entwickeln sich rasant und letztlich wird man sich hier am Stand der Technik zu orientieren haben.

4.2.3 Vorbehalt menschlicher Entscheidung

Eine rechtspolitisch umstrittene Vorgabe der DSGVO limitiert für die EU die Einsatzmöglichkeiten von maschinellem Lernen, nämlich der sog. Vorbehalt menschlicher Entscheidungen (Art. 22 DSGVO). Diese Vorschrift verbietet im Prinzip, dass autonome Systeme vollständig allein über Menschen entscheiden (einschließlich

Profilbildung), wenn diese Entscheidungen rechtliche Wirkung entfalten oder in vergleichbarer Hinsicht erheblich beeinträchtigen können. Beispielsweise wäre es nicht möglich, einem autonomen System, welches die Leistung von Mitarbeitenden misst, die alleinige Entscheidung über die Kündigung von diesen zu überlassen. Dahinter steht der Gedanke, dass kein Mensch bloßes Objekt einer Künstlichen Intelligenz werden soll. Der Einsatz von maschinellem Lernen im Unternehmenskontext muss daher betrieblich im Einklang mit dem Vorbehalt menschlicher Entscheidungen organisiert werden. Typischerweise muss das System so lange in einem Wartestatus verbleiben („pending"), bis der weitere Fortgang durch einen Menschen angestoßen wird. Auch wird in der Rechtsliteratur gefordert, dass der letztentscheidende Mensch über einen sinnvollen Entscheidungsspielraum auf Basis der relevanten Informationen verfügen muss, also nicht nur als Förmelei den Knopf betätigt. Für eine Reihe von Branchen werden hierdurch Automatisierungshürden gesetzt.

Zunächst gilt der Vorbehalt menschlicher Entscheidungen aber nur, wenn die Entscheidung vollständig von dem autonomen System getroffen wird. Bereitet das System auch mit maschinellem Lernen lediglich eine menschliche Entscheidung vor (entscheidungsunterstützende Systeme), ist dies ohne Weiteres zulässig. Man spricht auch von „automated decision making" gegenüber „computer aided decision making". Eine offenbare Schwäche dieser Vorschrift springt einem förmlich als rechtspolitische Kritik ins Auge: Für den beabsichtigten Schutz ist wenig gewonnen, wenn eine Maschine die Entscheidung vorbereitet und ein Mensch in blindem Vertrauen ohne weitere Prüfung massenhaft Freigaben klickt. Dies gilt besonders, wenn die letztentscheidende Person gar nicht über die Informationen verfügt, die maschinelle Entscheidung nachzuvollziehen. Die Annahme, ein Mensch werde besser als eine KI entscheiden, erscheint in vielen Kontexten zweifelhaft.

Sodann ist ein Vorbehalt menschlicher Entscheidungen nur notwendig, wenn die Entscheidung rechtliche Wirkung entfaltet (d. h. Rechte / Pflichten begründet oder in Rechtspositionen eingreift) oder die Betroffenen in ähnlicher Weise beeinträchtigt. Wann Letzteres der Fall ist, ist noch nicht im Detail geklärt. Die ähnliche Beeinträchtigung muss derart erheblich sein, dass sie Rechtswirkungen ähnelt, indem sie die wirtschaftliche oder persönliche Entfaltung der Person stört. Mögliche Beispiele sind die Versagung eines Kredits, Nichtbegründung eines Vertragsverhältnisses oder die Aufnahme in eine „black list". Diskutiert werden auch – hier erkennt man die mögliche Reichweite dieser Vorschrift – personalisierte Werbung und personalisierte Preisgestaltung aufgrund von Profiling / Online-Verhalten.

Schließlich kann es ausnahmsweise Rechtfertigungen für vollautomatisierte Entscheidungen geben, etwa wenn die Betroffenen wirksam eingewilligt haben, wenn die Entscheidung für den Abschluss oder die Erfüllung eines Vertrags mit der betroffenen Person erforderlich ist oder wenn eine spezialgesetzliche Erlaubnis vorliegt. Man sieht: Hier lassen sich durchaus Lösungen entwickeln.

Fallbeispiel: Boss Machine
Medien berichteten aufgrund eines geleakten Dokuments über eine Praxis von Amazon in den Fulfillment-Centern in den USA. Dort werde eine künstliche Intelligenz eingesetzt, die die Arbeitsleistung der Mitarbeiter:innen überwacht und im Falle geringer Produktivität automatisiert Abmahnung und Kündigungen ausspricht. Supervisor:innen könnten diesen Prozess allerdings außer Kraft setzen (Lecher 2019).

Liegt eine Rechtfertigung vor, müssen gleichwohl angemessene Maßnahmen getroffen werden, um die Rechte und Freiheiten sowie die berechtigten Interessen der betroffenen Person zu wahren. Erhöhte Anforderungen gelten auch hier wieder für personenbezogene Daten besonderer Kategorien (Art. 9 DSGVO), wie z. B. Herkunft oder Gesundheits- und biometrische Daten.

4.3 Verträge und Willenserklärungen

Können Systeme des maschinellen Lernens eigene Willenserklärungen abgeben (z. B. automatisch schlecht performende Mitarbeiter:innen kündigen) oder Verträge schließen (z. B. Teile bei Zulieferungsunternehmen einkaufen)? Insbesondere mit Blick auf ein „Internet der Dinge" stellt sich die Frage, wieweit Geschäftsvorgänge auch auf rechtlicher Ebene automatisiert werden können.

Nach geltendem Recht ist diese Frage einfach zu beantworten: Es ist nicht die Maschine, die Erklärungen abgibt oder Verträge schließt, sondern die natürliche Person oder die wirksam vertretene juristische Person dahinter. Willenserklärungen setzen Rechts- und Geschäftsfähigkeit voraus. Über solche verfügt die Maschine nicht, so ausgefeilt das System auch sein mag. Auch wenn die Künstliche Intelligenz zu einem gewissen Grade autonom entscheidet, wird diese Entscheidung stets der dahinterstehenden Person zugerechnet, entweder als Erklärende oder als Erklärungsempfängerin. Dies gilt auch, wenn die Person dahinter die konkrete von der Maschine erstellte Erklärung gar nicht kennt, sondern nur allgemein das maschinelle System verantwortet. Nicht anders verhält es sich bei sog. „Smart Contracts", die anders als es die Bezeichnung suggeriert, keine Verträge zwischen Systemen, sondern ebenfalls Verträge der dahinterstehenden Personen sind, die lediglich automatisiert abgewickelt werden. Empfängt eine künstliche Intelligenz eine Willenserklärung, so ist diese zugegangen, ganz als ob eine Willenserklärung in den Briefkasten der empfangenden Person geworfen wird. Für Arbeitgeberkündigungen und viele andere Willenserklärungen mit besonderer Bedeutung gelten allerdings Schriftformerfordernisse. Sie werden nur mit einer eigenhändigen Unterzeichnung der:des Erklärenden wirksam und können daher nicht von einer KI abgegeben werden.

Je autonomer ein System darüber entscheiden darf, welche Verträge mit welchem Inhalt geschlossen werden sollen, desto größer ist die Gefahr für das Unternehmen, durch Verträge gebunden zu werden, die es eigentlich gar nicht wollte. Auch dies ist im Prinzip nichts Neues, wenn man beispielsweise an die gelegentlich auftretenden Erdrutschverluste denkt, die im Aktien-Hochfrequenzhandel entstehen, wenn verschiedene Handelssysteme unvorhergesehen miteinander reagieren. Dieses Risiko lässt sich einerseits technisch durch Weiterentwicklung der Algorithmen, aber auch durch Verifikationsprozesse lösen, etwa, wenn Mitarbeiter:innen die Willenserklärungen des Systems vor Abgabe freigeben müssen (was freilich dem Zweck des Systems zuwiderlaufen kann).

Rechtspolitisch wird diskutiert, ob bei Systemen mit hohem Autonomiegrad gesetzlicher Anpassungsbedarf besteht und mit einer Fortentwicklung der Rechtsgeschäftslehre solche Systeme in die Lage versetzt werden sollen, eigene Rechtsgeschäfte tätigen zu können, die dann nicht mehr automatisch der Person dahinter zugerechnet werden. Dies ist aus vielerlei Gründen aber noch Zukunftsmusik.

4.4 Haftungsfragen

Die Tatsache, dass maschinelles Lernen menschliche Fähigkeiten weit übertreffen kann, bedeutet nicht, dass solche Systeme keine Fehler machen und immer perfekt funktionieren. Es gibt typische haftungsträchtige Risiken des maschinellen Lernens:

- Fehler, Qualitätsmängel oder Schieflagen in den Trainingsdaten, die auf das Entscheidungsverhalten des Systems durchschlagen (einschließlich solcher Effekte, die durch die unvermeidbaren Generalisierungen eines Modells verursacht werden).
- Fehlerhafte Schlussfolgerungen des Systems (einschließlich des erwähnten Clever-Hans-Effekts).

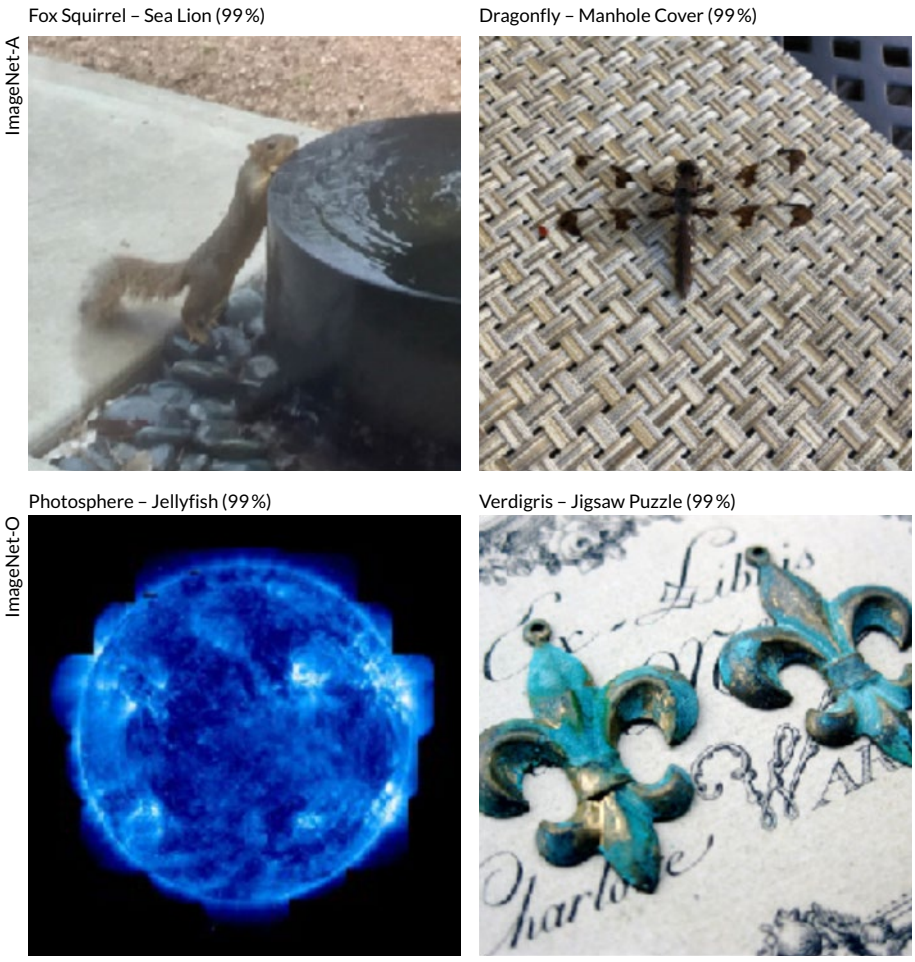

ABBILDUNG 9:
Natural Adversarial Examples
Natural Adversarials sind Fotos von realen Gegenständen, die, wenn man sie als Trainingsdaten verwendet, die Erkennungsleistung des Systems wesentlich kompromittieren. Sie illustrieren die Fehleranfälligkeit einer Bilderkennung durch maschinelles Lernen.
Quelle: Hendrycks et al. 2020

- Fehler und Ungeeignetheit der zugrunde liegenden Algorithmen / Entscheidungslogik.
- Sicherheitslücken und Manipulationen.
- Hardwarefehler oder defekte Sensoren.

Aus diesem Grunde stellt sich eine Reihe von Haftungsfragen. Haften Entwickler:innen / Hersteller:innen einer KI für deren Arbeitsergebnisse? Haften Hersteller:innen von Produkten mit integrierter künstlicher Intelligenz (z. B. ein autonom fahrendes Fahrzeug) für ihre Produkte? In welchem Umfang haften Anwender:innen des Systems?

4.4.1 Überblick über das einschlägige Haftungsrecht

Die Diskussionen über die Regulierung von autonomen Fahrzeugen hat zu einer Flut von rechtspolitischen Beiträgen zur Anpassung des Haftungsrechts geführt. Dies sollte aber nicht den Blick dafür versperren, dass es gegenwärtig ein umfassendes Haftungsrecht gibt, welches ohne Weiteres auf alle Haftungskonstellationen im Zusammenhang mit maschinellem Lernen Anwendung findet. Die folgende Darstellung geht nicht allen in Betracht kommenden Haftungskonstellationen nach, sondern beschränkt sich auf die juristischen Stellschrauben, mit denen die Haftungssituation beim Einsatz maschinellen Lernens in unternehmerischer Perspektive optimiert werden kann. Dabei sind drei im Prinzip nebeneinander anwendbare Haftungsregime zu berücksichtigen: **vertragliche Haftung**, **Produzentenhaftung** und **Produkthaftung**.

4.4.1.1 Vertragliche Haftung

Im Zusammenhang mit maschinellem Lernen ist eine Vielzahl von Vertragskonstellationen denkbar: Softwareentwicklungsverträge mit Entwickelnden, Dienst- oder Werkverträge mit Anbietenden von KI-relevanten Dienstleistungen, Kauf von Systemlösungen, Mietverträge mit Anbietenden von Cloudspeichern, Kauf von Produkten mit integrierter KI durch Verbraucher:innen usw. All diese Verträge enthalten Aussagen darüber, wer unter welchen Umständen für was haftet. Entweder enthalten die Verträge zwischen den Parteien ausdrücklich verhandelte Klauseln oder es gelten ansonsten die gesetzlichen Regelungen des Vertragstyps, beispielsweise die gesetzliche Kaufgewährleistung.

So unterschiedlich die verschiedenen Konstellationen auch sind, so gibt es zwei haftungsrelevante Stellschrauben, die allen Verträgen gemein sind, nämlich **Umfang der geschuldeten Leistung** und **Haftungsausschlussklauseln**.

Der Umfang der geschuldeten Leistung schlägt unmittelbar auf die Haftung durch, wenn die Leistung nicht so erbracht wird, wie es vertraglich geschuldet wird. Am besten lässt sich das an einem Beispiel zeigen: Ein:e Hersteller:in möchte ein System zur medizinischen Tumorerkennung auf radiologischen Bildern als Dienstleistung („KI as a Service") vertreiben. In der ersten Variante wird vertraglich versprochen, es handele sich um ein System mit einer garantierten Erkennungsrate von 96 Prozent. In der zweiten Variante wird lediglich versprochen, die Software auf einem beschriebenen Trainingsstand zur Verfügung zu stellen ohne jegliche Versprechen zu Funktionsweise und Leistungsfähigkeit. Liegt die Erkennungsrate des Systems nun tatsächlich deutlich unter 96 Prozent, erkennt man leicht: In der ersten Variante haftet der:die Hersteller:in wegen eines Mangels des Systems, in der zweiten Variante nicht. Das ist mit Umfang der versprochenen Leistungen (juristisch: Haftung wegen Pflichtverletzung) gemeint.

Was bedeutet das im Unternehmenskontext? Man muss sich bei Vertragsverhandlung genau überlegen, was man bereit ist, vertraglich zu versprechen oder zu akzeptieren. Die Frage, welche Aussagen man über ein System tätigt (auch in begleitender Marketingkommunikation), oder welche man als Kunde bzw. Kundin zwingend verlangt, ist unter Umständen ein „trade-off" zwischen einer Haftungsminimierung einerseits und der Vermarktbarkeit des Produkts andererseits. Je weniger man verspricht, desto geringer ist das Haftungsrisiko, aber desto uninteressanter ist das Angebot auch möglicherweise für die Kundschaft. Umgekehrt gilt aus der Perspektive der Kunden und Kundinnen: Was nicht an Leistungen versprochen wird, kann man auch nicht verlangen und hat keine Rechte diesbezüglich.

In der Vertragsgestaltung jedenfalls ist die Beschreibung der vertraglichen Leistung, der Systemeigenschaften, der Mitwirkungspflichten der Vertragspartner:innen usw. ein Fall für professionelle rechtsberatende Vertragsgestaltung. Das liegt auch daran, dass sich die geschuldeten Solleigenschaften eines Systems nicht nur aus ausdrücklichen Vertragsklauseln ergeben, sondern unter Umständen durch Auslegung des Vertrags ermittelt werden müssen, weil dieser unklar oder unvollständig ist. Wird in einem Vertrag beispielsweise für ein Assistenzsystem für Fahrzeuge das Wort Autopilot verwendet, stellt sich sofort die Frage, welchen Grad von Sicherheit man mit dieser Wortwahl impliziert. Manche Solleigenschaften werden auch stillschweigend in einen Vertrag hineininterpretiert. In einem Kaufvertrag über einen Neuwagen beispielsweise ist normalerweise nicht ausdrücklich festgeschrieben, dass das Fahrzeug tatsächlich auch fahren können soll. Trotzdem ist diese Eigenschaft geschuldet, weil Autos dazu da sind, dass sie gefahren werden können. All diese Grundsätze gelten auch bei Verträgen über Systeme mit maschinellem Lernen. Die Vertragsgestaltung hat die Aufgabe, die Haftungsrisiken für die unterschiedlichen Problemszenarien bei maschinellem Lernen klar und rechtssicher in dem Vertrag zu allokieren. Im IT-Bereich ist es gebräuchlich, in Abweichungen von den gesetzlichen Regelungen eigene Vertragsklauseln für Klassen von Fehlern mit eigenen Folgen zu vereinbaren. Es ist zu erwarten, dass sich solche Fehlerklassen in Zukunft auch für das maschinelle Lernen herauskristallisieren.

Die zweite Stellschraube sind vertragliche Regelungen über den Haftungsmaßstab, also Klauseln, in der die Vertragsparteien die Haftung einseitig oder gegenseitig reduzieren oder ausschließen. Haftungsausschlüsse kommen in der Praxis häufig vor. Ein Haftungsausschluss ist jedoch ein scharfes Schwert, das die wirtschaftliche Risikoverteilung in einem Vertrag drastisch verschieben kann. Auch gibt es zahlreiche gesetzliche Grenzen für die Wirksamkeit von Haftungsausschlüssen, beispielsweise bei einem Verkauf an Verbraucher:innen (§ 476 BGB) und vor allem aus dem AGB-Recht. In AGBs lässt sich die Haftung für Schäden an Leben, Körper, Gesundheit und bei grobem Verschulden nicht ausschließen (§ 309 Nr. 7 BGB), was beispielsweise für Bilderkennung im medizinischen Kontext besonders relevant ist. Haftungsklauseln sind bei Individualverträgen im Übrigen Verhandlungssache und bei allen Verträgen im Zusammenhang mit Leistungen des maschinellen Lernens (mit ihren rechtlichen Grenzen) zu bedenken. Ein Allheilmittel gegen Haftungsrisiken sind sie nach deutschem Recht jedoch nicht.

4.4.1.2 Delikts- / Produzentenhaftung

Neben etwaiger vertraglicher Haftung haftet auf Schadensersatz, wer ganz allgemein in zurechenbarer Weise rechtswidrig und schuldhaft (vorsätzlich oder fahrlässig) Rechtsgüter (Gesundheit, Eigentum, Persönlichkeitsrechte usw.) eines anderen verletzt. Dies gilt für jedermann, also auch Entwickler:innen, Hersteller:innen, KI-Anbieter:innen oder Verwender:innen, ebenso für Halter:innen eines autonomen Fahrzeugs. Untechnisch gesprochen: Können diese Personen eigentlich etwas dafür, wenn ein System mit maschinellem Lernen eine autonome Entscheidung trifft, die zu einer Schädigung führt? Dies ist dann der Fall, wenn die Person eine sog. Verkehrspflicht verletzt hat. Solche Verkehrspflichten beruhen nach der Rechtsprechung auf folgendem Gedanken: Wer eine Gefahrenquelle kontrolliert, eine gefährliche Tätigkeit ausübt oder eine Gefahrenquelle geschaffen hat, der ist verpflichtet, alle Maßnahmen zu treffen, die ein umsichtiger und verständiger, in vernünftigen Grenzen vorsichtiger Mensch für notwendig und ausreichend hält, um andere vor Schaden zu bewahren. Das bedeutet: Wer eine KI herstellt, vertreibt oder verwendet, hat in vernünftigem Rahmen dafür Sorge zu tragen, dass andere nicht geschädigt werden. Unterbleiben solche Maßnahmen und kommt es zu einer Schädigung, entstehen Haftungsansprüche.

Was nun genau zur Erfüllung einer Verkehrspflicht zu tun ist, hängt von den Umständen des Einzelfalls ab. Die Rechtsprechung berücksichtigt die Intensität der Gefahr, die Wahrscheinlichkeit des Schadenseintritts sowie berufsbezogene Fähigkeiten. Plakativ formuliert: Eine KI in der Tumordiagnose muss sorgfältiger betrieben und abgesichert werden als ein Witzegenerator zu Unterhaltungszwecken. Ebenso ist eine Spracherkennung für eine Textverarbeitung weniger gefährlich als ein autonom fahrendes Fahrzeug oder eine Herz-Rhythmus-Analyse durch eine Smartwatch.

Nicht erforderlich sind Maßnahmen, die völlig unverhältnismäßig sind oder die lediglich absolut lebensfremde Risiken abfedern würden. Ebenso muss man nicht Sorge für solche Personen tragen, die mit der KI oder einem KI-erzeugten Produkt unbefugt umgehen (Ausnahme: Produkte für Kinder). Entscheidend ist aber: Verkehrspflichten können nicht nur dann verletzt werden, wenn das System an sich fehlerhaft ist. Auch wenn das eingesetzte System nach dessen eigenen Logik fehlerfrei trainiert wurde, bestimmungsgemäß funktioniert, aber für den gewählten Einsatz untauglich ist oder nicht ordnungsgemäß und risikoangemessen bedient wird, kann es zur Haftung kommen. Würde ein Krankenhaus beispielsweise eine gerade neu entwickelte KI zur Tumordiagnose einsetzen und auf jegliche Kontrolle und Überwachung durch erfahrene Ärzte und Ärztinnen verzichten, wäre dies bei einer schadensverursachenden Fehleinschätzung eines an sich bestimmungsgemäß laufenden Systems als eine Verletzung der Verkehrspflicht durch das Krankenhaus zu werten. Unter Umständen könnte sogar die Verkehrssicherungspflicht bestehen, zur Validierung redundante Systeme mit gleichen Input-Daten einzusetzen und weitere mögliche technische Sicherungsvorkehrungen einzurichten. Ähnlich kann eine Verkehrssicherungspflicht technische Lösungen erfordern, wie z. B. die jederzeitige Eingriffsmöglichkeit eines Fahrenden bei autonomen Fahrzeugen oder ein manuelles Freigabeerfordernis für Updates an Systemen, die im laufenden Einsatz kontinuierlich mit weiteren Daten weiterlernen sollen. Welche konkreten Anforderungen an die Herstellung und die Verwendung von KI in verschiedenen Anwendungsfeldern bestehen, wird die Rechtsprechung im Zweifel zukünftig in Fallgruppen herausbilden. Spezielle Anforderungen gibt es überdies in bestimmten Branchen, etwa im Arznei- oder Lebensmittelrecht.

Besonders geregelt sind die Verkehrspflichten für Produkthersteller:innen (sog. Produzentenhaftung), die entsprechend sowohl für Hersteller:innen von KI-Lösungen als auch für Produzierende von Produkten mit integrierter KI gelten. Generell haben Hersteller:innen im Rahmen des technisch Möglichen und wirtschaftlich Zumutbaren alle erforderlichen Vorkehrungen zu treffen, um zu verhindern, dass Dritte durch ihre Produkte geschädigt werden. Die Reichweite solcher Pflichten bestimmt die Rechtsprechung durch eine Interessenabwägung und berücksichtigt dabei auch, inwieweit die Geschädigten sich selbst hätten schützen können und müssen. Zum Teil werden die Verkehrspflichten durch bereichsspezifische Gesetze näher konkretisiert, etwa durch das Geräte- und Produktsicherheitsgesetz.

In der Produzentenhaftung unterscheidet man zwischen den folgenden Fehlerarten:

- Ein **Konstruktionsfehler** liegt vor, wenn ein Produkt schon von seiner Konzeption her nicht den berechtigten Sicherheitserwartungen von durchschnittlichen Nutzern und Nutzerinnen entspricht. Ein solcher liegt jedenfalls vor, wenn es ein Produktdesign gibt, das den Schaden nicht verursacht hätte. Im Detail kann die Bestimmung eines Konstruktionsfehlers kompliziert sein. Wann ist ein System, ein Algorithmus, eine Software fehlerhaft? Zwei Aspekte können hier besonders wichtig werden: Zum einen bestimmt der von den Herstellenden festgelegte bestimmungsgemäße Verwendungszweck des Systems, ob in einem bestimmten Anwendungsfall ein Konstruktionsfehler vorliegt. Ist der Schaden außerhalb des

Verwendungszwecks aufgetreten, liegt ein Fehlgebrauch der Geschädigten vor, für den Hersteller:innen nicht voll haften. Je enger der bestimmte Verwendungszweck ist, desto geringer ist das Haftungsrisiko. Man kennt diese Strategie von Elektrogeräteherstellenden, die in beigefügten, mitunter seitenlangen Booklets Verwendungszweck und Sorgfaltspflichten für den Umgang mit dem Gerät definieren. Zum anderen werden die umfangreichen oben dargestellten Anforderungen der Datenschutzbehörden (➔2.3) relevant. Können Geschädigte zeigen, dass diese Anforderungen nicht berücksichtigt wurden, etwa im Trainingsvorgang, spricht dies dem ersten Anschein nach für ein fehlerhaftes System.

- Ein **Fabrikationsfehler** liegt vor, wenn es bei Herstellung einzelner Stücke zu einer planwidrigen Abweichung von der Designvorlage für das Produkt kommt. Zu solchen Effekten könnte es kommen, wenn ein Produktionsprozess durch eine KI autonom gesteuert und modifiziert wird.
- **Organisationsfehler** sind Mängel der Arbeitsabläufe und Produktionsprozesse, die es verhindern, dass Fehler durch Kontrollen entdeckt und frühzeitig beseitigt werden. Organisationsfehler kommen in Betracht, wenn ein Unternehmen kein dediziertes Qualitätsmanagement für Entwicklung, Produktion und Oursourcing oder keine Datenschutzressourcen für das Projekt einsetzt.
- **Instruktionsfehler** sind im Kern unzureichende Anleitungen der Nutzer:innen über den sachgerechten Umgang mit dem Produkt sowie fehlende Informationen über mögliche Gefahren des Produkts.
- **Produktbeobachtungsfehler** sind anzunehmen, wenn Hersteller:innen (u. U. auch Anwender:innen) nach Marktstart des Produkts versäumen, das Produkt in angemessenem Umfang laufend zu überwachen (z. B. durch Auswertung des Kundenfeedbacks), ob sich bei der praktischen Verwendung des Produkts Risiken für die Nutzer:innen ergeben. Unter Umständen müssen Hersteller:innen mit Warnungen oder Rückruf reagieren. Sie können sich nicht dadurch entlasten, darauf zu verweisen, sie hätten mit den Produktrisiken nichts zu tun, diese seien autonom durch das System verursacht. Schon wegen des von der Datenschutzaufsicht zum Schutz von Diskriminierungen geforderten Monitorings wird die Produktbeobachtung für maschinelles Lernen eine große Rolle spielen.

Gewissermaßen eine Schwachstelle im System der Produzentenhaftung ist der durch die geschädigten Anspruchsteller:innen zu erbringende Nachweis, dass die:der Hersteller:in den eingetretenen Schaden durch Verletzung einer der dargestellten Verkehrssicherungspflichten kausal verursacht hat. Bei multikausalen Zusammenhängen, bei der eine Vielzahl von Teilen, Software (teilweise Open-Source-Software in Kombination mit proprietärem Code) und Daten in laufenden und nicht immer transparenten Prozessen zusammenkommt, wird es immer schwieriger, Schadensbeiträge zu identifizieren und zu beweisen, was sich oftmals zulasten der Geschädigten auswirkt.

Produzentenhaftung greift nur bei Verschulden, also bei Vorsatz oder Fahrlässigkeit. Es wäre ein Missverständnis zu meinen, den Hersteller:innen / Anwender:innen einer KI sei pauschal kein Fahrlässigkeitsvorwurf zu machen, wenn das System autonom zu einer fehlerhaften Entscheidung gelangt. Ein Verschulden liegt vor, wenn fahrlässig ein fehlerhaft arbeitendes System hergestellt wurde. Ein Verschulden kann aber auch vorliegen, wenn selbst bei fehlerfreiem System

vorsätzlich oder fahrlässig eine der dargestellten Verkehrssicherungspflichten verletzt wird.

Liegen die Voraussetzungen der Produzentenhaftung vor, haften Hersteller:innen uneingeschränkt auf Ersatz der Vermögensschäden und bei Verletzungen von Leben, Körper, Gesundheit und Persönlichkeitsrechten zudem auf Ersatz der immateriellen Schäden (Schmerzensgeld).

Was können Hersteller:innen tun, um möglichst einen Vorwurf der Verletzung von Verkehrssicherungspflichten zu entkräften? Der Kern ist ein dokumentiertes Qualitätsmanagement, das sich in weiten Zügen mit den datenschutzrechtlichen Anforderungen überschneidet, die ja letztlich auch Fehler des maschinellen Lernens verhindern sollen (➔ 2.3). Dazu gehört es auch, den Stand der Technik zu verfolgen und aktuelle Tools / Methoden / Sicherungsmechanismen einzusetzen. Hinzu kommt die klar kommunizierte Abgrenzung zwischen bestimmungsgemäßer Verwendung des Systems und Fehlgebrauch der Anwender:innen / Kunden und Kundinnen. Überdies können – bei gefahrgeneigten Projekten – Zertifizierungen, Gütesiegel und freiwillige Audits die dokumentierte Übereinstimmungen mit einschlägigen DIN-Normen, technischen Standards und einschlägigen Ethikrichtlinien einen sorgsamen Umgang belegen. Gleiches gilt für eine freiwillige Superversion durch externe Beratergremien / Beiräte. Gegenwärtig arbeiten beispielsweise der TÜV-Verband mit dem Bundesamt für Sicherheit in der Informationstechnik (BSI) (Vd TÜV 2019) an Prüfschritten für KI-Anwendungen sowie das Fraunhofer IAIS mit dem BSI an einem Projekt zur KI-Zertifizierung (Fraunhofer IAIS 2020).

4.4.1.3 Produkthaftung

Neben vertraglichem Schadenersatz und Ansprüchen aus Produzentenhaftung kommt schließlich noch eine Haftung nach dem Produkthaftungsgesetz (ProdHaftG) in Betracht. Anders als bei der Produzentenhaftung haften Hersteller:innen verschuldensunabhängig, wenn ein von ihnen in Verkehr gebrachtes Produkt einen Fehler aufweist und dieser Fehler dazu führt, dass ein Mensch getötet, sein Körper oder seine Gesundheit verletzt oder eine Sache, die für den privaten Gebrauch oder Verbrauch bestimmt ist, beschädigt wird. Der Haftungsmaßstab geht damit sehr weit, nicht einmal Fahrlässigkeit ist erforderlich. Als Korrektiv für die strenge Haftung gilt eine Deckelung auf einen Höchstbetrag von 85 Millionen Euro für einen Produktfehler, der sich massenhaft auswirkt. Im Falle von Sachbeschädigungen sieht das Gesetz zudem eine Selbstbeteiligung der Geschädigten in Höhe von 500 Euro vor. Zur Klarstellung: Dieser Deckel gilt nur für die Haftung nach dem Produkthaftungsgesetz, nicht für die oben dargestellte Produzentenhaftung.

Anknüpfungspunkt ist auch hier wieder die Verletzung eines Rechtsguts durch ein fehlerhaftes Produkt. Als „Produkt“ zählen Hardware und physische Produkte, nicht jedoch reine Software, auch wenn sich dies wegen einer Novellierung des EU-Rechtsrahmens bald ändern mag. Damit scheidet Produkthaftung für die reine Software eines maschinellen Lernens aus. Als „Produkt“ zählen allerdings solche

physischen Produkte, in die eine KI integriert ist („embedded software“) und vollständige Systemlösungen.

Ein Fehler liegt vor, wenn das Produkt nicht die Sicherheit bietet, die unter Berücksichtigung aller Umstände, insbesondere seiner Darbietung, des typischen Gebrauchs und des Zeitpunktes des In-Verkehr-Bringens berechtigterweise erwartet werden kann. Ähnlich wie in der Produzentenhaftung wird auch hier wieder die Zweckbestimmung des Produkts durch die Hersteller:innen relevant. Allerdings haften die Hersteller:innen nicht, dies ist für maschinelles Lernen relevant, für sog. Entwicklungsfehler, wenn der Fehler nach dem Stand der Technik in dem Zeitpunkt, in dem die Hersteller:innen das Produkt in den Verkehr brachten, nicht erkannt werden konnte (§ 1 II Nr. 2 ProdHaftG). Zum Teil wird vertreten, diese Ausnahme gelte für Fehler, die erst aufgrund weiteren Trainings während des Einsatzes in der Blackbox entstanden sind und damit bei Marktstart nicht erkennbar waren. Diese Auffassung ist jedoch umstritten. Die Gegenauffassung sieht auch in diesen Fällen eine von Anfang an fehlerhafte Programmierung als gegeben an. Rechtsprechung gibt es hierzu bislang jedoch nicht.

4.4.2 Sonstige Maßnahmen zum Umgang mit Haftungsrisiken

Im unternehmerischen Kontext können Haftungsrisiken für maschinelles Lernen durch weitere Maßnahmen minimiert werden. Zu nennen ist etwa die gesellschaftsrechtliche Auslagerung des Systems in eine Betreibergesellschaft in einer Rechtsform mit Haftungsbeschränkung, z. B. eine GmbH. Zudem kommen Versicherungslösungen in Betracht. Inwieweit Versicherungsschutz durch eine Haftpflicht, Produkthaftpflicht, Unfallversicherung oder gar eine Versicherung über Cyberrisiken besteht, ist im Einzelfall mit den Versichererungsunternehmen zu klären.

4.4.3 „E-Personen“ als Mittel der Haftungsvermeidung

Rechtspolitisch wurde vom EU-Parlament vorgeschlagen, Künstliche Intelligenzen ab einem gewissen Reifegrad an Autonomie als sog. „E-Person“ mit einer Rechtsfähigkeit auszustatten. Eine KI könnte dann genauso wie z. B. eine Gesellschaft mit beschränkter Haftung (GmbH) oder Aktiengesellschaft (AG) Eigentum erwerben, Verträge schließen, haften und verklagt werden. Da „E-Personen“ wie z. B. autonome Fahrzeuge aber typischerweise über kein eigenes Vermögen verfügen dürften, könnte das Haftungsproblem über Haftungsfonds oder (Pflicht)Versicherungen gelöst werden. Diese Überlegungen sind letztlich als eine elaborierte Haftungsvermeidungsstrategie der Hersteller:innen zu werten, mit der Haftungsrisiken beim Einsatz von KI auf die Allgemeinheit externalisiert würden. Das Konzept der „E-Person“ wird viel diskutiert, hat aktuell politisch aber keine Priorität.

4.5 Arbeitsrechtliche Aspekte

Die Einführung von Lösungen des maschinellen Lernens im betrieblichen Kontext kann verschiedene arbeitsrechtliche Belange berühren und vor allem Beteiligungsrechte des Betriebsrats auslösen. Die arbeitsrechtlichen Bezüge von maschinellem Lernen können im vorliegenden Rahmen nur angerissen werden.

Ein sicherlich zu adressierender Gesichtspunkt ist eine mögliche Angst der Beschäftigten vor Jobverlust, wenn Aufgaben im Rahmen der betrieblichen Wertschöpfung in Zukunft von einer KI übernommen werden. Hier unterscheidet sich die Rechtslage aber nicht von anderen betrieblichen Rationalisierungsvorhaben. Arbeitgeber:innen entscheiden grundsätzlich frei über den Einsatz von digitalen Arbeitsprozessen. Resultieren daraus betriebsbedingte Kündigungen, finden die allgemeinen Regeln hierfür ohne Weiteres Anwendung, also ggf. das Kündigungsschutzgesetz (KSchG) sowie das Erfordernis der Betriebsratsanhörung gemäß § 102 Betriebsverfassungsgesetz (BetrVG). Auch auf kollektiver Ebene hat der Betriebsrat Beteiligungsrechte, soweit die Belange der Beschäftigten betroffen sind: Zunächst besteht ein Beratungsrecht nach § 90 BetrVG, wenn das System in Zusammenhang mit Arbeitsverfahren, Arbeitsabläufen oder Arbeitsplatzgestaltungen steht. Bei wesentlichen Betriebsänderungen (z. B. grundlegenden Änderungen der Betriebsorganisation oder Einführung grundlegend neuer Arbeitsmethoden) können nach § 111 ff. BetrVG Interessenausgleich und Sozialplan zu verhandeln sein. Da maschinelles Lernen nur für spezifische Aufgabenstellungen eingesetzt wird, sind Fälle einer wesentlichen Betriebsänderung allerdings nur schwer vorstellbar.

Ein zweiter wichtiger Gesichtspunkt ist das Mitbestimmungsrecht des Betriebsrats bei Einführung und Anwendung von technischen Einrichtungen, die geeignet sind, das Verhalten oder die Leistung von Arbeitnehmern und Arbeitnehmerinnen zu überwachen (§ 87 I Nr. 6 BetrVG). Die Mitarbeiterkontrolle mag in der Regel zwar nicht der Primärzweck des Systems maschinellen Lernens sein. Aber wenn das System die oben dargestellten Anforderungen an die Datensicherheit beim Umgang mit personenbezogenen Daten erfüllen soll, dann muss es eine ordnungsgemäße Zugriffsverwaltung und ein Berechtigungskonzept für alle mit dem System arbeitenden Mitarbeiter:innen geben. Die hierdurch erzeugten Logs mit Log-in-Daten und -Zeiten u. Ä. erlauben eine Kontrolle der Mitarbeiter:innen und können damit dieses Beteiligungsrecht auslösen. Es kommt nicht darauf an, ob diese Daten aus Sicht der Arbeitgeber:innen der Überwachung dienen sollen, sondern allein darauf, ob diese Daten hierzu objektiv geeignet wären. Da dieses Mitbestimmungsrecht erzwingbar ist, könnte es zu einer Situation kommen, bei der der Betriebsrat bis zu einer Einigung die Einführung des Systems blockiert. Empfehlenswert ist daher, sich entweder frühzeitig zu einigen oder (unter Wahrung der Datensicherheit) die Überwachungsmöglichkeit mit einer technischen Lösung zu unterbinden.

Werden durch das System personenbezogene Daten verarbeitet, ist darauf zu achten, dass im Rahmen der technisch-organisatorischen Maßnahmen (➔ 2.3.9)

die Mitarbeiter:innen, die Zugriff auf das System haben, auf das Datengeheimnis verpflichtet werden. Hierzu haben die Datenschutzbehörden ein Muster (DSK 2018) bereitgestellt.

Vor allem im Bereich der Robotik und industriellen Fertigung können Aspekte des Arbeitsschutzes betroffen sein. Hier gilt das Mitbestimmungsrecht des Betriebsrats über Regelungen zu Unfallverhütung (§ 87 I Nr. 7 BetrVG, flankierend § 90 BetrVG) sowie das Mitbestimmungsrecht bei besonders belastender Arbeitsplatzgestaltung (§ 91 BetrVG). Für Hersteller:innen von Robotik-Systemen gelten Anforderungen des Produktsicherheitsgesetzes einschließlich der hierzu erlassenen Verordnungen. Vor Einsatz im Betrieb kann eine Gefährdungsbeurteilung erforderlich sein (§ 4 I Betriebssicherheitsverordnung, BetrSichV). Im Einzelnen kann auf die Berufsgenossenschaften und die gesetzliche Unfallversicherung verwiesen werden, die zu diesem Themenfeld Handlungsleitfäden bereithalten.

Die Einführung eines KI-Systems kann Fortbildungen nötig machen. In diesem Themenfeld bestehen ebenfalls Beteiligungsrechte des Betriebsrats, die es verhindern sollen, dass Arbeitgeber:innen vorschnell Bestandsmitarbeiter:innen ohne Qualifikationsbemühungen kündigen und sich stattdessen das Know-how neu rekrutieren (§§ 96, 97 BetrVG).

Arbeitsrechtlich komplizierter sind die Spezialfälle, in denen maschinelles Lernen unmittelbar gegenüber Beschäftigten als „HR-Tool“ eingesetzt wird (z. B. automatisiertes Recruiting, „People Management“, „HR Analytics“) oder ein solches System die Rolle von Vorgesetzten einnimmt, auf die Arbeitgeber:innen ihr Weisungsrecht delegieren. Diese Fälle lösen zahlreiche Fragen des Beschäftigtendatenschutzes, der Diskriminierungsverbote, des Vorbehalts menschlicher Entscheidung (Art. 22 DSGVO), des Mitbestimmungsrechts und nicht zuletzt der Ethik im Umgang mit Mitarbeitenden aus. Diese Fragen sind aufgrund ihres Umfangs und ihrer Bedeutung gesondert zu betrachten.

5 Fazit

Wie so oft ist die Technik auch auf dem Gebiet der KI anderen relevanten Faktoren weit voraus. Technologien des maschinellen Lernens zu beschaffen oder gar selbst zu entwickeln, wird vielen Wirtschaftsteilnehmenden möglich sein. Weit schwieriger ist die Klärung von rechtlichen Anforderungen oder die Einschätzung von wirtschaftlichen Implikationen.

Die Komplexität der Rechtslage wird abschreckend wirken und viele vor allem kleinere Marktteilnehmer:innen zumindest auf den ersten Blick überfordern. Sie ist vor allem darauf zurückzuführen, dass unzählige Regelungsaspekte noch offen sind oder weiterer Anpassungen bedürfen. Die Politik ist aufgefordert, das Recht zügig weiterzuentwickeln und in diesem Zuge die regulativen Vorgaben an die Wirtschaft zu präzisieren und zu vereinfachen. Manche Grundgedanken und Ziele des Datenschutzrechts stehen mit der für KI-Technologien erforderlichen massenhaften Datenverarbeitung in diametralem Widerspruch. Es mag sich erweisen, dass es zur Lösung dieses Dilemmas – sowohl Datenschutz als auch Innovation sind politisch erwünscht – neuer Ansätze bedarf. Das gilt gleichermaßen für die immaterialgüterrechtlichen Fragen.

Unternehmen können derweil die Komplexitäten durch Prozesse, kluge und nachhaltige Strategien wie Datenverwendungskonzepte reduzieren. Rechtliche Aspekte sollten dabei stets frühzeitig mitgedacht werden, um Investitionen in Geschäftsmodelle oder technische Ansätze zu vermeiden, die letztlich an Compliance-Problemen scheitern. Auch ein Dialog mit den Datenschutzbehörden kann sich empfehlen, um Rechtsunsicherheit zu mindern. Im Übrigen ist auf eine sorgfältige Vertragsgestaltung zu achten, soweit über KI-Lösungen Vereinbarungen geschlossen werden. Gerade im Hinblick auf die oft unklare Rechtslage sind präzise Vertragstexte von großer Bedeutung. Geht es hierin um technische Spezifika, wie etwa bei der Leistungsbeschreibung einer selbstlernenden Softwarelösung, müssen sowohl juristische als auch technische Expertise in die Vertragsgestaltung einfließen.

Generell bedeutet Rechtsunsicherheit nicht, dass man nicht handeln kann oder sollte. Sie bedeutet erst einmal nur, dass Rechtsfragen als Teil des Risikomanagements behandelt werden müssen. Aus dieser Perspektive gibt es zumindest zwei Wege, mit Rechtsunsicherheit umzugehen: Man kann sie konstruktiv nutzen oder destruktiv vermeiden. Positiv betrachtet eröffnet Rechtsunsicherheit Spielräume. Sie trifft immer alle Beteiligten, Nutzer:innen wie Rechtsinhaber:innen, Datenschutzverantwortliche wie -behörden. Im Rahmen sinnvoller Risikoabwägung kann man die Spielräume für sich nutzen. Wer sich von der Rechtsunsicherheit jedoch abschrecken lässt, überlässt anderen das Feld. Das mag sich als großer Fehler erweisen. Technologien des maschinellen Lernens dauerhaft zu vermeiden, wird häufig schlicht keine Option sein. Ohne Zweifel werden sie Wirtschaftszweige aller Art nachhaltig beeinflussen und verändern.

Verzeichnisse

Literatur

AP – Associated Press Frankreich (o. J.). „AI@A". https://www.ap.org/discover/artificial-intelligence (Download 21.09.2020).

Bayerisches Landesamt für Datenschutzaufsicht (2018a). „Formulierungshilfe für einen Auftragsverarbeitungsvertrag nach Art. 28 Abs. 3 DS-GVO". https://www.lda.bayern.de/media/muster_adv.pdf (Download 21.09.2020).

Bayerisches Landesamt für Datenschutzaufsicht (2018b). „Liste der Verarbeitungstätigkeiten, für die eine DSFA durchzuführen ist". https://www.lda.bayern.de/media/dsfa_muss_liste_dsk_de.pdf (Download 29.10.2020).

Beining, Leonie (2019). *Wie Algorithmen verständlich werden. Ideen für Nachvollziehbarkeit von algorithmischen Entscheidungsprozessen für Betroffene*. November 2019. Hrsg. Stiftung Neue Verantwortung und Bertelsmann Stiftung. Berlin. https://www.bertelsmann-stiftung.de/fileadmin/files/BSt/Publikationen/GrauePublikationen/Wie_Algorithmen_verstaendlich_werden_final.pdf (Download 21.9.2020).

berlinvalley.com (2018). „DeepBach lernt zu komponieren". 27.6.2018. https://berlinvalley.com/deepbach-lernt-zu-komponieren/ (Download 29.10.2020).

Bitkom (2020a). „Anonymisierung und Pseudonymisierung von Daten für Projekte des maschienellen Lernens". https://www.bitkom.org/sites/default/files/2020-10/201002_lf_anonymisierung-und-pseudonymisierung-von-daten.pdf (Download 13.01.2021).

Bitkom (2020b). „Unternehmen tun sich noch schwer mit Künstlicher Intelligenz". Pressemitteilung. 8.6.2020. https://bitkom.org/Presse/Presseinformation/Unternehmen-tun-sich-noch-schwer-mit-Kuenstlicher-Intelligenz (Download 17.9.2020).

Bitkom (2018). „Machine Learning und die Transparenzanforderungen der DS-GVO. Leitfaden". https://bitkom.org/sites/default/files/file/import/180926-Machine-Learning-und-DSGVO.pdf (Download 21.9.2020).

Bitkom (2017). „Mustervertragsanlage. Auftragsverarbeitung i. S. d. Art. 28 Abs. 3 Datenschutz-Grundverordnung (DS-GVO)". https://www.bitkom.org/sites/default/files/file/import/170515-Auftragsverarbeitung-Anlage-Mustervertrag-online.pdf (Download 21.9.2020).

Brühl, Jannis (2020). „Gesichtserkennung. Clearview AI lässt sich Kundenliste stehlen". *Süddeutsche Zeitung* 27.2.2020. https://www.sueddeutsche.de/digital/gesichtserkennung-clearview-kunden-polizei-1.4823352 (Download 21.9.2020).

Bundesamt für Sicherheit in der Informationstechnik (2020a). „IT-Grundschutz". https://www.bsi.bund.de/DE/Themen/ITGrundschutz/itgrundschutz_node.html (Download 21.9.2020).

Bundesamt für Sicherheit in der Informationstechnik (2020b). „ISO 27001 Zertifizierung auf Basis von IT-Grundschutz". https://www.bsi.bund.de/DE/Themen/Zertifizierung undAnerkennung/Managementsystemzertifizierung/Zertifizierung27001/GS_Zertifizierung_node.html (Download 21.9.2020).

Bundesministerium der Justiz und für Verbraucherschutz (2020). „Diskussionsentwurf des Bundesministeriums der Justiz und für Verbraucherschutz. Entwurf eines Zweiten Gesetzes zur Anpassung des Urheberrechts an die Erfordernisse des digitalen

Binnenmarktes“. 24.6.2020. https://www.bmjv.de/SharedDocs/Gesetzgebungsverfahren/Dokumente/DiskE_II_Anpassung%20Urheberrecht_digitaler_Binnenmarkt.pdf?__blob=publicationFile&v=2 (Download 27.10.2020).

Bundesregierung (2018). „Strategie Künstliche Intelligenz der Bundesregierung“. https://www.ki-strategie-deutschland.de/files/downloads/Nationale_KI-Strategie.pdf (Download 17.9.2020).

Datatilsynet – The Norwegian Data Protection Authority (2018). „Artificial intelligence and privacy. Report. January 2018“. https://www.datatilsynet.no/globalassets/global/english/ai-and-privacy.pdf (Download 21.9.2020).

Der Hessische Beauftragte für Datenschutz und Informationsfreiheit (2020). „Hinweise und Muster DS-GVO“. https://datenschutz.hessen.de/infothek/hinweise-und-muster-ds-gvo (Download 21.9.2020).

DSK – Datenschutzkonferenz (2020). „Das Standard-Datenschutzmodell (SDM)“. https://www.datenschutzzentrum.de/sdm/ (Download 21.9.2020).

DSK – Datenschutzkonferenz (2019a). „Entschließung der 97. Konferenz der unabhängigen Datenschutzaufsichtsbehörden des Bundes und der Länder . Hambacher Erklärung zur Künstlichen Intelligenz. 3. April 2019“. https://www.datenschutzkonferenz-online.de/media/en/20190405_hambacher_erklaerung.pdf (Download 14.9.2020).

DSK – Datenschutzkonferenz (2019b). „Positionspapier der DSK zu empfohlenen technischen und organisatorischen Maßnahmen bei der Entwicklung und dem Betrieb von KI-Systemen“. 6.11.2019. https://www.datenschutzkonferenz-online.de/media/en/20191106_positionspapier_kuenstliche_intelligenz.pdf (Download 14.9.2020).

DSK – Datenschutzkonferenz (2018). „Kurzpapier Nr. 19. Unterrichtung und Verpflichtung von Beschäftigten auf Beachtung der datenschutzrechtlichen Anforderungen nach der DS-GVO“. https://www.datenschutzkonferenz-online.de/media/kp/dsk_kpnr_19.pdf (Download 22.9.2020).

DSK – Datenschutzkonferenz (2017). „Kurzpapier Nr. 5. Datenschutz-Folgenabschätzung nach Art. 35 DS-GVO. 24.7.2017”. https://www.datenschutzkonferenz-online.de/media/kp/dsk_kpnr_5.pdf (Download 21.9.2020).

Europäische Kommission (2020a). „White Paper on Artificial Intelligence: a European approach to excellence and trust“. COM(2020) 65 final.19.2.2020. https://ec.europa.eu/info/files/white-paper-artificial-intelligence-european-approach-excellence-and-trust_en (Download 17.9.2020).

Europäische Kommission (2020b). „European data strategy“ COM(2020) 66 final. 19.2.2020. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020DC0066&from=EN (Download 17.9.2020).

Europäische Kommission (2020c). „Assessment List for Trustworthy Artificial Intelligence (ALTAI) for self-assessment“. https://ec.europa.eu/digital-single-market/en/news/assessment-list-trustworthy-artificial-intelligence-altai-self-assessment (Download 4.11.2020).

Europäische Kommission (2019a). „A Definition of AI: Main Capabilities and Disciplines. Definition developed for the purpose of the AI HLEG’s deliverables“. https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=56341 (Download 17.9.2020).

Europäische Kommission (2019b). „EU artificial intelligence ethics checklist ready for testing as new policy recommendations are published. Shaping Europe’s digital future“. 26.6.2019. https://ec.europa.eu/digital-single-market/en/news/eu-artificial-intelligence-ethics-checklist-ready-testing-new-policy-recommendations-are (Download 21.9.2020).

Europäische Kommission (2018a). „Communication Artificial Intelligence in Europe“. 25.4.2018. https://ec.europa.eu/digital-single-market/en/news/communication-artificial-intelligence-europe (Download 17.9.2020).

Europäische Kommission (2018b). „Coordinated Plan on Artificial Intelligence“. 7.12.2018. https://ec.europa.eu/digital-single-market/en/news/coordinated-plan-artificial-intelligence (Download 17.9.2020).

Europäisches Parlament (2020). „Künstliche Intelligenz: Parlament will faire und sichere Nutzung für Verbraucher“. Pressemitteilung. 12.2.2020. https://www.europarl.europa.eu/news/de/press-room/20200206IPR72015/kunstliche-intelligenz-parlament-will-faire-und-sichere-nutzung-fur-verbraucher (Download 21.9.2020).

Europäisches Patentamt (2020). „EPO publishes grounds for its decision to refuse two patent applications naming a machine as inventor“. 28.1.2020. https://www.epo.org/news-events/news/2020/20200128.html (Download 21.9.2020).

Europäische Union (2016). „Amtsblatt der Europäischen Union. Verordnung (EU) 2016/679 des Europäischen Parlaments und des Rates vom 27. April 2016 zum Schutz natürlicher Personen bei der Verarbeitung personenbezogener Daten, zum freien Datenverkehr und zur Aufhebung der Richtlinie 95/46/EG (Datenschutz-Grundverordnung)“. https://eur-lex.europa.eu/legal-content/DE/TXT/PDF/?uri=CELEX:32016R0679 (Download 17.9.2020).

Fischer, Martin (2018). „Künstliche Intelligenz als Gefahr: Menschheit muss ich auf Regeln einigen“. *heise online* 13.6.2018. https://www.heise.de/newsticker/meldung/Kuenstliche-Intelligenz-als-Gefahr-Menschheit-muss-sich-auf-Regeln-einigen-4077950.html (Download 17.9.2020).

Fraunhofer IAIS – Fraunhofer-Institut für Intelligente Analyse- und Informationssysteme (2020). „KI-Zertifizierung“. https://www.iais.fraunhofer.de/de/kompetenzplattform-ki-nrw/ki-zertifizierung.html (Download 21.9.2020).

Fraunhofer-Institut für Nachrichtentechnik und Heinrich-Hertz-Institut (HHI) (2019). „Paper bei Nature Communications erschienen: Wissenschaftler stellen KI-Systeme auf den Prüfstand". https://www.hhi.fraunhofer.de/presse-medien/nachrichten/2019/wie-intelligent-ist-kuenstliche-intelligenz.html (Download 17.9.2020).

Friedman, Colby, und Helen Nissenbaum (1996). „Bias in Computer Systems". https://nissenbaum.tech.cornell.edu/papers/biasincomputers.pdf (Download 27.10.2020).

Hendrycks, Dan, Kevin Zhao, Steven Basart, Jacob Steinhardt und Dawn Song (2020). „Natural Adversarial Examples". https://arxiv.org/abs/1907.07174. (Download 4.11.2020).

hollyherndon.com/proto (2019). „Proto". https://www.hollyherndon.com/proto (Download 21.09.2020).

ICDPPC – International Conference of Data Protection & Privacy Commissioners (2018). „Declaration on ethics and data protection in artificial intelligence. 40th International Conference of Data Protection and Privacy Commissioners". 23.10.2018. https://edps.europa.eu/sites/edp/files/publication/icdppc-40th_ai-declaration_adopted_en_0.pdf (Download 21.9.2020).

ico. Information Commissioner's Office (2017). „Big data, artificial intelligence, machine learning and data protection". https://ico.org.uk/media/for-organisations/documents/2013559/big-data-ai-ml-and-data-protection.pdf (Download 21.9.2020).

ico. Information Commissioner's Office und Alan Turing Institute (2020a). „Explaining decisions made with AI". https://ico.org.uk/for-organisations/guide-to-data-protection/key-data-protection-themes/explaining-decisions-made-with-ai/ (Download 4.11.2020).

ico. Information Commissioner's Office (2020b). „Guidance on AI and data protection". https://ico.org.uk/for-organisations/guide-to-data-protection/key-data-protection-themes/guidance-on-ai-and-data-protection/ (Download 2.11.2020).

InfoCuria Rechtssprechung (2020). „Urteil des Gerichtshofs (Große Kammer)". 16.7.2020. http://curia.europa.eu/juris/document/document.jsf?text=&docid=228677&pageIndex=0&doclang=DE&mode=req&dir=&occ=first&part=1 (Download 21.9.2020).

iRights info (2019). „EuGH kassiert das sechs Jahre alte Presse-leistungsschutzrecht". 12.9.2019. https://irights.info/artikel/eugh-kassiert-das-sechs-jahre-alte-presseleistungsschutzrecht/29686 (Download 26.10.2020)

Krafft, Tobias D., und Katharina A. Zweig (2019). *Transparenz und Nachvollziehbarkeit algorithmenbasierter Entscheidungsprozesse. Ein Regulierungsvorschlag aus sozioinformatischer Perspektive.* 22.1.2019. Hrsg. Verbraucherzentrale Bundesverband e. V. Berlin. (Auch online unter https://www.vzbv.de/sites/default/files/downloads/2019/05/02/19-01-22_zweig_krafft_transparenz_adm-neu.pdf (Download 21.9.2020).

Lecher, Colin (2019). „How Amazon automatically tracks und fires warehouse workers for ‚productivity'. *The Verge* 25.4.2019. https://www.theverge.com/2019/4/25/18516004/amazon-warehouse-fulfillment-centers-productivity-firing-terminations (Download 4.11.2020).

McKinsey Global Institute (2018). „Notes from the AI frontier. Modeling the impact of AI on the world economy". Discussion Paper September 2018. https://www.mckinsey.com/~/media/McKinsey/Featured%20Insights/Artificial%20Intelligence/Notes%20from%20the%20frontier%20Modeling%20the%20impact%20of%20AI%20on%20the%20world%20economy/MGI-Notes-from-the-AI-frontier-Modeling-the-impact-of-AI-on-the-world-economy-September-2018.ashx (Download 17.9.2020).

Plattform Lernende Systeme (2019). „Künstliche Intelligenz in Deutschland. KI-Landkarte". https://www.plattform-lernende-systeme.de/ki-in-deutschland.html (Download 17.9.2020).

pwc – PricewaterhouseCoopers GmbH Wirtschaftsprüfungs-gesellschaft (2018). „Auswirkungen der Nutzung von künstlicher Intelligenz in Deutschland". https://www.pwc.de/de/business-analytics/sizing-the-price-final-juni-2018.pdf (Download 17.9.2020).

Schmeiss, Jessica, und Nicolas Friederici (2020). *Demystifying AI. Was steckt hinter KI-Unternehmen in Deutschland?* Hrsg. Alexander von Humboldt Institut für Internet und Gesellschaft gGmbH. Berlin. https://www.hiig.de/wp-content/uploads/2020/02/Demystifying_AI_web-1.pdf (Downlad 17.9.2020).

ssrn.com (o. J.). „Suchbegriff Artificial Intelligence". https://www.ssr.com/index.cfm/en

Thefader.com und Emilie Friedlander (2019). „How Holly Herndon and her AI baby spawned a new kind of folk music". https://www.thefader.com/2019/05/21/holly-herndon-proto-ai-spawn-interview (Download 29.10.2020).

Vd TÜV (2019) „TÜV-Verband und BSI arbeiten bei Künstlicher Intelligenz zusammen". 5.11.2019. https://www.vdtuev.de/en/vdtuev-startseite/news/KI-Expertengruppe (Download 4.11.2020).

VDMA (2019). „Wertschöpfung aus Maschinendaten 4.0 – Einführung und Motivation". https://ea.vdma.org/viewer/-/v2article/render/28872350 (Download 21.9.2020).

Wachter, Sandra, Brent Mittelstadt und Chris Russell (2018). „Counterfactual Explanations without opening the black box: Automated decisions and the GDPR". *Harvard Journal of Law & Technology* (31) 2. 841 ff. https://arxiv.org/ftp/arxiv/papers/1711/1711.00399.pdf (Download 21.9.2020).

Wikipedia (2020). „Algorithmic Bias". Bearbeitungsstand 17.10.2020. https://en.wikipedia.org/wiki/Algorithmic_bias (Download 27.10.2020).

WIPO – World Intellectual Property Organisation (2019). „Draft Issue Paper on intellectual property policy and artificial intelligence. WIPO Conversation on Intellectual Property (IP) and Artificial Intelligence (AI)". 13.12.2019. https://www.wipo.int/edocs/mdocs/mdocs/en/wipo_ip_ai_2_ge_20/wipo_ip_ai_2_ge_20_1.pdf (Download 21.9.2020).

YouTube.com (2018). „Holly Herndon & Jlin (feat. Spawn) – Godmother (Official Video)". https://www.youtube.com/watch?v=sc9OjL6Mjqo (Download 29.10.2020).

Weiterführende Literaturempfehlungen

Algo.Rules (2020). *Praxisleitfaden zu den Algo.Rules. Orientierungshilfen für Entwickler:innen und ihre Führungskräfte.* Hrsg. iRights.Lab und Bertelsmann Stiftung. Berlin. https://www.bertelsmann-stiftung.de/fileadmin/files/alg/Algo.Rules_Praxisleitfaden.pdf (Download 21.9.2020).

Beining, Leonie, und Carla Hustedt (2019). „Raus aus der Black Box. Algorithmen für alle nachvollziehbar machen – aber wie?" 21.11.2019. https://algorithmenethik.de/2019/11/21/raus-aus-der-black-box-algorithmen-fuer-alle-nachvollziehbar-machen-wie/ (Download 21.9.2020).

BMWi – Bundesministerium für Wirtschaft und Energie (Hrsg.) (2019). *Künstliche Intelligenz und Recht im Kontext von Industrie 4.0.* Berlin. https://www.plattform-i40.de/PI40/Redaktion/DE/Downloads/Publikation/kuenstliche-intelligenz-und-recht.pdf?__blob=publicationFile&v=4 (Download 21.9.2020).

Chen, Angela (2020). „Can an AI be an inventor? Not yet. But some campaigners are pushing for the rules to change". *MIT Technology Review* 8.1.2020. https://www.technologyreview.com/s/615020/ai-inventor-patent-dabus-intellectual-property-uk-european-patent-office-law/ (Download 22.9.2020).

Conrad, Sebastian Conrad (2018). „Künstliche Intelligenz und die DSGVO – Ausgewählte Problemstellungen". *Kommunikation & Recht (K&R).* 741–746.

Davis, Dave (2018). „How AI and copyright could work". 9.1.2018. https://techcrunch.com/2018/01/09/how-ai-and-copyright-would-work/?guccounter=2 (Download 22.9.2020).

Denga, Michael (2018). „Deliktische Haftung für künstliche Intelligenz. Warum die Verschulungshaftung des BGB auch künftig die bessere Schadensausgleichsordnung bedeutet". *Computer und Recht* (Band 34) 2. 69. https://doi.org/10.9785/cr-2018-0203 (Download 12.10.2020).

Europäische Kommission (2019). *Liability for Artificial Intelligence and other emerging digital technologies*. Report form the Expert Group on Liability and New Technologies – New Technology Formation". Brüssel. https://ec.europa.eu/transparency/regexpert/index.cfm?do=groupDetail.groupMeetingDoc&docid=36608 (Download 21.9.2020).

European Data Protection Board (2020). „Europäischer Datenschutzausschuss – 31. Plenartagung: Einrichtung einer Taskforce zu TikTok, Antwort an MdEP zur Verwendung von Clearview AI durch Strafverfolgungsbehörden, Antwort an die ENISA-Beratungsgruppe, Antwort auf den Offenen Brief von NOYB". 10.6.2020. https://edpb.europa.eu/news/news/2020/thirty-first-plenary-session-establishment-taskforce-tiktok-response-meps-use_de (Download 29.10.2020).

Garpentin, Justin (2019). „Die Erosion der Vertragsgestaltungsmacht durch das Internet und den Einsatz Künstlicher Intelligenz". *Neue Juristische Wochenschrift* 4/2019. 181–185.

Gausling, Tina (2019). „Künstliche Intelligenz im digitalen Marketing". *Zeitschrift für Datenschutz (ZD)* 8. 335–340.

Gausling, Tina (2018). „Künstliche Intelligenz und DSGVO". *DSRI-Tagungsband 2018*. 519–544.

Gervais, Daniel (2020). *The Machine as Author. Legal Studies Research Paper Series 19–35*. Iowa Law Review, Vol. 105. Nashville TN: Vanderbilt University Law School. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3359524 (Download 21.9.2020).

Gervais, Daniel (2019). „Can Machines be Authors". *Kluwer Copyright Blog* 21.5.2019. http://copyrightblog.kluweriplaw.com/2019/05/21/can-machines-be-authors/ (Download 22.9.2020).

González Otero, Begoña und João Petro Quintais (2018). „Before the Singularity: Copyright and the Challenges of Artificial Intelligence". *Kluwer Copyright Blog* 25.9.2018. http://copyrightblog.kluweriplaw.com/2018/09/25/singularity-copyright-challenges-artificial-intelligence/ (Download 21.9.2020).

Grävemeyer (2020). „Wie sich KI-Entscheidungen überprüfen lassen". *heise.de* 9.3.2020. https://www.heise.de/ct/artikel/Wie-sich-KI-Entscheidungen-ueberpruefen-lassen-4665982.html (Download 21.9.2020).

Graf von Westphalen, Friedrich (2019). „Haftungsfragen beim Einsatz Künstlicher Intelligenz in Ergänzung der Produkthaftungs-RL 85/374/EWG". *Zeitschrift für Wirtschaftsrecht* 19. 889–894.

Hartmann, Felix (2019). „Diskriminierung aus der Black Box – Neue Herausforderungen durch KI-gestützte Personalentscheidungen". *Europäische Zeitschrift für Arbeitsrecht (EuZA)* 2019. 421–422

Hecker, Dirk, Inga Döbel, Ulrike Petersen, André Rauschert, Velina Schmitz und Angelika Voss (2017). *Zukunftsmarkt Künstliche Intelligenz. Potenziale und Anwendungen*. Hrsg. Fraunhofer-Allianz Big Data. St. Augustin. http://publica.fraunhofer.de/eprints/urn_nbn_de_0011-n-4976615.pdf (Download 21.9.2020).

Herberg, Maximilian (2018). „Künstliche Intelligenz und Recht", *Neue Juristische Wochenschrift* (NJW). 39. 2825–2828.

Hustedt, Carla (2019). „Algorithmen-Transparenz. Was steckt hinter dem Buzzword?". 6.5.2019. https://algorithmenethik.de/2019/05/06/algorithmen-transparenz-was-steckt-hinter-dem-buzzword/ (Download 21.9.2020).

Kamarinou, Dimitra, Christopher Millard und Jatinder Singh (2016). „Machine Learning with Personal Data". Queen Mary School of Law Legal Studies Research Paper No. 247/2016. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2865811 (Download 21.9.2020).

Kaulartz, Markus und Tom H. Braeglemann (2020). *Rechtshandbuch Artificial Intelligence und Machine Learning*. München: C.H. Beck in Gemeinschaft mit Vahlen.

Kreutzer, Ralf T. und Marie Sirrenberg (2019). *Künstliche Intelligenz verstehen. Grundlagen – Use Cases – unternehmenseigene KI-Journey*. Wiesbaden: Springer Gabler.

Kroll, Joshua A., Joanna Huey, Solon Barocas, Edward W. Felten, Joel R. Reidenberg, David G. Robinson und Harlan Yu (2017). „Accountable Algorithmus". 165 University of Pensylvania. Law Review, first page 633. https://scholarship.law.upenn.edu/penn_law_review/vol165/iss3/3/ (Download 21.9.2020).

Lapuschkin, Sebastian, Stephan Wäldchen, Alexander Binder, Grégoire Montavon, Wojciech Samek und Klaus-Robert Müller (2019). „Unmasking Clever Hans predictors and assessing what machines really learn". *nature communications* 11.3.2019. https://doi.org/10.1038/s41467-019-08987-4 (Download 15.9.2020).

Martini, Mario, Jonas Botta, David Nink und Michael Kolain (2020). *Automatisch erlaubt? Fünf Anwendungsfälle algorithmischer Systeme auf dem juristischen Prüfstand.* Bertelsmann Stiftung, Gütersloh. https://www.bertelsmann-stiftung.de/fileadmin/files/BSt/Publikationen/GrauePublikationen/Automatisch_erlaubt_final.pdf (Download 21.9.2020).

O'Neil, Cathy (2016). *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. A New York Times Notable Book

Pieper, Fritz-Ulli (2019). „Wenn Maschinen Verträge schließen: Willenserklärungen beim Einsatz von Künstlicher Intelligenz. Praxis im Immaterialgüter- und Wettbewerbsrecht". *GRUR-Prax* 13. 298–300

Platts, Rachel (2019). „Conference Report: ‚Artificial Intelligence: Challenges for Intellectual Property Law'". *The IPKat* 7.11.2019. https://ipkitten.blogspot.com/2019/11/conference-report-artificial.html (Download, 21.9.2020).

Reusch, Philipp (2019). „Künstliche Intelligenz und Produkthaftung". *Kommunikation & Recht (K&R)* 7–8. Beilage. 20–21.

Rich, Emily und Giles Pratt (2018). „AI: who owns the output?". *Freshfields Bruckhaus Deringer* 24.8.2018. https://digital.freshfields.com/post/102f11c/ai-who-owns-the-output (Download 21.9.2020).

Schindler, Stephan (2019). „Künstliche Intelligenz und (Datenschutz-)Recht", *ZD-Aktuell* 10. 06647.

Schwarze, Roland (2019). „Die Zukunft der Betriebsverfassung". *Recht der Arbeit* (72) März / April. 114–118.

Söbbing, Thomas (2019). „Deep Learning: Wenn künstliche Intelligenz lernt, kann das durchaus rechtliche Relevanz haben". *Kommunikation & Recht (K&R) 2019*. 164–169.

von Lewinski, Kai, und Raphael de Barros Fritz (2018). „Arbeitgeberhaftung nach dem AGG infolge des Einsatzes von Algorithmen bei Personalentscheidungen". *Neue Zeitschrift für Arbeitsrecht* 10. 620–624.

Weber, Robert, Alexander Kiefner und Stefan Jobst (2018). „Künstliche Intelligenz und Unternehmensführung". *Neue Zeitschrift für Gesellschaftsrecht* 29. 1131–1136.

Weber, Stefan (2018). „Eine Millionen Bücher mit automatisch erzeugten Texten". *heise.de* 5.11.2018. https://www.heise.de/tp/features/Eine-Million-Buecher-mit-automatisch-erzeugten-Texten-4209972.html (Download 21.9.2020).

Wieder, Clemens (2018). „Datenschutzrechtliche Betroffenenrechte bei der Verarbeitung von personenbezogenen Daten mittels künstlicher Intelligenz". *DSRI-Tagungsband 2018*. 505-518

Zech, Herbert (2019). „Künstliche Intelligenz und Haftungsfragen". *Zeitschrift für die gesamte Privatrechtswissenschaft* 2. 198–219.

Zweig, Katharina, Sarah Fischer und Konrad Lischka (2018). *Wo Maschinen irren können. Fehlerquellen und Verantwortlichkeiten in Prozessen algorithmischer Entscheidungsfindung. Arbeitspapier.* Hrsg. Bertelsmann Stiftung. Gütersloh. https://www.bertelsmann-stiftung.de/fileadmin/files/BSt/Publikationen/GrauePublikationen/WoMaschinenIrrenKoennen.pdf (Download 21.9.2020).

Gesetze und Verträge

AGG – Allgemeines Gleichbehandlungsgesetz: https://www.gesetze-im-internet.de/agg/AGG.pdf

ArbnErfG – Gesetz über Arbeitnehmererfindungen: https://www.gesetze-im-internet.de/arbnerfg/ArbnErfG.pdf

BDSG – Bundesdatenschutzgesetz: https://www.gesetze-im-internet.de/bdsg_2018/BDSG.pdf

BGB – Bürgerliches Gesetzbuch: https://www.gesetze-im-internet.de/bgb/BGB.pdf

BetrSichV – Verordnung über Sicherheit und Gesundheitsschutz bei der Verwendung von Arbeitsmitteln: https://www.gesetze-im-internet.de/betrsichv_2015/BetrSichV.pdf

BetrVG – Betriebsverfassungsgesetz: https://www.gesetze-im-internet.de/betrvg/BetrVG.pdf

DSGVO – Datenschutz-Grundverordnung: https://eur-lex.europa.eu/legal-content/DE/TXT/HTML/?uri=CELEX:32016R0679

GeschGehG – Gesetz zum Schutz von Geschäftsgeheimnissen: https://www.gesetze-im-internet.de/geschgehg/GeschGehG.pdf

Internationales Abkommen über den Schutz der ausübenden Künstler, der Hersteller von Tonträgern und der Sendeunternehmen vom 26.10.1961 (Rom-Abkommen): https://www.admin.ch/opc/de/classified-compilation/19610224/201204230000/0.231.171.pdf

KSchG – Kündigungsschutzgesetz: https://www.gesetze-im-internet.de/kschg/BJNR004990951.html

PatentG – Patentgesetz: https://www.gesetze-im-internet.de/patg/PatG.pdf

ProdHaftG – Gesetz über die Haftung für fehlerhafte Produkte: https://www.gesetze-im-internet.de/prodhaftg/ProdHaftG.pdf

ProdSG – Gesetz über die Bereitstellung von Produkten auf dem Markt (Produktsicherheitsgesetz): https://www.gesetze-im-internet.de/prodsg_2011/ProdSG.pdf

UrhG – Gesetz über Urheberrecht und verwandte Schutzrechte: https://www.gesetze-im-internet.de/urhg/UrhG.pdf

Abbildungen und Tabelle

Autoren

Dr. iur Till Kreutzer ist Rechtsanwalt, Rechtswissenschaftler und Publizist. Er ist Mitgründer und geschäftsführender Partner der Rechtsanwaltskanzlei iRights.Law sowie Mitgründer und Herausgeber von iRights.info, dem mehrfach prämierten (u. a. Grimme-Online-Award 2006) Internetportal für Verbraucher und Kreative zum Urheberrecht in der digitalen Welt.

Till Kreutzer ist Mitglied im Fachausschuss „Kommunikation und Information" der Deutschen UNESCO Kommission (DUK). Er ist assoziiertes Mitglied des Leibniz-Instituts für Medienforschung (Hans-Bredow-Institut) und Mitglied des „Instituts für Rechtsfragen der Freien und Open Source Software" (ifrOSS). Er ist zudem Mitglied des Fachausschusses Urheber- und Medienrecht der GRUR sowie Repräsentant Deutschlands im Creative Commons Global Network Council. Bei den Urheberrechtsreformen in der Informationsgesellschaft war und ist Till Kreutzer vielfach auf nationaler sowie EU-Ebene als geladener Sachverständiger für Regierungen und Parlamente tätig.

Prof. Dr. Per Christiansen, M.Sc. (LSE) ist Rechtsanwalt und Hochschullehrer für Wirtschaftsrecht an der FOM Hochschule für Oekonomie & Management in Hamburg.

Er ist seit über 20 Jahren in verschiedenen Funktionen in der Internet-Branche tätig. Vor seiner Tätigkeit als Hochschullehrer arbeitete er 11 Jahre für die AOL-Gruppe in Deutschland, zuletzt als Chef-Justiziar und Personalleiter. Für die Dauer des Unternehmenszusammenschlusses von AOL und Time Warner war er in einem konzernweiten Legal Team mit urheber- und lizenzrechtlichen Fragestellungen sowie mit den hierin begründeten Konflikten der Geschäftsmodelle der einzelnen Medienkanäle befasst. Als Partner der Kanzlei iRights.Law betreut er in- und ausländische Unternehmen der Internet-Branche bei der Gestaltung von Internet-Produkten.

Impressum

Bertelsmann Stiftung
Carl-Bertelsmann-Straße 256
33311 Gütersloh
www.bertelsmann-stiftung.de

Verantwortlich:
Dr. Ole Wintermann

Konzept:
Dr. Ole Wintermann, Birgit Wintermann

Redaktion:
Dorothee Kubitza, Birgit Wintermann

Autoren:
Dr. Till Kreutzer, Prof. Dr. Per Christiansen

Lektorat:
Rudolf Gajdacz

Gestaltung:
Dietlind Ehlers

Februar 2021

Kontakt

Bertelsmann Stiftung
Carl-Bertelsmann-Straße 256
33311 Gütersloh

Birgit Wintermann
Programm Unternehmen in der Gesellschaft
Telefon 05241-8181289
birgit.wintermann@bertelsmann-stiftung.de

www.bertelsmann-stiftung.de

iRIGHTS law

BertelsmannStiftung